

Research Article

Bridge Displacement Measurement Using the GAN-Network-Based Spot Removal Algorithm and the SR-Based Coarse-to-Fine Target Location Method

Shanshan Yu ¹ and Jian Zhang ^{2,3}

¹School of Civil Engineering and Architecture, Jiangsu University of Science and Technology, Zhenjiang 212100, China

²School of Civil Engineering, Southeast University, Nanjing 210096, China

³Jiangsu Key Laboratory of Engineering Mechanics, Southeast University, Nanjing 210096, China

Correspondence should be addressed to Jian Zhang; jian@seu.edu.cn

Received 2 March 2023; Revised 6 August 2023; Accepted 12 August 2023; Published 4 September 2023

Academic Editor: Lucia Faravelli

Copyright © 2023 Shanshan Yu and Jian Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Image-based bridge displacement measurement still suffers from certain limitations in outdoor implementation. Each of these limitations was addressed in this study. (1) The laser spot is difficult to identify visually during the object distance (OD: mm) measurement using a laser rangefinder, which makes the scale factor (SF: mm/pixel) calibration tricky. To overcome this issue, a stereovision-based full-field OD measurement method using only one camera was suggested. (2) Sunlight reflected by the water surface during the measurement causes light spot interference on the captured images, which is not conducive to target tracking. A network for light spot removal based on a generative adversarial network (GAN) is designed. To obtain a better image restoration effect, the edge prior was novelly designed as the input of a shadow mask-based semantic-aware network (S^2 Net). (3) A coarse-to-fine matching strategy combined with image sparse representation (SR) was developed to balance the subpixel location precision and efficiency. The effectiveness of the above innovations was verified through algorithm evaluation. Finally, the integrated method was applied to the vibration response monitoring of a concrete bridge impacted by the traffic load. The image-based measurement results show good agreement with those of the long-gauge fiber Bragg grating sensors and lower noise than that of the method before improvement.

1. Introduction

Calibration of the scale factor and the target location are two critical steps in image-based bridge displacement measurement technology [1–3]. However, the convenience and reliability of traditional techniques used to perform these two steps need to be improved.

The scale factor (SF) is used to convert image displacement (pixel) into real displacement (mm) [4]. As the object distance (OD)-based method does not need a reference object having a known size near the measured target, it is widely used. For this method, two parameters, OD and the camera angle, are measured. At present, techniques to correct the effects of camera angles have been well studied [4, 5]. However, OD measurement relies on the laser

rangefinder, and the related challenges have not been sufficiently researched. On the one hand, target sections are difficult to identify directly in the field of view (FOV), especially when the bridge bottom is curved. On the other hand, the indicator point of a laser rangefinder is difficult for human eyes to capture, especially during the daytime. In addition, laser ranging can only obtain the OD of sparse points. Consequently, displacement conversion cannot be carried out if the measured distance does not correspond to the real targets. To avoid the above issues, visual ranging has been widely used. However, the working distance of an integrated depth camera is too small to be applied to the bridge structures considered in this study. Given the above analysis, a stereovision-based full-field OD measurement method was proposed. To improve the practicability in the

field measurement, only one camera and two monopods are used in fact, which is different from the traditional binocular stereovision method. Scale conversion uses the actual distance between two shooting positions because it is easier to achieve than placing a ruler in the field of view.

Furthermore, the images of a bridge bottom for displacement calculation are often degraded by light spots reflected from the water surface, reducing the matching reliability of the region of interest (ROI). Therefore, it is necessary to first restore the image. The light spot problem encountered in this task is similar to the shadow phenomenon [6–8] in text image processing. In addition to traditional image optimization methods [9–12], deep learning (DL) has received immense attention in the past few years. As one of the most popular models in deep learning, generative adversarial network (GAN)-based models [13, 14] have been widely used in image restoration. Recently, a shadow mask-based semantic-aware network (S²Net) [15] was proposed and showed a better restoration effect [16, 17]. For instance, nonshadow regions were kept unchanged when filtering shadows. More importantly, much attention was paid to artifacts around the shadow edge. Accordingly, S²Net was also used for spot removal in this study. However, to achieve equilibrium in the training process of GAN, although the quality of the restored image is significantly improved, the gradient is smooth, and detail clarity is lost. Inspired by the idea in DeepSemanticfaceNet [18], it is proposed to take the edge image of the original image as input and redesign the loss function.

Using the restored images, both the accuracy and speed of the target location need to be addressed. Widely used template-matching (TM) techniques [19, 20] cannot meet real-time requirements and will fail when features on the structure surface are sparse. Feature-based methods [21, 22] have fewer demands regarding the texture and take less time. By contrast, the sparse representation (SR)-based target tracking method [23] shows higher efficiency and robustness when image quality is affected. However, when the scene is complex or the real-time map is very small, there are usually many similar areas. In such cases, the mapping position of nonzero items is very scattered, which can easily lead to inaccurate positioning. To mitigate this issue, a distance-weighted sparse representation algorithm [24] was developed. In addition, the larger the dictionary size, the faster the sparse representation and the lower the positioning accuracy. Therefore, a coarse-to-fine matching strategy was designed to guarantee both speed and accuracy.

The remainder of this paper is structured as follows: The research framework and the algorithmic innovations, which cover three blocks for bridge displacement measurements, are introduced in Section 2. The effectiveness of the improved algorithm is evaluated in Section 3. Then, the engineering application to a concrete bridge is introduced in Section 4. Finally, this research work is summarized in Section 5.

2. Proposed Method

A convenient and robust camera-based method for bridge displacement measurement was proposed. The framework is shown in Figure 1. First, the distance from the target to the camera needs to be measured; then, SF is determined [4] and used to convert the pixel displacement into physical displacement. Then, image sequences are collected before and after deformation, and the target region is continuously positioned in these sequences to get the pixel displacement. As the effect of the light spot is considered, image preprocessing is required. In addition, to balance positioning speed and accuracy, a coarse-to-fine matching method based on sparse image representation was proposed. As the camera can continuously capture the image of the whole bridge, the displacements at multiple control sections of the whole bridge can be extracted synchronously. Object distance measurement, image preprocessing, and target location are the three key steps that affect the accuracy of measurement results. Corresponding innovative works were carried out, and the principles were introduced as follows.

2.1. Principle of Stereovision-Based Ranging Using Only One Camera. Conventional binocular stereovision requires two fixed cameras. However, to minimize hardware costs, only one camera was used in this study. As shown in Figure 2(a), the internal parameter matrix \mathbf{K} of the camera is kept constant and can be calibrated in advance [25]. After focusing, the lens was locked, and then, two images of the same scene were shot at different stations. The coordinates of the matching point pair are obtained based on the feature point-matching algorithm SURF-BRISK [26]. Here, l/r in the parameter's subscript indicates that the parameter is for the left/right camera. The projection equation corresponding to two images can be written as

$$\begin{cases} s_l p_l = \mathbf{M}_l X_w = (\mathbf{M}_{l1} & m_l) X_w, \\ s_r p_r = \mathbf{M}_r X_w = (\mathbf{M}_{r1} & m_r) X_w, \end{cases} \quad (1)$$

where $X_w = (x_w \ y_w \ z_w)^T$ is the world coordinate of the target point, s_l and s_r are the scale parameters of the left and right cameras, p_l and p_r are the coordinates of the projection points in the left and right images, and \mathbf{M}_l and \mathbf{M}_r are the projection matrices of two cameras.

