



Research Article

Complex Background Removal Method for Video of an Inclined Cable Based on UAV Detection of Structural Vibration Frequencies

Shengli Li,^{1,2} Zihao Cao ,^{1,2} Jinke Li ,^{1,2,3} Panjie Li,^{1,2} Bin Xu,^{1,2} and Wudi Gao⁴

¹School of Civil Engineering, Zhengzhou University, Zhengzhou 450001, China

²Henan Province Engineering Research Center of Safety and Life Extension of Prestressed Cable Structure, Zhengzhou 450000, China

³Henan Jiaoyuan Engineering Technology Group Co., Ltd., Zhengzhou 450000, China

⁴School of Civil Engineering, Dalian University of Technology, Dalian 116024, China

Correspondence should be addressed to Jinke Li; lijinke_zzu@zzu.edu.cn

Received 13 January 2023; Revised 22 December 2023; Accepted 31 December 2023; Published 18 January 2024

Academic Editor: Sara Casciati

Copyright © 2024 Shengli Li et al. 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.

The safety of inclined cables is fundamental to the integrity of cable-stayed bridges. The vibrational frequencies of these cables form the foundation for assessing the cable force. Traditional contact measurement methods necessitate the installation of sensors on each cable, incurring substantial costs. In scenarios where camera placement adjacent to an inclined cable is impractical, noncontact approaches such as video capture via unmanned aerial vehicles prove effective. However, unmanned aerial vehicle-captured videos present a challenge due to their complex background, impeding cable feature recognition. In our study, we initially utilized the Region Growing algorithm for background subtraction. To enhance this method, we integrated it with the unique structural characteristics of cables, leading to the creation of the RGv2 algorithm. This novel algorithm offers increased processing speed and improved accuracy. Furthermore, we combined our method with empirical mode decomposition for effective detection of cable frequency characteristics. We also implemented a hybrid method, combining the K-Means and line segment detector algorithms with empirical mode decomposition. Compared to deep learning techniques for background subtraction, our proposed method demonstrates superior computational efficiency and promising potential for measuring vibrational frequencies of inclined cables.

1. Introduction

Inclined cables constitute integral elements within the load-bearing structures of cable-stayed bridges. The precise measurement of the vibration frequency of these inclined cables holds great significance [1, 2] for cable force assessment [3]. Currently, methods for measuring the vibration of inclined cables can be categorized into two primary domains [4]: contact measurement methods and noncontact measurement methods.

Most contact measurement methods rely on the utilization of accelerometers for monitoring purposes [5, 6]. Nevertheless, structural health monitoring systems employing contact monitoring methods necessitate

a substantial number of connection wires due to the extended distances and numerous inclined cables involved, resulting in elevated costs [7]. Wireless sensors [8] offer a notable reduction in the number of needed connection wires but introduce the challenge of potential wireless data loss.

Alternatively, noncontact measurement methods, such as laser Doppler techniques [9] and microwave remote sensing [10], can achieve high-precision measurements. However, the equipment associated with these methods is relatively costly and demands specialized operational expertise [11]. In contrast, noncontact measurement methods founded on computer vision algorithms offer a more cost-effective and user-friendly

approach [11]. As early as 1998, Gehle and Masri [12] employed a video camera to capture footage for measuring the vibration frequency of cables. Subsequently, Guo et al. [13] integrated deep learning methods with traditional optical flow techniques to assess the vibration frequency of an inclined cable. Kim et al. [14] conducted a comparative study evaluating accelerometers and various smart devices under diverse weather conditions, affirming the measurement accuracy of computer vision algorithms. Zhao et al. [15] employed a smartphone to record vibrations in an inclined cable and utilized a computer vision algorithm to determine the vibration frequency.

Noncontact measurements using smartphones or industrial cameras offer increased efficiency and cost savings in comparison to contact measurements [4]. However, when applying these methods to photograph large civil engineering structures, challenges may arise in camera setup and the identification of suitable angles for video capture [16]. The advent of unmanned aerial vehicles (UAVs) has presented a novel solution. UAVs, capable of swiftly capturing video footage of sizable structures, find myriad applications in civil engineering, including 3D reconstruction of buildings, dams, and bridges [17–19], assessment of seismic damage in buildings [20], long-term monitoring of slope displacements [21], and detection of structural cracks [22–24].

Moreover, UAVs have been harnessed for measuring structural vibrations. These methodologies concentrate on tracking alterations in natural features or artificial markers on a structure to derive structural displacement. For instance, Weng et al. [25] combined the optical flow method with perspective transformation to identify the displacement of a supertall building from video footage captured by a UAV. Hoskere et al. [26] employed an optical flow method to discern the vibrations of a pedestrian suspension bridge from UAV-captured video. Khadka et al. [27] utilized a UAV to capture video footage of a wind turbine model and applied a digital image correlation method to evaluate the structural integrity of the turbine blades. Tian et al. [16], utilizing a UAV to record video of an inclined cable, employed a line segment detector (LSD) to determine the vibration frequency of the cable.

However, when a UAV captures video footage of an inclined cable, it invariably records images of the surrounding landscape, including mountains, rivers, and urban structures situated behind the inclined cable. The presence of complex background imagery can introduce interference in the images of inclined cable [16], potentially leading to inaccuracies in the analysis results of computer vision algorithms. In the domain of computer vision, methods for distinguishing between foreground targets and background images can be categorized into traditional algorithms and deep learning algorithms.

Traditional algorithms typically employ videos acquired through stationary cameras and analyze multiframe images to distinguish static backgrounds from moving foreground objects. To eliminate static backgrounds, a common approach involves applying filters to the results [28], such as median filtering [29] and Frame Difference [30]. However, these methods exhibit limited effectiveness when dealing with dynamic backgrounds. The proposed method, which entails

using a UAV to capture cable videos, introduces a dynamically shifting background, albeit with relatively minor movements compared to the displacement of the cables.

In recent years, an increasing number of researchers have recognized the significance of background removal and have undertaken relevant investigations. For instance, Wei and Peng [31] proposed a block Frame Difference method and conducted experiments in various scenarios. This algorithm succeeded in removing most dynamic backgrounds of sailing ships but occasionally misclassified certain areas of the sea surface as foreground. Additionally, other background modeling techniques, such as the hybrid Gaussian model [32] and Vibe's algorithm [33], calculate the video's background by tracking pixel intensity changes from frame to frame, achieving the extraction of moving objects. Nevertheless, these methods encounter difficulties in eliminating the complex backgrounds encountered in cable videos captured by UAVs. Among these approaches, Vibe and Frame Difference yield results with numerous erratic straight lines, as observed in Figures 1 and 2.

In contrast, deep learning methods discern the foreground from the background in images through the training of convolutional neural networks [34–36]. However, the accuracy of background removal using deep learning methods relies on the chosen deep learning models and image datasets [37–39]. Aside from selecting an appropriate deep learning model, the quality of the image dataset significantly impacts recognition. Consequently, the collection of an extensive image dataset, comprising a minimum of 1000 images, becomes crucial. When applied to cable region recognition in this study, the necessity arises to capture and annotate a substantial number of cable images. It is worth noting that labeling data entail substantial manual effort [38]. In cases where the image dataset lacks comprehensiveness, the model trained with images from a specific bridge may exhibit limited generalization, potentially resulting in inaccuracies in cable background removal if the cable features differ.

The Region Growing algorithm [40] represents a region-based image segmentation technique capable of gradually expanding and merging small regions based on predefined rules. It has found applications in various domains. For instance, Wei et al. [30] applied the Region Growing algorithm to identify road cracks, while Shao et al. [41] used it to segment roofs from UAV-captured images. Lin [42] employed a modified Region Growing algorithm for automatic detection of remote sensing images, and Lu et al. [43] utilized it for segmenting abdominal CT images.

However, when evaluating the suitability of the Region Growing algorithm for displacement detection in cable structure analysis, several challenges come to the forefront. The algorithm inherently lacks robust edge detection capabilities, a critical requirement for precise delineation in such analyses. Furthermore, when combined with algorithms such as LSD, it results in noticeably extended processing times. Additionally, during motion, cables often overlap with backgrounds having similar grayscale values, leading to segmentation errors when relying solely on grayscale values of neighboring pixels. To address these

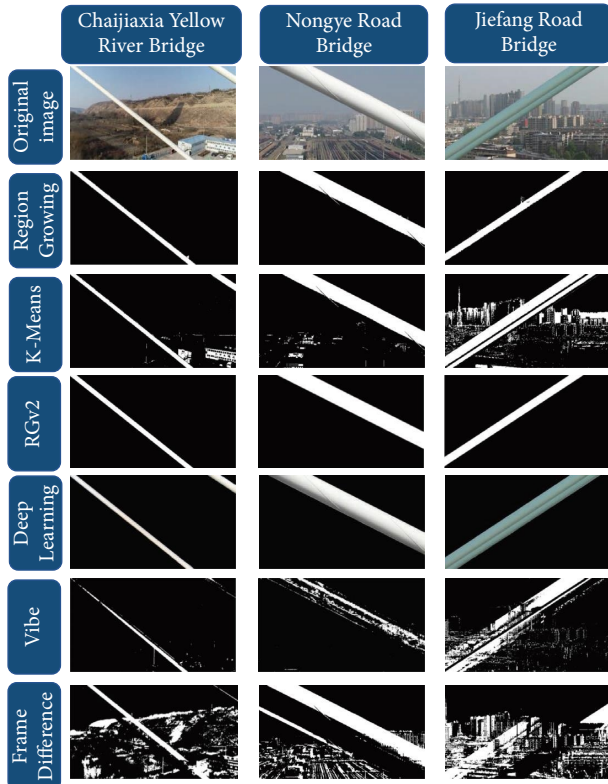


FIGURE 1: The effects of different background removal algorithms.

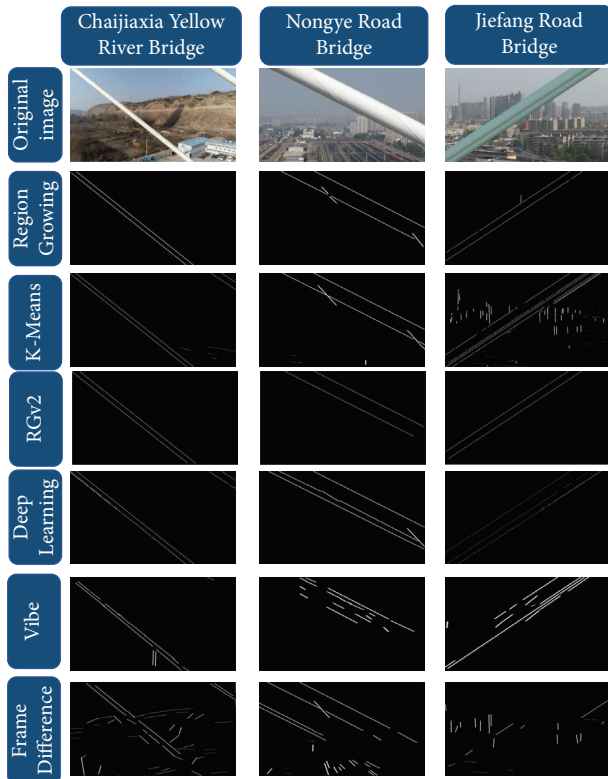


FIGURE 2: The cable edge detection results after background removal.

challenges, this study introduces RGv2, an enhanced iteration of the Region Growing algorithm rigorously optimized for structural cable system analysis. RGv2 not only exhibits markedly improved accuracy in segmenting cable structures compared to its conventional counterpart but also excels in directly extracting displacement information from cables during the segmentation process. In terms of efficiency, RGv2 accomplishes its objectives with only one-third of the processing time needed by the traditional combination of the Region Growing algorithm and LSD, representing a substantial advancement in processing speed. Furthermore, RGv2 adopts a more comprehensive growth approach, enhancing its accuracy in managing cable structure scenarios. A detailed exposition of RGv2's mechanism is provided in Section 2.1.2 of this paper.

If a substantial color disparity exists between foreground and background images, the removal of the background can be achieved through the application of a clustering algorithm such as K-Means [44]. It has found applications in various domains. For instance, Ding [45] utilized K-Means to extract dominant colors from an image, while Zhang et al. [46] adapted K-Means for hyperspectral image classification. In this study, RGv2, the Region Growing algorithm, and K-Means are employed based on the unique image characteristics of inclined cables to discern cables from complex backgrounds.

Unlike capturing video with a stationary camera, UAVs are adept at recording video footage of inclined cables. Nevertheless, videos captured by UAVs inherently contain relative displacements of both the tested structure and the UAV itself. The absolute displacement of the cable can be defined as the discrepancy between the displacement recorded during UAV hover shooting and the UAV's absolute displacement, as illustrated in Figure 3.

Presently, significant research attention is directed towards addressing this issue, and methods for mitigating or eliminating UAV motion can be categorized into three principal approaches:

- (1) *Utilizing Inertial Measurement Devices* [47–49]. This methodology entails measuring the UAV's motion using either the UAV's flight data or additional devices such as gyroscopes, GPS, and accelerometers. Mathematical models are then constructed using these supplementary data to compensate for the UAV's movement. However, this method necessitates the incorporation of additional devices alongside a consumer-grade UAV, leading to an increase in monitoring costs.
- (2) *Employing Stationary Objects as References* [25, 50, 51]. UAV motion is determined by tracking feature points within stationary backgrounds or through a template matching algorithm [51]. Subsequently, UAV motion can be nullified through photogrammetry techniques. Nevertheless, the success of this method relies on high-quality backgrounds, as it may yield errors in cases where background quality is compromised.

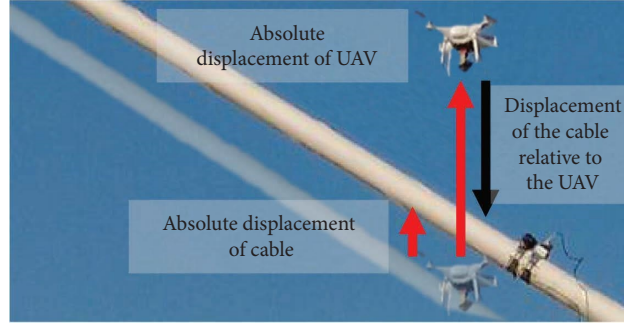


FIGURE 3: Demonstration of the relationship between absolute and relative displacement.

(3) *Leveraging Frequency Characteristics.* some literature [26, 52] mentions the use of high-pass filters to eliminate low-frequency components of UAV motion from raw data. Nevertheless, this approach necessitates the definition of a cutoff frequency and can be relatively complex. Conversely, empirical mode decomposition (EMD) [53] represents an adaptive decomposition technique that simplifies the process and breaks down the original data into multiple intrinsic mode functions (IMFs). EMD effectively removes the low-frequency components, leaving behind the high-frequency aspects of the data. EMD and its extended algorithms have already been applied in a variety of fields. For instance, these applications span across medical [54], engineering [36, 55, 56], and mechanical [57–60] fields. In this paper, EMD is employed to decompose displacement data collected by a UAV.

