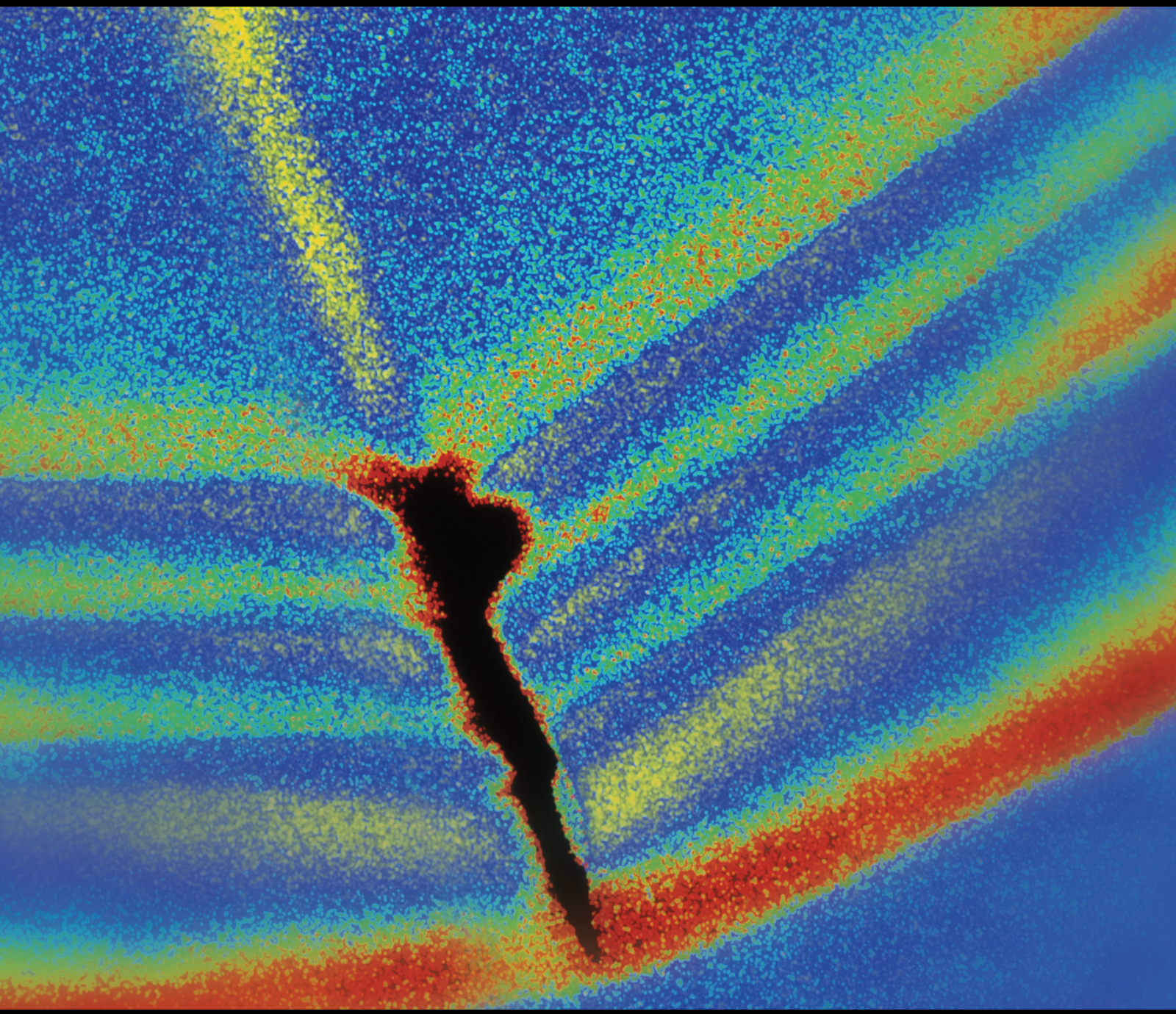
Shock and Vibration

Vibration-Based Health Monitoring of Mechanical Systems 2021

Lead Guest Editor: Franco Concli
Guest Editors: Claudio Sbarufatti and Lorenzo Maccioni



Vibration-Based Health Monitoring of Mechanical Systems 2021

Shock and Vibration

Vibration-Based Health Monitoring of Mechanical Systems 2021

Lead Guest Editor: Franco Concli
Guest Editors: Claudio Sbarufatti and Lorenzo Maccioni

Chief Editor

Associate Editors

Academic Editors

Contents

Hindawi
Shock and Vibration
Volume 2022, Article ID 5478274, 12 pages
https://doi.org/10.1155/2022/5478274

Research Article

Multidomain Feature Fusion Network for Fault Diagnosis of Rolling Machinery

Dewei Yang,[1,2] **Kefa Zhou**,[1,2] **Feng Qi**,[1,3] **and Kai Dong**[1,4]

[1]*Nanjing Hydraulic Research Institute, Nanjing 210029, China*
[2]*Country Dam Safety Management Center of the Ministry of Water Resources, Nanjing 210029, China*
[3]*Najing R&D Hydro-Information Technology Co., Ltd, Nanjing 210029, China*
[4]*Sichuan University, Chengdu 610065, China*

Correspondence should be addressed to Dewei Yang; dwyang@nhri.cn

Received 13 August 2021; Accepted 15 March 2022; Published 12 April 2022

Academic Editor: Chengwei Fei

Mechanical vibration constitutes a valuable cue for performing fault diagnosis as it is directly related to the transient regime of rolling machinery. This study establishes a multidomain feature fusion network (MFFN) to extract and fuse multidomain features through a novel multistream architecture. Three primary features are simultaneously extracted from the time, frequency, and time-frequency domains. Then, highly representative features are extracted via three convolutional branches in one- or two-dimensional spaces. A novel squeeze-connection-excitation (SCE) module is proposed to adaptively fuse features in the three domains. The advantage offered by the proposed method is that it can leverage cues from the raw vibration signal, resulting in accurate fault diagnosis. Experimental results comprehensively demonstrate and analyze the high accuracy and generalization achieved by this MFFN-based fault diagnosis method.

1. Introduction

Rolling machinery is a foundational element in industrial infrastructures. Machinery faults are the main factors that significantly affect equipment and production safety. Intelligent fault diagnosis of rolling machinery has been a topic of interest in studies concerning vibration-based health monitoring of mechanical systems [1]. Previously, fault diagnosis was realized through a combination of traditional signal processing methods, such as Fourier and wavelet transforms (WTs), and shallow learning techniques, such as support vector machine (SVM) [2] and Bayes classifiers [3]. In general, these methods are physically analyzable; however, they provide an inadequate representation of faults, which may result in a low diagnosis accuracy. This problem has motivated the development of deep learning-based methods, such as deep belief networks (DBNs) [4], stacked autoencoders (SAEs) [5], convolutional neural networks (CNNs) [6], and long short-term memory (LSTM) [7]. The high representability offered by deep learning methods significantly improves fault diagnosis accuracy.

Recently, multistream architectures are being used for fault diagnosis. In contrast to single-stream architectures, multistream architectures can represent faults in terms of multiple aspects; thus, they can achieve further enhancements in the representability of intrinsic characteristics of machinery faults. This property may further improve the performance of fault diagnosis methods. However, current multistream architectures primarily focus on the multiscale characteristic of raw vibration signals [8] and ignore the various physical properties observed in multiple domains. A novel multistream architecture that can extract and fuse multidomain features is desirable to facilitate accurate fault diagnosis.

This study proposes a novel multidomain feature fusion network (MFFN) for fault diagnosis. To this end, three one-dimensional (1D) and two-dimensional (2D) convolutional streams are designed and combined to construct the multistream architecture. Two 1D streams manage the data in the time and frequency domains, while a 2D stream extracts the time-frequency feature. At the backend joint, three representative features are fused by the squeeze-connection-

excitation (SCE) module. Finally, the fused features are used for fault classification. The contributions of this study to fault diagnosis include the following:

(i) A novel multistream architecture that can process multidomain features in an organized and comprehensive manner

(ii) A novel SCE module that can adaptively fuse multidomain features

(iii) A novel feature type that offers high representability of fault patterns and improves the decision-making capabilities of fault diagnosis

The remainder of this paper is organized as follows. Section 2 reviews related works. The overall framework of MFFN is presented in Section 3. Section 4 describes the proposed MFFN-based fault recognition method. Experimental comparisons and analysis are discussed in Section 5. Section 6 provides the summary and conclusions.

2. Related Works

Various fault diagnosis methods have been proposed to classify faults in mechanical systems. These methods generally collect vibration signals as the source data because vibrations directly relate to the transient state of running elements. Various existing shallow learning models, such as the hidden Markov model [9], k-nearest neighbors [10], and SVMs [11], have been applied in fault classification. Recently, deep learning-based methods have demonstrated excellent performance in fault diagnosis. The advantage afforded by deep learning methods is the high representability of faults. For example, the DBN model [4], which is a typical probabilistic generative model, has been introduced to solve the problems of nonlinear dynamics and discrete failure patterns. However, experimental results have revealed that DBN architectures are susceptible to overfitting. The SAE method, which is a popular deep learning-based fault diagnosis method, can incrementally learn new samples without a retraining process [5].

Another key issue in fault diagnosis is that feature extraction. Previously, temporal and frequency analyses were the two main approaches toward fault feature extraction [12]. However, they cannot represent the temporal variation of a vibration signal accurately [13]. This problem has been solved via methods including short-time Fourier transforms [13], Wigner Ville distributions [14], and WTs [15]. Among these, WT is the most practical because its relaxed structure can decompose signals with varying temporal resolutions. Moreover, WT can produce 2D feature maps such that successful image classification methods can be transformed into fault diagnosis methods.

Deep learning-based fault diagnosis has attracted considerable attention recently [16]. The advantage of deep learning lies in its excellent ability to abstract signals by performing layer-wise nonlinear calculations, thereby enabling the deeper layer to generate more representative features. This encourages the utilization of various deep learning methods in fault recognition. The DBN is one of the most widely used deep learning methods because it can adapt to a wide range of problems, including those of nonlinear dynamics and discrete failure patterns [17]. To leverage valuable cues for fault diagnosis, an adaptive spatiotemporal feature learning architecture with multiple measurements was proposed [18]. Subsequently, the generalization of deep learning architecture was considered. For example, a domain generalization-based hybrid diagnosis network was established, which could be deployed in unseen working conditions instead of in real-world working conditions [19]. Moreover, a domain adversarial transfer network has been evaluated for application in fault diagnosis, wherein a transfer learning mechanism can be implemented to enhance the generalization of deep learning-based fault diagnosis [20]. Recently, a novel convolutional neural network is established to diagnose faults from small samples. Based on the domain adaption, this method won success when the vibration data are not available in abundant [21]. Different from this previous strategy, our study in this paper aims to solve another problem in fault diagnosis—feature representation and fusion.

Also, other types of signal have been introduced in faults diagnosis. For example, the thermographic information has been utilized in fault diagnosis of ventilation in BLDC motors [22, 23]. In contrast to the vibration signal, the thermographic signals provide additional informative clues which help to increase the accuracy of the fault diagnosis. Moreover, the thermographic signal is relatively simple in contrast to the vibration signal, such that it can better identify the fault types. However, the main drawback of the thermographic signal-based strategy lies in that it is commonly hysteretic to reflect the machinery statement. In practice, it is observed that the temperature significantly increases after a while of the fault occurrence. Alternatively, the acoustic signal has been investigated in the field of fault diagnosis [24]. The main advantage of the acoustic signal-based strategy lies in the noncontact measurement that we can efficiently deploy the acoustic sensors to diagnose faults. However, the acoustic signal is likely affected by the environmental noises. As the result, noise removal is the main issue of the acoustic signal processing.

This study aims to leverage valuable cues from multiple domains for fault diagnosis. To this end, a novel MFFN is proposed. This network comprises three streams that can comprehensively extract highly representative features in multiple domains, such as the temporal, frequency, and time-frequency domains. The MFFN can obtain more valuable cues for fault diagnosis than those of current single-stream and multistream architectures. Moreover, the novelty of the proposed architecture lies in its ability to adaptively fuse 1D and 2D features using the SCE block.

3. Proposed Fault Diagnosis Scheme

3.1. MFFN. To achieve high fault representability, this study proposes a novel multistream architecture for extracting and classifying three types of features in the temporal, time-frequency, and frequency domains. The block diagram of MFFN is shown in Figure 1. The sliding window block is

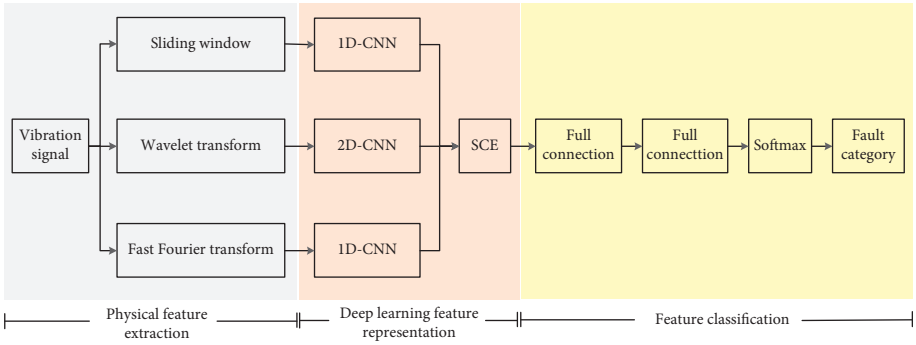


Figure 1: Block diagram of multidomain feature fusion network.

applied to segment the vibration signal into sequence vectors at the first stream. Time-frequency features are extracted using WT at the second stream. The third stream extracts frequency features via fast Fourier transform (FFT). These primary features are subsequently enhanced in terms of their representability through layer-wise convolutional calculations. Finally, these highly representative features are adaptively fused via the SCE module. The backend classifier is established using two fully connected layers and a Softmax calculation block. The data used in this study were obtained from public datasets. The fault diagnosis platform consists of a motor, torque transducer/encoder, dynamometer, and control electronics [21]. The reason for selecting these datasets lies in that they provide a baseline to fairly evaluate and compare different methods.

3.2. Primary Feature Extraction. The primary feature extraction process is shown in Figure 2. The sliding window is used for extracting temporal features. L denotes the window length, and M is the sliding step. The frequency spectrum is extracted via the following FFT:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi/Nkn}, \quad k = 0, 1, \ldots, N-1, \tag{1}$$

where N denotes the length of the signal segmentation.

A limitation of FFT is that it analyses the frequency spectrum pattern of the vibration signal exclusively from a global perspective; therefore, it is not suitable for an amplitude-modulated or nonstationary signal. This drawback can be addressed by wavelet package transform (WPT), which is a time-frequency analysis method that can analyze vibration signals with flexible temporal resolutions at both high and low frequencies [15]. Therefore, the WT is operated with a wavelet packet tree that decomposes a signal into several levels of wavelet packets. A three-layer wavelet packet tree is used in our method. As a result, eight sub-bands are obtained, and the energy value of each sub-band signal can be calculated through the following equation:

$$E_n^j = \int \left| C_n^j(t) \right|^2 \mathrm{d}t = \sum_{k=1}^{N_i} \left| x_j^k \right|^2, \tag{2}$$

where $C_n^j(t)\,(n = 3, \quad j = 0, 1, \ldots, 7)$ is the reconstructed sub-band signal, N_i is the length of the reconstructed signal, and $x_j^k\,(k = 1, 2, \ldots, r)$ is the amplitude of the j^{th} reconstructed signal. The energy spectrum feature of a sub-band signal can be presented through a normalized value, as shown in the following:

$$\left[\frac{E_4^0}{\sqrt{E_r}}, \frac{E_4^1}{\sqrt{E_r}}, \ldots, \frac{E_4^7}{\sqrt{E_r}} \right], \tag{3}$$

where E_r is the square root of the summed square values of the sub-band signal energy and is expressed as follows:

$$E_r = \sqrt{\sum_{j=1}^{M} \left(E_n^j \right)^2}. \tag{4}$$

3.3. Extraction and Fusion of Highly Representative Features. In the first and third streams, two 1D-CNNs are connected to the primary feature extractor to manage features in the temporal and frequency domains, while a 2D-CNN is connected to the WT module to process the feature in the time-frequency domain. These highly representative features are then fused at the backend joint by the SCE module. In general, there are four successive phases in the SCE module, namely, squeeze, connection, excitation, and reweight, as shown in Figure 3. The feature matrices of the three domains are the input to the SCE module. The temporal, frequency, and time-frequency feature matrices are presented as follows:

$$\begin{aligned} X &= \begin{bmatrix} x_{11} & \cdots & x_{1C_1} \\ \vdots & \ddots & \vdots \\ x_{L_1 1} & \cdots & x_{L_1 C_1} \end{bmatrix}_{L_1 \times C_1}, \\ Y &= \begin{bmatrix} y_{11} & \cdots & y_{1C_2} \\ \vdots & \ddots & \vdots \\ y_{L_2 1} & \cdots & y_{L_2 C_2} \end{bmatrix}_{L_2 \times C_2}, \\ Z &= \begin{bmatrix} z_{11} & \cdots & z_{1C_3} \\ \vdots & \ddots & \vdots \\ z_{L_3 1} & \cdots & z_{L_3 C_3} \end{bmatrix}_{L_3 \times C_3}, \end{aligned} \tag{5}$$

where L_1, L_2, and L_3 are the feature dimensions and C_1, C_2, and C_3 are the number of feature channels.

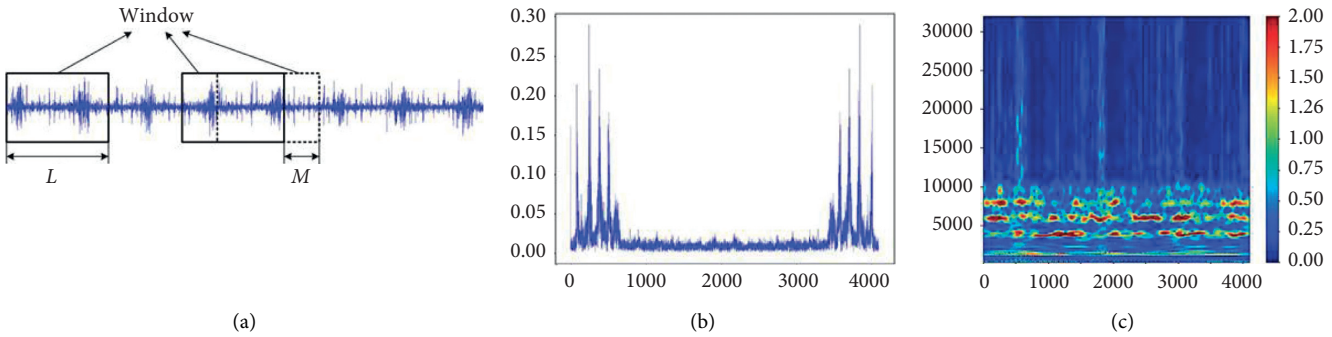


Figure 2: Primary feature extraction: (a) temporal feature, (b) frequency feature, and (c) time-frequency feature.