Let the 3×3 part of the projection matrix coordinates be \mathbf{M}_{l1} and \mathbf{M}_{r1} , and the remaining 3×1 part be m_l and m_r . Setting $X = (x_w \ y_w \ z_w)$, equation (1) can be rewritten as

$$\begin{cases} s_l p_l = \mathbf{M}_{l1} X + m_l, \\ s_r p_r = \mathbf{M}_{r1} X + m_r. \end{cases} \quad (2)$$

By eliminating X from the above equation,

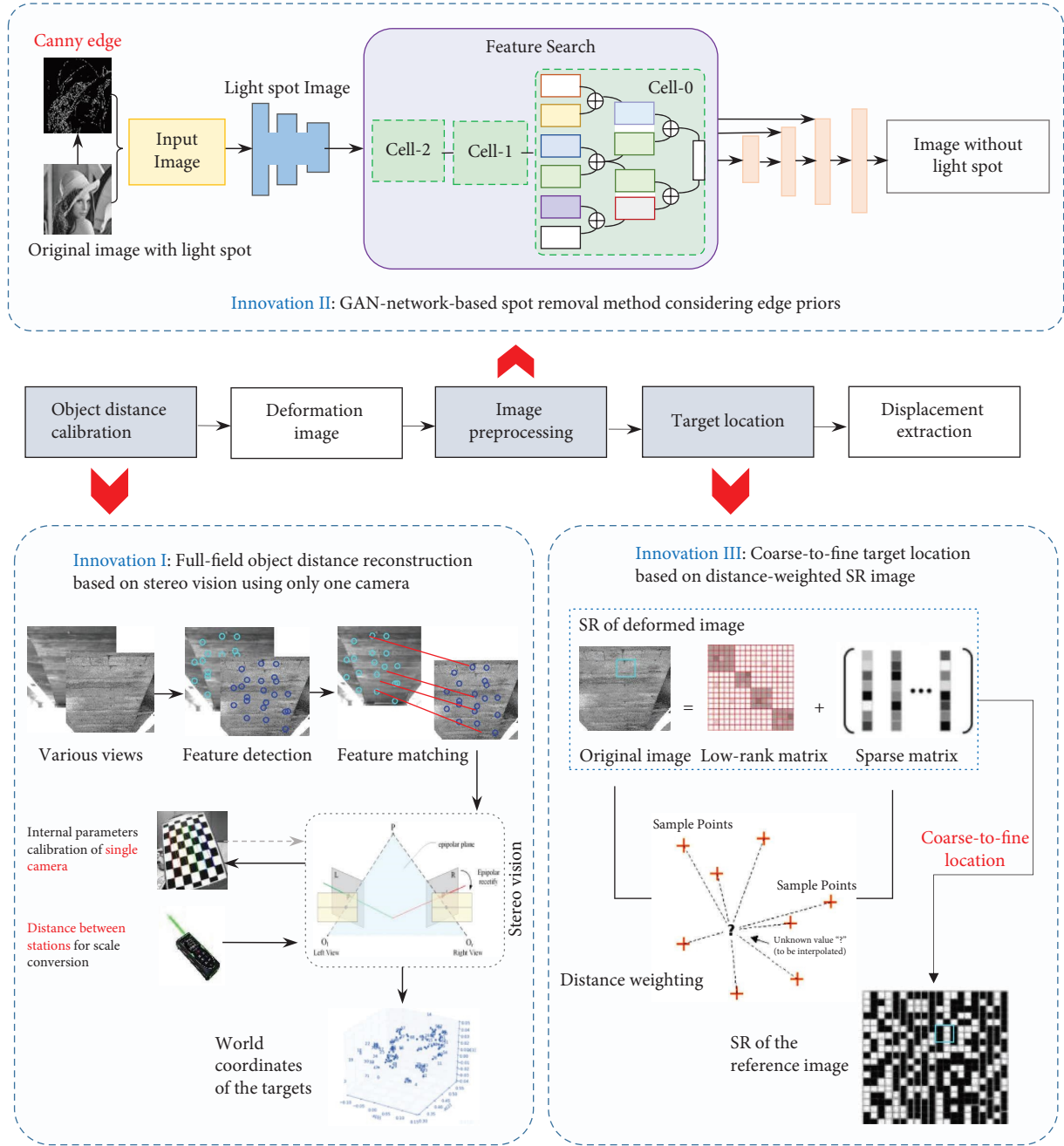


FIGURE 1: Overview of the proposed bridge displacement measurement method.

$$s_r p_r - s_l M_{r1} M_{l1}^{-1} p_l = m_r - M_{r1} M_{l1}^{-1} m_l. \quad (3)$$

As both sides of the above equation are three-dimensional vectors, s_l and s_r can be eliminated to obtain the relationship between p_l and p_r , which are just the polar constraint. Let the left side of equation (3) be m , and then, equation (3) can be transformed into

$$m = m_r - M_{r1} M_{l1}^{-1} m_l. \quad (4)$$

The antisymmetric matrix of m is denoted as $[m] \times$. As $m \times [m] \times = 0$, equation (5) can be derived:

$$[m] \times (s_r p_r - s_l M_{r1} M_{l1}^{-1} p_l) = 0. \quad (5)$$

Multiplying p_r^T by the left side, $s_r p_r^T [m] \times p_r - s_l p_r^T M_{r1} M_{l1}^{-1} p_l = 0$ can be obtained. As $[m] \times p_r$ is perpendicular to p_r^T , $s_r p_r^T [m] \times p_r = 0$. Then, the linear relationship between p_l and p_r can be described as

$$p_r^T [m] \times M_{r1} M_{l1}^{-1} p_l = 0. \quad (6)$$

Let $F = [m] \times M_{r1} M_{l1}^{-1}$; then,

$$p_r^T F p_l = 0, \quad (7)$$

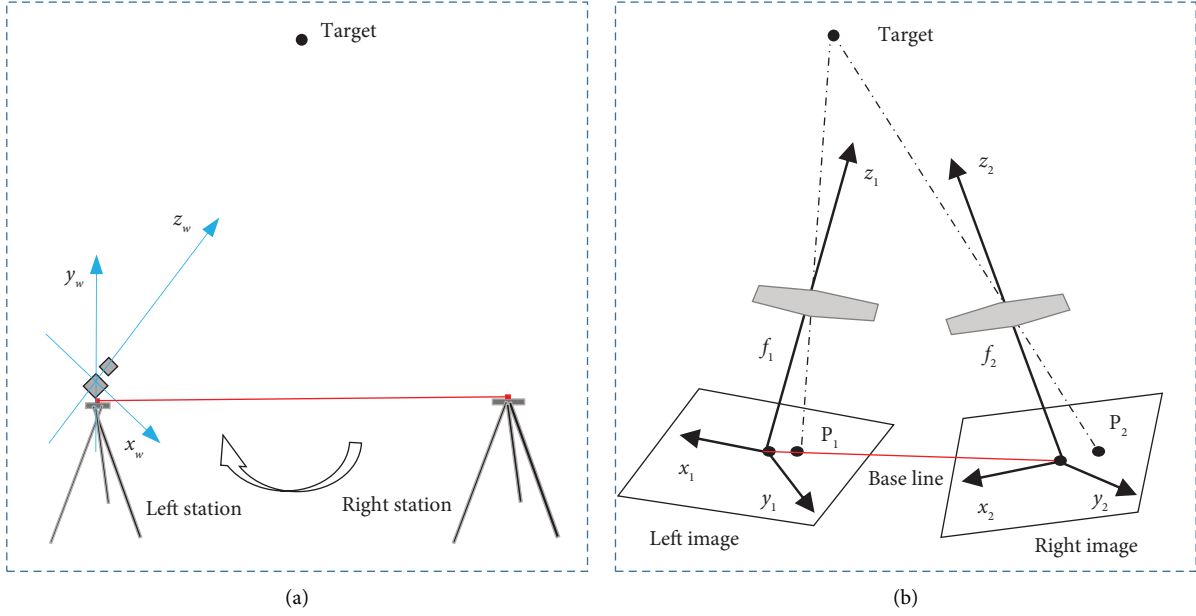


FIGURE 2: Stereovision-based object distance calibration: (a) field operation diagram; (b) stereovision model.