The primary focus of this paper is to utilize the proposed RGv2 algorithm for the purposes of background removal and cable tension detection. Additionally, a series of algorithms based on Region Growing and K-Means have also been employed to achieve these objectives. In analyzing the dynamic characteristics of the inclined cable, this paper employs EMD to reduce the influence of the UAV’s own vibration. The paper is organized as follows, as shown in Figure 4. Section 2 introduces the computer vision algorithms used for background removal and displacement extraction. Additionally, this section describes the experiment conditions at the Chaiji Xia Yellow River Bridge and experimental equipment used. Section 3 presents the effects of the background removal algorithms on time and frequency domains. Additionally, this section analyzes the effect of EMD and estimates the cable force of the inclined cable. The results show that the processing method proposed in this paper can accurately identify vibration frequencies of the inclined cable.

2. Proposed Approach

A video of an inclined cable’s vibration captured by a UAV may have complex background images, which will affect the recognition of the cable edge features. The misidentification of cable edge could seriously affect the accuracy of displacement time history. In this paper, according to the image

characteristics of the inclined cable, three algorithms are used to remove the inclined cable’s background, respectively.

The proposed method is structured into two distinct segments. The initial segment involves the processing of video footage capturing the vibrations of an inclined cable recorded by a UAV. The subsequent segment focuses on converting the obtained displacement time history, derived from the first segment, into the frequency domain for comprehensive analysis. A schematic representation of the proposed method is shown in Figure 5.

During the image processing phase, to assess the effectiveness of various methodologies, we implemented four distinct approaches. The first approach harnessed our newly developed RGv2 algorithm, which served the dual purposes of background removal and displacement detection. The second method employed the Region Growing algorithm for background subtraction, followed by the application of LSD to determine the cable displacement data. In the third approach, the K-Means algorithm was applied for background elimination and subsequently integrated with LSD for the determination of displacement information. Finally, as a control method, we directly utilized LSD to calculate displacement information without any prior background removal, thus providing a baseline for comparative evaluation against the other techniques.

In the subsequent phase involving frequency domain processing, the displacement data undergo analysis through EMD, facilitating the extraction of cable vibration frequencies utilizing fast Fourier transform (FFT). It is noteworthy that all the computational procedures described can be efficiently executed using MATLAB.

2.1. Three Background Removal Algorithms for Inclined Cable

2.1.1. Region Growing Algorithm. The exterior surface of an inclined cable is enveloped with a polyethylene sheathing, rendering the surface predominantly white in appearance. Within the video footage captured by a UAV, there is minimal variation in the image intensity across its surface. Consequently, the entirety of the inclined cable region can be effectively outlined using the Region Growing algorithm [40]. The original Region Growing method includes either 8-region or 4-region expansion. For instance, considering the 8-connected domain, the algorithm compares the gray

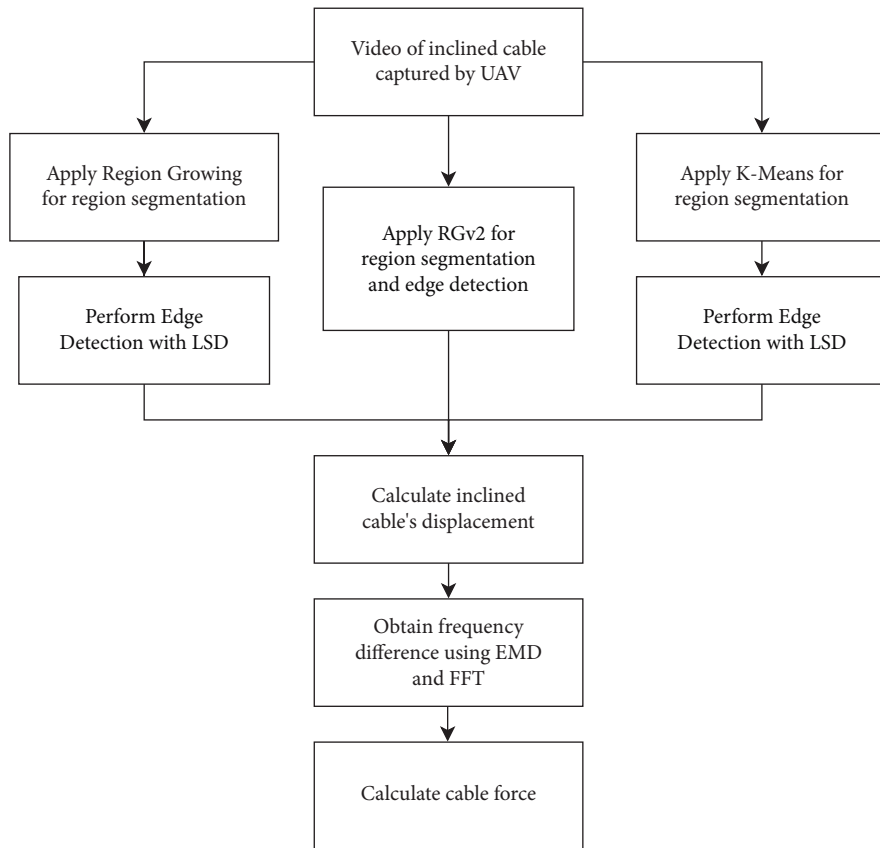


FIGURE 4: Workflow of this paper.

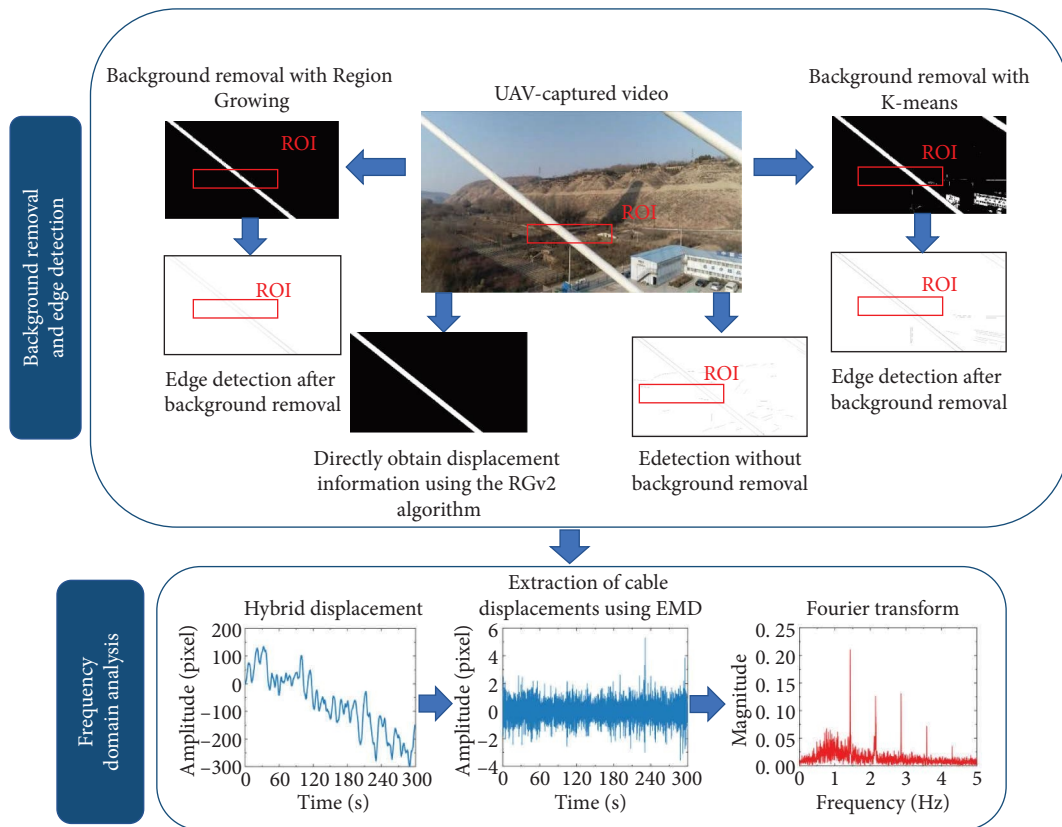


FIGURE 5: A framework of the proposed approach.

threshold of a seed point with that of its eight adjacent points. Points exhibiting a gray difference below the gray threshold are earmarked for potential expansion. As depicted in Figure 6, points highlighted in blue represent preselected seed points. If the grayscale threshold is set to 30, the algorithm can ultimately extend to include the points indicated in green.

After removing the background image, the LSD algorithm [61] is deployed to deduce the straight lines defining the edges of the inclined cable.

In the growing process of Region Growing, the gray threshold determines the acceptable tolerable of gray value difference. As shown in Figure 7, for backgrounds where the gray values are too similar, if they are too close to each other, it is difficult to obtain good results even by adjusting the gray threshold.

2.1.2. RGv2 Algorithm. In pursuit of superior results, taking into account the cable's specific attributes, we have introduced a more efficient growth rule, which concurrently allows for the direct extraction of linear information pertaining to the cable's edges.

The precise growth rules are outlined as follows: initially, we segment the images into clusters predicated on the inclination of the cable. As illustrated in Figure 8, the cluster where the seed point resides (indicated in blue) is designated the seed cluster. Subsequently, we assess the average gray value between the two neighboring clusters. If the value is less than the grayscale threshold, the growth process is initiated. If it exceeds the grayscale threshold or the image boundaries, meeting the termination criteria, the growth is then halted. Ultimately, all segments marked in green can be expanded. This procedural insight is elucidated in Figure 9. The linear attributes of the cable's edge can be discerned through the positioning of the seed cluster and the frequency of adjacent expansions, obviating the necessity for cable edge detection. This method results in a notable optimization of the processing time.

In comparison to the previous Region Growing and LSD algorithms, our proposed RGv2 algorithm demonstrates swifter processing times and heightened accuracy. A comprehensive comparative analysis with additional algorithms is provided in Section 3 of this paper.

2.1.3. K-Means Algorithm. In instances where the background image of the inclined cable contains limited white regions, cluster algorithms such as K-Means can be effectively employed to differentiate between the inclined cable and the background image. The steps involved in K-Means clustering can be summarized as follows.

In a color image, each individual pixel comprises three components: red (R), green (G), and blue (B). The R, G, and B components of all pixel points collectively form a three-dimensional sample space. Through the utilization of K-Means, pixel points within the image that closely resemble the color of the inclined cable are grouped into a single cluster, while other colors are distributed across multiple

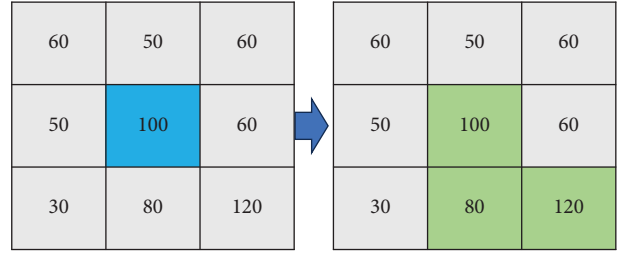


FIGURE 6: Traditional Region Growing algorithm principal demonstration.

clusters. Following the clustering process, only the results of the initial cluster are retained, leading to the effective removal of the majority of the complex background.

Following the removal of the background, the edge information pertaining to the cable is computed utilizing the LSD algorithm, ultimately yielding the displacement data for the cable. This sequence of operations is depicted in Figure 10.

2.2. EMD. To enhance the precision of identifying vibration frequencies in inclined cables and reduce the influence of the UAV's motion, this study employs EMD to analyze the displacement time history. EMD, which can decompose the time history into multiple empirical modes called IMFs [53], represents a nonsmooth time-history decomposition method.

Although the selection of high and low frequencies remains somewhat subjective, this study endeavors to design a method for the automatic selection of suitable IMFs. This approach draws inspiration from the technique employed by Zhang and Wei for estimating high-frequency noise boundaries [62], involving the creation of an evaluation system for identifying IMFs with distinct frequency peaks. Furthermore, it is influenced by the research of Yoon et al. [63], which asserts that UAV motion predominantly occurs between 0 and 0.5 Hz. The entire selection process is illustrated in Figure 11.

Initially, this method conducts spectral analysis on the signal and compares the energy within the low-frequency range (0–0.5 Hz) with that in other frequency ranges. During the initial filtering step, if the proportion of low-frequency energy is substantial, it indicates that the given IMF predominantly reflects the UAV's flight motion characteristics. Consequently, these IMFs are deemed for exclusion. Conversely, when the proportion of low-frequency energy is relatively small, it suggests that interference from the UAV on the signal is relatively limited, making these IMFs suitable for further refinement and analysis.

To select the IMFs capable of reflecting the cable's dynamic characteristics, the standard deviation and spectral area with respect to the x -axis are also utilized to describe the data. For the spectral data in this research, a smaller standard deviation implies reduced interference and a more pronounced peak frequency. Additionally, the area enclosed with x -axis of the frequency spectral serves as another indicative measure. A smaller area corresponds to a more



FIGURE 7: The effect of tuning parameters in the Region Growing algorithm. (a) The gray threshold is set to 0.25; (b) the gray threshold is set to 0.20.

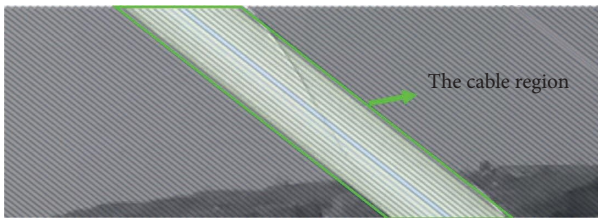


FIGURE 8: RGv2 principal demonstration.