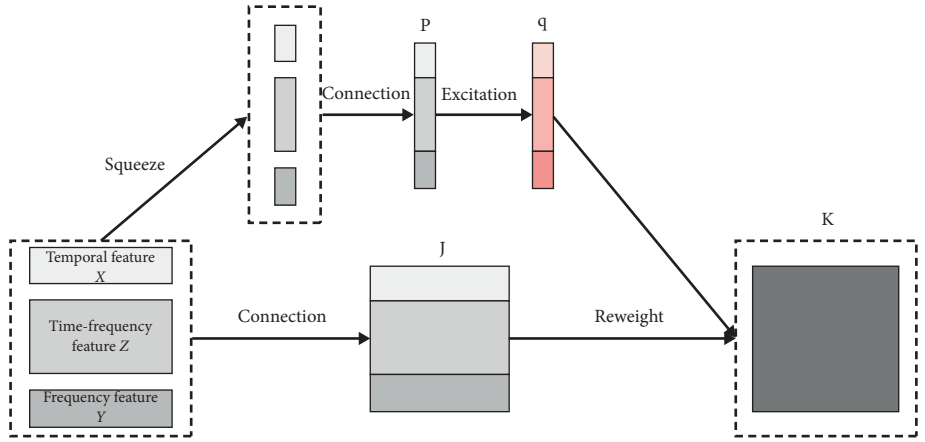


Figure 3: SCE module.

3.4. Squeeze. A pooling operation is applied to squeeze the feature matrix. As a result, three 1D feature vectors are generated to present the feature matrix across three domains, shown as follows:

$$l_j = F_{sq}(X) = \frac{1}{L_1}\sum_{i=1}^{L_1} x_{ij}, \quad j = (1,2,\ldots,C_1), \mathbf{l} = \left(l_1, l_2, \ldots, l_{C_1}\right),$$

$$m_j = F_{sq}(Y) = \frac{1}{L_2}\sum_{i=1}^{L_2} y_{ij}, \quad j = (1,2,\ldots,C_2), \mathbf{m} = \left(m_1, m_2, \ldots, m_{C_2}\right),$$

$$n_j = F_{sq}(Z) = \frac{1}{L_3}\sum_{i=1}^{L_3} z_{ij}, \quad j = (1,2,\ldots,C_3)\mathbf{n} = \left(n_1, n_2, \ldots, n_{C_3}\right). \tag{6}$$

3.5. Connection. The feature vectors are fused by the following concatenation operation:

$$\begin{aligned}\mathbf{p} &= F_c(\mathbf{l}, \mathbf{m}, \mathbf{n}) \\ &= \left(l_1, l_2, \ldots, l_{C_1}, \quad m_1, m_2, \ldots, m_{C_2}, n_1, n_2, \ldots, n_{C_3}\right) \\ &= (p_1, p_2, \ldots, p_C),\end{aligned} \tag{7}$$

$$\mathbf{J} = F_c(\mathbf{X}, \mathbf{Y}, \mathbf{Z}) = \begin{bmatrix} j_{11} & \cdots & j_{1C} \\ \vdots & \ddots & \vdots \\ j_{L1} & \cdots & j_{LC} \end{bmatrix}_{L\times C}.$$

3.6. Excitation. Multilayer mapping is performed to achieve excitation, as follows:

$$\mathbf{q} = F_{\text{ex}}(\mathbf{p}, \mathbf{W}) = \sigma\left(\mathbf{W}_2\delta\left(\mathbf{W}_1\mathbf{p}\right)\right). \tag{8}$$

In the above equation, σ is the sigmoid function, δ is the ReLU activation function, and $\mathbf{W}$, $\mathbf{W}_1$, and $\mathbf{W}_2$ are the full-connection weights.

3.7. Reweight. The learned weight is added to feature channels to generate the weighted feature for final classification:

$$\mathbf{K} = F_r(\mathbf{J}, \mathbf{q}) = \mathbf{J} \times \mathbf{q}. \tag{9}$$

The advantage of the *reweight* calculation is similar to that of the global average pooling operation in the *squeeze* process that can generate channel-wise statistics. Subsequently, this global information is embedded by the *excitation* process to generate the channel descriptor $\mathbf{q}$, which comprehensively captures channel-wise dependencies. As a result, the most important feature can be emphasized by multiplying the feature channels with the channel descriptor. In this regard, SCE blocks intrinsically introduced dynamics conditioned on the input, thereby helping boost feature discriminability of specific fault patterns [5].

4. MFFN-Based Fault Diagnosis

The MFFN-based fault diagnosis method comprises four modules. The first module performs primary data processing, wherein the physical significance of the vibration signal is presented with respect to the temporal, frequency, and time-frequency features. In the second module, the high-representation features are extracted through layer-wise mapping. Adaptive feature fusion is realized in the third module, wherein the most credible factor is enhanced, while feature redundancy is reduced considerably. Finally, the fourth module is designed for fault classification, wherein a shallow architecture is established with two fully connected layers and a Softmax calculation block. The details of our proposed MFFN are presented in Table 1.

5. Experimental Evaluation and Discussion

5.1. Setup. To evaluate the performance of the proposed MFFN, experiments were conducted on defective bearing datasets provided by the Case Western Reserve University Bearing Data Center (CWRU dataset) [25], Jiangnan University (Jiangnan dataset) [26], and Paderborn University (Paderborn dataset) [27]. The bearing system platform in the Case Western Reserve University Bearing Data Center includes a 2 HP motor, torque transducer, dynamometer, and load motor. The vibration signal was collected via an accelerometer at a sampling frequency of 12 kHz. In addition to the normal state, nine categories of fault state data were included in this database: single-point faults with sizes of 0.007, 0.14, and 0.021 were individually identified on the inner race (IR), outer race (OR), and rolling elements (REs), respectively. For each state, 120,000 samples were collected in 10 s. The data from Jiangnan University include four categories of running states: the normal state and fault states separately seeded on the bearing at IR, OR, and RE. All data were collected at a 50 kHz sampling frequency at rolling speeds of 600, 800, and 1000 rpm. For the normal state, 1800 samples were randomly collected, while 600 samples were collected for each fault state. The data from the Paderborn University were provided via measurements concerning six healthy and 26 damaged bearings at IR and OR. All data were collected at a 64 kHz sampling frequency at rolling speeds of 900 and 1500 rpm. For each state, 256,0000 samples are collected in 4 s. The training and testing samples for experimental evaluations are shown in Tables 2 to 4.

5.2. Model Pretraining and Fine-Tuning. An adequate number of epochs in the training period is important for model training. Excessive epochs may result in overfitting, while the learning outcome may be poor in the case of insufficient epochs. Figure 4 illustrates the training times of the three datasets and reveals that that the MFFN can converge rapidly on all three datasets. Thirty iterations are sufficient for model learning with the CWRU and Jiangnan datasets, while 25 iterations are required for model learning with the Paderborn dataset.

5.3. MFFN-Based Fault Diagnosis. Confusion matrices were utilized for evaluating the performance of the proposed MFFN. Figure 5(a) presents the confusion matrix for the CWRU dataset. This result demonstrates that our fault diagnosis method is highly accurate. Only two samples were erroneously classified; the rest were identified correctly. The classification results regarding the Jiangnan dataset are satisfactory (Figure 5(b)); only one sample was misclassified. A similar outcome was observed in the results on the Paderborn dataset (Figure 5(c)) with one error.

5.4. Comparison against Existing Deep Learning Methods. We evaluated the proposed MFFN by comparing it to state-of-the-art methods. The temporal-, frequency-, and time-frequency feature-based methods are comprehensively catalogued for experimental comparison. For example, the 1D-CNN was used to classify 1D temporal features [28] and denoted as "TF + 1D-CNN." Furthermore, temporal feature + WDCNN (TF + WDCNN) [29], frequency feature + 1D-CNN (FF + 1D-CNN) [30], frequency feature + SDAE (FF + SDAE), time-frequency feature + 2D-CNN (TFF + 2D-CNN), and time-frequency feature + VGG16 (TFF + VGG16) [31] were included in the experimental comparison. Each dataset was divided into 10 subsets for experimental evaluation with respect to working conditions. Figure 6 reveals that no salient performance variation was observed among the 10 evaluations for MFFN; the maximum differences among evaluations were 0.08% on the CWRU dataset, 0.12% on the Jiangnan dataset, and 0.13% on the Paderborn dataset. Thus, experimental analysis demonstrates the stability of our MFFN compared with that of the other methods, which exhibit lesser model stability owing to significant performance variations across tested subsets.

The average accuracies of the compared fault diagnosis methods are listed in Table 5. Two observations can be made based on this table. First, the feature in the time-frequency domain outperforms the temporal- and frequency-domain features. This is because the time-frequency domain feature can identify details of the frequency spectrum of the vibration signal, which facilitates an improved fault diagnosis. Second, fusing features in multiple domains is preferable for fault diagnosis. This is because the feature fusion results can represent machinery faults from multiple aspects and allow more valuable cues to be leveraged for fault diagnosis. As a result, intraclass fault differences are enlarged, while interclass clustering is enhanced, which theoretically explains the better performance of the proposed MFFN.

5.5. Visualization. Aiming to comprehensively understand the benefits of our proposed MFFN, the t-SNE technique was applied to reduce the dimensionality of the learned features

Table 1: Parameters of MFFN.

Module	Name	Size/step/number	Parameter size	Output size
1D-CNN	Input_1	—	0	(None,4096,1)
	ResBlock_1	3/1/16	912	(None,4096,16)
	ResBlock_2	3/1/16	1568	(None,4096,16)
	ResBlock_3	3/1/4	441	(None,4096,4)
	Max_Pooling	2/1/-	0	(None,2048,4)
2D-CNN	Input_2	—	0	(None,128,128,3)
	Conv2D_1	30/5/256	691200	(None,20,20,256)
	Conv2D_2	6/2/256	2359552	(None,8,8,256)
	Inception_1	(1,3,5,7)/1/32	688768	(None,8,8,128)
	Reshape_1	—	0	(None,2048,4)
SCE	Global_Average_Pooling_1	—	0	(None,4)
	Global_Average_Pooling_2	—	0	(None,4)
	Global_Average_Pooling_3	—	0	(None,4)
	Concatenate_1	—	0	(None,2048,12)
	Concatenate_2	—	0	(None,12)
	Dense_1	6/-/-	78	(None,6)
	Dense_2	12/-/-	84	(None,12)
	Multiply	—	0	(None,2048,12)
Classifier	Flatten_1	—	0	(None,24576)
	Dense_3	100/-/-	2457700	(None,100)
	Dense_4	10,4,3/-/-	1010,404,303	(None,10),(None,4),(None,3)

Table 2: Samples in the CWRU dataset.

Category	Normal	Outer race			Inner race			Rolling element		
Size	—	0.007	0.014	0.021	0.007	0.014	0.021	0.007	0.014	0.021
Label	0	1	2	3	4	5	6	7	8	9
Training	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200
Validation	400	400	400	400	400	400	400	400	400	400
Testing	400	400	400	400	400	400	400	400	400	400

Table 3: Samples in the Jiangnan dataset.

Category	Normal	Outer race	Inner race	Rolling element
Label	0	1	2	3
Training	1200	1200	1200	1200
Validation	400	400	400	400
Testing	400	400	400	400

Table 4: Samples in the Paderborn dataset.

Category	Normal	Outer race	Inner race
Label	0	1	2
Training	1200	1200	1200
Validation	400	400	400
Testing	400	400	400

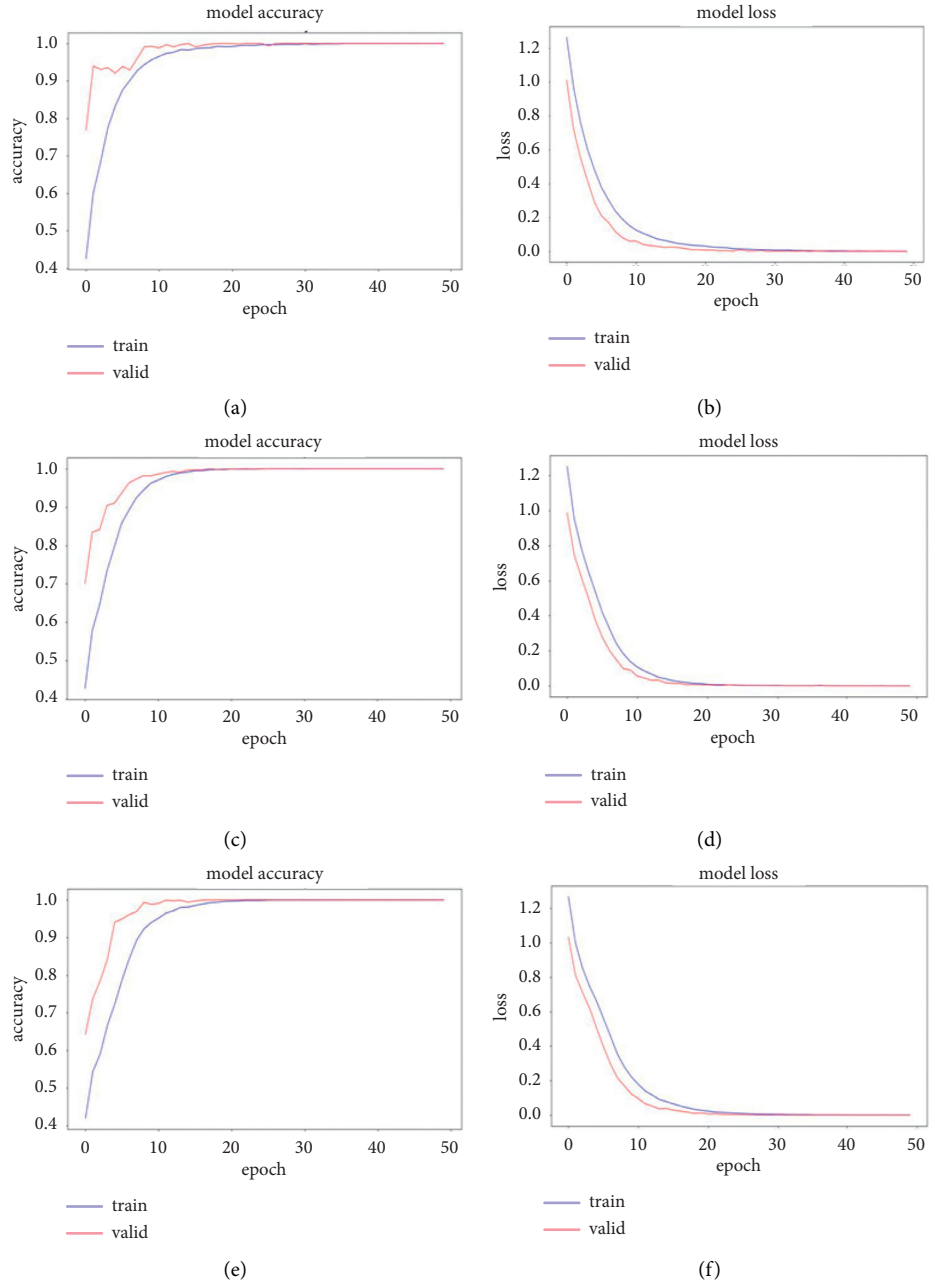


Figure 4: Training-time investigation for CWRU, Jiangnan, and Paderborn datasets. CWRU: (a) accuracy and (b) loss; Jiangnan: (c) accuracy and (d) loss; Paderborn: (e) accuracy and (f) loss.

to two for facilitating map generation. The resulting 2D feature maps are shown in Figure 6, wherein different colors represent various fault or normal categories. As shown in Figure 7, after MFFN feature learning, a fault-category clustering effect is observed in contrast to the raw distribution, along with linear margins between fault categories. This result is desirable and enables simpler classification. This further demonstrates that using the MFFN architecture can significantly improve the accuracy of fault diagnosis.

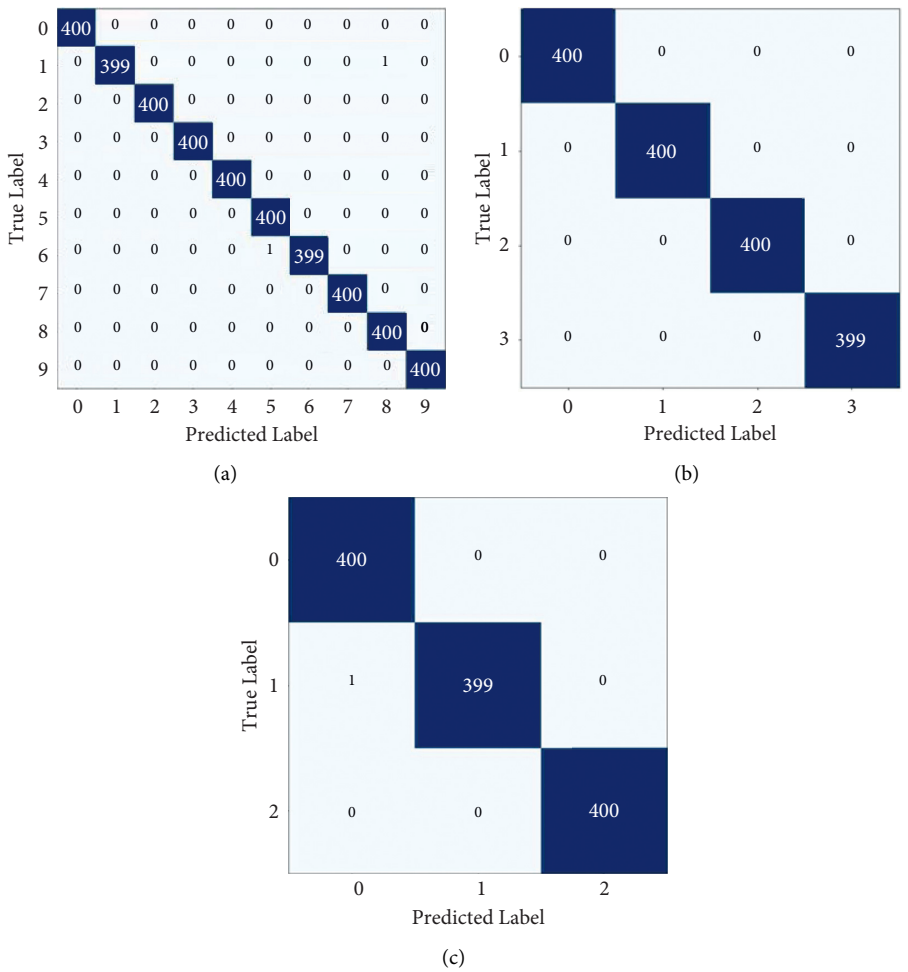


Figure 5: Confusion matrices for (a) CWRU, (b) Jiangnan, and (c) Paderborn datasets.