where \mathbf{F} is the fundamental matrix. The eight-point method was adopted to solve \mathbf{F} . The intrinsic matrix \mathbf{E} reflects the relation of the space points in different coordinate systems (equation (8)) which can be obtained through \mathbf{K} and \mathbf{F} (equation (9)):

$$p_r^T \mathbf{E} p_l = 0, \quad (8)$$

$$\mathbf{E} = \mathbf{K}_r^T \mathbf{F} \mathbf{K}_l, \quad (9)$$

where \mathbf{E} is only related to the camera motion, and in the sense of a nonzero factor difference, it can be expressed as

$$\mathbf{E} = \mathbf{R} [\mathbf{R}^T \quad \mathbf{t}]. \quad (10)$$

The rotation matrix \mathbf{R} and the translation vector $\mathbf{t} = (t_x, t_y, t_z)$ can be recovered by performing SVD decomposition of \mathbf{E} , and then, the two camera matrices \mathbf{M}_l and \mathbf{M}_r are obtained. Finally, the spatial coordinates are reconstructed through the triangulation method using the projection matrix. However, due to the lack of real-scale information, the three-dimensional result here is dimensionless. The conventional method is to place a measuring scale of known length in the field of view. But this is suitable for the test scenario focused in this study. Therefore, it is proposed to realize the scale conversion by the distance between two shooting stations, which is easier to measure. The specific instructions are as follows.

According to the principle of relative orientation, the distance T (mm) between the optical centers (Figure 2(b)) is

$$T = \sqrt{T_x^2 + T_y^2 + T_z^2}, \quad (11)$$

where T_x , T_y , and T_z are the components of T in three directions. The relationship between them and the elements of \mathbf{t} is $T_x : T_y : T_z = t_x : t_y : t_z$. As T can be measured in

advance, then the relative orientation relationship and 3D coordinates with a real scale will be obtained. z_w is just OD used for calculating the displacement conversion factor [4].

2.2. Proposed GAN-Based Light Spot Removal Method considering Edge Priors. The architecture of the GAN-based light spot removal method is shown in Figure 3. It consists of two parts, the first network is for light detection and the second is for light spot removal. The purpose of multitask learning is achieved by taking the output of the first network as the input of the second network. To impress the interference of the light spot on target positioning, an end-to-end spot mask-based semantic-aware network (S²Net) [15] for light spot removal was adopted in this study. Using the guidance of semantic prior from the spot masks, spot-mask-based semantic transformation (SST) transfers statistical information from nonspot features to spot features, and the nonspot features were kept intact.

As the importance of edge information to image restoration has been proved [27], the author proposed to use the edge image of the degraded image as additional network input for deep training. The Canny() function of OpenCV in Python was called for edge extraction to construct a new dataset, called GOPRO Gauss edge. The restored clear image can then be expressed as

$$L = B + G(B, E(B)), \quad (12)$$

where B is the degraded image and $E(B)$ is the edge image of B . Through the edge constraint, more prior information is provided, and more attention is paid to the key features of the structure. The image gradient will be clearer and more detailed. The loss function of the constructed network consists of four parts.

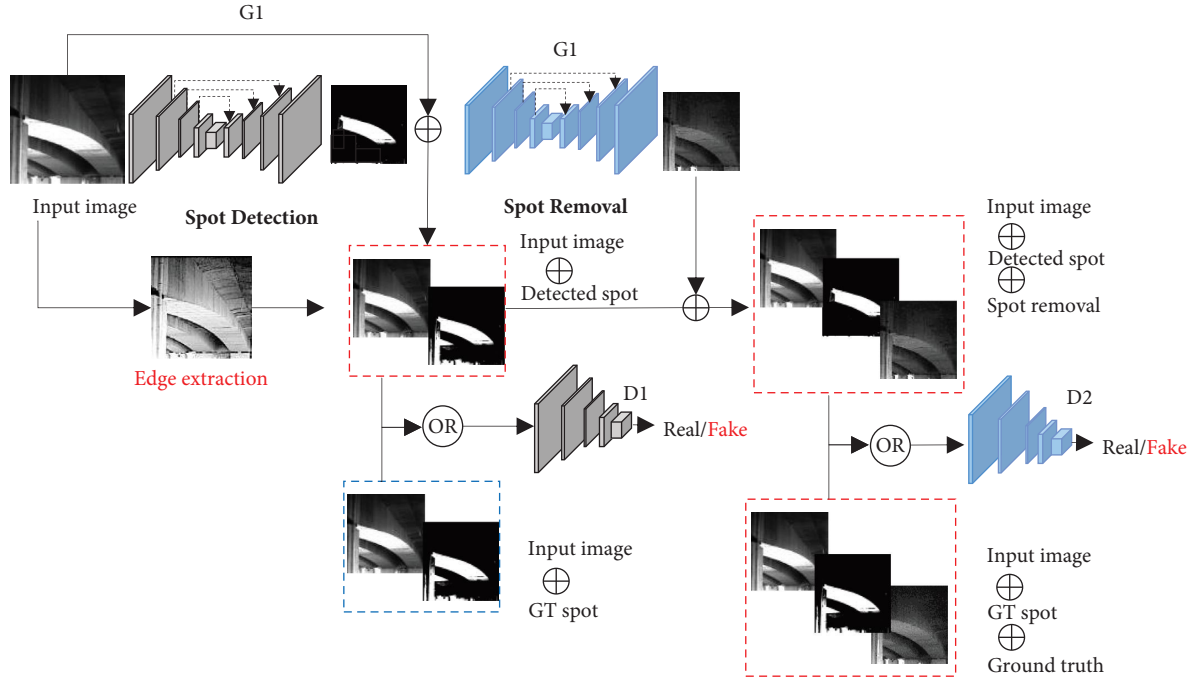


FIGURE 3: Light spot removal method considering edge priors.

2.2.1. *Edge Loss* \mathcal{L}_{edg} . With the introduction of $E(B)$ in the training stage, to ensure that $E(B)$ can be adequately considered and to avoid the possible ringing phenomenon, the edge loss was designed as

$$\mathcal{L}_{\text{edg}} = \sum_{i=1}^{i=k} \|E(B) \odot (L_i - S_i)\|_1, \quad (13)$$

where \odot represents the dot product. S_i , L_i , and B are the input clear image, the restored clear image, and the blurred image corresponding to the input clear image, respectively. The generator based on GAN is shown in Figure 3.

2.2.2. *Scale Loss* \mathcal{L}_{sca} . Three scales (256×256 , 128×128 , and 64×64) were set in this network, and the low-scale output was fused to the larger scale by upsampling. The loss of each scale is \mathcal{L}_1 actually, corresponding to the pixel difference between the restored clear image L and the real clear image S :

$$\mathcal{L}_{\text{sca}} = \sum_{i=1}^{i=k} \frac{1}{c_i w_i h_i} \|L_i - S_i\|_1, \quad (14)$$

where k represents multiple scales and $k=3$ in this method and c_i , w_i , and h_i are the channel number, width, and height of the input image, respectively.

2.2.3. *Adversarial Loss* \mathcal{L}_{adv} . To solve the problem that the GAN network training does not easily converge, deep convolution is used to generate the optimized formula of the adversarial network DCGAN [28]. The adversarial loss is only applied to the output of the last scale:

$$\mathcal{L}_{\text{adv}} = -\log D(G(B, E(B))), \quad (15)$$

where G stands for the generator and D represents the discriminator.

Furthermore, because of the existence of the skip connection, \mathcal{L}_{adv} can be written as

$$\mathcal{L}_{\text{adv}} = -\log D(L). \quad (16)$$

2.2.4. *Perceptual Loss* \mathcal{L}_{per} . Perceptual loss [29] helps enhance the detail of the image and is also used only for the last output:

$$\mathcal{L}_{\text{per}} = \sum \|\phi_1(L) - \phi_1(S)\|_2^2, \quad (17)$$

where $\phi_i(x)$ is the first layer of a characteristic pattern. The perceptual loss layer used in this study is the same as that in the study by Shen et al. [27]. The perceptual loss was then calculated through the pooling layers, i.e., Pool2 and Pool5.