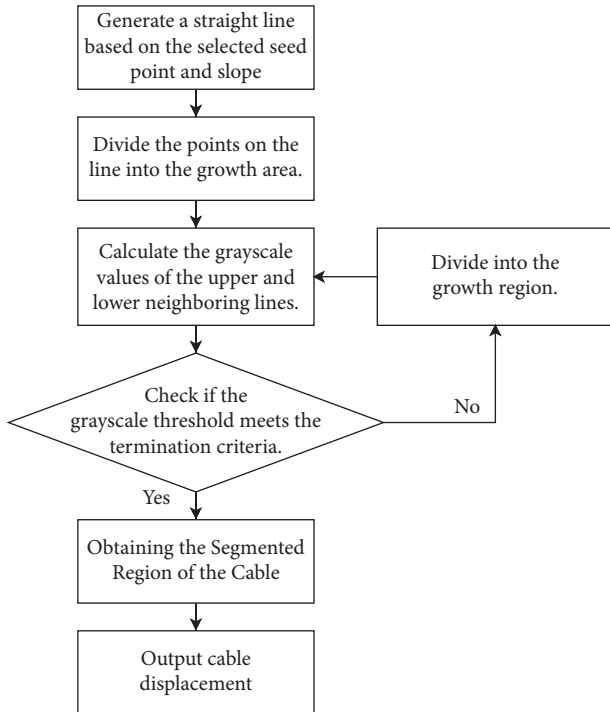


FIGURE 9: The process of RGv2.

prominent frequency peak and less interference. Notably, the proportion of low-frequency energy exhibits a similar trend to the standard deviation and the area enclosed by the x -axis. Hence, smaller values for these three indicators

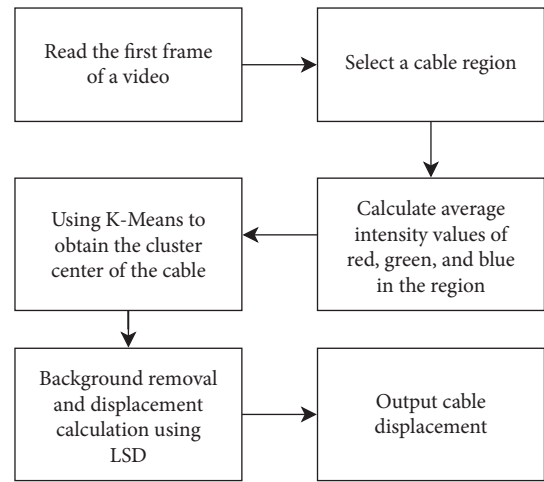


FIGURE 10: Detailed flowchart of the proposed K-Means algorithm.

signify that the IMF is better suited for capturing the cable’s dynamic characteristics. Consequently, there is a need to formulate an index that can encapsulate these features.

Before constructing the index, data normalization is a prerequisite. Due to the diverse frequency ranges of the IMFs decomposed by EMD, the Fourier transformed spectral diagrams exhibit significant variations along the Y -axis. To enable a uniform comparison across all IMFs, the frequency domain data of each IMF are initially normalized using formula (1), where NM denotes the normalization result.

$$NM = \frac{IMF - \min(IMF)}{\max(IMF) - \min(IMF)} \quad (1)$$

Given the disparate distribution ranges of each IMF within the frequency domain, the analysis is concentrated on the 0–10 Hz range. Taking into account the aforementioned considerations, the standard deviation, area, and low-frequency energy ratio of the frequency domain separately using X_n^m , the subscript “ n ” represents the IMF number, while “ m ” pertains to the three indices: standard deviation, area, and low-frequency energy ratio.

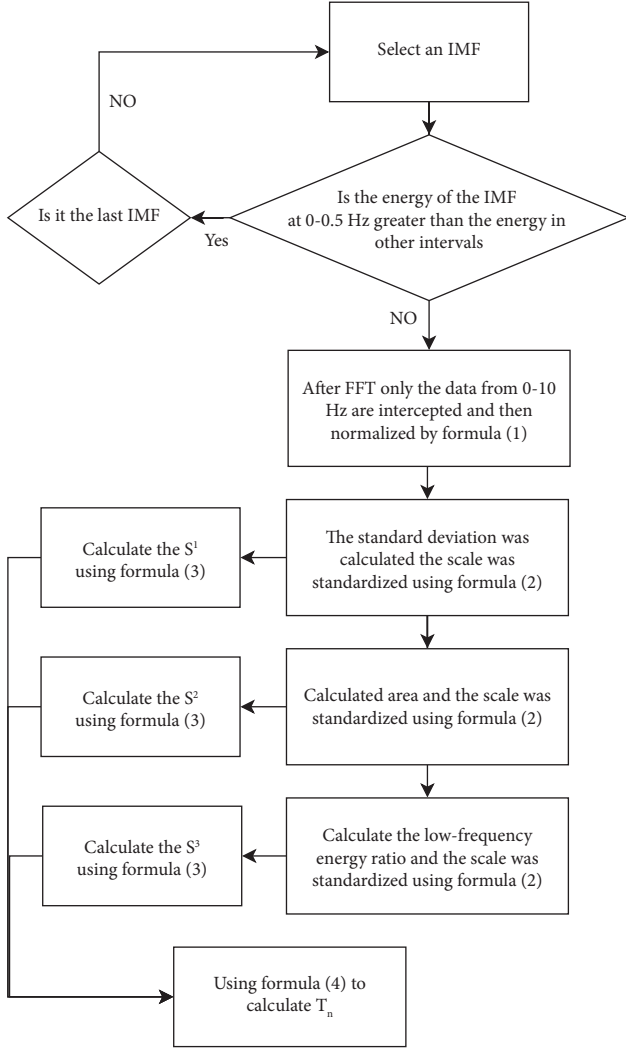


FIGURE 11: The IMF selection process flowchart.

Due to the varying scales of the three indices, namely, standard deviation, area, and low-frequency energy ratio, as depicted in Figure 11, if X_n^m is used directly to make judgments, some indices will be decisive, while others will be dispensable. However, each X_n^m is expected to have the same weight. Consequently, the maximum value of each index is set to 1, and the remaining values are scaled accordingly, as demonstrated in formula (2). This adjustment yields a more balanced evaluation of standard deviation, area, and low-frequency energy.

$$A_n^m = \frac{X_n^m}{\max(X_n^m)} \quad m = 1, 2, 3. \quad (2)$$

As the objective is to construct an index that attains higher values when the standard deviation, area, and low-frequency energy ratio are minimized, formula (3) is devised for this purpose. The final score, as denoted in formula (4), is derived by summing all three S_n^m together. A higher scoring of the T_n value indicates that the respective IMF exhibits

more pronounced frequency peaks. This methodology has been rigorously validated and successfully applied in Section 3.4 of the study.

$$S_n^m = 1 - A_n^m \quad m = 1, 2, 3, \quad (3)$$

$$T_n = \sum_1^m S_n^m \quad m = 1, 2, 3. \quad (4)$$

2.3. Experimental Approach of Capturing an Inclined Cable by a UAV. To validate the efficacy of the background removal techniques presented in this manuscript, we conducted experiments on an inclined cable situated on a bridge spanning the Yellow River in Lanzhou. As depicted in Figure 12(a), our experimental procedure involved artificially exciting one of the inclined cables, followed by capturing video footage using a UAV. Simultaneously, an accelerometer was affixed to the inclined cable to record vibration data for comparative analysis.

For our experimentation, we selected the third inclined cable on the northeastern side of the bridge, counting from the top (as illustrated in Figure 12(a)). The primary parameters of this inclined cable are summarized in Table 1. Since the bridge had not yet been opened to traffic, we employed a rope-based artificial excitation method to induce vibrations in the inclined cable. Once the amplitude reached a stabilizing point, we ceased the excitation, allowing the inclined cable's vibrations to gradually attenuate. Throughout the experiment, we utilized a DJI Phantom 4 Pro UAV equipped with a 1-inch 20-megapixel image sensor, and the camera operated at a frame rate of 60 frames per second.

To expand our dataset for the validation of the methodologies outlined in Section 3.4, we also employed UAVs to capture data from two other cable-stayed bridges, namely, the Nongye Road Bridge and Jiefang Road Bridge. These bridges are located in Zhengzhou, China. The distinguishing features of the Jiefang Road Bridge include light green cables and closer alignment of two rows of cables, enabling the UAV to capture two overlapping cables in a single frame.

3. Results and Discussion

3.1. Analysis of the Three Proposed Algorithms. This section compares the effect of the three algorithms on background removal and their respective time consumption. As shown in Figure 13(a), to make the comparison clearer, the result of the RGv2 is green and the K-Means is blue, while the result of the Region Growing algorithm is changed to red, as shown in Figure 13.

In Figure 13, the area covered by the RGv2 algorithm is noticeably larger than that of the traditional Region Growing algorithm. This is due to the different growth mechanisms and gray threshold values adopted by the two. In fact, if the Region Growing algorithm were to use a higher gray threshold at this point, it would lead to extensive

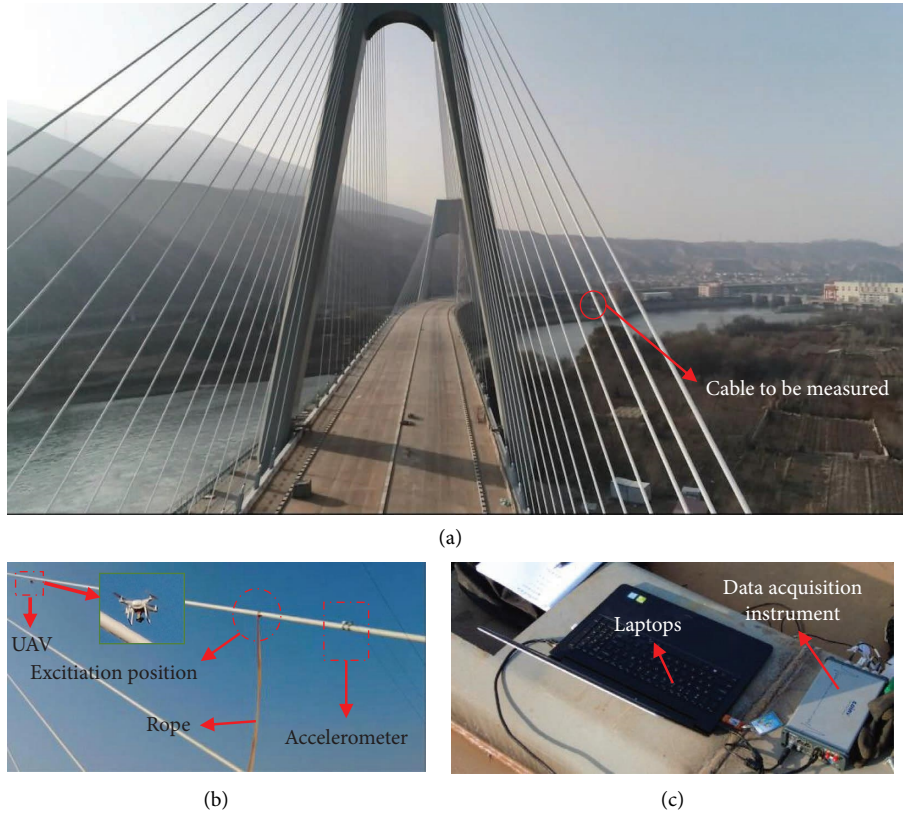


FIGURE 12: The Yellow River bridge in Lanzhou: (a) location of the inclined cable to be measured; (b) deployment of the accelerometer and the UAV; (c) data acquisition instrument.

TABLE 1: Main parameters of the measured inclined cable.

Length (m)	Outer diameter (m)	Unit mass (kg/m)
165.2430	0.1330	68.6000

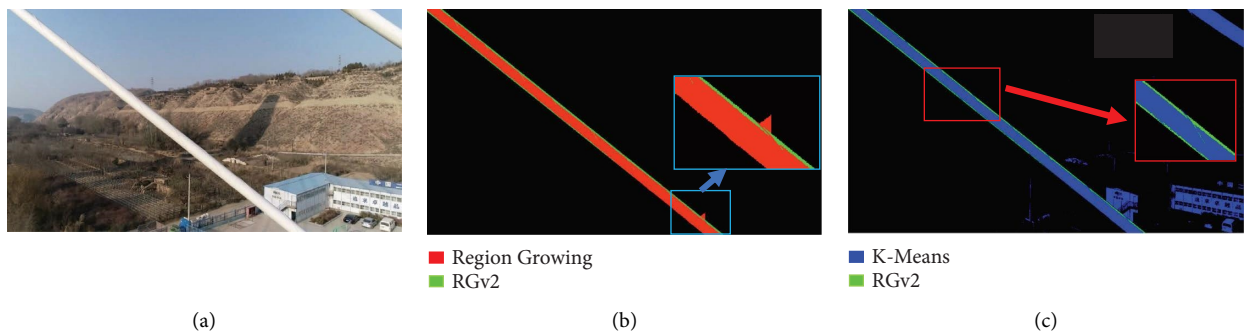


FIGURE 13: Comparison of background removal results.

misjudgment like that shown in Figure 7(a). Therefore, the Region Growing algorithm can only use a relatively conservative gray threshold. In fact, even using a smaller gray threshold inevitably leads to the growth of an area that does not belong to the cable. In addition, the outer part of the cable is wrapped in a PE sheath in a spiral, which results in many curved lines on the cable surface that differ in gray value from the cable itself, as shown in Figure 1. These

curved lines, which the Region Growing algorithm cannot grow, will interfere with the subsequent edge detection. These are the reasons for the poor robustness of the Region Growing algorithm as shown in Figures 14 and 15. In contrast, the RGv2 algorithm proposed in this study judges and grows based on the gray values of a straight line, which can avoid many local minor issues and thus grow a more complete cable. K-Means also has similar issues.

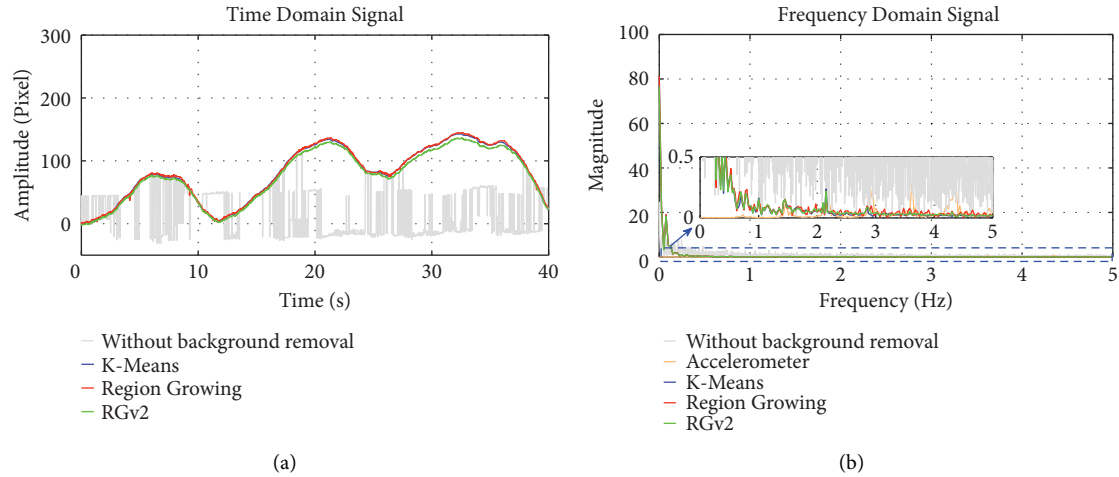


FIGURE 14: Calculation results for the short-time condition: (a) displacement time history; (b) frequency domain.