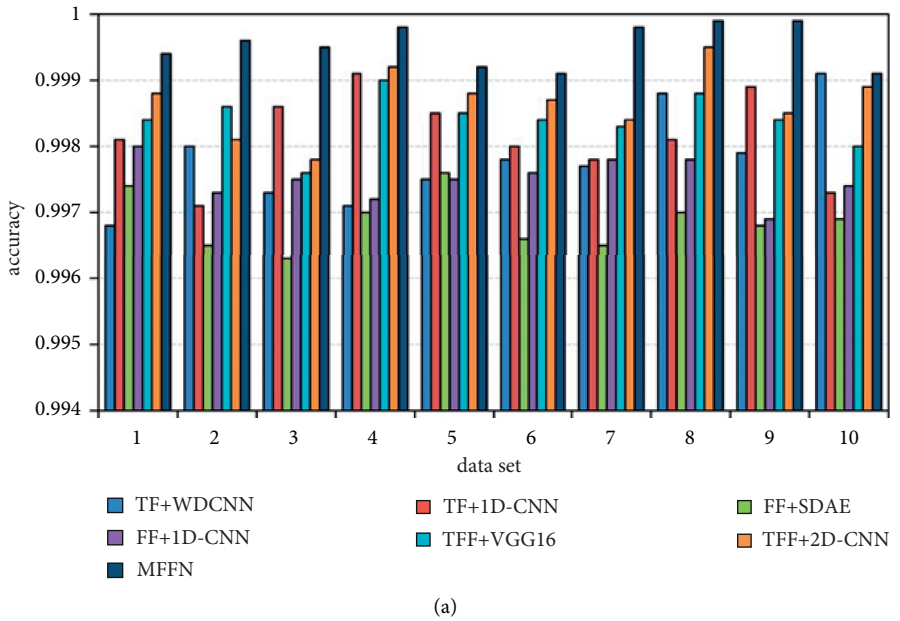


Figure 6: Continued.

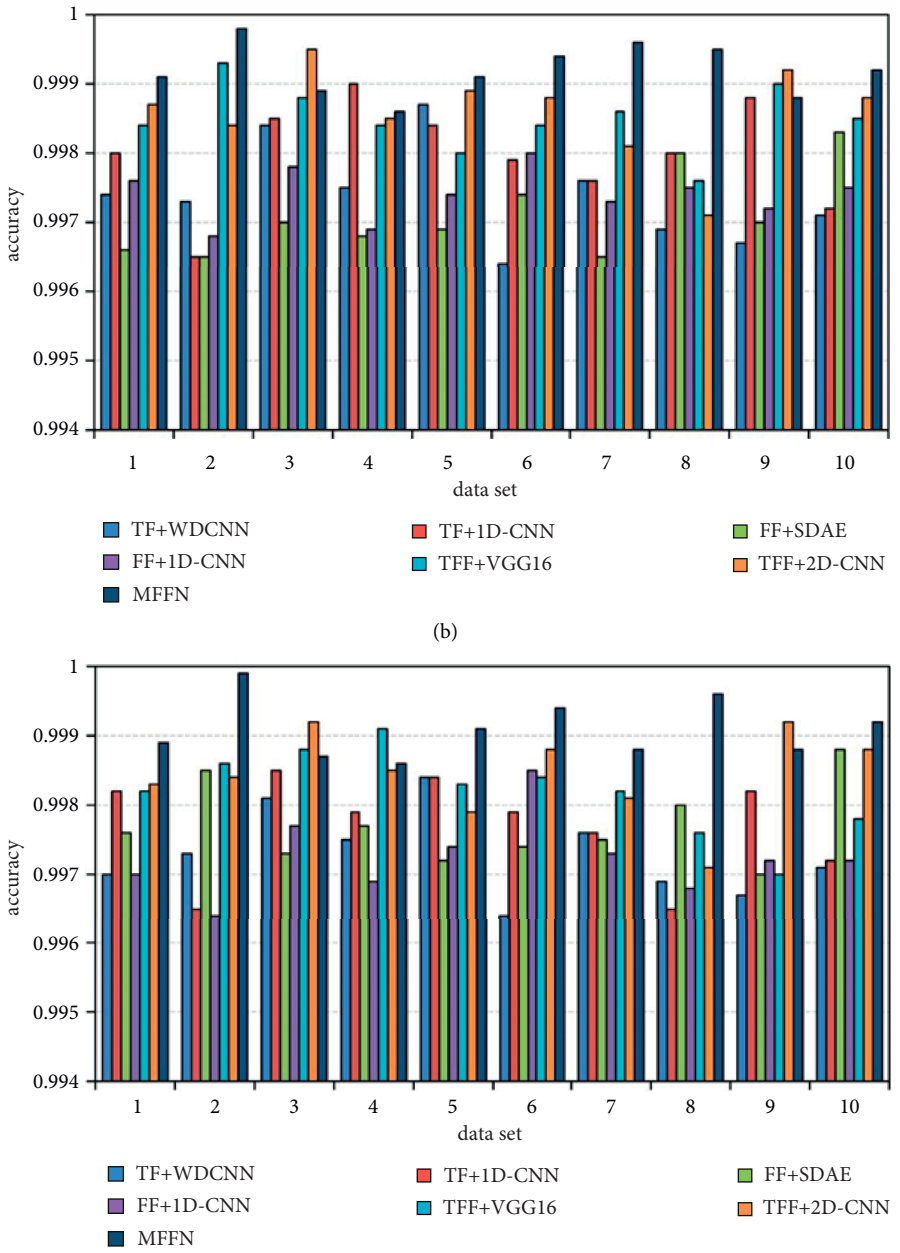


Figure 6: Experimental comparisons on (a) CWRU, (b) Jiangnan, and (c) Paderborn datasets.

Table 5: The average performance in different domains.

Input	Model	Average accuracy		
		CWRU (%)	Jiangnan (%)	Paderborn (%)
TF	WDCNN	99.78	99.74	99.73
	1D-CNN	99.81	99.80	99.77
FF	SDAE	99.69	99.71	99.78
	1D-CNN	99.75	99.74	99.72
TFF	VGG16	99.84	99.85	99.82
	2D-CNN	99.87	99.86	99.84
TF + FF + TFF	MFFN	99.95	99.92	99.91

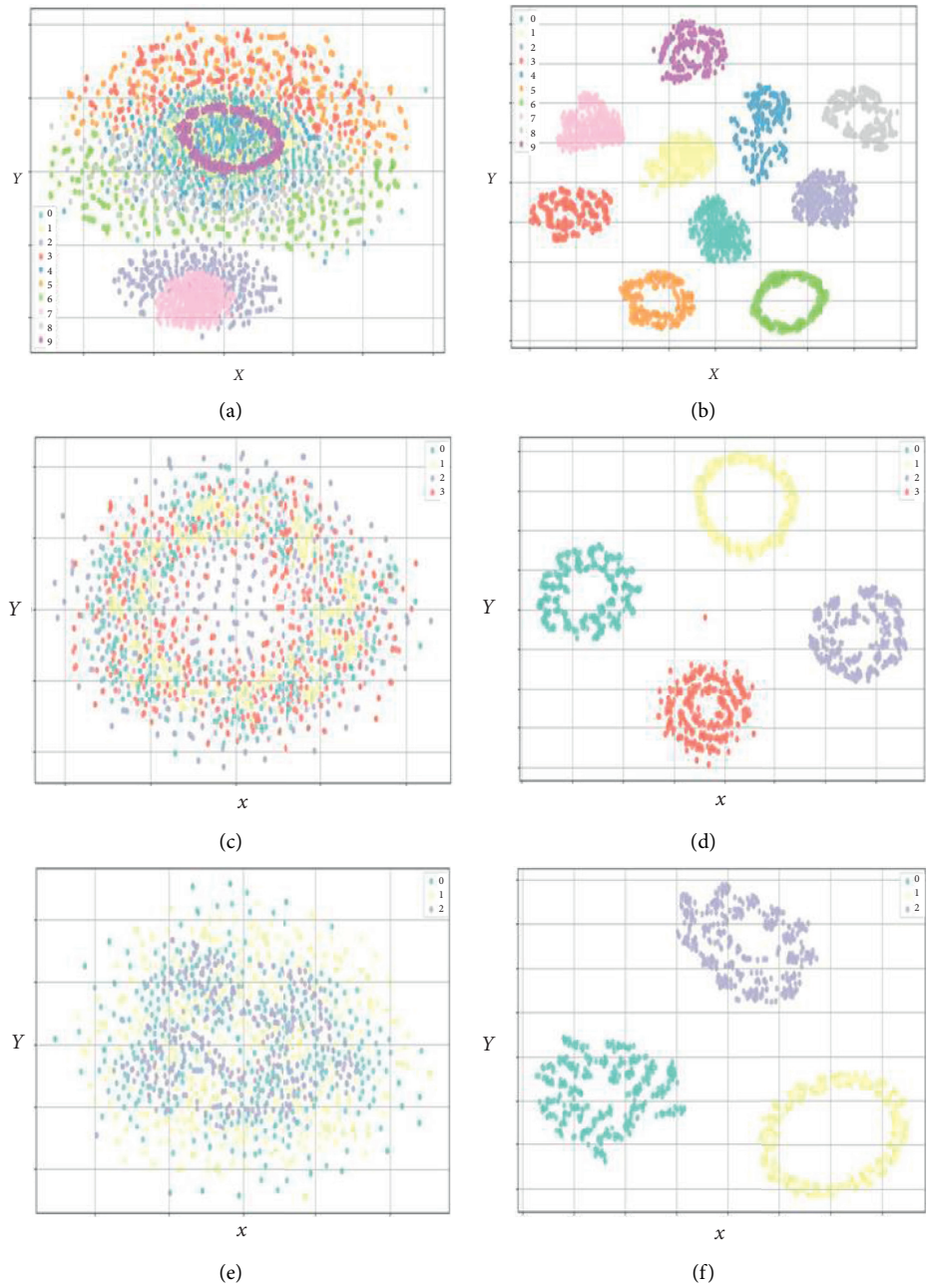


Figure 7: Feature visualization with (a) raw data on the CWRU dataset, (b) learned features on the CWRU dataset, (c) raw data on the Jiangnan dataset, (d) learned features on the Jiangnan dataset, (e) raw data on the Paderborn dataset, and (f) learned features on the Paderborn dataset.

6. Conclusions

A novel MFFN-based fault diagnosis method is proposed. The proposed MFFN can fuse features in different domains, such as the temporal, frequency, and time-frequency domains. Sufficient cues are comprehensively leveraged through the deep learning process of MFFN. The main contribution of MFFN is that it can improve the representability of faults, leading to a significant improvement in the accuracy of fault diagnosis. Theoretically, features in multiple domains depict faults from multiple perspectives, which are complementary in physical significance. Moreover, the importance of features in multiple domains varies with respect to the tasks on hand. Intrinsically, our proposed MFFN adopts a feature fusion strategy using adaptive weights. Features extracted in multiple domains are weighted and fused, leading to a comprehensive utilization of their advantages. Consequently, MFFN achieves higher accuracy compared with existing architectures.

Using our proposed MFFN, exceptional accuracy can be achieved, enabling its utilization in many practical

applications. In our future work, we will train the MFFN model to handle more signal types, such as thermal imaging and acoustic data, which contain much more valuable features for diagnosing faults. Moreover, we will evaluate MFFN in real-world applications, especially in online fault diagnosis.

Data Availability

The experiment data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study was funded by the National Key Research and Development Program of China (2018YFC0407106) and the National Natural Science Foundation of China (51909174 and 61801169).

References

[1] S. Zhang, S. Zhang, B. Wang, and T. G. Habetler, "Deep learning algorithms for bearing fault diagnostics-A comprehensive review," *IEEE Access*, vol. 8, pp. 29857–29881, 2020.

[2] Q. Shi and H. Zhang, "Fault diagnosis of an autonomous vehicle with an improved SVM algorithm subject to unbalanced datasets," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 7, pp. 6248–6256, 2020.

[3] Y. Hu, F. Cui, and X. Tu, "Bayesian estimation of instantaneous speed for rotating machinery fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 9, pp. 8842–8852, 2021.

[4] X. Zhao, M. Jia, and Z. Liu, "Semisupervised graph convolution deep belief network for fault diagnosis of electormechanical system with limited labeled data," *IEEE Transactions on Industrial Informatics*, vol. 17, pp. 5450–5460, 2020.

[5] W. Sun, S. Shao, R. Zhao, R. Yan, X. Zhang, and X. Chen, "A sparse auto-encoder-based deep neural network approach for induction motor faults classification," *Measurement*, vol. 89, pp. 171–178, 2016.

[6] S. Tang, S. Yuan, and Y. Zhu, "Convolutional neural network in intelligent fault diagnosis toward rotatory machinery," *IEEE Access*, vol. 8, pp. 86510–86519, 2020.

[7] J. Lei, C. Liu, and D. Jiang, "Fault diagnosis of wind turbine based on Long Short-term memory networks," *Renewable Energy*, vol. 133, pp. 422–432, 2019.

[8] G. Jiang, H. He, J. Yan, and P. Xie, "Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 4, pp. 3196–3207, 2019.

[9] H. Ocak and K. A. Loparo, "HMM-based fault detection and diagnosis scheme for rolling element bearings," *Journal of Vibration and Acoustics*, vol. 127, no. 4, pp. 299–306, 2005.

[10] Q. Wang, Y. B. Liu, X. He, S. Y. Liu, and J. Liu, "Fault diagnosis of bearing based on KPCA and KNN method," *Advanced Materials Research*, vol. 1491, pp. 986-987, 2014.

[11] X. Zhang, Y. Liang, and J. Zhou, "A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM," *Measurement*, vol. 69, pp. 164–179, 2015.

[12] A. Krishnakumari, M. Saravanan, M. Ramakrishnan, S. M. Ponnuri, and R. Srinadh, "Vibration condition monitoring of spur gear using feature extraction of EMD and hilbert-huang transform," *Intelligent Manufacturing and Energy Sustainability*, vol. 169, pp. 133–145, 2020.

[13] L. Zhang, Y. Li, L. Dong et al., "Gearbox fault diagnosis using multiscale sparse frequency-frequency distributions," *IEEE Access*, vol. 9, pp. 113089–113099, 2021.

[14] J. A. Rosero, L. Romeral, J. A. Ortega, and E. Rosero, "Short-circuit detection by means of empirical mode decomposition and wigner-ville distribution for PMSM running under dynamic condition," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 11, pp. 4534–4547, 2009.

[15] J. Li, H. Wang, X. Wang, and Y. Zhang, "Rolling bearing fault diagnosis based on improved adaptive parameterless empirical wavelet transform and sparse denoising," *Measurement*, vol. 152, Article ID 107392, 2020.

[16] H. Li, G. Hu, J. Li, and M. Zhou, "Intelligent fault diagnosis for large-scale rotating machines using binarized deep neural networks and random forests," *IEEE Transactions on Automation Science and Engineering*, vol. 19, pp. 1–11, 2021.

[17] B. Cai, L. Huang, and M. Xie, "Bayesian networks in fault diagnosis," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2227–2240, 2017.

[18] T. Han, C. Liu, L. Wu, S. Sarkar, and D. Jiang, "An adaptive spatiotemporal feature learning approach for fault diagnosis in complex systems," *Mechanical Systems and Signal Processing*, vol. 117, pp. 170–187, 2019.

[19] T. Han, Y.-F. Li, and M. Qian, "A hybrid generalization network for intelligent fault diagnosis of rotating machinery under unseen working conditions," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–11, 2021.

[20] Z. Chen, G. He, J. Li, Y. Liao, K. Gryllias, and W. Li, "Domain adversarial transfer network for cross-domain fault diagnosis of rotary machinery," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 11, pp. 8702–8712, 2020.

[21] A. Kumar, G. Vashishtha, C. P. Gandhi, Y. Zhou, A. Glowacz, and J. Xiang, "Novel convolutional neural network (NCNN) for the diagnosis of bearing defects in rotary machinery," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–10, 2021.

[22] A. Glowacz, "Thermographic fault diagnosis of ventilation in BLDC motors," *Sensors*, vol. 21, no. 21, Article ID 7245, 2021.

[23] A. Glowacz, "Ventilation diagnosis of angle grinder using thermal imaging," *Sensors*, vol. 21, no. 8, Article ID 2853, 2021.

[24] A. Glowacz, R. Tadeusiewicz, S. Legutko et al., "Fault diagnosis of angle grinders and electric impact drills using acoustic signals," *Applied Acoustics*, vol. 179, Article ID 108070, 2021.

[25] Case western Reserve university, "Case western Reserve university bearing data centre website[DB/OL]," 2017, https://csegroups.case.edu/bearingdatacenter/pages/download-data-file.

[26] K. Li, "Jiangnan university bearing dataset [DB/OL]," 2018, https://mad-net.org:8765/explore.html.

[27] Paderborn University, "Paderborn University Website," 1972, https://mb.uni-paderborn.de/en/kat/main-research/datacenter/bearing-datacenter/data-sets-and-download.