To sum up, the loss function of DeepEdgeGAN is

$$\mathcal{L} = \mathcal{L}_{\text{sca}} + \lambda_{\text{adv}} \mathcal{L}_{\text{adv}} + \lambda_{\text{edg}} \mathcal{L}_{\text{edg}} + \lambda_{\text{per}} \mathcal{L}_{\text{per}}, \quad (18)$$

where the weight values are $\lambda_{\text{adv}} = 1 \times 10^{-4}$, $\lambda_{\text{edg}} = 10$, and $\lambda_{\text{per}} = 5 \times 10^{-5}$, respectively [27].

GAN-based methods are prone to generate artifacts, and optimization is underconstrained due to unstable data acquisition. Since large-scale high-quality datasets [15] are publicly available, we adopt the strategy of training on paired data. Large and diverse training datasets can give the trained model better generalizability. The input image is resized to 256×256 , and the minibatch size is set to 8 for training. The initial learning rate is 0.0001. The learning rate is reduced by

the ‘‘poly’’ policy [30], with a power of 0.9. We trained 600 epochs for each network.

2.3. Coarse-to-Fine Location Algorithm Based on Sparse Representation. At present, image restoration mainly focuses on target detection reliability, while positioning accuracy is not given enough consideration, especially the displacement measurement to be realized in this study. In addition, existing image matching still suffers from issues of low matching efficiency, high time complexity, and computation intensiveness. Combining with SR, a distance-weighted image-matching method from coarse-to-fine was proposed.

For an image x , it can be reexpressed as

$$x = D\alpha. \text{ s.t. } \|\alpha\|_0 \leq k, \quad (19)$$

where $D = \{d_1, d_2, \dots, d_n\} \in \mathbb{R}_{m \times n}$ is the dictionary and $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ is the sparse vector matrix. Referring to the conclusion in the research by Donoho [31], the solution of equation (19) is

$$\hat{\alpha} = \arg \min_{\alpha} \|D\alpha - X\|_2^2 + \lambda \|\alpha\|_1, \quad (20)$$

where $\lambda \geq 0$. For image matching, after obtaining the dictionary D of a real-time image x , the sparse vector α can be obtained using equation (20). *arg min* stands for the argument of the minimum. The atom of D corresponding to $\max(\alpha)$ represents the position pp (\hat{m}, \hat{n}) of the real-time reference x in S . $\max(\alpha)$ is the largest element of α . This can be expressed using the following formula:

$$\begin{aligned} pp &= (\hat{m}, \hat{n}) \\ &= \text{map}(\max(\alpha)), \end{aligned} \quad (21)$$

where $\text{map}()$ is the position identity mapping function. The positioning error is

$$\begin{aligned} PD &= |pt - pp| \\ &= |m - \hat{m}| + |n - \hat{n}|, \end{aligned} \quad (22)$$

where $pt(m, n)$ is the real pixel location of x in S .

It is preferable to use as many identical atoms as possible to express the real-time graph. Considering the spatial location constraint between the real image and the reference image, a distance constraint operator ω was introduced to ensure that atoms near the real location were given more emphasis during the sparse expression of real-time images. Then, the nonzero term of the sparse vector α is constrained to be near the location pp as much as possible, that is, to ensure that similar candidate positioning regions have similar coefficient values.

Based on the above analysis, a sparse representation algorithm based on distance weighting was developed. Then, the solution of equation (19) becomes

$$\hat{\alpha} = \arg \min_{\alpha} \|D\alpha - x\|_2^2 + \lambda \|\omega \circ \alpha\|_1, \quad (23)$$

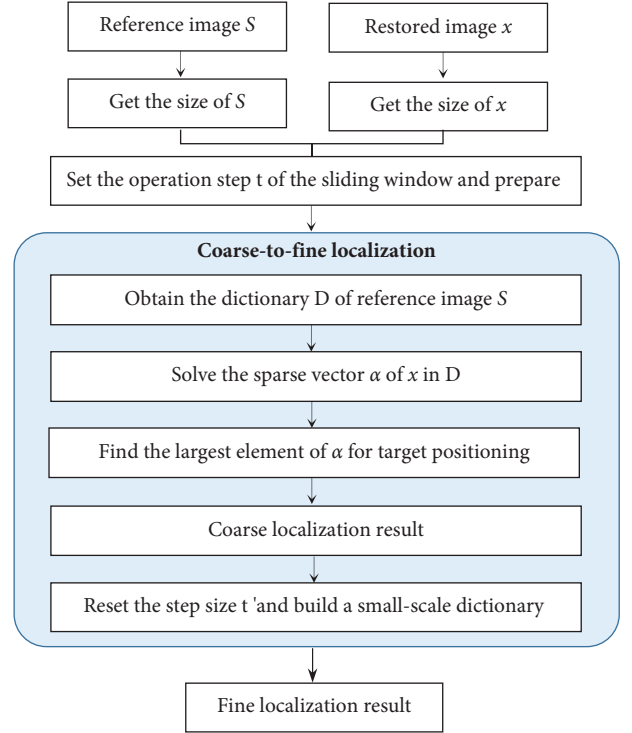


FIGURE 4: Flowchart of the proposed target localization method based on the weighted SR and coarse-to-fine matching strategy.

where ω is the distance constraint operator, representing the Euclidean distance between x and D , ‘‘ \circ ’’ denotes the operation at the pixel level, and $\lambda \|\omega \circ \alpha\|_1$ is the distance constraint of SR. With the distance constraint ω , the sparse expression is further sparse, while ensuring that similar atoms have approximate sparse coefficient values.

A small length t (or the size of the sliding window, unit: pixel) can get higher localization accuracy but lower matching efficiency. Then, a coarse-to-fine matching strategy was designed, as shown in Figure 4. First, an initial step length t_1 was set to construct a dictionary for coarse matching. Then, a smaller step length, $t_2 = 1$ pixel, was set to construct a new dictionary for fine localization.

3. Algorithm Tests for Image Restoration and Localization

The images of real bridges are used for algorithm tests through two criteria.

3.1. Restoration Effects. In this study, the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [32] were set to evaluate restoration effects. Then, 500 images and corresponding edge images were input into the proposed restoration model considering the edge prior. The values of PSNR and SSIM were then calculated.