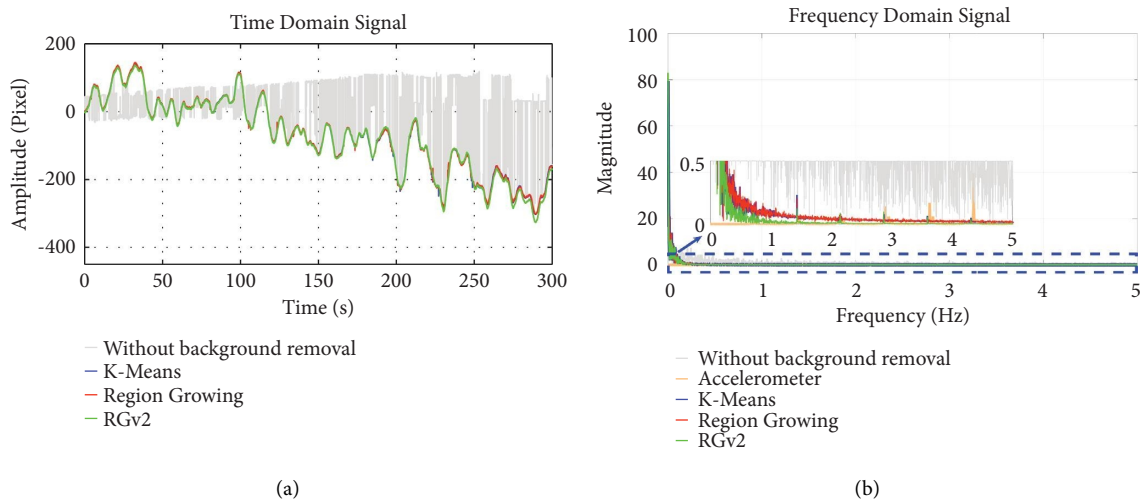


FIGURE 15: Calculation results of the long-time conditions: (a) displacement time history; (b) frequency domain.

We employ MATLAB’s integral “profile” command to determine the execution times associated with RGv2, Region Growing, and K-Means. To facilitate a precise comparison of the durations for each method, we conducted processing on a personal computer, specifically focusing on the time expended in tasks such as image retrieval, background removal, displacement computation, and the output of displacement data. As outlined in Table 2, the “Other Processing” category includes time allocation for activities such as image retrieval and displacement data output, while the “Total Time” category represents the cumulative time necessary to process 100 frames.

Table 2 provides a clear depiction of the superior time efficiency of the RGv2-based method, requiring only one-third of the time compared to the Region Growing-based approach.

3.2. Analysis of the Results of Different Background Removal Algorithms. To quantitatively assess the efficacy of various background removal algorithms, we employ the mean intersection over union (MIOU) and Dice coefficient (Dice) metrics to evaluate the accuracy and consistency of image segmentation methods. Both MIOU and Dice yield values within the range of 0 to 1, where a value closer to 1 indicates a higher degree of overlap between the segmentation result and the ground truth, signifying superior performance [64, 65].

In this section, in addition to employing RGv2, Region Growing, and K-Means algorithms, we introduce deep learning, Vibe [33], and Frame Difference algorithms for comparative analysis. Among these, the Vibe algorithm initially models the backgrounds and subsequently removes them, while the Frame Difference method relies on pixel

TABLE 2: The statistics of algorithm time (unit: s).

	Background removal and displacement calculation	Other processing	Total time
Method based on K-Means	24.578	1.710	26.288
Method based on Region Growing	16.100	1.963	18.063
Method based on RGv2	4.087	1.811	5.898

differences between frames to eliminate backgrounds. The effect of these diverse methods on background removal in varying environments is illustrated in Figure 1, accompanied by the corresponding evaluation indices presented in Table 3.

Upon analyzing the evaluation metrics and Figure 1, it becomes evident that the Frame Difference method exhibits suboptimal performance. This approach operates under the assumption that moving objects display significant pixel value disparities compared to the background. However, in UAV-captured videos, both the pixel values between the cable and the background may exhibit substantial variations, and some background areas might undergo pixel value changes across consecutive frames. Consequently, the Frame Difference method struggles to deliver satisfactory results under these conditions. Moreover, while the Vibe algorithm demonstrates a degree of adaptability to dynamic backgrounds, it faces challenges in accurately segmenting moving cables and is hindered by extended computation times. In contrast, RGv2 proves effective when a pronounced color contrast exists between the cables and the background. However, when the cable and background colors closely resemble each other, as observed in the case of the Jiefang Road Bridge, the results tend to be less favorable. Our proposed RGv2 algorithm, which places greater emphasis on overall grayscale differences and avoids significant local issues, yields higher scores in image segmentation.

Furthermore, RGv2, along with the Region Growing and K-Means algorithms, exhibits certain lighting dependencies. For instance, in the scenario of the Nongye Road Bridge with the sun directly above the cable, UAV photographs of the cables display noticeable light-dark transitions. This may lead to inaccurate recognition of the darker portions of the cables during image segmentation. Nevertheless, background removal in the upper section remains effective, allowing for cable displacement determination through tracking the straight line of their upper edge.

The deep learning results in Figure 1 were chosen from a selection of images that exhibited relatively good performance for comparison with other algorithms. In reality, the efficacy of deep learning methods is contingent upon the size of their datasets. If applied to a different bridge, there would likely be a substantial decrease in accuracy, as illustrated in Figure 16. In fact, even when examining the same bridge, variations in the shooting angle or lighting conditions can result in reduced accuracy, as illustrated in Figure 17. While it is capable of identifying the cable region, this method falls short in differentiating between

multiple cables in scenarios such as cable-stay bridges with overlapping cables. This shortcoming results in errors in cable edge recognition, as evidenced in Figures 1, 16, and 17. In contrast, our proposed RGv2 algorithm is adept at exclusively recognizing the specified cables, a capability clearly illustrated in Figure 17. Furthermore, when two cables are in close proximity, it becomes impractical to employ the region of interest (ROI) approach to analyze a single cable independently. The edge information from multiple independently moving cables can significantly interfere with the subsequent calculation of cable displacement.

In fact, besides RGv2, the results of the line segment detection after background removal by other algorithms should also be evaluated. This is because the cable edge detection results are related to the displacement recognition accuracy directly. The straight lines of cable edges can be obtained by algorithms such as Hough transform and LSD. As Hough transform's computational efficiency is relatively low when the picture is complicated, the LSD algorithm was used to track the cable edge (Figure 2). To demonstrate the clarity of the edge information obtained by RGv2, the results of the line segment detection after background removal by RGv2 are also presented. However, in reality, RGv2 can determine the information of edge straight lines through the location of seed points and the number of times it grows around, without the need for external edge detection algorithms.

It can be observed that ineffective background removal leads to cluttered and numerous straight lines. This significantly disrupts the subsequent process of calculating cable displacement, particularly using the Vibe and Frame Difference method. In contrast, the Region Growing and K-Means algorithms demonstrate more consistent performance in detecting cable edges, notwithstanding some spiral lines on the cable's surface. This may impact the calculation of cable vibrations. Moreover, although these two algorithms have removed most of the background, some remnants still detected by the LSD could potentially interfere with subsequent calculations. Meanwhile, the RGv2 algorithm yields more concise and clear edge detection results.

It is noteworthy that the deep learning method may struggle with cable edge detection. This limitation arises from the difficulty in completely separating the cable from the background, as evident in the zoomed image in Figure 18. The deep learning method would leave a narrow background, which could also cause interference in LSD detection as shown in Figure 2.

TABLE 3: MIOU and Dice of different methods.

	Chaijiaxia Yellow River Bridge		Nongye Road Bridge		Jiefang Road Bridge	
	MIOU	Dice	MIOU	Dice	MIOU	Dice
Region Growing	0.9075	0.9515	0.7130	0.8324	0.6722	0.8040
K-Means	0.5127	0.6778	0.6168	0.7630	0.3544	0.5233
RGv2	0.9368	0.9673	0.7229	0.8392	0.7870	0.8808
Deep Learning	0.7383	0.8495	0.9716	0.9856	0.5915	0.7433
Vibe	0.1986	0.3314	0.1625	0.3215	0.3465	0.5652
Frame Difference	0.0068	0.0135	0.4923	0.6598	0.1436	0.2512



FIGURE 16: Application of deep learning methods to other bridge scenarios.

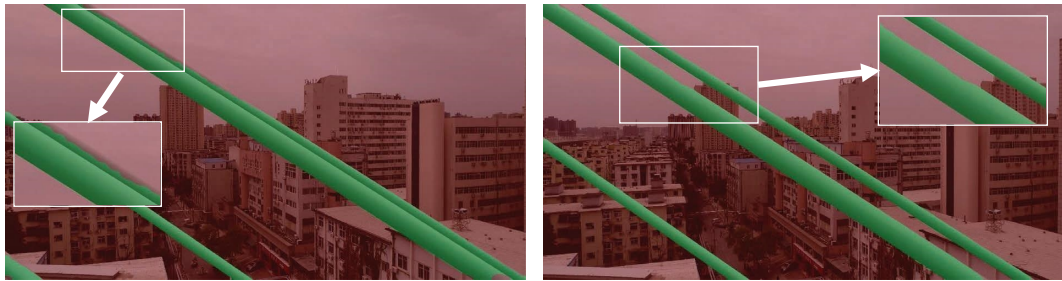


FIGURE 17: Results for the Jiefang Road Bridge with changes in lighting or shooting angle.

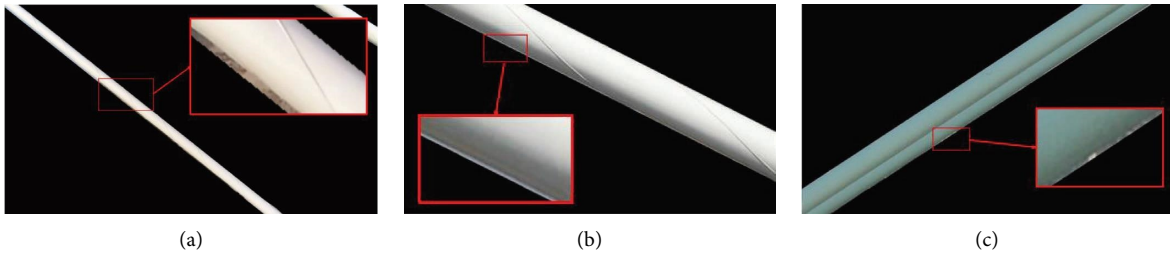


FIGURE 18: Partial details of the results of deep learning methods. (a) Detailed view of Chaijiaxia Yellow River Bridge; (b) detailed view of Nongye Road Bridge; (c) detailed view of Jiefang Road Bridge.

3.3. Analysis of the Inclined Cable's Vibration Frequency.

To test the reliability of the above two background removal methods, a short-time condition of 40 seconds and a long-time condition of 300 seconds were collected by the UAV for displacement time history identification. Afterwards, the frequency domain diagram is obtained by FFT for comparison.

3.3.1. Result of the Short-Time Condition. Figure 14(a) presents the displacement time history results for the short-time condition. Notably, the results obtained without employing background removal exhibit numerous abrupt value changes. This phenomenon arises due to the presence of interference lines in the background when background removal is not applied. In contrast, the results obtained using

the three background removal algorithms display closer proximity, characterized by improved continuity and robustness in the detection results. Furthermore, no abrupt changes occur in the monitoring results. The slight discrepancies among the results obtained with the three background removal algorithms can be attributed to variations in the identification of inclined cable edges, as depicted in Figure 13(b).

Figure 14(b) exhibits the FFT results of the data presented in Figure 14(a). To facilitate comparison with the accelerometer data and account for the low energy of the low-order modes recorded by the accelerometer, the results obtained with background removal algorithms have been attenuated in the frequency domain, as demonstrated in Figure 14(b).

In summary, a substantial disparity is observed between the results obtained with and without background removal. The results lacking background removal exhibit complex and multifaceted frequency patterns. Conversely, the results obtained with background removal clearly manifest 3–5 discernible peaks when the frequency surpasses 1 Hz. The frequency differences among these peaks align with the vibration characteristics of the inclined cable. This underscores the crucial nature of processing the original video using background removal methods prior to edge detection of the inclined cable.

3.3.2. Result of the Long-Time Condition. Figure 15(a) presents the displacement time history and frequency domain analysis for an extended period of 300 s. Remarkably, most of the observations made under the long-time condition align with those from the short-time condition. The long-time condition also yields favorable detection results when compared to the short-time condition, underscoring the robustness of the proposed method.

In Figure 15(b), the displacement time history obtained from UAV video captures exhibits higher energy at low frequencies, gradually diminishing as frequencies increase. In contrast, the accelerometer data exhibit the inverse trend, with energy increasing as the frequency rises. Notably, the Region Growing and K-Means algorithms can detect three vibration frequencies, while RGv2 can detect five vibration frequencies. This highlights the superior performance of RGv2 in frequency domain analysis.

However, it is important to note that both short-time and long-time conditions feature complex frequency patterns with elevated energy in the low-frequency range, which can complicate the automatic detection of vibration frequencies. To address this challenge, we employ EMD to analyze the obtained displacement time history in subsequent steps.

3.3.3. Comparison with the Accelerometer. The instability inherent to a hovering UAV during video capture introduces an additional influence on measurement results. Notably, the primary energy source related to the UAV's motion consists primarily of a low-frequency component [66], resulting in elevated and complex energy levels below 1 Hz

in the frequency domain (as depicted in Figure 15(b)). Consequently, the proposed method is unable to measure the first vibration frequency of the inclined cable. However, this limitation does not hinder the manual identification of higher-order vibration frequencies.

As indicated in Table 4, three background removal methods are capable of observing 3–5 frequencies, closely aligning with the accelerometer's results. Notably, the proposed methods exhibit consistency with the accelerometer results in measuring modes 2, 3, and 4 of the inclined cable, as shown in Table 4. Furthermore, Table 5 provides a comparison of the frequency differences obtained by the three methods relative to the accelerometer.

Table 5 reveals that the RGv2 method demonstrates mean relative errors of 0.89% and 0.71% across the two working conditions. In contrast, the Region Growing method exhibits mean relative errors of 1.83% and 1.03% for the same conditions, while the K-Means method presents mean relative errors of 3.43% and 0.71%. Consequently, the RGv2 method stands out by delivering superior results in terms of accuracy and consistency.

3.4. Vibration Frequency Analysis after EMD. To enhance the accuracy of detecting the vibration frequency of the inclined cable while mitigating the influence of the UAV's inherent vibration, EMD is employed to analyze the displacement time history under both short-time and long-time conditions. The decomposition process of the long-term condition using the RGv2 algorithm is illustrated in Figure 19(a). In this process, the original data are decomposed into 8 IMFs and one residual, with the IMFs organized in descending order of frequency. As demonstrated in Figures 14(b) and 15(b), the UAV's motion exhibits relatively low frequencies, resulting in frequency overlap between the UAV's motion and the inclined cable's motion in the low-frequency range.

To evaluate the EMD data of the three cables on the three different bridges, the evaluation metrics discussed in Section 2.2 are employed and the results are presented in Table 6. This table facilitates the identification and selection of the most suitable IMFs. For instance, the highest scores among the three datasets correspond to IMF2, IMF1, and IMF2, respectively. As illustrated in Figure 19, it becomes evident that IMF2, IMF1, and IMF2 yield the most favorable results for the three cases, confirming the accuracy of the method proposed in this paper. The optimal IMF is visually represented in Figure 20.