[28] W. Zhang, F. Zhang, W. Chen, Y. Jiang, and D. Song, "Fault state recognition of rolling bearing based fully convolutional

network," *Computing in Science & Engineering*, vol. 21, pp. 132–147, 2018.

[29] W. Zhang, G. Peng, C. Li, Y. Chen, and Z. Zhang, "A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals," *Sensors*, vol. 17, no. 2, pp. 425–446, 2017.

[30] C. Lu, Z.-Y. Wang, W.-L. Qin, and J. Ma, "Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification," *Signal Processing*, vol. 130, pp. 377–388, 2017.

[31] M. T. Pham, J.-M. Kim, and C. H. Kim, "Accurate bearing fault diagnosis under variable shaft speed using convolutional neural networks and vibration spectrogram," *Applied Sciences*, vol. 10, no. 18, pp. 6385–6397, 2020.

Hindawi
Shock and Vibration
Volume 2022, Article ID 1098337, 9 pages
https://doi.org/10.1155/2022/1098337

Research Article

Vibration Measurement of a Metal Sheet Using Single-Camera Digital Image Correlation with Projection Components

Z. H. Liang and L. Yue

College of Mechanical & Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Correspondence should be addressed to L. Yue; yuelinme@nuaa.edu.cn

Received 16 October 2021; Accepted 15 January 2022; Published 15 February 2022

Academic Editor: Lorenzo Maccioni

Digital image correlation has emerged as a popular method for the dynamic performance measurement of metallic and polymer sheets, owing to the benefits of being a noncontact, full-field, and high-precision method. Two or more high-speed cameras are required for full-field vibration measurements with three-dimensional digital image correlation, which is generally costly. Perpendicular view to the specimen surface is conventional in two-dimensional digital image correlation, and the out-of-plane displacement is regarded as a part of systematic errors. In this study, a single view method was implemented with no complex optical settings. The full-field vibration displacement of the metal sheet was measured with projection components, and the first four orders of displacement modes were identified. Finite element analysis and traditional experimental modal analysis were then implemented to validate the effectiveness and accuracy of the proposed approach. The results show that the dynamic parameters, including the natural frequencies and mode shapes, were well consistent. Meanwhile, there is a significant difference in the length of mode shape vectors. The number of measurement points in the proposed method is 2016, which is far more than the number of measurement points in the traditional experimental modal analysis. This would be convenient and beneficial for damage identification towards thin-wall parts including turbine blade with the continuum hypothesis of mode shapes and a single-camera DIC system. It is worth noting that this is effective with conditions of small deformation vibration and no rigid-body rotation.

1. Introduction

In the past decade, the digital image correlation (DIC) method was widely studied and employed for deformation, shape, and motion measurements in various applications [1–4]. The field of mechanical performance testing of new materials and structures is revolutionary, such as additive manufacturing (AM) alloys [5], shape memory alloys (SMA) [6], and carbon fibre reinforced plastics (CFRP) [7]. In addition, it is also a new path to the crack near-field measurement [8, 9] and model updating [10]. With the development of high-speed camera performance, such as achieving maximum frame rates at full resolution, an increasing number of structural dynamic characteristics can be measured through the high-speed digital image correlation (HS-DIC) method [11, 12]. The dynamic performance measurement of metallic and polymer sheets is increased by the benefits of being a noncontact, full-field, and high-precision method [13–15]. It also benefits from the inverse-compositional Gauss–Newton (IC-GN) algorithm and parallel computing which greatly increases the computational efficiency [16–18]. However, the cost of a high-speed camera is usually in excess of $10 000, and this measurement often requires two or more cameras to capture the full-field vibration measurement through the stereo-vision method. Such a high price is close to the laser Doppler vibrometer (LDV) system. The DIC and LDV methods both have their own strengths in full-field vibration measurement [19, 20]. The essential difference between the two methods is that DIC adopts the time-frozen method and records a frame containing full-field displacement information for each sample, while LDV is generally limited to the velocity of a single point in each sample [21].

To overcome this significant drawback of HS-DIC, a series of investigations aimed at single-camera stereo-DIC systems [22–26] have been published, while the time synchronization errors caused by triggers between different cameras are eliminated, especially high-speed rotating parts measurement of the HS-DIC are potential and valuable [27]. Single-camera stereo-DIC techniques have gained increasing attention owing to their advantages of cost effectiveness, compactness, and avoidance of additional camera synchronization.

Two types of single-camera stereo-DIC systems have been proposed in recent studies. The first one adopts a biprism and a set of planar mirrors to split the single view into two different views. The two views are captured in the same CCD/CMOS sensor, which is generally divided equally by the centerline. Owing to hardware constraints, the maximum resolution of most high-speed cameras ranges from 1 to 4 megapixels, such as 1280 px × 800 px. This is an extremely poor resolution of strain field measurement based on the DIC method when it is divided into two halves, usually 800 px × 640 px or even less. Compared with the megapixel resolutions of low-speed cameras, the number of full-field grid in ten megapixel images is 10 times larger than above. For example, Genovese et al. [23] presented a single-camera stereo-DIC system that uses a biprism in front of the camera objective to split a single sensor into two views. Pankow et al. [24] devised a single-camera stereo-DIC system to record images at high speeds using a series of mirrors. Yu and Pan [25] proposed a single-camera high-speed stereo-DIC system using a four-mirror adapter for full-field 3D vibration measurement. One common element across these methodologies was that all these methods were confronted with low-resolution and stable accessory platforms.

The second type of single-camera stereo-DIC system adopts different optical bandpass filters to gather the two different views into the same CCD/CMOS color sensor. This method is questionable when improving the spatial resolution of the high-speed camera, considering that a monochrome camera is replaced by a color camera. This is primarily because there are many camera sensor capability penalties that should be considered in color image acquisition, such as sensor pixel size, sensitivity, and dark current, to balance the spatial resolution and image noise. For example, Yu and Pan [26] designed a single-camera stereo imaging apparatus, and the specimen surface was recorded using a high-speed color CMOS camera, with the images consisting of blue and red channels, which were from two different optical paths. A specialized camera, called a 3CCD color camera, was also trialled in a single-camera stereo-vision system. Hijazi et al. [28] demonstrated the feasibility of using a low-cost 3CCD color camera and recorded six frame sequences at frame rates up to 20 kHz, proving that it was suitable for application in quantitative and transient full-field measurements, such as DIC and particle image velocimetry (PIV). Yu and Pan [29] adopted a single 3CCD color camera for full-field shape, motion, and deformation measurements, thus avoiding sacrificing the spatial resolution of the camera sensor, with the aid of a specially designed color separation device using a beam splitter and two optical bandpass filters.

The two types of single-camera stereo-DIC system mentioned above are not limited to static and steady-state vibration measurement but also include transient response measurement as long as the frame rate and image quality are sufficient. Moreover, a new single-camera method for steady-state vibration measurement was proposed newly. Gorjup et al. [30] utilized a moving high-speed camera, extending the image-based vibration measurement method. Spatial small harmonic motion can be identified in the frequency domain. The properties of the stationary vibration response of linear, time-invariant mechanical structures are leveraged to produce full-field 3D operating deflection shape measurements using only a single monochrome high-speed camera. The disadvantage is that three or more different views are required, which signifies that at least three images are processed. In addition, multiple perspectives must have the same field of view (FOV), which is a limitation for specimen shapes. Earlier, Quan et al. [31] attempted at using the traditional 2D-DIC method and local displacement gradient feature in the image coordinate system for 3D displacement measurement. The robustness is at the risk of the filtering size and signal-noise ratio, especially for a small angle between the local normal vector of specimen surface and optical axis. Essentially, the core issue is also how to separate the component of out-plane displacement from the image coordinate system under a series of assumptions such as a linear approximation of deformation mapping with first-order shape function.

Although quite a few single-camera DIC methods have been published, various contributions devoted to metal sheet or similar composite structure vibration analysis have been proposed by several teams based on traditional 3D-DIC with two high-speed cameras. Huňady and Hagara [14] realized a huge efficiency improvement using the enhanced frequency function (EFRF) instead of FRF in modal parameter estimation procedure, with hundreds or thousands of output degrees of freedom in full-field modal analysis being necessary. Chang et al. [32] followed the adaptive geometric moment descriptor (AGMD) and combined it with K-SVD and Gram–Schmidt orthonormalization (GSO) to achieve data compression of displacement maps. Passieux et al. [33] developed a new regularized DIC method for time-dependent measurements to improve the space field uncertainties and achieve a trade-off between the frame rate and spatial resolution. Bharadwaj et al. [15, 34] migrated the strain expansion-reduction approach (SERA) from the traditional finite element analysis (FEA) to experimental mechanics. The full-field strain mode shapes from the FEA were instead obtained from the DIC results. Therefore, a metal sheet vibration measurement was also implemented in the current work.

In this study, the cumbersome settings for single-camera stereo-DIC systems, such as biprisms, mirror adapters, or optical bandpass filters, were removed completely. At the same time, it breaks the traditional methodology in 2D-DIC perpendicular to the specimen surface. Furthermore, using

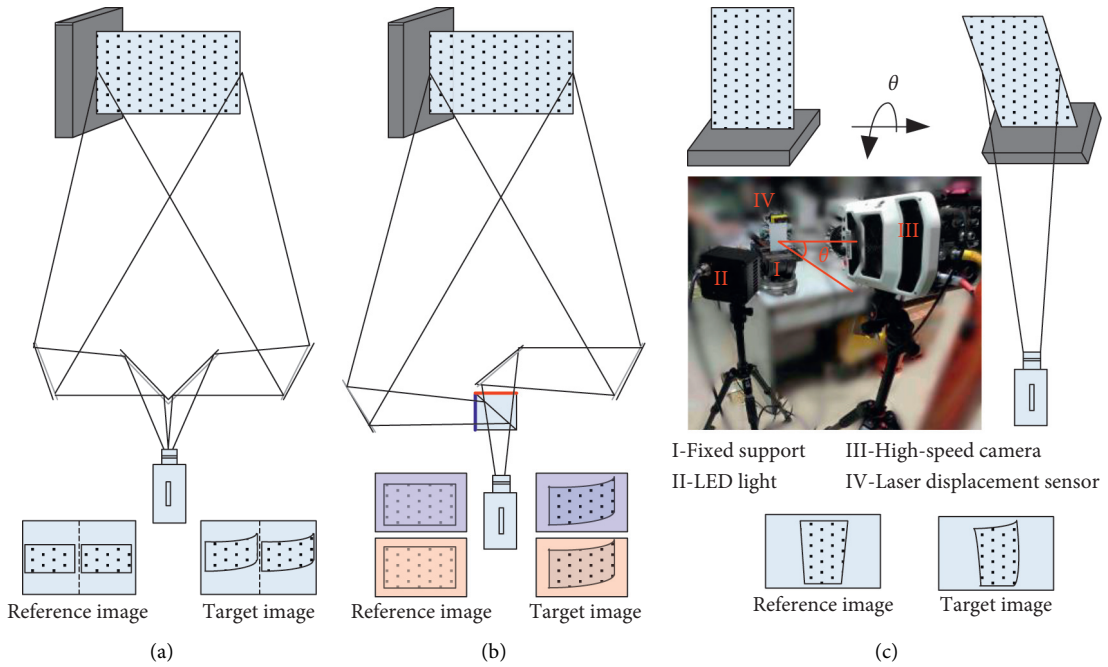


FIGURE 1: Schematics of experimental setup for single-camera DIC vibration measurement: (a) the scheme of splitting the sensor into two halves in [24, 25]; (b) the scheme of splitting the different color channels in [26]; (c) the simple and convenient experimental setup in the current work.

projection components, the full-field vibration displacement of the metal sheet was measured with a single monochrome high-speed camera, which is subject to the classical pin-hole model and projection model. The first four order displacement mode shapes were identified and analyzed. To validate the effectiveness and accuracy of the proposed method, FEA simulations and traditional experimental modal analysis (EMA) were implemented.

2. Experimental Setup and Measuring Principles

Figure 1 schematically shows the three types of single-camera DIC experimental setups for full-field vibration measurements. A four-mirror adaptor with a monochrome camera scheme in Figure 1(a) and a set of optical bandpass filters with color camera scheme in Figure 1(b) were proposed in [24–26], which required a stable environment to hold these optical components, such as an air-floating platform in the laboratory. The scheme proposed in this study was specifically designed to have no complex optical arrangement or projection component and is shown in Figure 1(c). Although the vibration deformation is strictly a 3D curved surface, the principal vibrating directions of arbitrary point on the surface were approximately the same due to normal vectors, almost parallel with regard to the metal sheet or other small-curvature thin-wall parts. The tilt angle, θ, was between the optical axis of the camera and the normal direction of the metal sheet.

The main objectives were as follows: (1) converting the projection component from the in-plane displacement to the vibration direction and (2) converting the pixel unit to an actual physical unit, such as a millimeter, according to the different spatial resolutions at different horizontal heights.

As shown in Figure 1(c), the DIC experimental setup primarily consisted of four parts: (I) a bench vice for fixed support; (II) a Phantom v2511 high-speed camera with a Nikon lens with a focal length of 85 mm; (III) a high-intensity LED light, with a power supply up to 250 W, which gave consideration both to enough luminance in low exposure time and sufficient depth-of-view in small aperture; and (IV) a laser displacement sensor (MTI Corporation, LTS-50-10) which was used to validate the single-point displacement and EMA. The specimen used was a rectangular stainless-steel sheet with dimensions of 138 mm × 80 mm and thickness of 0.5 mm. The depth of the clamped part between the jaws was 18 mm, which gave a vibration measurement region of 120 mm × 80 mm ($H \times W$), as shown in Figure 2(a). The weight of the specimen was 43.5 g after covering a layer of water transfer speckles on one side. Meanwhile, the weight of the bench vice was ~ 15 kg to achieve a reliable fixed boundary. In Figure 2(b), a real image from the HS camera shows the effect of tilt angle, where there was little width at the tip and narrowness at the root. Owing to the tilt angle, θ, the different horizontal heights have different object distances to the camera, corresponding to different spatial resolutions, with two different spatial resolutions of height approximately given by δ_{tip} and δ_{root} in (1). It is worth noting that the middle position is described in the physical coordinates rather than being described in the image coordinates. Thus, the middle position was not in the center of the image but slightly closer to the root. The enlarged four corners in Figure 2(c) showed a satisfactory imaging quality within a certain range of depth of field

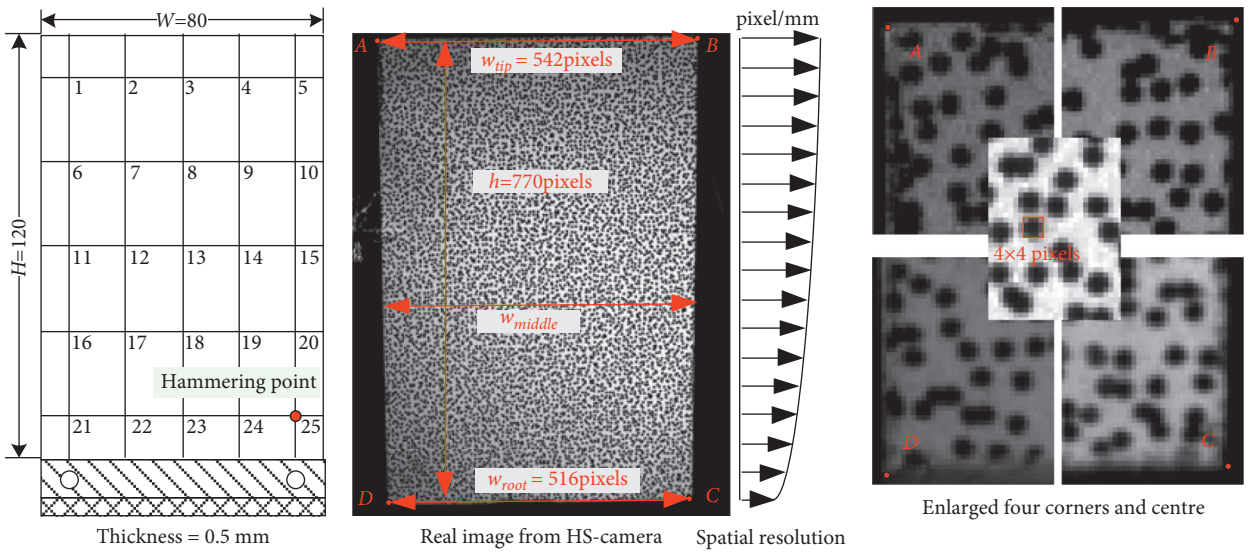


FIGURE 2: Specimen: (a) the schematic of metal sheet size, hammering point, and vibration pickup nodes layout for EMA validation; (b) the real image from HS camera and the schematic of different spatial resolution in different horizontal heights; (c) enlarged four corners and center showing a satisfactory imaging quality in different object distances and illumination uniformity.

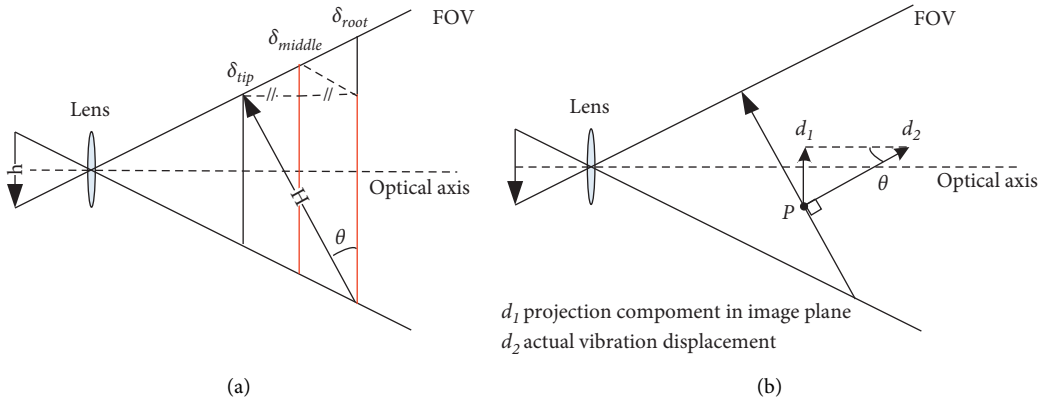


FIGURE 3: Measurement principle: (a) the different spatial resolution in different object distance and specific geometric relations; (b) the displacement projection component from actual vibration to image.