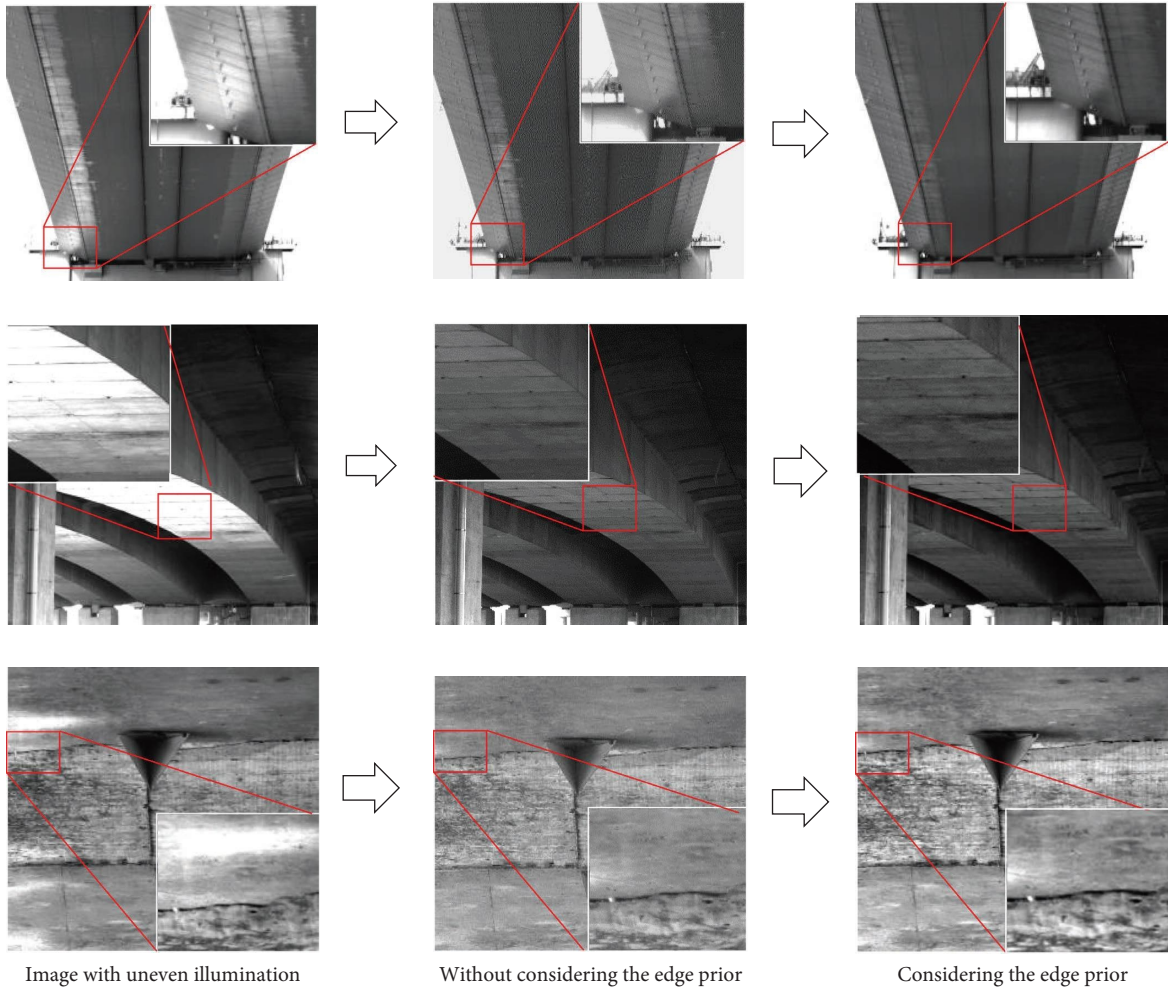


FIGURE 5: Comparison of image restoration results.

For comparison, the above operations were also performed in the training model without considering the edge prior.

From Figure 5, it is obvious that the restored image considering the edge has a higher image quality and clearer details. As presented in Table 1, it can be seen that the restoration effect is further improved after the introduction of the edge prior. This shows that the introduction of the edge prior plays a major role in promoting the restoration effect. This is because edge information provides the network with a clearer optimization direction in the learning process and a better constraint on the image structure in the restoration process.

3.2. Matching Accuracy. The original images disturbed by light spots were used as reference images, whose size was 2048×2048 . The real-time images were from the restoration images, and the size was 50×50 . A total of 100 small images were randomly selected from each restored large image as real-time images. The average matching localization results of these 300 small images were counted as the final result. The matching methods include template matching (TM), the original method, subsequently improved methods, and the

TABLE 1: Comparison of image restoration results considering edge prior and without edge prior.

Methods	PSNR (DB)	SSIM
[15]	29.5274	0.8946
Ours	33.0936	0.9335

proposed method. Their implementation details are shown in Table 2. The red word indicates that this step takes an improved approach. For the proposed method, the step sizes are $t_1 = 5$ and $t_2 = 1$. The positioning error was calculated through equation (22).

In this study, the average of the positioning error was used as the matching accuracy, and its unit is a percentage (%), which represents the proportion of the real-time image meeting the positioning error in the total test images under the current setting of PD. For example, the positioning accuracy of TM is 43.27 when $PD \leq 1$, indicating that 43.27% of 300 real-time images have pixel error less than or equal to 1 pixel.

Two conclusions can be drawn from Table 2. First, In contrast to TM, the matching accuracy of the original method has been noticeably improved. In the condition of $PD \leq 5$, positioning accuracy is improved by 56.66%. Second,

TABLE 2: Localization result of different methods.

Methods	PD ≤ 0.1	PD ≤ 1	PD ≤ 2	PD ≤ 5
Template matching	27.83	43.27	35.00	35.17
Original method				
[15] + parse expression + coarse matching	23.83	52.17	87.33	91.83
"[15] + edge prior" + sparse expression + coarse matching	24.07	53.91	89.85	92.96
[15] + weighted sparse expression + coarse matching	23.65	53.52	87.98	91.84
[15] + sparse expression + coarse-to-fine matching	70.44	86.27	90.16	92.58
Proposed method				
"[15] + edge prior" + weighted sparse expression + coarse-to-fine matching	76.79	90.46	92.83	94.67

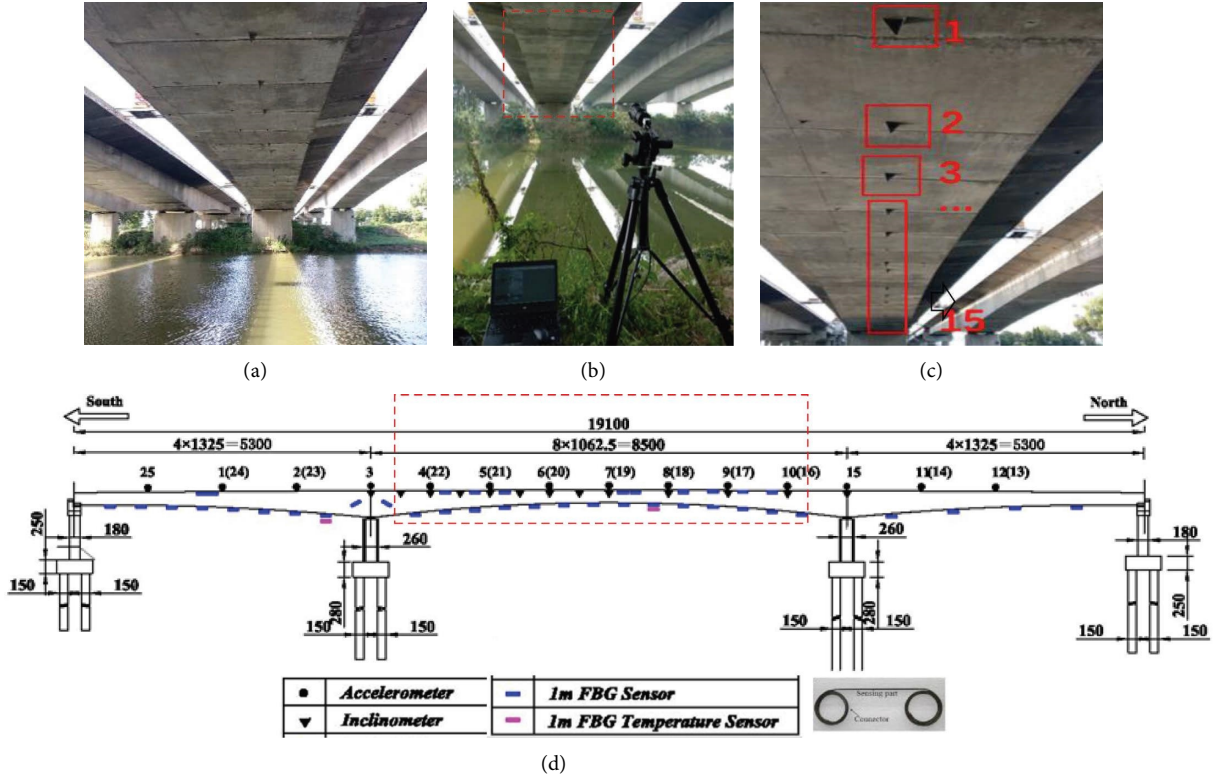


FIGURE 6: Survey of real bridge displacement measurement based on the image and FBG: (a) real bridge to be measured; (b) system layout; (c) captured image; (d) layout of FBG.

comparing a series of optimized methods with the original method, it is found that with the introduction of the edge information and distance weighting algorithm, positioning accuracy can be improved to some extent. However, it is still not enough for the subpixel measurement requirement. The coarse-to-fine search strategy makes the most significant contribution to subpixel positioning accuracy. With the improved method, matching accuracy was substantially improved.