The optimal IMF is selected for both the long-term and short-term conditions of the Chaijiaxia Yellow River Bridge using this approach. Comparing the calculated results with accelerometer data (Figure 21), it is apparent that the frequency peaks become more distinct and prominent with the application of this processing method.

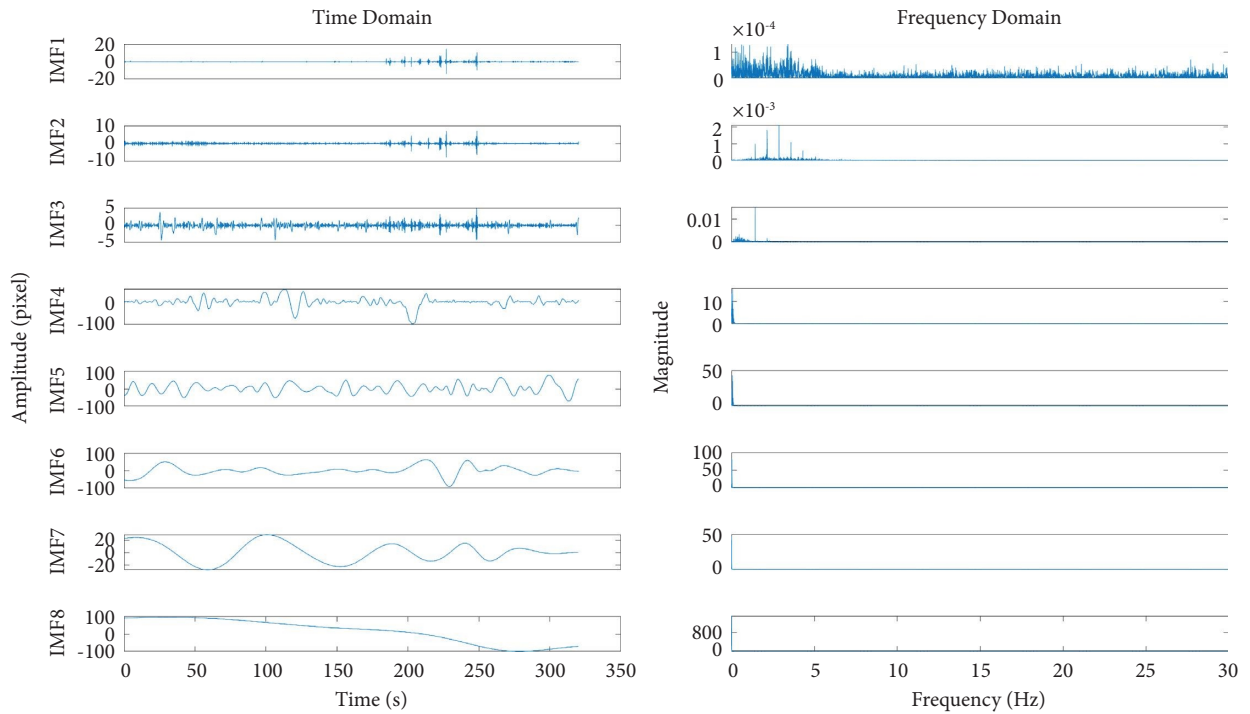
The frequencies obtained after implementing the proposed approach are summarized in Tables 7 and 8. In comparison to Table 5, this approach notably reduces the relative error in frequency disparities for all three methods. Furthermore, it aids K-Means in identifying a greater number of frequency peaks in the short working conditions.

TABLE 4: Detected vibration frequencies (unit: Hz).

Condition	Method	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6
Short-time condition	Accelerometer	—	1.4553	2.1662	2.8966	3.6242	4.3518
	K-Means	—	1.4486	2.1479	2.8472	—	—
	Region Growing	—	1.4467	2.1452	2.8685	—	—
	RGv2	—	1.4480	2.1470	2.8709	3.5949	4.3189
Long-time condition	Accelerometer	—	1.4570	2.1680	2.8952	3.6250	4.3522
	K-Means	—	1.4326	2.1599	2.8684	3.5863	4.3073
	Region Growing	—	1.4326	2.1537	2.8653	—	—
	RGv2	—	1.4318	2.1477	2.8664	3.5878	4.3064

TABLE 5: The detected frequency differences.

Condition	Method	1-2 (Hz)	2-3 (Hz)	3-4 (Hz)	4-5 (Hz)	5-6 (Hz)	Mean relative error
Short-time condition	Accelerometer	—	0.7109	0.7305	0.7276	0.7276	—
	K-Means	—	0.6993	0.6993	—	—	3.43%
	Region Growing	—	0.6985	0.7233	—	—	1.83%
	RGv2	—	0.6990	0.7239	0.7240	0.7240	0.89%
Long-time condition	Accelerometer	—	0.7110	0.7272	0.7298	0.7272	—
	K-Means	—	0.7273	0.7085	0.7179	0.7210	0.71%
	Region Growing	—	0.7211	0.7116	—	—	1.03%
	RGv2	—	0.7159	0.7187	0.7214	0.7186	0.71%



(a)
FIGURE 19: Continued.

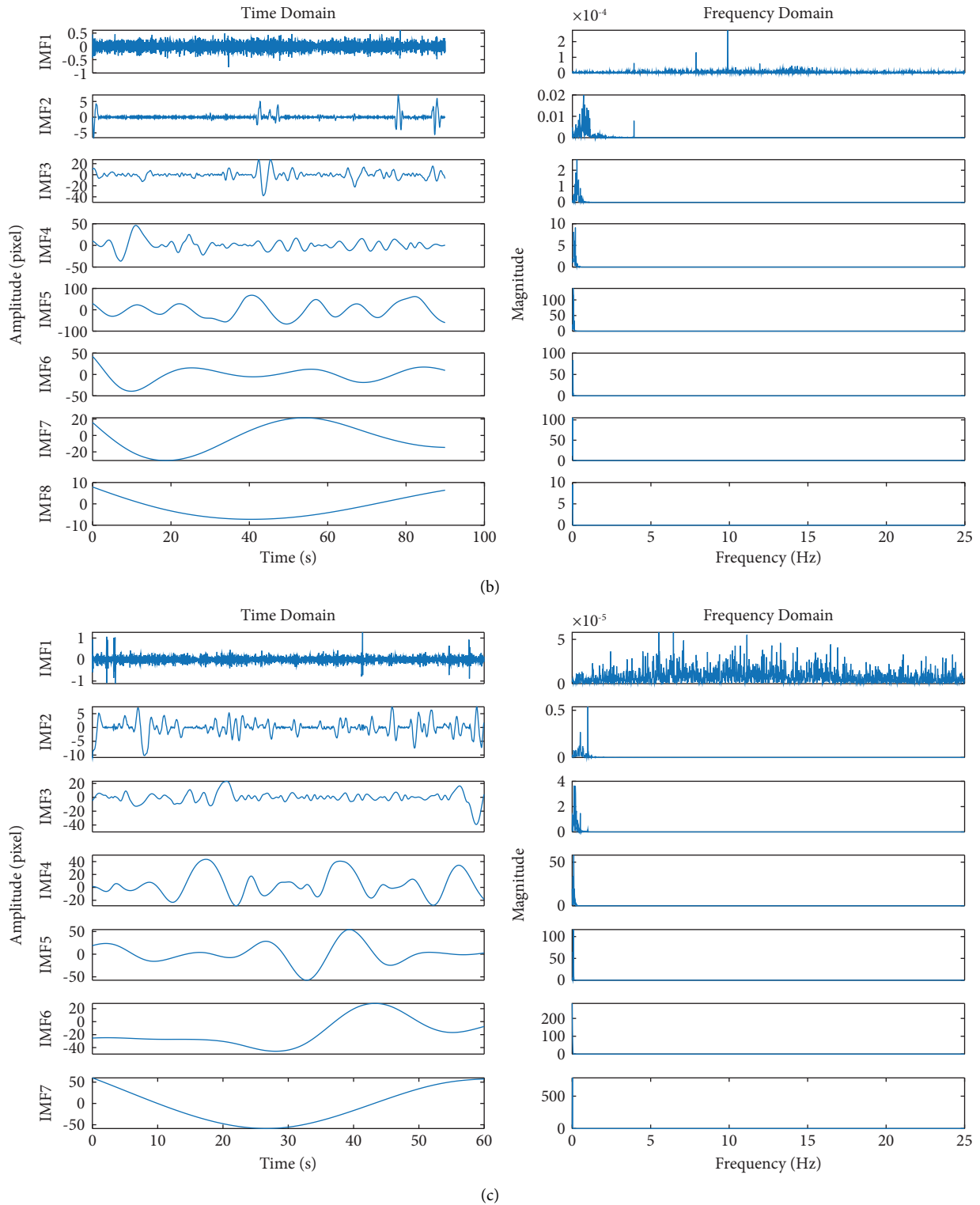


FIGURE 19: Decomposition results of EMD. (a) Data of Chaijiaxia Yellow River Bridge; (b) data of Nongye Road Bridge; (c) data of Jiefang Road Bridge.

In summary, EMD, which is well suited for handling data fluctuations induced by a UAV's own vibrations, proves to be a superior method for accurately extracting the vibration frequencies of inclined cables.

To better illustrate the effectiveness of the proposed UAV motion filtering method, this article compares some filtering methods similar to EMD. Based on the EMD algorithm, scholars have proposed various improved

TABLE 6: Calculated T_n using the proposed method.

	IMF1	IMF2	IMF3
Chaijiaxia Yellow River Bridge	0.7849	2.1580	1.5629
Nongye Road Bridge	1.3023	0.2718	—
Jiefang Road Bridge	0.9255	1.3712	—

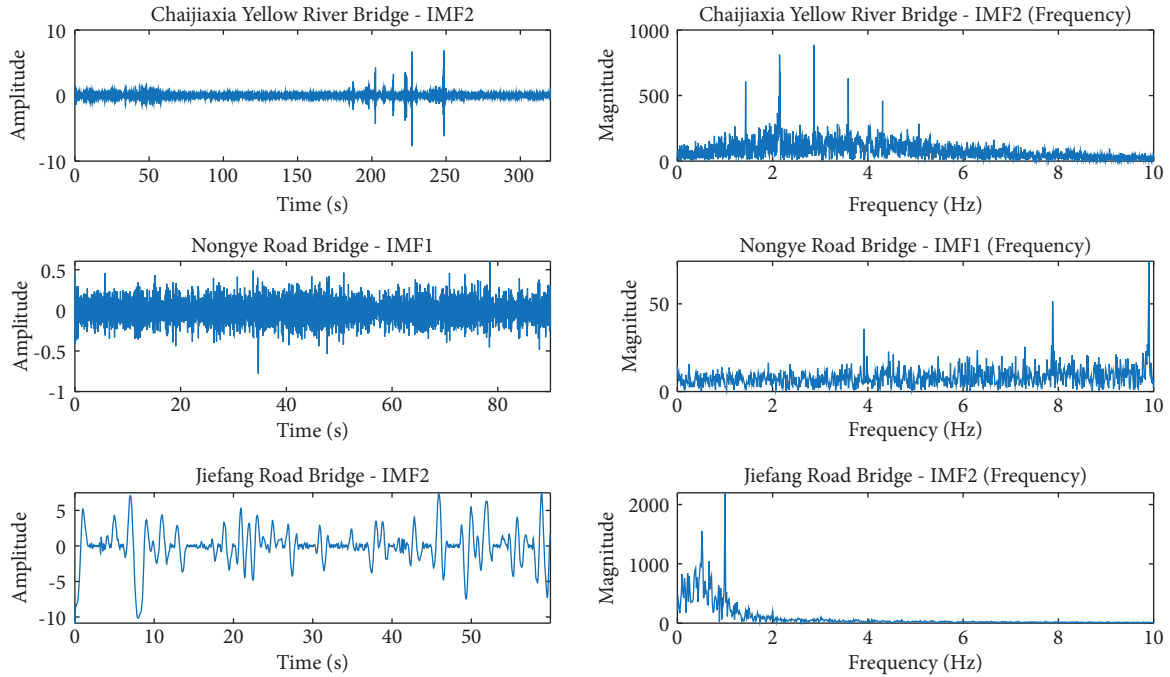


FIGURE 20: The superimposed results of the highest rated IMFs using EMD.

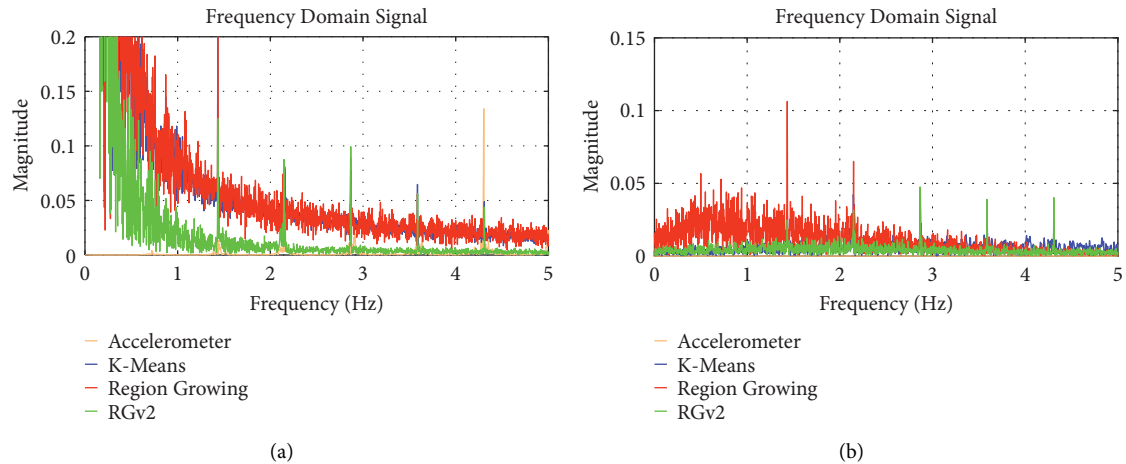


FIGURE 21: Comparison before and after EMD for the long-time condition: (a) before using the proposed method; (b) after using the proposed method.

algorithms, such as the ensemble empirical mode decomposition (EEMD) [67] algorithm and the variational mode decomposition (VMD) algorithm. EEMD is an improved method where white noise is added to the original signal before each EMD iteration, but it also has the limitation of large computational complexity and