(DOF) and illumination uniformity, which benefited largely from the sufficient light intensity and small aperture. Meanwhile, a speckle is framed in a 4 px × 4 px box in the enlarged center part, illustrating that the water transfer speckle size was suitable for this experimental setup (recommended in 3 px ~ 5 px).

$$\begin{cases} \delta_{\text{tip}} = \dfrac{w_{\text{tip}}}{W} \approx \dfrac{6.7750\text{pixel}}{\text{mm}}, \\ \delta_{\text{root}} = \dfrac{w_{\text{root}}}{W} \approx \dfrac{6.4500\text{pixel}}{\text{mm}}. \end{cases} \tag{1}$$

In addition, the image resolution was 800 px × 640 px at 2000 frames per second (FPS), with an exposure time of 20 μs. There were 8101 images recorded for each excitation at around 4.05 seconds. It should be noted that the conventional local-DIC tracking algorithm was employed, including an initial value estimation and zero-mean normalized sum of the squared difference (ZNSSD) criteria optimization with a first-order shape function. There were 9801 points within the trapezium region of interest (ROI). A grid spacing of ~2 mm and a subset size of 31 px × 31 px were set for data processing.

The bold arrows in Figure 3(a) represent the metal sheets, with the capital letter H representing the object size in millimeters and the small letter h representing the pixel size in the image. Three different spatial resolutions represent the three different object distance planes along the optical axis. δ_{middle} was defined in the precise middle of δ_{tip} and δ_{root}, as shown in (2). According to geometric similarity, the tilt angle, θ, was derived using (3) and (4), in this case. Combined with the projection relationship in Figure 3(b), the first step was to convert the in-plane displacement to the vibration direction. The vibration response was then derived

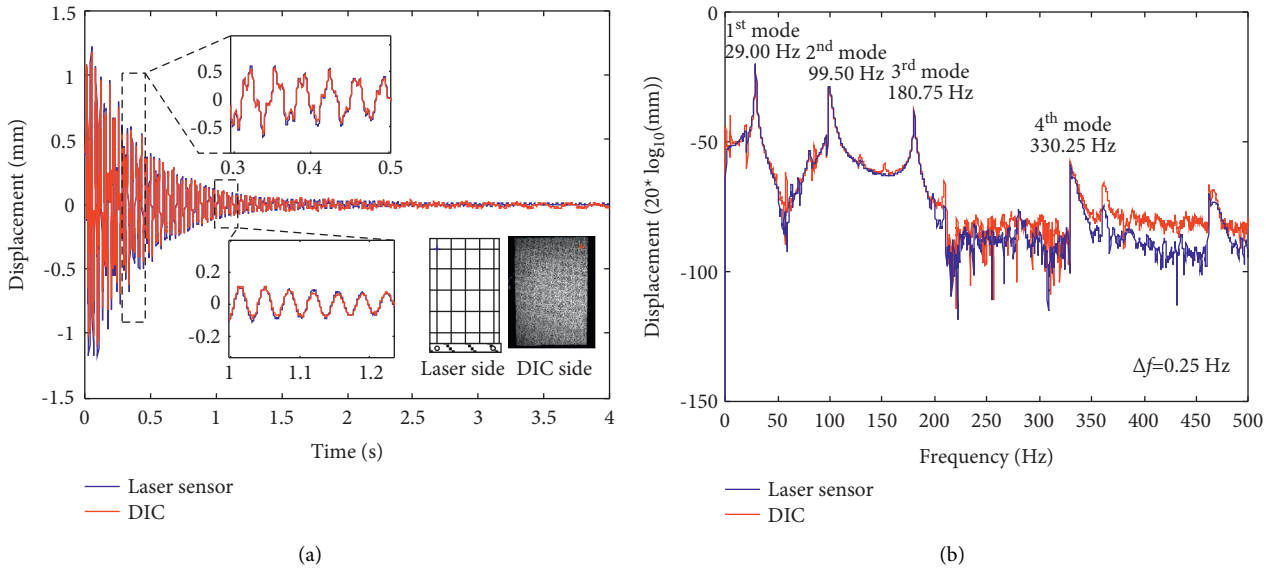


FIGURE 4: Laser and DIC measurement results of a single-point response of impulse excitation: (a) time domain; (b) frequency domain.

from the DIC results in (5), where δ_y denotes the local spatial resolution, d_1 represents the DIC tracking results in the y-direction in pixels, and d_2 represents the vibration response associated with d_1. The vibration response, D, in millimeters was defined from d_2 and δ_y in (5).

$$\delta_{\text{middle}} = \frac{\left(\delta_{\text{tip}} + \delta_{\text{root}}\right)}{2} \approx \frac{6.6125\text{pixel}}{\text{mm}}, \tag{2}$$

$$\delta_{\text{middle}} = \frac{h}{H \cdot \cos\theta}, \tag{3}$$

$$\theta = \arccos\left(\frac{h}{H \cdot \delta_{\text{middle}}}\right) \approx 13.98^{\circ}, \tag{4}$$

$$D = \frac{d_2}{\delta_y} = \frac{d_1}{\delta_y \cdot \sin\theta}. \tag{5}$$

The origin of the pixel coordinates was located at the top-left corner of the image. In this case, the y-coordinate of tip line AB was around $y_{\text{tip}} = 10\text{px}$, and the root line CD was around $y_{\text{root}} = 780\text{px}$. The variable upper bound integral expression of the local spatial resolution can then be written as (6) under the boundaries of $\delta_{y=10} = \delta_{\text{tip}}$ and $\delta_{y=780} = \delta_{\text{root}}$.

$$\delta(y) = \delta_{\text{tip}} + \left(\delta_{\text{root}} - \delta_{\text{tip}}\right)\frac{1}{H}\int_{10}^{y}\frac{1}{\delta(y)}\,\mathrm{d}y, \tag{6}$$

when solved as the following equation:

$$\delta(y) = \sqrt{ky + b}, \quad y \in [1, 800], \tag{7}$$

substituting the upper and lower boundary conditions and solving the coefficients $b \approx 45.96$ and $k \approx -5.58 \times 10^{-3}$. Therefore, the metal sheet vibration displacement was mapped from the DIC tracking results using (5) and (7).

3. Results and Discussion

3.1. Measurement Results of Proposed Method. As shown in Figure 4, a single-point displacement response was compared between the laser displacement sensor and DIC results in the time and frequency domains. The position of the measuring point was near the tip corner on both sides of the metal sheet, as shown in Figure 4(a). In the time-domain diagram, the two displacement response curves were generally consistent, with a redundant low-frequency component in the DIC results of approximately 5 Hz, which was considered as the shake of the camera system. The influence of this low-frequency component in the data process was ignored because it was far less than the modal frequencies, and more details about the influence of the camera rotation can be found in [35]. Through the fast Fourier transform (FFT) and logarithm operation, the displacement amplitude spectra are presented in Figure 4(b). The first four modal frequencies were picked with ranges from 29.00 Hz to 330.25 Hz, and the frequency resolution is 0.25 Hz.

The full-field displacement response under impulse excitation is shown in Figure 5, and several vibration modes were excited. These nine frames represent a series of deformations in 0.04 s intervals, which was little more than a period of the first-order vibration modal. Comparing $T = 0.005s$ and $T = 0.040s$, the duration from $T = 0.005s$ to $T = 0.040s$ went through a period of the first-order vibration modal approximately, which was the first bending mode. It went through approximately half a period of the second-order modal vibration, which was the first torsion mode. The modal analysis results in Section 3.2 provide more accurate mode shapes of the single edge fixed metal sheet. In Figure 5, the clamped edge was

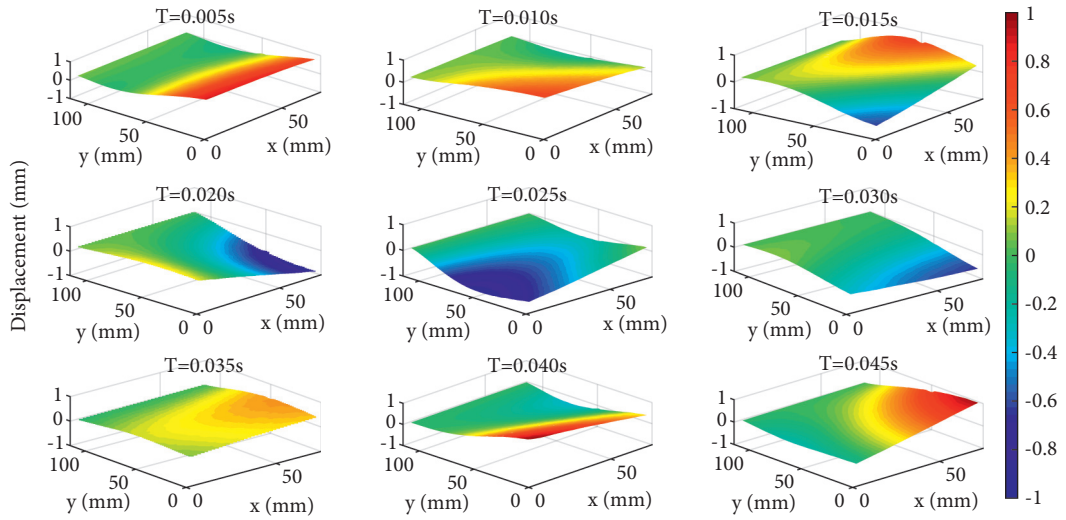


Figure 5: Full-field displacement of vibration response under the impulse excitation at a series of moments from $T = 0.005s$ to $T = 0.045s$.

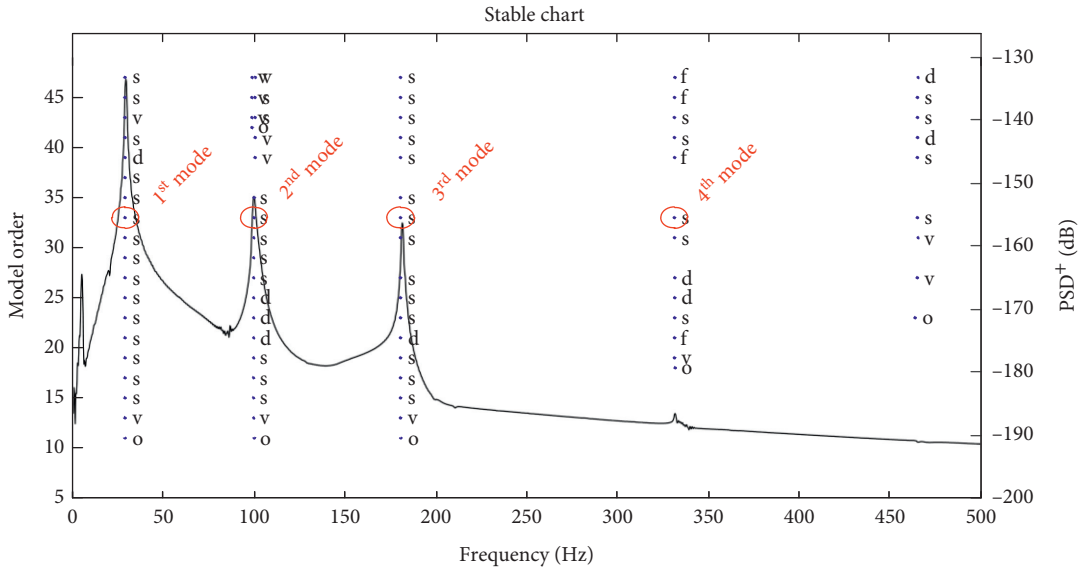


Figure 6: Stable chart of modal parameters identification using p-LSCF method. The original displacement data is from the proposed single-camera DIC results. The first four vibration modes' poles, which lie in the 33rd model order of the polynomial basis function, are picked.

around the line $y = 120$mm, where the displacement response is close to zero at every moment. At the other end, the maximum amplitude of the displacement response was approximately 1 mm.

3.2. Modal Analysis and Validation. In this section, the poly-reference least-squares complex frequency-domain method (p-LSCF) [36] was is employed for modal parameter identification of experimental data. Compared with the traditional EMA method, the DIC results have thousands of vibration receiver points and were a challenge for the parameter identification process due to the cost function size of the reduced linear least-squares estimates, detonated by the large frequency response function matrix (FRF). In a recent study, an approximated enhanced frequency response function (EFRF) was utilized to reduce the matrix size [14], computational time, and memory usage. In the current work, 2016 output channels of DIC results were implemented for modal parameter identification. The stable chart is shown in Figure 6, and the poles of the first four vibration modes, which lie in the 33rd order of the polynomial basis function, were selected.

The EMA and FEA validations were implemented. In this traditional EMA, there are 25 points for the laser displacement sensor successively, and the hammering point was invariable at point 25, as shown in Figure 2(a). In the simulation FE model, it was necessary to demonstrate some material parameters, such as the density, $\rho = 7.93 \times 10^3 \text{kg}/m^3$; Young's modulus, $G = 199$Gpa; and Poisson's ratio $\nu = 0.3$. The solid shell element dimension was ~1 mm. A total of 9801 nodes on the metal sheet were extracted in the modal analysis. The boundary condition

Table 1: The first four natural frequencies and damping ratios of the proposed DIC method, EMA, and FEA results.

	Single-camera DIC		EMA		FEA	Frequency difference between DIC and EMA (%)	Frequency difference between DIC and FEA (%)
	f_n (Hz)	ζ (%)	f_n (Hz)	ζ (%)	f_n (Hz)		
1st mode	28.83	1.09	28.52	0.51	28.94	1.09	0.38
2nd mode	99.81	0.83	96.26	0.91	97.52	3.69	2.35
3rd mode	180.49	0.33	178.94	0.33	179.87	0.87	0.34
4th mode	331.34	0.33	327.23	0.04	328.99	1.26	0.71

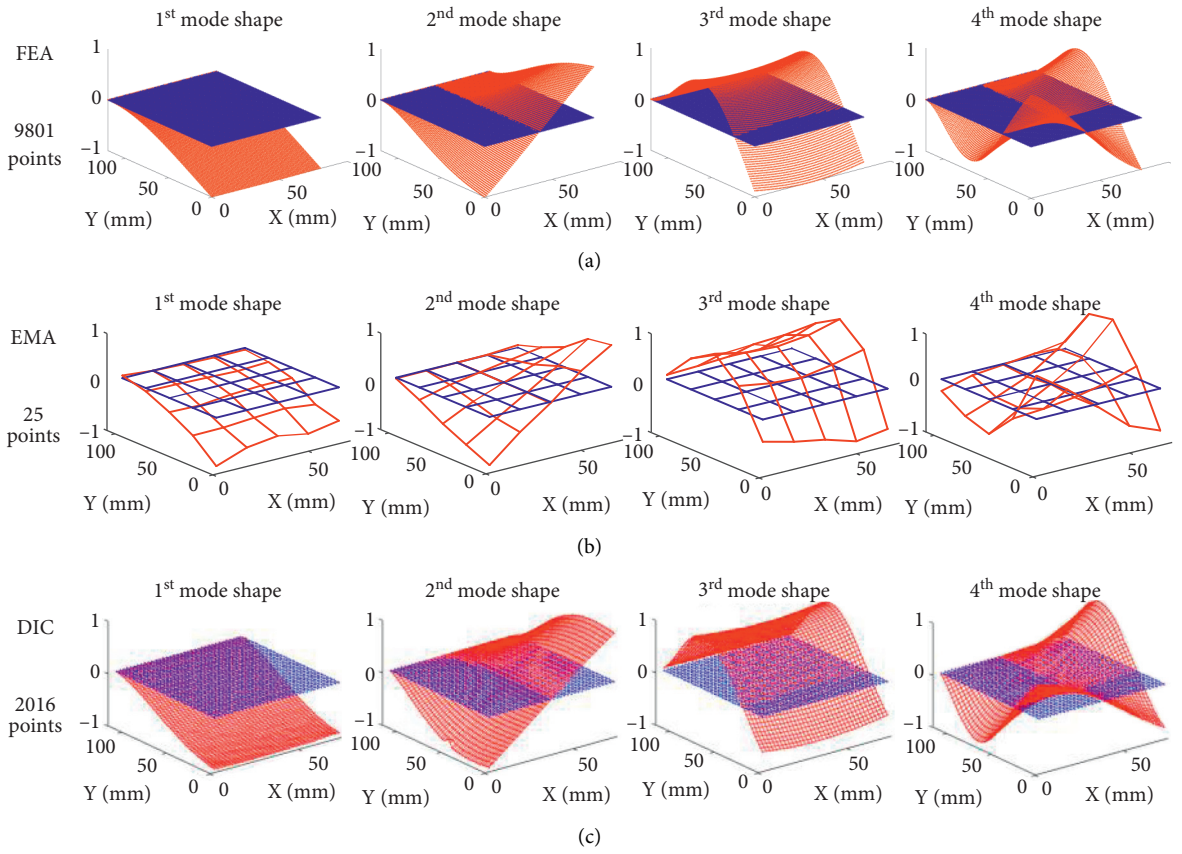


Figure 7: The first four mode shapes of the FEA, the EMA, and the proposed DIC method.