4. Real Bridge Application

As shown in Figure 6, the tested structure is a concrete bridge with three consecutive spans. The middle span with a length of 85 meters is the monitoring object. The contact-type deformation sensor, long-gauge fiber Bragg (FBG), has

been installed on the bridge, which can be used as a reference. FBG can sense the minor change in external physical quantities through the change in the wavelength of light. Deflection is inversely related through the conjugate beam method, and its principle is shown in Figure 7(b). The positive strain on the beam surface at section x is $\varepsilon(x)$ when deflection is $y(x)$. The angular displacement is $\theta(x)$. The height of the neutral axis is h_m . Q is the shearing force, and M is the bending moment, respectively. When the boundary ($x=0$) conditions of the virtual beam satisfy $\bar{Q}_0 = \theta_0$ and $\bar{M}_0 = y_0$, the angular displacement distribution $\bar{\theta}(x)$ of the real beam is equal to the shear distribution $\bar{Q}(x)$ of the virtual beam and the deflection distribution $y(x)$ of the real beam is equal to the moment distribution $\bar{M}(x)$ of the virtual beam. This can be expressed by the following equation:

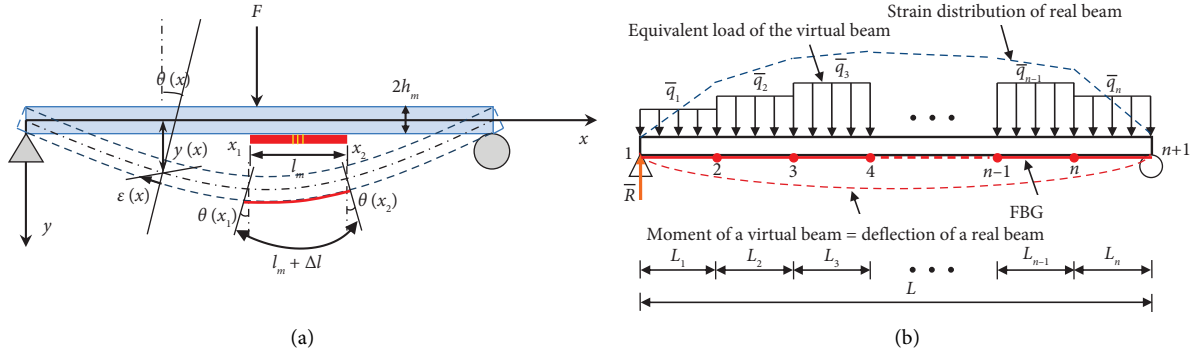


FIGURE 7: Deflection inversion principle of the bridge based on long labeled FBG. (a) Schematic diagram of bending deformation of Euler beam. (b) Schematic diagram of the conjugate beam method for calculating beam deflection by long-gauge strain.

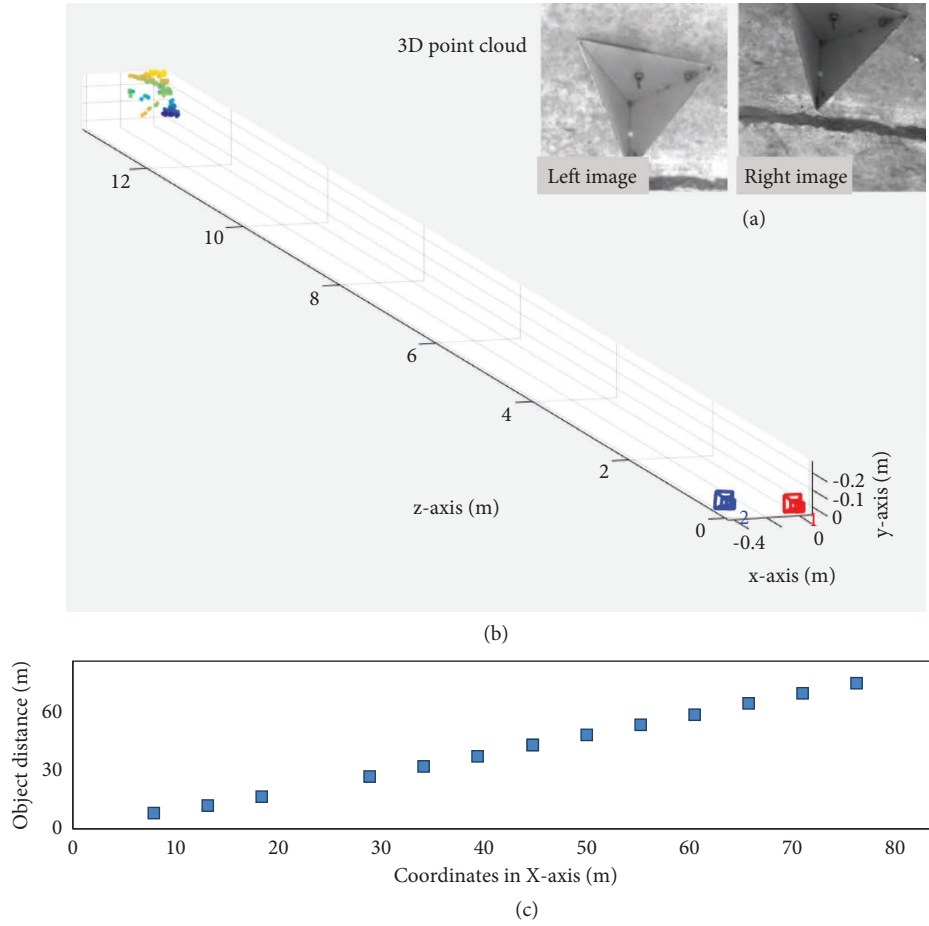


FIGURE 8: Object distance calibration results based on the stereovision method. (a) Images of two views. (b) 3D point cloud of one target. (c) Calibrated object distance and corresponding coordinate in X-axis of different targets.

$$\begin{aligned}
 \bar{Q}(x) &= \int_0^x \bar{q}(x) dx + \bar{Q}_0, \bar{M}(x) \\
 &= \int_0^x \left(\int_0^x \bar{q}(x) dx \right) dx + \bar{Q}_0 x + \bar{M}_0, \\
 \theta(x) &= \int_0^x \bar{q}(x) dx + \theta_0, y(x) \\
 &= \int_0^x \left(\int_0^x \bar{q}(x) dx \right) dx + \theta_0 x + y_0.
 \end{aligned} \tag{24}$$

Here, $\bar{q}(x) = -\varepsilon(x)/h_m$ “-” indicates that the parameter belongs to the virtual beam. Based on the long-scale sensor, the deflection distribution is not dependent on the load and stiffness distribution of the beam; rather, it only has an explicit linear relationship with the measured long-scale strain distribution. In addition, \bar{q}_m in the equation accounts for the change in neutral axis height, so the method can also be applied to beams with variable cross-sections.