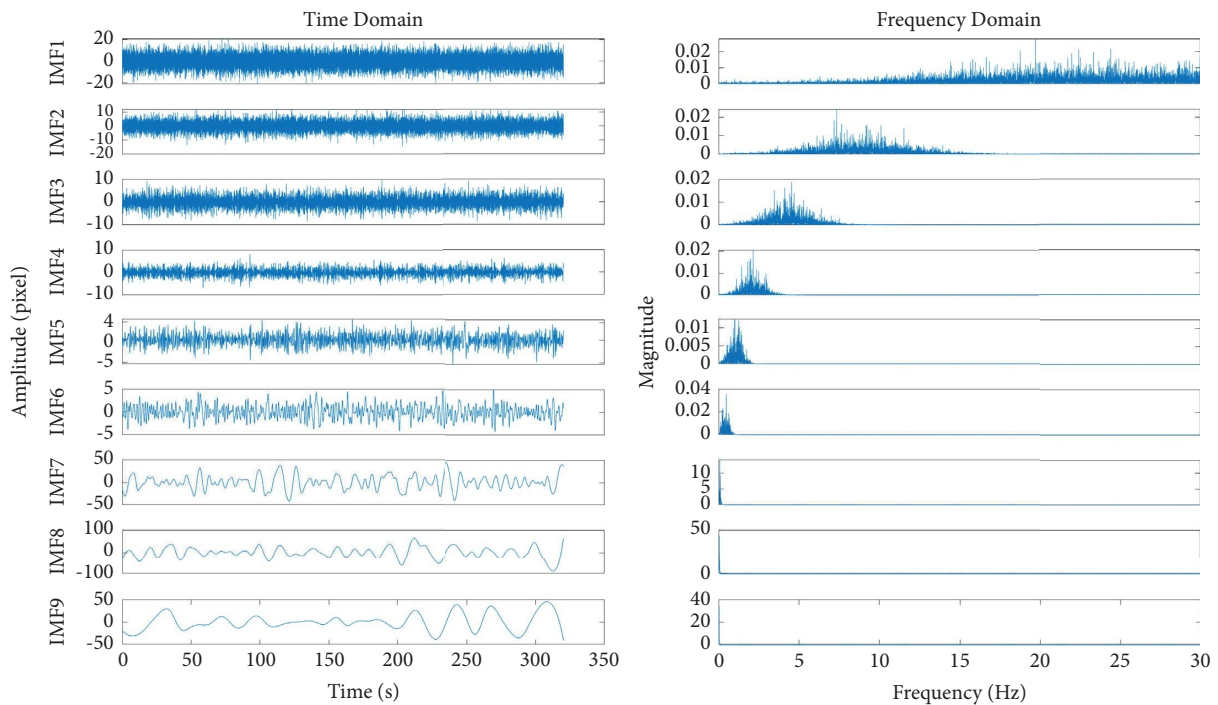
interference. As shown in Figure 22, the frequency peak after decomposition is unclear and will cause interference to the frequency difference calculation. Besides, VMD determines the frequency center and bandwidth of each IMF by iteratively searching for the optimal solution of the variational model. Thus, VMD could achieve signal

TABLE 7: Vibration frequencies after the proposed method (unit: Hz).

Condition	Method	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6
Short-time condition	Accelerometer	—	1.4553	2.1662	2.8966	3.6242	4.3518
	K-Means	—	1.4253	2.1625	2.8751	3.5878	4.3004
	Region Growing	—	1.4241	2.1486	2.8731	—	—
	RGv2	—	1.4241	2.1486	2.8731	3.5976	4.3221
Long-time condition	Accelerometer	—	1.4570	2.1680	2.8952	3.6250	4.3522
	K-Means	—	1.4331	2.1450	2.8662	3.5875	4.3087
	Region Growing	—	1.4331	2.1481	2.8662	—	—
	RGv2	—	1.4331	2.1481	2.8662	3.5875	4.3087

TABLE 8: Frequency differences after the proposed method.

Condition	Method	1-2 (Hz)	2-3 (Hz)	3-4 (Hz)	4-5 (Hz)	5-6 (Hz)	Mean relative error
Short-time condition	Accelerometer	—	0.7109	0.7304	0.7276	0.7276	—
	K-Means	—	0.7372	0.7126	0.7127	0.7126	0.11%
	Region Growing	—	0.7245	0.7245	—	—	0.05%
	RGv2	—	0.7245	0.7245	0.7245	0.7245	0.05%
Long-time condition	Accelerometer	—	0.7110	0.7272	0.7298	0.7272	—
	K-Means	—	0.7119	0.7212	0.7213	0.7212	0.68%
	Region Growing	—	0.7150	0.7181	—	—	1.00%
	RGv2	—	0.7150	0.7181	0.7213	0.7212	0.68%



(a)
FIGURE 22: Continued.

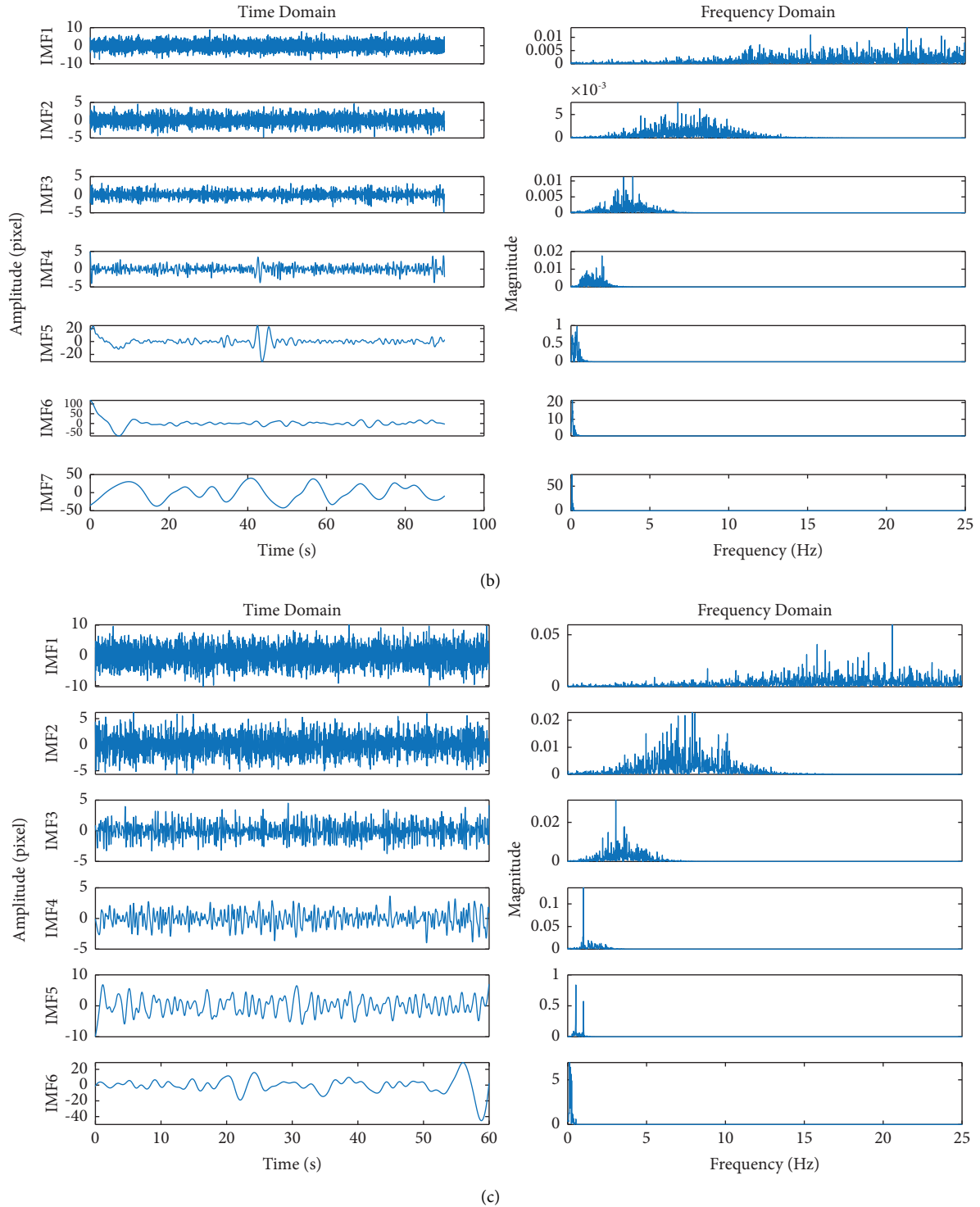
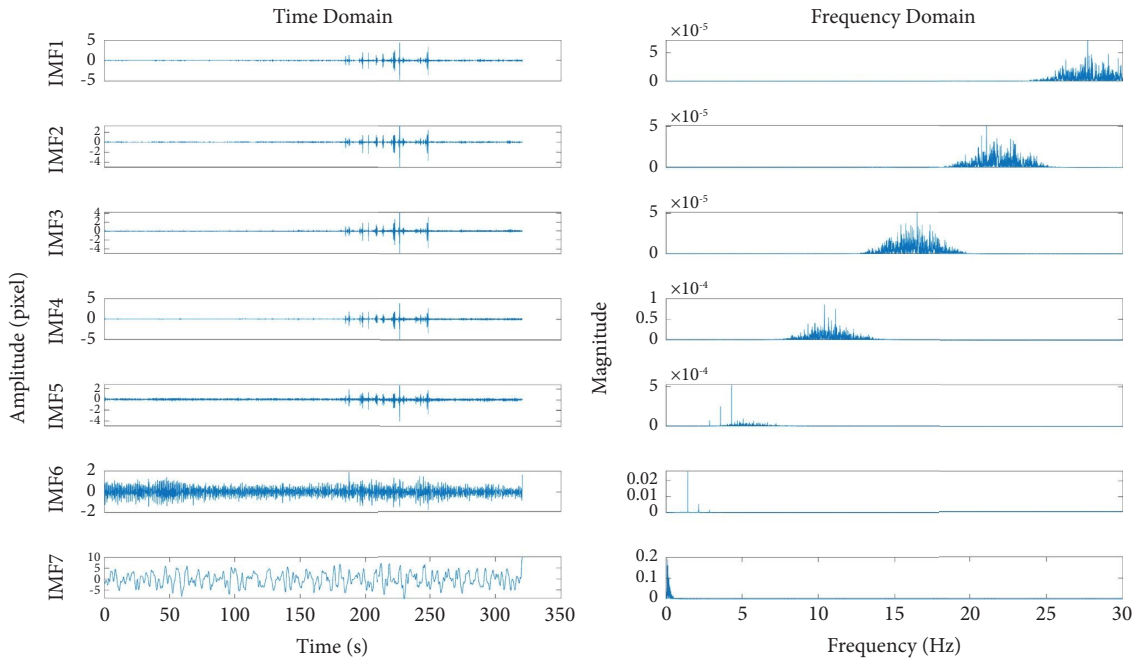


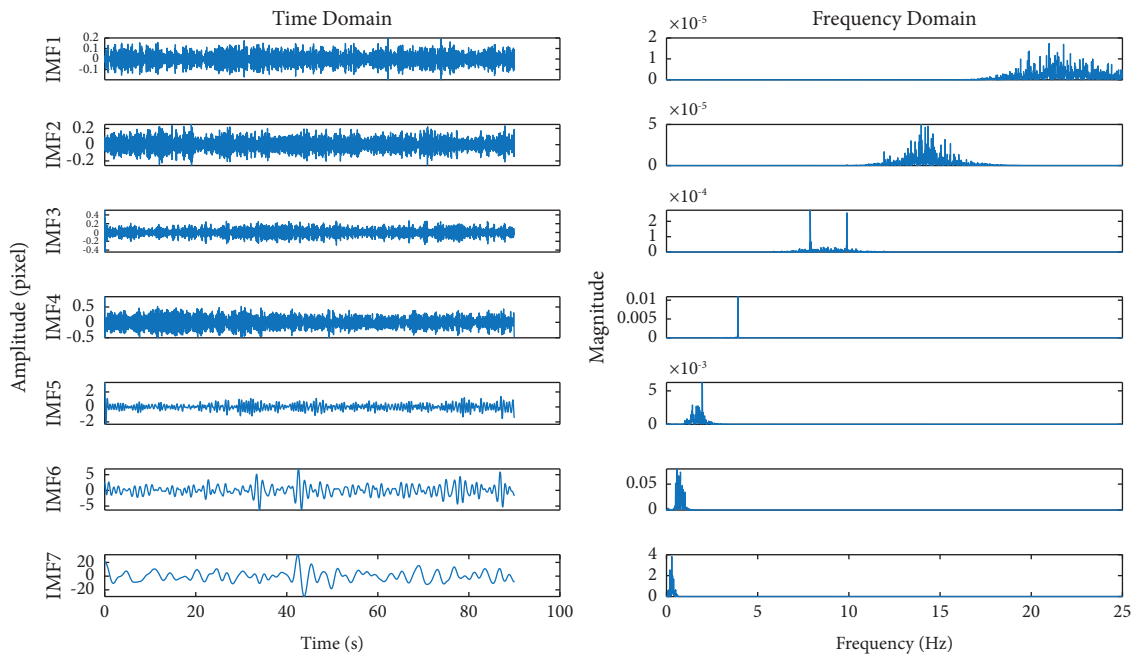
FIGURE 22: Decomposition results of EEMD. (a) Data of Chaijiaxia Yellow River Bridge; (b) data of Nongye Road Bridge; (c) data of Jiefang Road Bridge.

frequency domain division and effective separation of each IMF, as shown in Figure 23. In the decomposition of relative motion of cables, the frequency peaks obtained are the same as those obtained by EMD. However, VMD's decomposition number K lacks unified theoretical guidance and needs to be determined artificially [68]. They currently have numerous applications [69–72].

At the same time, in order to compare with EEMD and VMD, this study uses the average of Kullback–Leibler divergence (KLD) of all IMFs to evaluate the decomposition effect of each decomposition method. KLD calculates the relative entropy between two random signals from a probabilistic perspective and effectively quantifies their differences. The calculation results are shown in Table 9.



(a)



(b)

FIGURE 23: Continued.

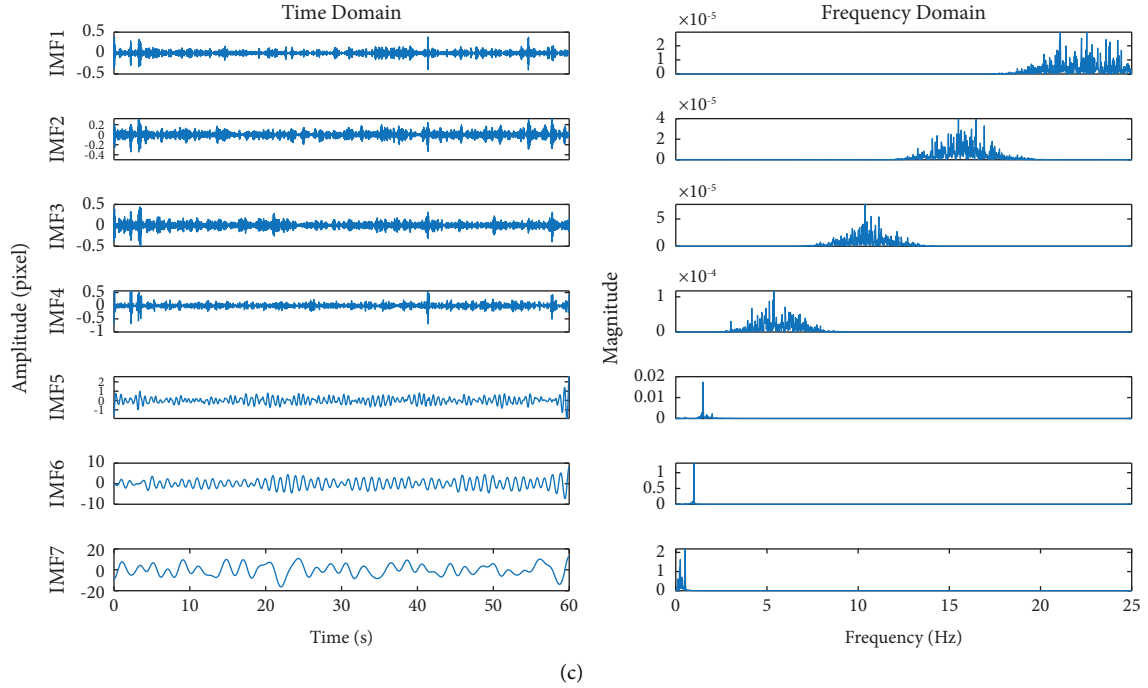


FIGURE 23: Decomposition results of VMD. (a) Data of Chaijiaxia Yellow River Bridge; (b) data of Nongye Road Bridge; (c) data of Jiefang Road Bridge.