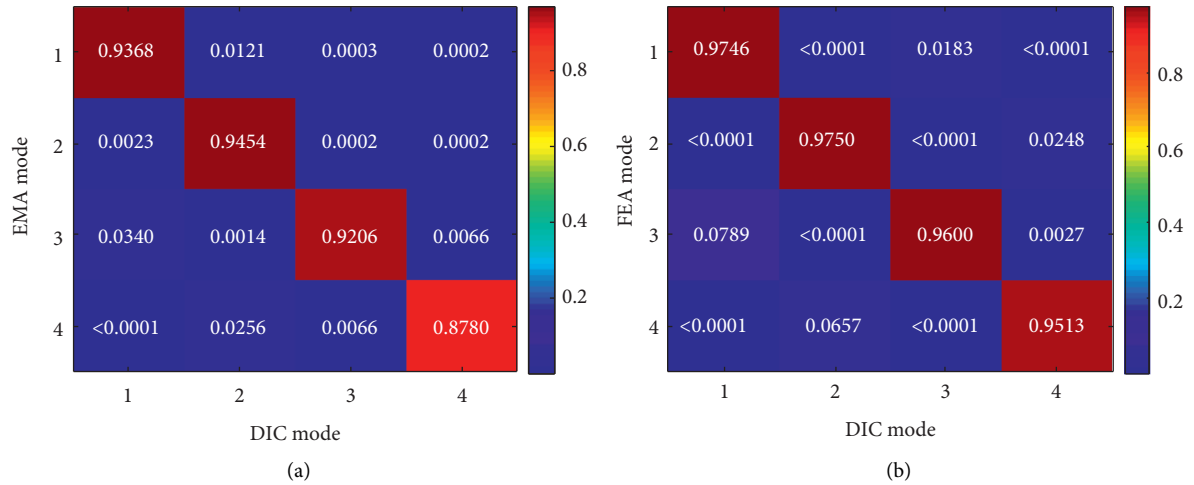


Figure 8: MAC of the first four mode shapes: (a) MAC of DIC and EMA results; (b) MAC of DIC and FEA results.

was unilateral fixation such as a cantilever plate. The first four natural frequencies identified from the proposed DIC method, EMA, and FEA results are listed in Table 1. The first four damping ratios of the proposed DIC method and EMA are also given, with the maximum value being only 1.09% in this small damping structure. In Table 1, the maximum frequency difference between the 2nd mode of DIC and EMA results was 3.69%. This result was slightly better than that of previous studies in [25], which were 4.36% and 10%, respectively, corresponding to the traditional accelerometer measurements and FEA results.

As shown in Figure 7, the first four mode shapes identified from the FEA, EMA, and DIC data are presented together. The blue grids represent the reference equilibrium position, and the red grids represent the mode shapes. All modal vectors were normalized in ±1. The first and third modes are bending, and the second and fourth modes are torsion. Different numbers of nodes correspond to different grid densities, and the EMA mode shapes were the sparsest among these three groups; meanwhile, the curved surfaces of the FEA and DIC mode shapes were smooth. According to the modal assurance criterion (MAC) results in Figure 8, the MAC of DIC mode shapes with FEA and that with EMA were both acceptable. Meanwhile, the MAC with FEA was slightly better than EMA, which benefits from the simple boundary conditions and structure of the metal sheet in the FE model.

4. Conclusions

With a projection component, the vibration displacement of the metal sheet was measured using a single high-speed camera, and the FEA and EMA were validated. Firstly, a single-point displacement response was compared between the laser displacement sensor data and DIC data in the time and frequency domains. Secondly, the modal analysis results, including natural frequencies, mode shapes, and MAC matrices, demonstrated consistent results among the three different methods.

In several single-camera DIC vibration measurement schemes, this method is simpler, more convenient, and more accurate for the vibration measurement of metal sheets because there is no need for other optical devices such as mirrors or prisms, and there is no use of a color camera instead of a monochrome camera. The monochrome camera has a better sensitivity and signal-to-noise ratio than those of the color camera, which is equipped with the same size of CCD/COMS sensor.

Furthermore, this is an alternative strategy to the full-field vibration measurement of small-curvature thin-wall parts and differs from the motion camera for the multiview method [30], which requires steady-state vibration and data processing in the frequency domain. In contrast, the full-field transient response in the time domain can be recorded and processed directly through this proposed method. It is worth noting that the full-field displacement response was projected from the physical coordinates to image coordinates, assuming that the displacement is a linear mapping from physical coordinates to image coordinates. In other words, it is based on the condition that the vibration amplitude is far less than the structure size in the FOV. For most nonflexible structures, this requirement is similar to that of a natural match.

Overall, regardless of the splitting optical path or motion camera for multiple views, the scheme was designed for three-dimensional reconstruction and following traditional stereo vision with a single camera. However, when it comes to small vibration measurements of small-curvature thin-wall parts, the proposed approach is a good choice. The full-field vibration vectors could be determined through the simple affine transformation or FE model, which is expected to be applied to damage detection with mode shape continuity and more dynamic performance measurements.

Data Availability

All data, models, and codes that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research program was funded by the project of the Research on the Engine Nonlinear Coupling Vibration Characteristics, the National Key Research & Development Program of China (no. 2016YFF0203304), and the Postgraduate Research & Practice Innovation Program of Jiangsu Province (no. KYCX18_0261).

References

[1] M. A. Sutton, J. J. Orteu, and H. W. Schreier, "Image correlation for shape, motion and deformation measurements: basic concepts, theory and applications," Springer, New York, NY, USA, 2009.

[2] B. Pan, K. Qian, H. Xie, and A. Asundi, "Two-dimensional digital image correlation for in-plane displacement and strain measurement: A Review Meas," *Science Technology*, vol. 20, no. 6, Article ID 062001, 2009.

[3] M. A. Sutton and F. Hild, "Recent advances and perspectives in digital image correlation," *Experimental Mechanics*, vol. 55, no. 1, pp. 1–8, 2015.

[4] B. Pan, "Digital image correlation for surface deformation measurement: historical developments, recent advances and future goals Meas," *Science and Technology*, vol. 29, no. 8, Article ID 082001, 2018.

[5] F. Nalli, L. Cortese, and F. Concli, "Ductile damage assessment of Ti6Al4V, 17–4PH and AlSi10Mg for additive manufacturing," *Engineering Fracture Mechanics*, vol. 241, Article ID 107395, 2021.

[6] M. M. Sherif, E. M. Khakimova, J. Tanks, and O. E. Ozbulut, "Cyclic flexural behavior of hybrid SMA/steel fiber reinforced concrete analyzed by optical and acoustic techniques," *Composite Structures*, vol. 201, pp. 248–260, 2018.

[7] J. Xu, W. X. Wang, Q. H. Han, and X. Liu, "Damage pattern recognition and damage evolution analysis of unidirectional

CFRP tendons under tensile loading using acoustic emission technology," *Composite Structures*, vol. 238, Article ID 111948, 2020.

[8] L. Maccioni and F. Concli, "Fracture locus of a CORTEN steel: finite Element calibration based on experimental results," *Procedia Structural Integrity*, vol. 24, pp. 738–745, 2019.

[9] B. V. Farahani, F. Q. de Melo, P. J. Tavares, and P. M. G. P. Moreira, "New approaches on the stress intensity factor characterization - review," *Procedia Structural Integrity*, vol. 28, pp. 226–233, 2020.

[10] K. Zaletelj, J. Slavic, and M. Boltezar, "Full-field DIC-based model updating for localized parameter identification," *Mechanical Systems and Signal Process*, vol. 164, Article ID 108287, 2022.

[11] T. J. Beberniss and D. A. Ehrhardt, "High-speed 3D digital image correlation vibration measurement: recent advancements and noted limitations," *Mechanical Systems and Signal Processing*, vol. 86, pp. 35–48, 2017.

[12] J. Baqersad, P. Poozesh, C. Niezrecki, and P. Avitabile, "Photogrammetry and optical methods in structural dynamics-a review," *Mechanical Systems and Signal Processing*, vol. 86, pp. 17–34, 2017.

[13] V. Srivastava and J. Baqersad, "A multi-view optical technique to extract the operating deflection shapes of a full vehicle using digital image correlation," *Thin Walled Structures*, vol. 145, Article ID 106426, 2019.

[14] R. Huňady and M. Hagara, "A new procedure of modal parameter estimation for high-speed digital image correlation Mech. Syst," *Signal Processing*, vol. 93, pp. 66–79, 2017.

[15] K. Bharadwaj, A. Sheidaei, A. Afshar, and J. Baqersad, "Full-field strain prediction using mode shapes measured with digital image correlation," *Measurement*, vol. 139, pp. 326–333, 2019.

[16] B. Pan, K. Li, and W. Tong, "Fast, robust and accurate digital image correlation calculation without redundant computations," *Experimental Mechanics*, vol. 53, no. 7, pp. 1277–1289, 2013.

[17] X. Shao, X. Dai, and X. He, "Noise robustness and parallel computation of the inverse compositional Gauss-Newton algorithm in digital image correlation," *Optics and Lasers in Engineering*, vol. 71, pp. 9–19, 2015.

[18] L. Zhang, T. Wang, Z. Jiang et al., "High accuracy digital image correlation powered by GPU-based parallel computing," *Optics and Lasers in Engineering*, vol. 69, pp. 7–12, 2015.

[19] D. A. Ehrhardt, M. S. Allen, S. Yang, and T. J. Beberniss, "Full-field linear and nonlinear measurements using continuous-scan laser Doppler vibrometry and high speed three-dimensional digital image correlation," *Mechanical Systems and Signal Processing*, vol. 86, pp. 82–97, 2017.

[20] P. L. Reu, D. P. Rohe, and L. D. Jacobs, "Comparison of DIC and LDV for practical vibration and modal measurements," *Mechanical Systems and Signal Processing*, vol. 86, pp. 2–16, 2017.

[21] P. Castellini, M. Martarelli, and E. P. Tomasini, "Laser Doppler Vibrometry: development of advanced solutions answering to technology's needs," *Mechanical Systems and Signal Processing*, vol. 20, no. 6, pp. 1265–1285, 2006.

[22] B. Pan, L. Yu, and Q. Zhang, "Review of single-camera stereo-digital image correlation techniques for full-field 3D shape and deformation measurement," *Science China Technological Sciences*, vol. 61, no. 1, pp. 2–20, 2018.

[23] K. Genovese, L. Casaletto, J. A. Rayas, V. Flores, and A. Martinez, "Stereo-Digital Image Correlation (DIC) measurements with a single camera using a biprism," *Optics and Lasers in Engineering*, vol. 51, no. 3, pp. 278–285, 2013.

[24] M. Pankow, B. Justusson, and A. M. Waas, "Three-dimensional digital image correlation technique using single high-speed camera for measuring large out-of-plane displacements at high framing rates," *Applied Optics*, vol. 49, no. 17, pp. 3418–3427, 2010.

[25] L. Yu and B. Pan, "Single-camera high-speed stereo-digital image correlation for full-field vibration measurement," *Mechanical Systems and Signal Processing*, vol. 94, pp. 374–383, 2017.

[26] L. Yu and B. Pan, "Full-frame, high-speed 3D shape and deformation measurements using stereo-digital image correlation and a single color high-speed camera," *Optics and Lasers in Engineering*, vol. 95, pp. 17–25, 2017.

[27] L. W. Kohlman, *Improvements in High Speed, High Resolution Dynamic Digital Image Correlation for Experimental Evaluation of Composite Drive System Components*, NASA Glenn Research Center, Cleveland, OH, USA, 2013, https://ntrs.nasa.gov/api/citations/20140004925/downloads/20140004925.pdf.

[28] A. Hijazi, A. Friedl, C. Cierpka, C. Kähler, and V. Madhavan, "High-speed imaging using 3CCD camera and multi-color LED flashes," *Measurement Science and Technology*, vol. 28, no. 11, Article ID 115401, 2017.

[29] L. Yu and B. Pan, "Color stereo-digital image correlation method using a single 3CCD color camera," *Experimental Mechanics*, vol. 57, no. 4, pp. 649–657, 2017.

[30] D. Gorjup, J. Slavič, and M. Boltežar, "Frequency domain triangulation for full-field 3D operating-deflection-shape identification," *Mechanical Systems and Signal Processing*, vol. 133, Article ID 106287, 2019.

[31] C. Quan, C. J. Tay, W. Sun, and X. He, "Determination of three-dimensional displacement using two-dimensional digital image correlation," *Applied Optics*, vol. 47, no. 4, pp. 583–593, 2008.

[32] Y.-H. Chang, W. Wang, T. Siebert, J.-Y. Chang, and J. E. Mottershead, "Basis-updating for data compression of displacement maps from dynamic DIC measurements," *Mechanical Systems and Signal Processing*, vol. 115, no. 15, pp. 405–417, 2019.

[33] J.-C. Passieux, R. Bouclier, and J. N. Périé, "A space-time PGD-DIC algorithm," *Experimental Mechanics*, vol. 58, no. 7, pp. 1195–1206, 2018.

[34] J. Baqersad and K. Bharadwaj, "Strain expansion-reduction approach," *Mechanical Systems and Signal Processing*, vol. 101, pp. 156–167, 2018.

[35] R. Balcaen, P. L. Reu, P. Lava, and D. Debruyne, "Influence of camera rotation on stereo-DIC and compensation methods," *Experimental Mechanics*, vol. 58, no. 7, pp. 1101–1114, 2018.

[36] P. Guillaume, P. Verboven, S. Vanlanduit, H. V. Auweraer, and B. Peeters, "A ploy-reference implementation of the least-squares complex frequency-domain estimator," 2003, https://www.researchgate.net/publication/265423092.

Hindawi
Shock and Vibration
Volume 2022, Article ID 2007022, 17 pages
https://doi.org/10.1155/2022/2007022

Research Article

Singularity-Free Adaptive Controller for Uncertain Hysteresis Suspension Using Magnetorheological Elastomer-Based Absorber

Hoa Thi Truong, Xuan Bao Nguyen, and Cuong Mai Bui

The University of Danang, University of Technology and Education, Danang, Vietnam

Correspondence should be addressed to Xuan Bao Nguyen; xuanbao233@gmail.com

Received 28 July 2021; Revised 8 December 2021; Accepted 22 December 2021; Published 11 January 2022

Academic Editor: Said Elias

The magnetorheological elastomer (MRE) is a smart material widely used in recent vibration systems. A system using these materials often faces difficulties designing the controller such as unknown parameters, hysteresis state, and input constraints. First, a model is designed for the MRE-based absorber to portray the behavior of MRE and predict the appropriate electric current supplied. The conventional adaptive controller often suffers from so-called control singularities. The singularity-free adaptive controller is proposed to eliminate the singularity with parametric uncertainty. The proposed controller consists of four components: an adaptive linearizing controller, a deputy adaptive neural network controller, an auxiliary part designed for the controller to overcome the input constraint problem, and a smooth switching algorithm used to exchange the takeover rights of the two controllers. Moreover, the controller is designed to obtain the stabilization of hysteretic state estimation for the vibration system. The adaptive algorithms are proposed to update the unknown system parameters and to observe the unmeasurable hysteretic state. Meanwhile, closed-loop system stability is comprehensively assessed. Finally, the simulation performed on a quarter-car suspension with an MRE-based absorber shows the proposed controller's efficiency.

1. Introduction

Semiactive vibrating systems using magnetorheological materials have become well known. In particular, the magnetorheological elastomers (MREs) used in semiactive controls have recently emerged as a new material for vibration control [1, 2]. The system can change the natural frequency by varying the stiffness of the material. These properties are attractive for many engineering applications such as vibration isolators and vibration absorbers [3–5]. For example, Gao et al. used the MRE as a semiactive vibration isolator to suppress the vibration [4]. The results showed that the natural frequency was adjustable by 3.9 Hz. This study introduces the MRE-based absorber to reduce the suspension system's vibration caused by road irregularities and onboard engines. Using the MREs, the system can adjust its own frequency to avoid resonances for different types of road and engine speeds. It is expected that the MRE-based absorber overcomes the limitations of the MR damper. There are many methods proposed to represent material properties in recent years [6–10]. Optimization algorithms are an effective method to determine model parameters. An innovative nonlinear model has been proposed for MRE, and an improved PSO algorithm has been designed to estimate the model's parameters [7]. An extreme machine learning method was proposed to predict the device's nonlinear (shear force) responses with applied current, displacement, and velocity level. The new swarm optimization method, called a binary coded discrete cat, was applied to select the optimal input and the number of neurons in the hidden layer for the network development [8]. The fruit fly optimization algorithm was used to determine the model parameters. A three-story standard building model under four standard earthquake excitations was tested to evaluate the model's effectiveness [9]. Artificial intelligence approaches, including linear and nonlinear regression analysis, adaptive neural fuzzy inference systems (ANFIS), and artificial neural network (ANN) techniques, are highly reliable methods for

predicting various nonlinear properties, which have been comprehensively analyzed in [10].

The MRE-based device needs a suitable controller to achieve efficiency in the vibration system. The vibration system using an MRE-based absorber is as effective as an active system without the need for large energy. In [11–15], semiactive controllers have been widely used in vibration systems, such as the sky-hook, ground-hook, fuzzy clipped on-off, and LPV approaches. These controllers do not consider the system's dynamics, so the controller does not guarantee stability in some cases. Many modern controllers have been proposed for the semiactive system, such as optimal control, adaptive control, and robust linear controller [16–18]. The adaptive control strategies ensure asymptotic stability with a small gain. However, singularities can occur, which causes a tremendous control force in these controllers. A common remedy is to limit the estimated parameter to a compact set with no specified singularity. The system parameters were bounded by the maximum and minimum values to ensure that the singularity does not occur [19, 20]. In recent years, adaptive intelligent control algorithms have achieved high efficiency in controlling complex, time-varying, and highly nonlinear civil structures [21]. These algorithms mainly work on the principles of soft computing methods and artificial intelligence. The adaptive neural network (ANN) controller has recently achieved high efficiency in controlling the system with unknown dynamics [22]. Optimization associated with multiple control devices is considered a difficult task. Rashid et al. [23] proposed an adaptive algorithm based on acceleration response combined with a displacement optimization algorithm for 5-stage steel frames. However, the ANN controller often requires large amounts of computation. The unknown dynamics have been approximated by the radial basis function where the weights are optimal. However, this method cannot identify the parameters of the system such as mass, stiffness, and damping coefficient. Therefore, control strategies encompassing all the aforementioned controller's advantages and eliminating drawbacks should be designed to yield high-quality performance.

The major challenge with the semiactive device is the control force limitation and hysteresis state. The force constraint is a complicated problem because the maximum force value depends on the displacement and velocity value. Consequently, the actuator is inadequate in the controller requirements. Actuator limitations need special attention in the controller design. Recent studies have also mentioned this problem in engineering systems [24]. Hysteresis is a fundamental phenomenon in engineering. The semiactive vibration system usually exhibits a stable hysteretic state. The Bouc–Wen hysteresis model (BWM) is widely used to represent the properties of MR materials which have attracted researchers to develop intelligent vibration systems [25, 26]. The model is flexible and can be adjusted for different hysteretic states. BWM, with its flexibility in shape control, has been used to describe asymmetric hysteresis loops. The parametric modeling approach includes spring, damping, and Bouc–Wen models represented by a mathematical function. The coefficients of this function can be determined by using an optimization technique. The parameter values are changed until the model's output force closely matches to the experimental output force. In contrast, nonparametric models are entirely based on the performance of a specific MR-based device, such as the neural network model and fuzzy model. These models are more flexible, but the physical relationship between modal parameters and hysteresis phenomena may not be explicitly maintained. These methods need large amounts of data and are performed in advance. We introduce a hysteresis observer to approximate the hysteresis state. The developed observer is expected to estimate the hysteresis quickly. The observer supports the controller to improve robustness against unmeasurable hysteresis. For practical applications, a novel controller is necessary to ensure the stability of a semiactive system.