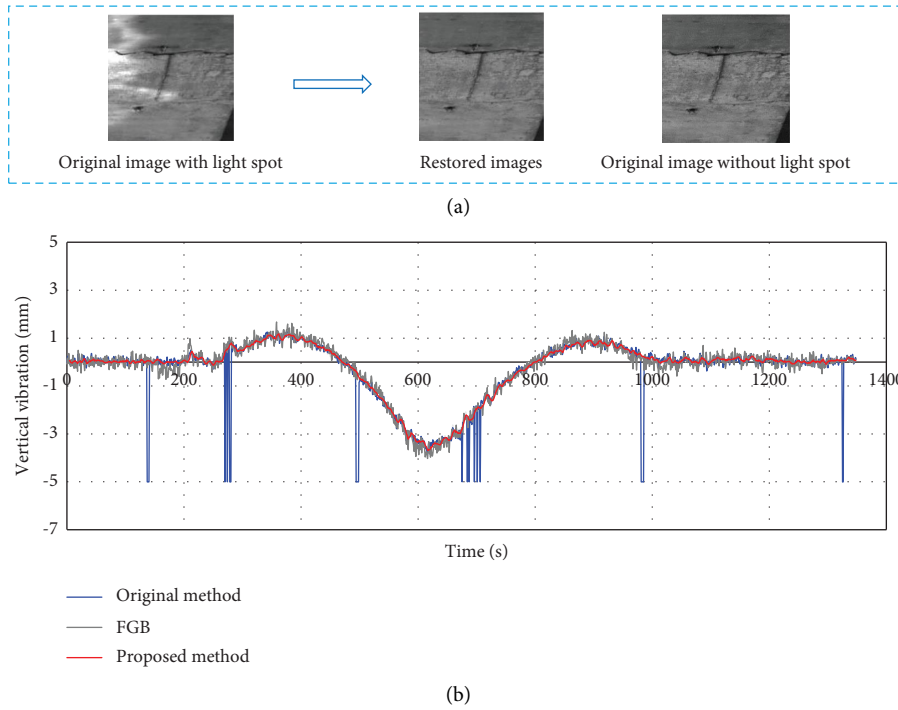


FIGURE 9: Effect of image restoration and its improvement in the displacement accuracy: (a) display of the restoration effect; (b) comparison of the measured displacement using different methods.

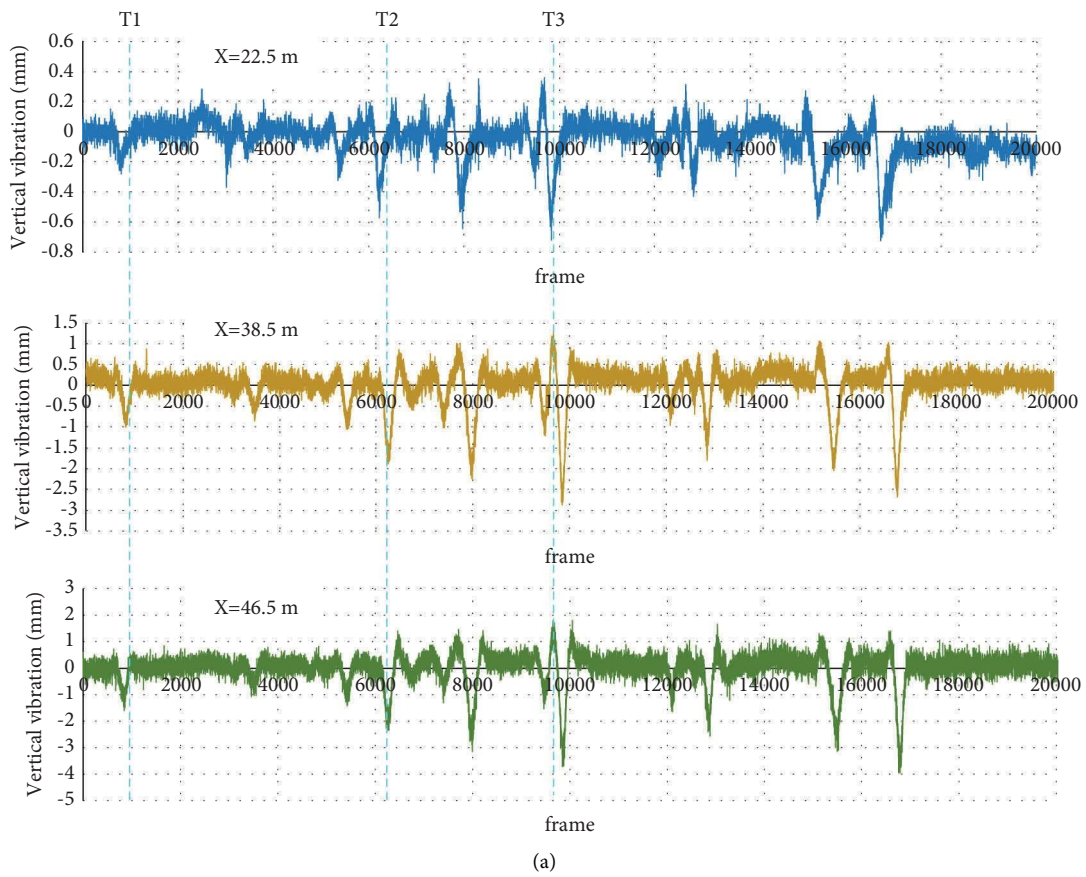


FIGURE 10: Continued.

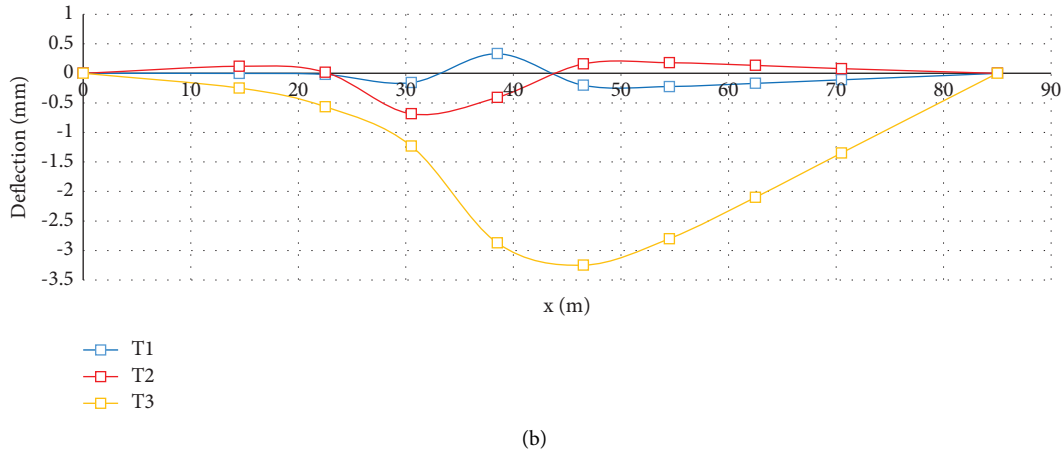


FIGURE 10: Displacement measurement results under normal traffic flow. (a) Deflection time history of the representative section. (b) Deflection line of the bridge at three different times (T1: 972frame; T2: 6252frame; T3: 9741frame).

For easier comparison, the location of the focused target of the camera was consistent with that of FBG. Two images (Figure 8(a)) of the target from different perspectives were used to obtain the 3D point cloud (Figure 8(b)). Then, the OD of each target was deduced for scale factor calculations, as shown in Figure 8(c). The lens used has a focal length of 100 mm. The acquisition frequency was 10 Hz. As shown in Figure 9(a), the captured images were affected by light spots. Then, with the light spot removal model, they were restored. Compared to the original image without a light spot at other times, it can be seen that the restored image is qualified. The original method, “[15] + sparse expression + coarse matching,” was used to extract the displacement from the original image, and the results are drawn using the blue line in Figure 9(b). To contrast against it, the improved method was applied to the restored images, and the displacement measurement result was described using the red line. Taking the results from FGB as a reference, it can be concluded that the abnormal data were effectively reduced by performing image restoration, and the noise level was controlled through the improved matching strategy. The displacement curves for several sections are shown in Figure 10(a). By extracting the deflections of different measuring points at the same time, the linearity of the deflection of the bridge can be obtained, as shown in Figure 10.

5. Conclusions

Two key problems, object distance measurement and light spot interference, which are often encountered in the process of camera-based structure displacement measurement, were studied. The main contributions of this study are as follows:

- (1) To overcome the limitations of the laser rangefinder, a fast and accurate object distance measurement method based on stereovision was suggested, which has the advantage of full-field multipoint synchronous calibration.