TABLE 9: KLD of different decomposition methods.

	Chaijiaxia Yellow River Bridge	Nongye Road Bridge	Jiefang Road Bridge
EMD	11.1903	3.4869	3.6537
EEMD	5.8947	3.8947	3.5865
VMD	17.8596	2.6892	6.1890

It can be found that the average KLD of EMD is generally smaller, which indicates that the decomposed distribution, in terms of shape or probability mass allocation, is relatively similar to the original distribution.

The IMF selection method proposed in Section 2.2 is also used in this section to the decomposition results of EEMD and VMD, and the highest rated IMFs are shown in Figures 24 and 25, respectively. However, due to the characteristics of EEMD decomposition, it has difficulties to clearly distinguish frequency peaks. Therefore, the IMF selection method cannot be effectively used in conjunction with EEMD. In contrast, VMD can clearly decompose modes into different IMFs. However, the highest-scoring IMF often only contains one frequency peak, making it impossible to obtain frequency differences. For this reason, the IMF selection method is also difficult to use in conjunction with VMD. Among them, the T_n scores of the results computed by EMD, EEMD, and VMD are shown in Tables 6, 10, and 11, respectively. Due to the fact that the frequency difference can be obtained solely based on the IMF with the highest T_n score, the proposed selecting IMF method is more applicable

when used in conjunction with EMD. The above analyses reflect the applicability of the IMF selection method and can be used to select the ideal IMFs automatically.

3.5. Cable Force Estimation. In this section, we utilize the average frequency differences obtained from various algorithms to compute the cable force using the following formula [16]:

$$T = 4mL^2 \left(\frac{\Delta f_{rs}}{s-r} \right)^2, \quad (5)$$

where m is the unit mass of the cable, L is the length of the cable, r and s are the modal orders of the correspondingly identified frequencies, and Δf_{rs} is the difference between the r th and s th frequencies. The results of this calculation are presented in Figure 26 and Table 12. It becomes evident that the cable force computed using the proposed method aligns closely with the values obtained via the accelerometer. This underscores the smaller error margin associated with the

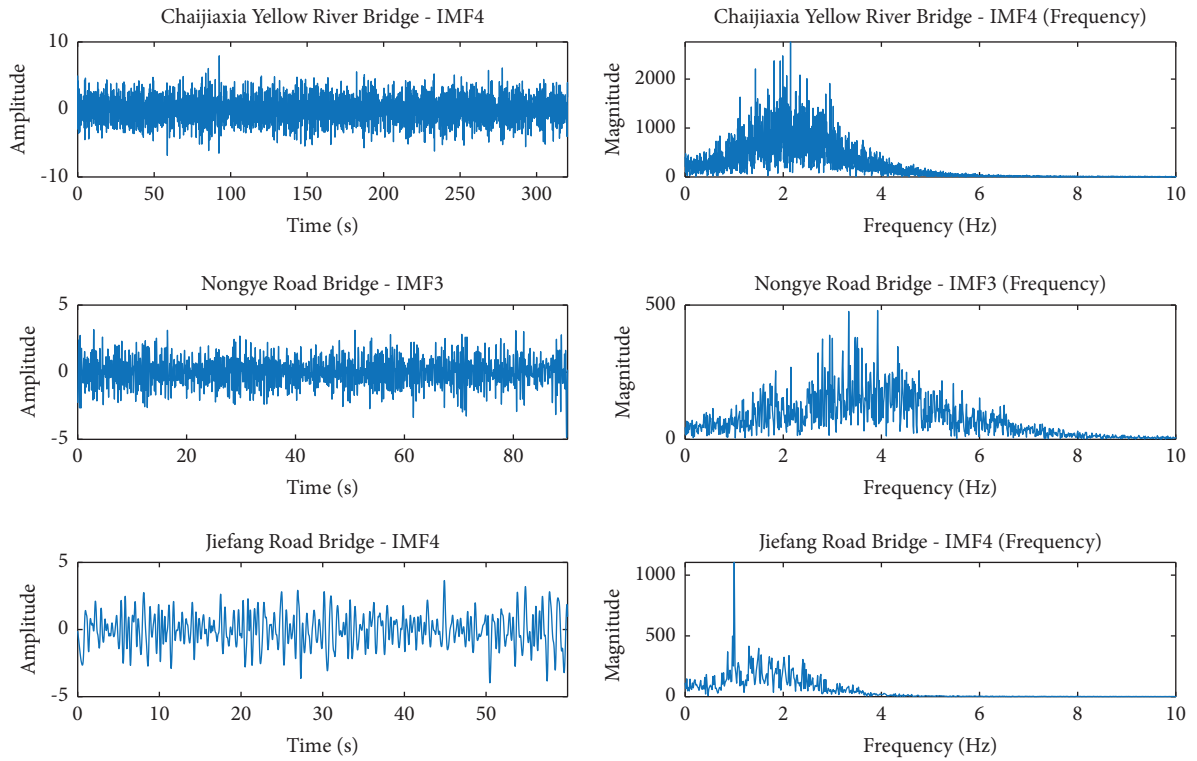


FIGURE 24: The superimposed results of the highest rated IMFs using EEMD.

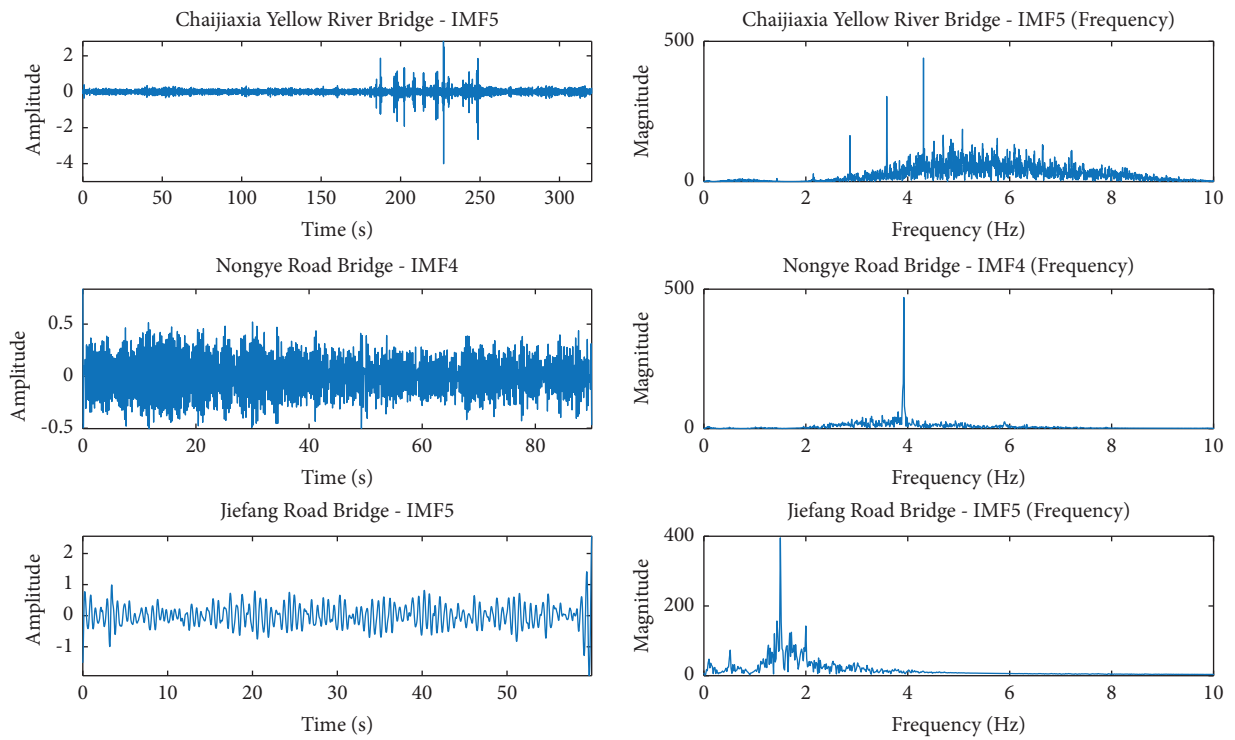


FIGURE 25: The superimposed results of the highest rated IMFs using VMD.

TABLE 10: Calculated T_n of EEMD using the proposed method.

	IMF1	IMF2	IMF3	IMF4	IMF5
Chaijiaxia Yellow River Bridge	0.7590	1.4296	1.3657	1.6206	1.0072
Nongye Road Bridge	0.3153	0.8222	1.0745	0.8824	—
Jiefang Road Bridge	1.3348	0.9375	1.6146	2.0456	1.5103

TABLE 11: Calculated T_n of VMD using the proposed method.

	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6
Chaijiaxia Yellow River Bridge	1.1621	1.1071	0.6932	1.2952	1.9930	1.7255
Nongye Road Bridge	0.7022	1.6458	1.5318	2.3726	1.1972	0.5315
Jiefang Road Bridge	0.3444	0.7410	1.3593	1.0046	1.7656	1.6641

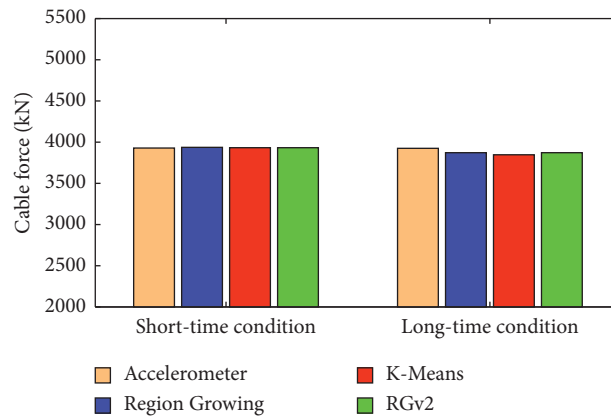


FIGURE 26: Diagram of the cable force comparison.

TABLE 12: Comparison of cable force estimation.

Condition	Method	Cable force (KN)	Absolute error (KN)	Mean relative error
Short-time condition	Accelerometer	3928.7766	0	0
	Region Growing	3937.1907	8.4141	0.21%
	K-Means	3932.8468	4.0702	0.10%
	RGv2	3932.8468	4.0702	0.10%
Long-time condition	Accelerometer	3925.2508	0	0
	Region Growing	3872.2841	52.9666	1.35%
	K-Means	3847.0094	78.2413	2.00%
	RGv2	3872.2841	52.9666	1.35%

RGv2 method proposed in this study during the final cable force computation.

4. Conclusion

This study utilizes UAV technology to record the vibrational motion of an inclined cable, capturing these data in video format. To extract the vibrational time history of the inclined cable, complex backgrounds within the video were eliminated using both a Region Growing algorithm and a K-Means algorithm. Due to various limitations in the Region Growing algorithm, an enhanced version named RGv2 was developed and applied for image segmentation and displacement calculation. Subsequently, the displacement time history was analyzed using EMD. The final step involved

employing the derived frequency differences to compute the cable force. The principal findings are as follows:

- (1) In the context of background removal and displacement calculation, the RGv2 algorithm demonstrates a higher accuracy and shorter processing time compared to the methods based on K-Means and Region Growing. In addition, the RGv2 algorithm achieves higher MIOU and Dice scores in background removal.
- (2) All three background removal algorithms successfully identify the 2nd–5th cable vibration frequencies, demonstrating an average relative error of less than 3.43%. However, RGv2 outperforms the others by maintaining the error below 0.89% and

consistently identifying five frequency peaks in both long and short working conditions.

- (3) Utilizing EMD, the study introduces a method for automatically selecting IMFs containing clearer peak frequency information and obtaining the frequency difference of the stay cable vibration. This approach enhances the accuracy of vibration frequency identification.
- (4) Cable force computation: by utilizing the frequency differences derived from the background removal methods and EMD, the relative error in estimating the cable force is limited to below 2%. Specifically, the cable force error calculated from the frequency differences detected by RGV2 remains within 1.35%.

Data Availability

The data used to support the findings of this study are included within the Supplementary Files.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors are grateful for the financial support from the National Natural Science Foundation of China (52108290) and Key Scientific and Technological Research Projects of Henan Province (212102310975 and 222102320436).