In this study, we proposed an innovative control method to overcome the singularity in the traditional adaptive controller. The controller aims to exploit the advantages of adaptive controllers and neural network controllers and eliminate the disadvantages of these controllers with a smooth switching mechanism. Consequently, the denominator part of the adaptive control formula is absorbed near zero to eliminate the singularity problem. The adaptive controller is temporarily disabled in the event of a singularity occurring. An adaptive neural network controller is introduced to take over the system to ensure system stability. The displacement response converges to zero using the proposed controller, and the output control value can be remarkably reduced near the singularity condition. Firstly, a model was designed for the MRE-based isolator using the Bouc–Wen model, and an inverse model was developed to predict the desired current. Next, the ANN controller is used to estimate the uncertainty nonlinearity, and an adaptive controller (e.g., sliding adaptive controller) is designed to override the approximation error. A smooth switching algorithm is introduced to observe the singularity and determine the control authority between the ANN controller and a conventional adaptive controller. The new strategy is expected to avoid singularities, small control force, and fast stability. The novel adaptive controller includes five components:

(i) A robust adaptive controller is designed to ensure system stability.

(ii) An ANN controller is designed as the temporary controller in the singularity.

(iii) A smooth switching is used to exchange the take-over rights of the two controllers.

(iv) An auxiliary controller is developed to overcome the input constraint.

(v) Adaptive laws provide online estimates of the uncertain parameters without bounds, and a hysteresis observer is proposed to support the controller.

2. Magnetorheological Elastomer (MRE)

2.1. Model of MRE-Based Absorber. Three main materials used to fabricate the MRE samples included the matrix silicon RTV (68%) of the brand Shin-Etsu, carbon iron

powder with 20 μm diameter (30%) of the brand BASF SG-BH, and silicone oil (2%). MRE samples' fabrication procedures like natural rubber synthesis consist of mixing, compressing, molding, and curing. Firstly, these components were mixed to form a homogeneous mixture for 12 minutes. The mixture was placed in a vacuum chamber to remove air bubbles inside the material for 30 minutes. Finally, the mixture was vulcanized in a mold under a magnetic field or without a magnetic field for 24 hours at room temperature (26 degrees Celsius). Anisotropic MRE samples were vulcanized in a magnetic field, while isotropic MRE samples were vulcanized without a magnetic field. We use $25 \times 25 \times 8$ mm cube samples of MRE materials for the experiment.

The MRE-based absorber is used in this study, whose properties depend on displacement, amplitude, frequency, and magnetic field. In particular, its stiffness increases significantly when the applied current is increased. Consequently, the absorber operates efficiently over a wide range of frequencies presented in the research.

An MRE model is necessary for vibration system design; the hysteresis force-displacement loop is a major challenge under different applied currents. In this study, the Bouc–Wen model was used to present the behavior of MRE as shown in Figure 1. The model consists of a Bouc–Wen component and a Maxwell component. In the Bouc–Wen model, the evolutionary variable z describes the hysteresis behavior. The force of the MRE-based absorber is given by

$$F_{MRE} = \alpha k_0 x + c_0 \dot{x} + (1 - \alpha) k_0 z, \tag{1}$$

$$\dot{z} = A\dot{x} - \beta |\dot{x}| |z|^{n-1} z - \gamma \dot{x} |z|^n, \tag{2}$$

where the linear stiffness force and purely hysteretic force are $\alpha k_0 x$ and $(1 - \alpha) k_0 z$, respectively. The coefficient, $\alpha \in (0, 1)$, represents the linearity level of the loop. The size and the shape of the hysteresis loops are determined by nondimensional parameters A, n, β, and γ as shown in equation (2). The parameter A has a significant influence on the force amplitude of the hysteresis, β and γ represent the shape of the hysteresis, and n is the order transition from linear to nonlinear state that was set to be one to reduce the amount of computation.

The variables of the model are approximated under the input current as follows [27, 28]:

$$k_0 = k_{0a} + k_{0b} I, \tag{3a}$$

$$c_0 = c_{0a} + c_{0b} I, \tag{3b}$$

$$\alpha = \alpha_a + a_b I, \tag{3c}$$

$$A = A_a + A_b I, \tag{3d}$$

$$\beta = \beta_a + \beta_b I, \tag{3e}$$

$$\gamma = \gamma_a + \gamma_b I. \tag{3f}$$

The model parameters were identified using a numerical optimization algorithm presented in Figure 2. The genetic algorithm is used to optimize the parameters of the Bouc–Wen model. The parameters were adjusted to fit the experimental data. Data were collected in many different cases (different frequency values, different current values, and different amplitude values). The fit values are listed in Table 1.

The MRE model, equation (1), is analyzed into three components: the viscous passive component, the active component F_a, and nonlinear hysteresis component Φ:

$$\begin{aligned} F_{MRE} &= \alpha_a k_{0a} x + c_{oa} \dot{x} + \left((\alpha_a k_{0b} + \alpha_b k_{0a}) I + \alpha_b k_{0b} I^2 \right) x + c_{0b} I \dot{x} + (1 - \alpha) k_0 z \\ &= \alpha_a k_{0a} x + c_{oa} \dot{x} + F_a + \Phi, \end{aligned} \tag{4}$$

where

$$F_a = \left((\alpha_a k_{0b} + \alpha_b k_{0a}) I + \alpha_b k_{0b} I^2 \right) x + c_{0b} I \dot{x}, \tag{5a}$$

$$\Phi = (1 - \alpha) k_0 z. \tag{5b}$$

2.2. Experimental Tests and Validation. An experimental schematic was set up, as shown in Figure 3. The shear displacements were conducted with the sinusoidal function where the amplitudes were set from 0.4 mm to 0.8 mm and the frequencies were adjustable from 1 Hz to 20 Hz. The experiment was performed with various values of amperages from 0 A to 4 A. The displacement-force responses are compared between the measurement data and the numerical model with different current inputs and at the low-frequency case (1 Hz), as shown in Figure 4(a). In this case, the viscosity is very low, while the effect of hysteresis is very apparent. It can be seen that the effect of hysteresis is significant even at very low frequencies. The Bouc–Wen model with the appropriate parameters portrays very well the nonlinear hysteresis behavior. The viscous behavior was shown when performed at 10 Hz in Figure 4(b). The hysteresis loops tend to become elliptic as the applied current increases. Numerical responses and experimental results achieved a good agreement. The numerical model still achieves high accuracy when applying different current levels compared to the experiments, as shown in Figure 5. Based on the measured data, the current-dependent Bouc–Wen hysteretic model has fit the MRE isolator's dynamic behavior. Because the thickness of the MRE sample is small (0.8 mm), the value of the performance amplitude is also small. We perform at medium and large amplitudes according to 0.4 mm (5% shear stress) and 0.8 mm (10% shear stress), respectively. The

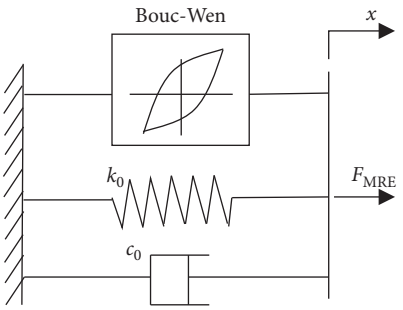


Figure 1: Schematic diagram of MRE-based isolator model.

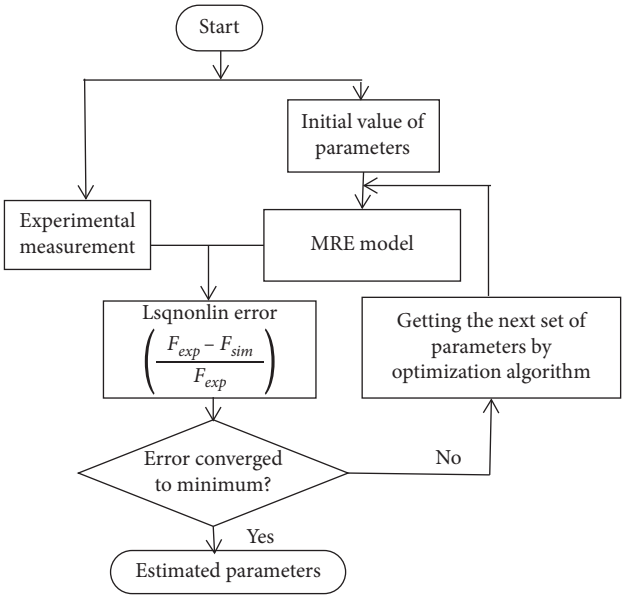
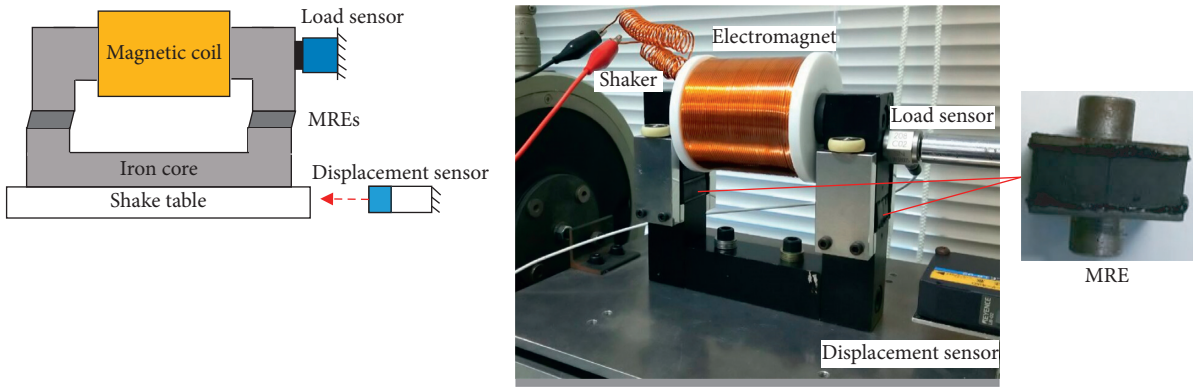


Figure 2: Flowchart of parameter identification using numerical optimization algorithm.

Table 1: Parameter values of MRE-based absorber using the Bouc–Wen model.

k_{0a}	c_{0a}	α_a	A_a	β_a	γ_a	k_{0b}	c_{0b}	a_b	A_a	β_a	γ_a
22	0.05	0.65	2	3.8	−1	12	0.01	0.05	0.2	0.33	0.3



(a) (b)

Figure 3: Experimental schematic for collecting force-displacement data under different current values: (a) schematic; (b) photo.

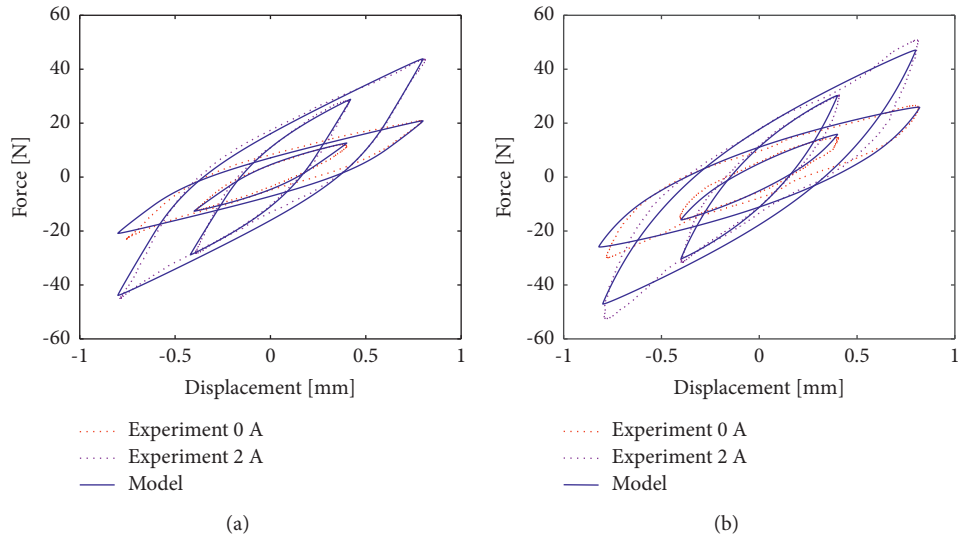


Figure 4: Force-displacement response under two displacement amplitude values (x_0 = 0.4 and 0.8) and two levels of applied current (I = 0 A (0 mT) and I = 2 A (218 mT)): (a) f = 1 Hz and (b) f = 15 Hz.

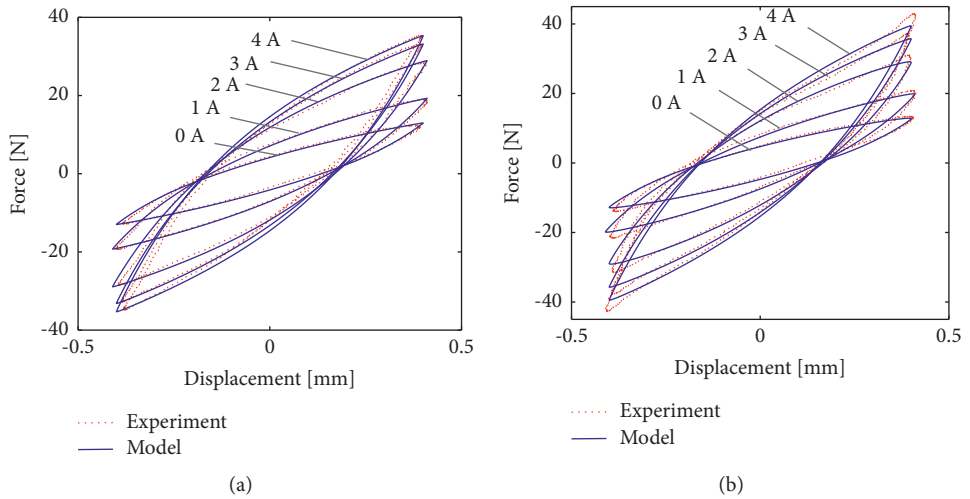


Figure 5: Force-displacement response under various applied currents (I = 0 A (0 mT), 1 A (109 mT), 2 A (218 mT), 3 A (325 mT), and 4 A (437 mT)): (a) f = 1 Hz and (b) f = 15 Hz.

isolator can perform up to 14% shear stress [14]. In the case of large amplitude performance, the MRE material sample thickness needs to be larger, and the magnetic system also needs to be enhanced to increase the system's efficiency.

2.3. Inversed Model. In practical applications, the inversed model is used to determine the input current/voltage for the isolator from the control force [17, 18, 29]. From equation (5a) and Table 1, the active force can be rewritten by variable amperage:

$$\begin{aligned} F_a &= (\alpha_a k_{0b} + \alpha_b k_{0a} + c_{0b}\dot{x})I + (\alpha_b k_{0b} x)I^2 \\ &= (8.9 + 0.01\dot{x})I + (0.6x)I^2. \end{aligned} \tag{6}$$

It is expected that the active force generated coincides with the control force, $F_a(t) = u(t)$. The input current is solved with the following equation and the electric current must be positive and $I \in [0, 4]$ ampere:

$$I^2(t) + \left(\frac{8.9 + 0.01\dot{x}(t)}{0.6x(t)}\right) I(t) - \frac{u(t)}{0.6x(t)} = 0, \tag{7}$$

where $u(t)$ is the control force determined by the proposed controller, $x(t)$ is measured by a displacement sensor, and velocity, $\dot{x}(t)$, is the first derivative of the displacement with respect to time.

The experiment was conducted under harmonic excitation. The applied current was adjusted to different values within the range of 0–4 A. The displacement data and force data were inputs of the MRE inverse model, as shown in Figures 6(a) and 6(b). The response current of the inverse model was compared with the measured current to evaluate the effectiveness of the model, as shown in Figure 6(c). The figure shows that the inverse model performed well in determining the current. The results demonstrate that the developed inverse model can convert the required control force into the value of current, which was fed to the MRE-based isolator.

3. Nonlinear Adaptive Control Design for Suspension Systems

3.1. A Quarter-Car Model Using MRE-Based Absorber. We consider the quarter-car model with MR elastomer as shown in Figure 7, and the system can be given by the following description.

The dynamic equations of the suspension system can be expressed as

$$m_s\ddot{x}_s + c_s(\dot{x}_s - \dot{x}_u) + k_s(x_s - x_u) + F_{MRE} = 0, \tag{8a}$$

$$m_u x_u + c_s(\dot{x}_u - \dot{x}_r) + k_s(x_u - x_s) + k_u(x_u - x_r) - F_{MRE} = 0. \tag{8b}$$

The absorber force, F_{MRE}, is modeled by equation (4) by using the Bouc–Wen model to describe the effect of the hysteresis, $\Phi = (1-\alpha)k_0 z$, $A_\Phi = (1-\alpha)k_0 A$. Let $u = F_a$ be a control input, and the sprung mass dynamics system equation (8a) can be rewritten as

$$m_s x_s + (c_s + c_{0a})(\dot{x}_s - \dot{x}_u) + (k_s + \alpha_a k_{0a})(x_s - x_u) + u + \Phi = 0, \tag{9a}$$

$$\dot{\Phi} = A_\Phi \dot{x} - \beta|\dot{x}|\Phi - \gamma\dot{x}|\Phi|^1. \tag{9b}$$

Assumption 1. The system parameters m_s, k_s, and c_s are uncertain and unbound. The component Φ represents the unmeasurable hysteresis. The control input is bounded by $[u_{\min}, u_{\max}]$.

Lemma 1 (see [30]). *For any $\epsilon > 0$ and $\eta \in R$, the inequality is introduced as*

$$0 \le |\eta| - \eta \tanh\left(\frac{\eta}{\epsilon}\right) \le \kappa\epsilon, \tag{10}$$

where $\kappa = 0.2785$ is the constant. To increase the smoothness of the system, the function sat(.) is replaced by tanh(.) in the robust controller.