- (2) To protect the image quality from random light spots, the edge information of the degraded image was designed to be used for deep learning to achieve a better restoration effect.
- (3) To balance the matching speed and accuracy when images are degraded, a distance-weighted coarse-to-fine matching strategy was developed combining sparse representation.
- (4) Through the algorithm test, it was found that the introduction of edge priors causes the restored image to have a higher signal-to-noise ratio and a higher structural similarity with the original image. Compared with the template-matching method without considering image restoration, the sparse representation-based matching method has higher matching accuracy using the restored images, and using the coarse-fine matching strategy, the accuracy will be further improved.
- (5) The integrated algorithms above were applied to a concrete bridge, and the vertical displacement impacted by the normal traffic load was monitored. Compared with the algorithms before improvement, the results of the proposed method are closer to those of the FGB sensor and the noise level is lower.

In conclusion, this study is beneficial in promoting image-based displacement measurement technology to adapt to the complex environment. Comprehensive attention will be given to other possible factors in the future.

Data Availability

The image data used to support the findings of this study are currently under embargo, while the research findings are commercialized.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The research presented was financially supported by the National Key Research and Development Program of China (2020YFC1511900).

References

- [1] S. Bhowmick and S. Nagarajaiah, "Identification of full-field dynamic modes using continuous displacement response estimated from vibrating edge video," *Journal of Sound and Vibration*, vol. 489, no. 7, Article ID 115657, 2020.
- [2] Y. Xu, J. M. W. Brownjohn, and F. Huseynov, "Accurate deformation monitoring on bridge structures using a cost-effective sensing system combined with a camera and accelerometers: case study," *Journal of Bridge Engineering*, vol. 24, no. 1, 2019.
- [3] S. Bhowmick and S. Nagarajaiah, "Spatiotemporal compressive sensing of full-field Lagrangian continuous displacement response from optical flow of edge: identification of full-field dynamic modes," *Mechanical Systems and Signal Processing*, vol. 164, Article ID 108232, 2022.
- [4] B. Pan, L. Tian, and X. Song, "Real-time, non-contact and targetless measurement of vertical deflection of bridges using off-axis digital image correlation," *NDT & E International*, vol. 79, pp. 73–80, 2016.
- [5] S. Yu, J. Zhang, and X. He, "An advanced vision-based deformation measurement method and application on a long-span cable-stayed bridge," *Measurement Science and Technology*, vol. 31, no. 6, Article ID 065201, 2020.
- [6] Y. H. Lin, W. C. Chen, and Y. Y. Chuang, "BEDSR-net: a deep shadow removal network from a single document image," in *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12902–12911, Seattle, WA, USA, June 2020.
- [7] C. Wei, W. Wang, and W. Yang, "Deep retinex decomposition for low-light enhancement," pp. 1186–1191, 2018, <https://arxiv.org/abs/1808.04560>.
- [8] Z. Tang, J. Li, J. Huang, Z. Wang, and Z. Luo, "Multi-scale convolution underwater image restoration network," *Machine Vision and Applications*, vol. 33, no. 6, 85 pages, 2022.
- [9] J. Stark, "Adaptive image contrast enhancement using generalizations of histogram equalization," *IEEE Transactions on Image Processing*, vol. 9, no. 5, pp. 889–896, 2000.
- [10] R. Kumar and A. Kumar Bhandari, "Luminosity and contrast enhancement of retinal vessel images using weighted average histogram," *Biomedical Signal Processing and Control*, vol. 71, Article ID 103089, 2022.
- [11] J. Chaki, "A two-fold fusion fuzzy framework to restore non-uniform illuminated blurred image," *Optik*, vol. 206, Article ID 164197, 2020.
- [12] J. Zhou, T. Yang, W. Ren, D. Zhang, and W. Zhang, *Underwater Image Restoration via Depth Map and Illumination Estimation Based on Single Image*, Optics Express, Glasgow, Scotland, UK, 2021.
- [13] J. Wang, X. Li, L. Hui, and J. Yang, "Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal," in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1788–1797, Salt Lake City, UT, USA, June 2018.
- [14] Y. Jiang, X. Gong, D. Liu et al., "Enlighten GAN: deep light enhancement without paired supervision," *IEEE Transactions on Image Processing*, vol. 30, pp. 2340–2349, 2021.
- [15] Q. Bao, Y. Liu, B. Gang, W. Yang, and Q. Liao, "S²Net: shadow mask-based semantic-aware network for single-image shadow removal," *IEEE Transactions on Consumer Electronics*, vol. 68, no. 3, pp. 209–220, 2022.
- [16] L. Qu, J. Tian, S. He, Y. Tang, and R. W. H. Lau, "Deshad-owNet: a multi-context embedding deep network for shadow removal," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4067–4075, IEEE, Honolulu, HI, USA, July 2017.
- [17] X. Hu, Y. Jiang, C. W. Fu, and P. A. Heng, "Mask-ShadowGAN: learning to remove shadows from unpaired data," in *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 2472–2481, Seoul, Korea (South), October 2019.
- [18] Z. Shen, W. S. Lai, and T. Xu, "Deep semantic face deblurring," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 8260–8269, IEEE, Salt Lake City, UT, USA, June 2018.
- [19] D. Feng, T. Scarangelo, M. Q. Feng, and Q. Ye, "Cable tension force estimate using novel noncontact vision-based sensor," *Measurement*, vol. 99, pp. 44–52, 2017.
- [20] Y. Xu, J. Zhang, and J. Brownjohn, "An accurate and distraction-free vision-based structural displacement measurement method integrating Siamese network based tracker and correlation-based template matching," *Measurement*, vol. 179, Article ID 109506, 2021.
- [21] T. Khuc and F. N. Catbas, "Completely contactless structural health monitoring of real-life structures using cameras and computer vision," *Structural Control and Health Monitoring*, vol. 24, no. 1, Article ID e1852, 2017.
- [22] S. Bhowmick, S. Nagarajaiah, and Z. Lai, "Measurement of full-field displacement time history of a vibrating continuous edge from video," *Mechanical Systems and Signal Processing*, vol. 144, Article ID 106847, 2020.
- [23] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.
- [24] Y. Li, L. Li, W. Li, Z. Fan, and N. Sang, "Blind deblurring of Gaussian blurred images by blurred edge image," in *Proceedings of the MIPPR 2019: Pattern Recognition and Computer Vision*, Article ID 11430, Wuhan, China, February 2020.
- [25] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [26] S. Yu and J. Zhang, "Fast bridge deflection monitoring through an improved feature tracing algorithm," *Computer-Aided Civil and Infrastructure Engineering*, vol. 35, no. 3, pp. 292–302, 2020.
- [27] Z. Shen, W. S. Lai, T. Xu, J. Kautz, and M. H. Yang, "Deep semantic face deblurring," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 8260–8269, IEEE, Salt Lake City, UT, USA, June 2018.
- [28] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, <https://arxiv.org/abs/1511.06434>.

- [29] J. Johnson, A. Alahi, and F. Li, "Perceptual losses for real-time style transfer and super-resolution," *European conference on computer vision*, vol. 9906, pp. 694–711, 2016.
- [30] W. Liu, A. Rabinovich, and A. C. B. Parzenet, "Looking wider to see better," 2015, <https://arxiv.org/abs/1506.04579>.
- [31] D. L. Donoho, "For most large underdetermined systems of linear equations, the minimal ℓ_1 -norm solution is also the sparsest solution," *Communications on Pure and Applied Mathematics*, vol. 59, no. 6, pp. 797–829, 2006.
- [32] Y. B. Tong, Q. S. Zhang, and Y. P. Qi, "Image quality assessing by combining PSNR with SSIM," *Journal of Image and Graphics*, vol. 33, 2006.