Supplementary Materials

The 1630928.f1.docx in the “Supplemental Files” contains the download link for the video of the cable-stayed bridge vibrations captured by our UAV. (*Supplementary Materials*)

References

- [1] L. Zhang, G. Qiu, and Z. Chen, “Structural health monitoring methods of cables in cable-stayed bridge: a review,” *Measurement*, vol. 168, Article ID 108343, 2021.
- [2] M. M. Alamdari, N. L. Dang Khoa, Y. Wang, B. Samali, and X. Q. Zhu, “A multi-way data analysis approach for structural health monitoring of a cable-stayed bridge,” *Structural Health Monitoring*, vol. 18, no. 1, pp. 35–48, 2019.
- [3] B. H. Kim and T. Park, “Estimation of cable tension force using the frequency-based system identification method,” *Journal of Sound and Vibration*, vol. 304, no. 3–5, pp. 660–676, 2007.
- [4] S. Sony, S. Laventure, and A. Sadhu, “A literature review of next-generation smart sensing technology in structural health monitoring,” *Structural Control and Health Monitoring*, vol. 26, no. 3, Article ID e2321, 2019.
- [5] S. Zhan, Z. Li, J. Hu, Y. Liang, and G. Zhang, “Model order identification for cable force estimation using a Markov chain Monte Carlo-based bayesian approach,” *Sensors*, vol. 18, no. 12, p. 4187, 2018.
- [6] Y. Yang, S. Li, S. Nagarajaiah, H. Li, and P. Zhou, “Real-time output-only identification of time-varying cable tension from accelerations via complexity pursuit,” *Journal of Structural Engineering*, vol. 142, no. 1, Article ID 4015083, 2016.
- [7] D. Jana and S. Nagarajaiah, “Computer vision-based real-time cable tension estimation in Dubrovnik cable-stayed bridge using moving handheld video camera,” *Structural Control and Health Monitoring*, vol. 28, no. 5, pp. 1–22, 2021.
- [8] S. Sim, J. Li, H. Jo et al., “A wireless smart sensor network for automated monitoring of cable tension,” *Smart Materials and Structures*, vol. 23, no. 2, Article ID 025006, 2013.
- [9] A. Cunha and E. Caetano, “Dynamic measurements on stay cables of cable-stayed bridges using an interferometry laser system,” *Experimental Techniques*, vol. 23, no. 3, pp. 38–43, 1999.
- [10] W. Zhao, G. Zhang, and J. Zhang, “Cable force estimation of a long-span cable-stayed bridge with microwave interferometric radar,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 35, no. 12, pp. 1419–1433, 2020.
- [11] S. Wangchuk, D. M. Siringoringo, and Y. Fujino, “Modal analysis and tension estimation of stay cables using non-contact vision-based motion magnification method,” *Structural Control and Health Monitoring*, vol. 29, no. 7, Article ID e2957, 2022.
- [12] R. W. Gehle and S. F. Masri, “Tracking the multi-component motion of a cable using a television camera,” *Smart Materials and Structures*, vol. 7, no. 1, pp. 43–51, 1998.
- [13] J. Guo, X. Wu, J. Liu et al., “Non-contact vibration sensor using deep learning and image processing,” *Measurement*, vol. 183, Article ID 109823, 2021.
- [14] S. Kim, B. Jeon, J. Cheung, S. Kim, and J. Park, “Stay cable tension estimation using a vision-based monitoring system under various weather conditions,” *Journal of Civil Structural Health Monitoring*, vol. 7, no. 3, pp. 343–357, 2017.
- [15] X. Zhao, K. Ri, and N. Wang, “Experimental verification for cable force estimation using handheld shooting of smartphones,” *Journal of Sensors*, vol. 2017, no. 1, Article ID 5625396, 13 pages, 2017.
- [16] Y. Tian, C. Zhang, S. Jiang, J. Zhang, and W. Duan, “Non-contact cable force estimation with unmanned aerial vehicle and computer vision,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 1, pp. 73–88, 2021.
- [17] S. Zhao, F. Kang, J. Li, and C. Ma, “Structural health monitoring and inspection of dams based on UAV photogrammetry with image 3D reconstruction,” *Automation in Construction*, vol. 130, Article ID 103832, 2021.
- [18] B. J. Perry, Y. Guo, R. Atadero, and J. W. van de Lindt, “Streamlined bridge inspection system utilizing unmanned aerial vehicles (UAVs) and machine learning,” *Measurement*, vol. 164, Article ID 108048, 2020.
- [19] A. Khaloo and D. Lattanzi, “Automatic detection of structural deficiencies using 4D hue-assisted analysis of color point clouds,” *Dynamics of Civil Structures*, vol. 2, pp. 197–205, 2019.
- [20] C. Xiong, Q. Li, and X. Lu, “Automated regional seismic damage assessment of buildings using an unmanned aerial vehicle and a convolutional neural network,” *Automation in Construction*, vol. 109, Article ID 102994, 2020.
- [21] M. Li, D. Cheng, X. Yang et al., “High precision slope deformation monitoring by UAV with industrial photogrammetry,” *Institute of Physics Conference Series: Earth and Environmental Science*, vol. 636, no. 1, Article ID 012015, 2021.
- [22] D. Choi, W. Bell, D. Kim, and J. Kim, “UAV-driven structural crack detection and location determination using convolutional neural networks,” *Sensors*, vol. 21, no. 8, p. 2650, 2021.

- [23] D. Kang and Y. Cha, "Autonomous UAVs for structural health monitoring using deep learning and an ultrasonic beacon system with geo-tagging," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 10, pp. 885–902, 2018.
- [24] B. Lei, Y. L. Ren, N. Wang, L. S. Huo, and G. B. Song, "Design of a new low-cost unmanned aerial vehicle and vision-based concrete crack inspection method," *Structural Health Monitoring*, vol. 19, no. 6, pp. 1871–1883, 2020.
- [25] Y. Weng, J. Shan, Z. Lu, X. Lu, and B. F. Spencer, "Homography-based structural displacement measurement for large structures using unmanned aerial vehicles," *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 9, pp. 1114–1128, 2021.
- [26] V. Hoskere, J. Park, H. Yoon, and B. F. Spencer, "Vision-based modal survey of civil infrastructure using unmanned aerial vehicles," *Journal of Structural Engineering*, vol. 145, no. 7, pp. 1–14, 2019.
- [27] A. Khadka, B. Fick, A. Afshar, M. Tavakoli, and J. Baqersad, "Non-contact vibration monitoring of rotating wind turbines using a semi-autonomous UAV," *Mechanical Systems and Signal Processing*, vol. 138, Article ID 106446, 2020.
- [28] S. Lee and C. Lee, "Low-complexity background subtraction based on spatial similarity," *European Association for Signal Processing Journal on Image and Video Processing*, vol. 2014, no. 1, p. 30, 2014.
- [29] M. Hung, J. Pan, and C. Hsieh, "Speed up temporal median filter for background subtraction," in *Proceedings of the 2010 First International Conference on Pervasive Computing, Signal Processing and Applications (PCSPA 2010)*, pp. 297–300, Harbin, China, September 2010.
- [30] L. Wei, S. Zhao-Yun, H. Ju, S. Ai-Min, and Z. Juan, "Pavement crack type judgment method based on three-dimensional pavement data," *China Journal of Highway and Transport*, vol. 28, no. 11, pp. 21–28, 2015.
- [31] H. Wei and Q. Peng, "A block-wise Frame Difference method for real-time video motion detection," *International Journal of Advanced Robotic Systems*, vol. 15, no. 4, 2018.
- [32] N. Katsarakis, A. Pnevmatikakis, Z. Tan, and R. Prasad, "Improved Gaussian mixture models for adaptive foreground segmentation," *Wireless Personal Communications*, vol. 87, no. 3, pp. 629–643, 2016.
- [33] O. Barnich and M. Van Droogenbroeck, "ViBe: a universal background subtraction algorithm for video sequences," *IEEE Transactions on Image Processing*, vol. 20, no. 6, pp. 1709–1724, 2011.
- [34] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, 2017.
- [35] A. Ji, X. Xue, Y. Wang, X. Luo, and W. Xue, "An integrated approach to automatic pixel-level crack detection and quantification of asphalt pavement," *Automation in Construction*, vol. 114, Article ID 103176, 2020.
- [36] C. Zhang, Y. Tian, and J. Zhang, "Complex image background segmentation for cable force estimation of urban bridges with drone-captured video and deep learning," *Structural Control and Health Monitoring*, vol. 29, no. 4, pp. 1–17, 2022.
- [37] V. Kamath and A. Renuka, "Deep learning based object detection for resource constrained devices: systematic review, future trends and challenges ahead," *Neurocomputing*, vol. 531, no. C, pp. 34–60, 2023.
- [38] J. Johnson, A. Karpathy, and L. Fei-Fei, "Densecap: fully convolutional localization networks for dense captioning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4565–4574, Seattle, WA, USA, June 2016.
- [39] D. K. Mahajan, R. B. Girshick, V. Ramanathan et al., "Exploring the limits of weakly supervised pretraining," in *Proceedings of the 15th European Conference on Computer Vision*, pp. 185–201, Munich, Germany, September 2018.
- [40] S. A. Hojjatoleslami and J. Kittler, "Region growing: a new approach," *IEEE Transactions on Image Processing*, vol. 7, no. 7, pp. 1079–1084, 1998.
- [41] J. Shao, W. Zhang, A. Shen et al., "Seed point set-based building roof extraction from airborne LiDAR point clouds using a top-down strategy," *Automation in Construction*, vol. 126, Article ID 103660, 2021.
- [42] T. K. Lin, "A Region-Growing adaptation-based algorithm for the smart detection of remote sensing images," *Advanced Materials Research*, vol. 787, pp. 932–937, 2013.
- [43] X. Lu, J. Wu, X. Ren, B. Zhang, and Y. Li, "The study and application of the improved Region Growing algorithm for liver segmentation," *Optik*, vol. 125, no. 9, pp. 2142–2147, 2014.
- [44] X. Y. Wang and Y. P. Bai, "The global Minmax k-means algorithm," *SpringerPlus*, vol. 5, no. 1, p. 1665, 2016.
- [45] M. Ding, "Quantitative contrast of urban agglomeration colors based on image clustering algorithm: case study of the Xia-Zhang-Quan metropolitan area," *Frontiers of Architectural Research*, vol. 10, no. 3, pp. 692–700, 2021.
- [46] B. Zhang, S. Li, C. Wu, L. Gao, W. Zhang, and M. Peng, "A neighbourhood-constrained k-means approach to classify very high spatial resolution hyperspectral imagery," *Remote Sensing Letters*, vol. 4, no. 2, pp. 161–170, 2013.
- [47] D. Ribeiro, R. Santos, R. Cabral et al., "Non-contact structural displacement measurement using unmanned aerial vehicles and video-based systems," *Mechanical Systems and Signal Processing*, vol. 160, Article ID 107869, 2021.
- [48] K. S. Hatamleh, O. Ma, A. Flores-Abad, and P. Xie, "Development of a special inertial measurement unit for UAV applications," *Journal of Dynamic Systems, Measurement, and Control*, vol. 135, no. 1, Article ID 11003, 2013.
- [49] Q. Yang and L. Sun, "A fuzzy complementary Kalman filter based on visual and IMU data for UAV landing," *Optik*, vol. 173, pp. 279–291, 2018.
- [50] G. Liu, C. He, C. Zou, and A. Wang, "Displacement measurement based on UAV images using SURF-enhanced camera calibration algorithm," *Remote Sensing*, vol. 14, no. 23, p. 6008, 2022.
- [51] X. Bai and M. Yang, "UAV based accurate displacement monitoring through automatic filtering out its camera's translations and rotations," *Journal of Building Engineering*, vol. 44, Article ID 102992, 2021.
- [52] P. Garg, F. Moreu, A. Ozdagli, M. R. Taha, and D. Mascareñas, "Noncontact dynamic displacement measurement of structures using a moving laser Doppler vibrometer," *Journal of Bridge Engineering*, vol. 24, no. 9, p. 0, 2019.
- [53] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

- [54] W. Mohguen and R. Bekka, "Comparative study of ECG signal denoising by empirical mode decomposition and thresholding functions," in *Proceedings of the 2019 6th International Conference on Electrical and Electronics Engineering*, pp. 126–130, Istanbul, Turkey, April 2019.
- [55] F. Xu, A. Chen, J. Liu, and D. Wang, "Identification of bridge flutter derivatives based on processed signal by EMD," *Journal of Vibration and Shock*, vol. 27, no. 5, pp. 144–147, 2008.
- [56] Y. F. Shao, C. Q. Miao, B. W. Li, and Q. D. Wu, "Simultaneous de-noising and enhancement method for long-span bridge health monitoring data based on empirical mode decomposition and fractal conservation law," *Measurement Science and Technology*, vol. 30, no. 6, p. 65103, 2019.
- [57] S. Sarangi, C. Biswal, B. K. Sahu, I. S. Samanta, and P. K. Rout, "Fault detection technique using time-varying filter-EMD and differential-CUSUM for LVDC microgrid system," *Electric Power Systems Research*, vol. 219, Article ID 109254, 2023.
- [58] X. Chen, X. Zhao, Y. Liang, and X. Luan, "Ocean turbulence denoising and analysis using a novel EMD-based denoising method," *Journal of Marine Science and Engineering*, vol. 10, no. 5, p. 663, 2022.
- [59] J. D. Zheng, M. X. Su, W. M. Ying, J. Y. Tong, and Z. W. Pan, "Improved uniform phase empirical mode decomposition and its application in machinery fault diagnosis," *Measurement*, vol. 179, Article ID 109425, 2021.
- [60] K. Shi-Tang, Z. Lin, S. Ya-Hui, and G. Yao-Jun, "Identification of high-order modes of vibration based on EMD and wavelet transform for wind-induced response," *Journal of Harbin Institute of Technology*, vol. 43, no. 4, pp. 75–80, 2011.
- [61] R. Grompone von Gioi, J. Jakubowicz, J. Morel, and G. Randall, "LSD: a line segment detector," *Image Processing On Line*, vol. 2, no. 4, pp. 35–55, 2012.
- [62] M. Zhang and G. Wei, "An integrated EMD adaptive threshold denoising method for reduction of noise in ECG," *Public Library of Science One*, vol. 15, no. 7, Article ID e0235330, 2020.
- [63] H. Yoon, V. Hoskere, J. Park, and B. Spencer, "Cross-Correlation-Based structural system identification using unmanned aerial vehicles," *Sensors*, vol. 17, no. 9, p. 2075, 2017.
- [64] T. Liu, Y. Lu, Y. Zhang, J. Hu, and C. Gao, "A bone segmentation method based on Multi-scale features fuse U2Net and improved dice loss in CT image process," *Biomedical Signal Processing and Control*, vol. 77, Article ID 103813, 2022.
- [65] M. Talha, F. A. Bhatti, S. Ghuffar, and H. Zafar, "ADU-Net: semantic segmentation of satellite imagery for land cover classification," *Advances in Space Research*, vol. 72, no. 5, pp. 1780–1788, 2023.
- [66] 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.
- [67] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 01, pp. 1–41, 2009.
- [68] F. Ding, Y. Xia, J. Tian, X. Zhang, and G. Hu, "An AVMD method based on energy ratio and deep belief network for fault identification of automation transmission device," *IEEE Access*, vol. 9, pp. 150088–150097, 2021.
- [69] D. Camarena-Martinez, J. P. Amezcua-Sanchez, M. Valtierra-Rodriguez, R. J. Romero-Troncoso, R. A. Osornio-Rios, and A. Garcia-Perez, "EEMD-MUSIC-Based analysis for natural frequencies identification of structures using artificial and natural excitations," *The Scientific World Journal*, vol. 2014, Article ID 587671, 587712 pages, 2014.
- [70] G. G. Amiri and E. Darvishan, "Damage detection of moment frames using ensemble Empirical Mode Decomposition and clustering techniques," *Korean Society of Civil Engineers Journal of Civil Engineering*, vol. 19, no. 5, pp. 1302–1311, 2015.
- [71] J. Z. Xin, Y. Jiang, J. T. Zhou, L. L. Peng, S. Y. Liu, and Q. Z. Tang, "Bridge deformation prediction based on SHM data using improved VMD and conditional KDE," *Engineering Structures*, vol. 261, Article ID 114285, 2022.
- [72] C. Xiong, M. Wang, Z. Shang, T. Liu, and Q. Shi, "Modal frequencies evaluation of a damaged bridge using RCVM algorithm based on sensor dynamic responses," *Measurement Science and Technology*, vol. 34, no. 10, Article ID 105013, 2023.