3.2. Problem Statement. For the semiactive suspension system, many problems need to be dealt with in the controller design, and in this study, we consider the following aspects.

(1) *Ride Comfort*. In semiactive suspension design, stabilizing the vertical displacement is the main task in the controller's design which absorbs the maximum force of passengers.

(2) *Uncertain Parameters*. The system parameters such as mass, stiffness, and damping coefficient are uncertain and unbound. Singularity may occur during parameter adaptation, which can cause enormous forces or a faulty controller. A new adaptive controller needs to be designed to overcome the issue.

(3) *Actuator Saturation*. The control force is just active in the first and third of the force-displacement quadrant using an MRE-based absorber. The value of the force is also constrained by the maximum value and the minimum value.

(4) *Hysteresis State*. Hysteresis is a major problem in the MR system. This is a nonmeasurable component that greatly affects the stability of the system.

3.3. Adaptive Control Design

(a) The sliding control is defined as

$$S = \dot{x}_s + \lambda x_s, \tag{11}$$

where $\lambda > 0$ is the gain constant.

The time derivative of the sliding function S is as follows:

$$\dot{S} = \ddot{x}_s + \lambda\dot{x}_s. \tag{12}$$

Dynamic system equation (9) is written in terms of S:

$$\dot{S} = -b_1 u - b_1\Phi - b_2 x_r - b_3\dot{x}_r + \lambda x_s, \tag{13}$$

where $b_1 = 1/m_s$ $b_2 = 1/m_s(k_s + \alpha_a k_{0a})$; $b_3 = 1/m_s(c_s + c_{0a})$, $x_r = x_s - x_u$; and $\dot{x}_r = \dot{x}_s - \dot{x}_u$. Considering the fact that the system parameters b_1, b_2, b_3, and m_s are unknown in advance, the hysteresis state Φ is an unmeasurable component. To solve this problem, the parameters are estimated using the controller. The adaptive control force is proposed as

$$u = \frac{1}{\hat{b}_1}\left(kS - \hat{b}_2 x_r - \hat{b}_3\dot{x}_r + \lambda x_s\right) - \hat{\Phi}, \tag{14}$$

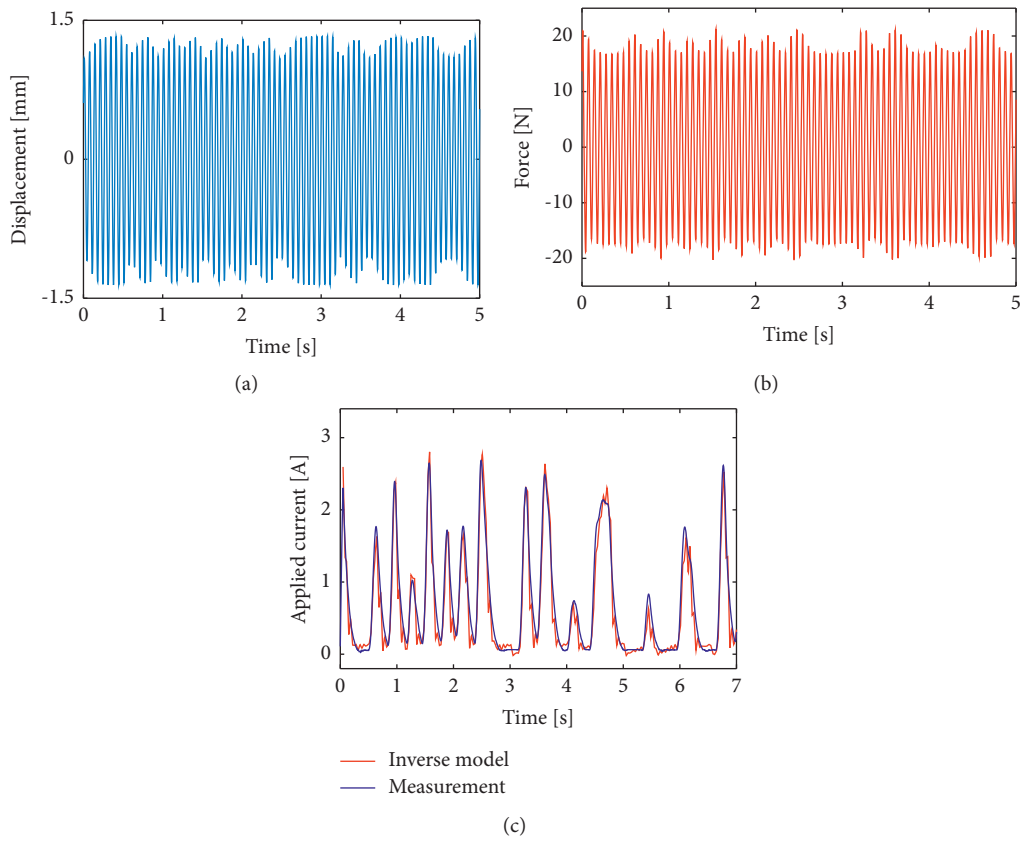


Figure 6: Applied current response using MRE inverse model: (a) displacement data, (b) force data, and (c) applied current response.

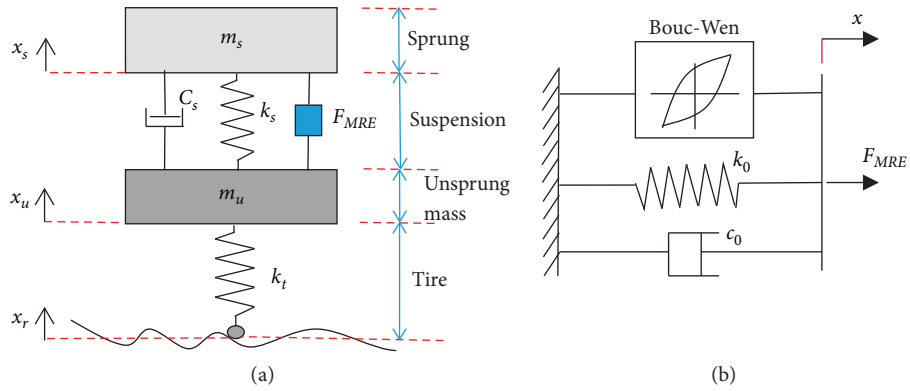


Figure 7: The novel suspension system using MRE-based absorber: (a) the suspension system and (b) the isolator model.

where $\widehat{b}_1$, $\widehat{b}_2$, and $\widehat{b}_3$ are the estimated values of the unknown model parameters b_1, b_2, and b_3, respectively, and k is a positive constant.

The controller is a significant dynamic variation in the plant. When the estimates parameter $\widehat{b}_1$ reach around zero or $\widehat{b}_2, \widehat{b}_3$ achieve large values, the control force becomes a large value; if $\widehat{b}_1 \longrightarrow 0$ or $\widehat{b}_2 \longrightarrow \infty$ or $\widehat{b}_3 \longrightarrow \infty$, then $u \longrightarrow \infty$. This problem will greatly affect the stability of the system [31].

(b) Adaptive neural network control.

We applied the radial basis function (RBF) neural network $W^T Z(S)$ that can estimate the function $f(S) = b_1^{-1}(-b_2 x_r - b_3 \dot{x}_r + \lambda x_s)$ with arbitrary accuracy, such that

$$b_1^{-1}\left(-b_2 x_r - b_3 \dot{x}_r + \lambda x_s\right) = W_c^T Z(S) + \varepsilon_c, \quad \forall x \in D, \tag{15}$$

where $W_c = [w_1, w_2, \ldots, w_N]^T$ is an optimal constant weight vector, $N > 1$ is the number of the neurons, $Z(S) = [z_1(S), z_2(S), \ldots, x_N(S)]^T$ is the RBF vector, and ε_c is error that is optimized by the vector W.

The weight vector W is updated to minimize ε_c on the compact set D:

$$W := \underset{W \in R}{\operatorname{argmin}} \left\{ \sup_{x \in D} \left| f(S) - W_c^T Z(S) \right| \right\}. \tag{16}$$

Assume that ε_c is bounded by $|\varepsilon_c| < \varepsilon^*$ with ε^* being an unknown positive constant. The Gaussian function, $z_i(x)$, is given by

$$z_i(S) = \exp\left[\frac{-(S - \gamma_i)^T (S - \gamma_i)}{\omega_i^2} \right], \tag{17}$$

where γ_i and ω_i represent the center and width of the function, respectively.

The adaptive neural network controller was designed, u_{ann}, for the nonlinear uncertain part of the suspension system:

$$u_{ann} = \widehat{W}_c^T Z(S) + \widehat{\varepsilon}_c + kS. \tag{18}$$

The controller has a capacity in predicting model nonsingularities on a compact set D and achieving a good performance in nonlinear identification. However, the controller takes up a lot of computation and takes a long time to process. The system parameters, such as stiffness, mass, and damping coefficient, cannot be identified by using this method.

3.4. Smooth Switching Adaptive Controller. A control strategy that encompasses all advantages of the controller mentioned above and eliminates the drawbacks is proposed in this study. The block diagram of the controller is shown in Figure 8. First, the smooth switching algorithm is introduced in this study to observe the singularity and to determine the authority of the two above controllers:

$$\begin{aligned} \eta(b) &= 1 - \exp\left(-\left(\frac{b}{\delta} \right)^2 \right), \; b \\ &= \min\left\{ \left|\widehat{b}_1\right|, \frac{1}{\left|\widehat{b}_2\right|}, \frac{1}{\left|\widehat{b}_3\right|} \right\}, \forall \widehat{b}_1, \widehat{b}_2, \widehat{b}_3 \in R, \end{aligned} \tag{19}$$

where b is the variable that causes the singularity and δ is the width of the corresponding transition. The switching algorithm has the following characteristics:

$$\text{C1.} 0 \le \eta(b) \le 1 \; \forall \widehat{b}_1 \in R, \tag{20a}$$

$$\text{C2.} \lim_{b \longrightarrow 0} \eta(b) = 0, \tag{20b}$$

$$\text{C3.} \lim_{b \longrightarrow 0} \eta(b) = 1, \tag{20c}$$

$$\begin{aligned} \text{C4.} \lim_{\widehat{b}_1 \longrightarrow 0} \frac{\eta(b)}{\widehat{b}_1} &= 0, \lim_{\widehat{b}_2 \longrightarrow \infty} \left[\eta(b) \times \widehat{b}_2 \right] \\ &= 0, \lim_{\widehat{b}_3 \longrightarrow \infty} \left[\eta(b) \times \widehat{b}_3 \right] = 0. \end{aligned} \tag{20d}$$

Furthermore, to support the controller, an observer was developed to estimate the hysteresis state Φ that can be described by

$$\dot{\widehat{\Phi}} = A_\Phi \dot{x} - \beta |\dot{x}| \widehat{\Phi} - \gamma \dot{x} |\widehat{\Phi}| + \vartheta, \tag{21}$$

where ϑ is the observer dynamic component suggested later. Suppose the observation error is defined as $\widetilde{\Phi} = \widehat{\Phi} - \Phi$, and the observation misalignment is determined as follows:

$$\dot{\widetilde{\Phi}} = \dot{\widehat{\Phi}} - \dot{\Phi} = -\beta |x| \widetilde{\Phi} - \gamma \dot{x} \left(|\widehat{\Phi}| - |\Phi| \right) + \vartheta. \tag{22}$$

A switching adaptive control algorithm is proposed as follows:

$$u_c = \eta u_a + (1 - \eta) u_{ann} - \widehat{\Phi}, \tag{23}$$

where

$$u_a = \left(\widehat{b}_1 \right)^{-1} \left(kS - \widehat{b}_2 x_r - \widehat{b}_3 \dot{x}_r + \lambda x_s \right), \tag{24}$$

$$u_{ann} = \widehat{W}_c^T Z(S) + \widehat{\varepsilon}_c + kS, \tag{25}$$

where $\widehat{b}_1$, $\widehat{b}_2$, and $\widehat{b}_3$ are estimated values of the unknown model parameters b_1, b_2, and b_3, respectively. The error responses were defined as $\widetilde{b}_1 = \widehat{b}_1 - b_1$, $\widetilde{b}_2 = \widehat{b}_2 - b_2$, and $\widetilde{b}_3 = \widehat{b}_3 - b_3$. The hysteresis observer $\widehat{\Phi}$ is developed for the proposed controller.

The force of MRE-based absorber is limited by the maximum and minimum values [32]. The input control force of the system is satisfied with the following requirements:

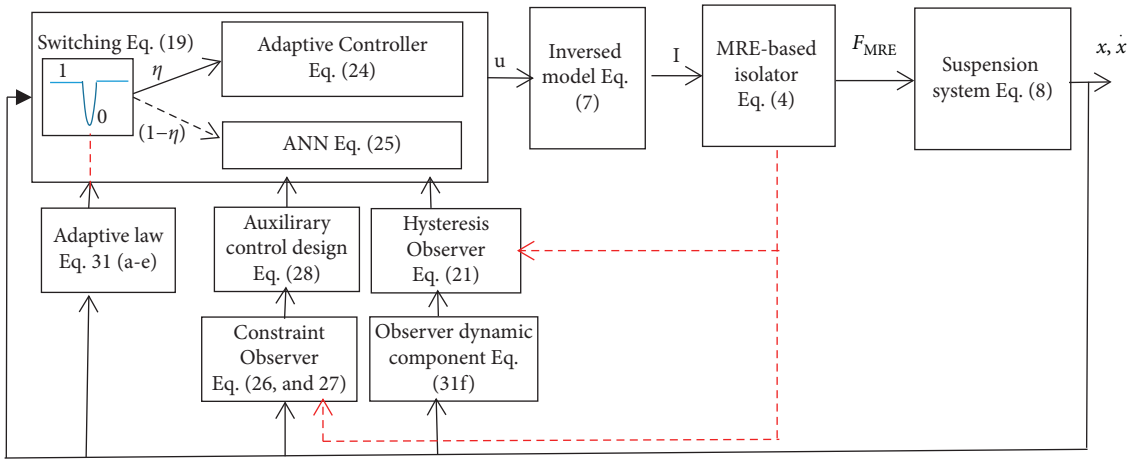


Figure 8: Block diagram of the proposed controller.

$$u = \begin{cases} u_{\max} & if \quad u_c > u_{\max} \\ u_{\min} & if \quad u_c < u_{\min} \\ u_c & if \quad u_{\max} < u_c < u_{\max} \end{cases}, \tag{26}$$

and

$$\delta = \begin{cases} u_c - u_{\max} & if \quad u_c > u_{\max} \\ u_c - u_{\min} & if \quad u_c < u_{\min} \\ 0 & if \quad u_{\max} < u_c < u_{\max} \end{cases}, \tag{27}$$

where $u_{\max}$ and $u_{\min}$ are determined by using equation (5a), $u_{\max} = F_a(I_{\max})$ if $F_a > 0$, $u_{\min} = F_a(I_{\max})$ if $F_a < 0$, the direction of F_a depends on x_r, δ is the amount of value that exceeds the limits of the controller, ξ is a regulator to ensure the system is stable, u is the actual force, and u_c is the desired control force. The absorber force fails to meet the control force due to actuator limitations in many cases. In this study, the following auxiliary design system is proposed to regulate the phenomenon:

$$\dot{\xi} = \begin{cases} -k_\lambda \xi - \dfrac{S\delta + |\xi|}{\xi} + \tanh\left(\dfrac{\xi}{\epsilon}\right) & |\xi| \geq \mu \\ 0 & |\xi| \leq \mu \end{cases}, \tag{28}$$

where $\xi \in R$ is an auxiliary design system state, $K_\lambda \in R^+$, and μ is a small positive value.

The auxiliary controller u_{au} that satisfies the constraint of MRE isolator is added as

$$u_{au} = -\delta - \xi. \tag{29}$$

Finally, the controller proposed in this study consists of four components including the adaptive controller u_a, the adaptive neural network controller u_{ann}, the auxiliary controller u_{au}, and the smooth switching $\eta(b)$:

$$u = \eta u_a + (1 - \eta)u_{ann} + u_{au} - \hat{\Phi}. \tag{30}$$

The updated laws for the parameters are proposed as follows:

$$\dot{\hat{b}}_1 = c_1 \frac{\eta}{\hat{b}_1} S\left(kS - \hat{b}_2 x_r - \hat{b}_3 \dot{x}_r + \lambda x_s\right), \tag{31a}$$

$$\dot{\hat{b}}_2 = c_2 \eta S x_r, \tag{31b}$$

$$\dot{\hat{b}}_3 = c_3 \eta S \dot{x}_r, \tag{31c}$$

$$\dot{\hat{W}}_c = c_4 (1 - \eta) S Z(S), \tag{31d}$$

$$\dot{\hat{\varepsilon}}_c = c_5 (1 - \eta) S. \tag{31e}$$

The dynamic component of the hysteresis state can be regulated as

$$\vartheta = -c_6 S. \tag{31f}$$

Remark 1. Regarding property characteristic C4, we see that the singularity is eliminated, which means $\hat{b}_1 \longrightarrow 0$, $\hat{b}_2 \longrightarrow \infty$, $\hat{b}_3 \longrightarrow \infty$, $\lim_{\hat{b}_1 \longrightarrow 0} \eta(b)/\hat{b}_1 = 0$, $\lim_{\hat{b}_2 \longrightarrow \infty} [\eta(b) \times \hat{b}_2] = 0$, and $\lim_{\hat{b}_3 \longrightarrow \infty} [\eta(b) \times \hat{b}_3] = 0$. As a result, adaptive control signal equation (24) and adaptive signal equation (31a) are bounded. Hence, the singularity is totally avoided. Furthermore, the smooth switching algorithm ensures continuous signals. The chattering is also reduced by switching.

4. Stability Analysis

Theorem 1. *Consider the vibration system ((8a) and (8b)) with the sliding function given by (9a) and (9b) under the novel adaptive controller (30) and the updated laws (31a)–(31e) such that all signals are bounded and the system is stable.*

Proof. Lyapunov function candidate is selected as

$$V = \frac{1}{2}S^2 + \frac{1}{2c_1}\tilde{b}_1^2 + \frac{1}{2c_2}\tilde{b}_2^2 + \frac{1}{2c_3}\tilde{b}_3^2 + \frac{1}{2c_4}b_1\tilde{W}_c^2 + \frac{1}{2c_5}b_1\tilde{\varepsilon}_c^2 + \frac{1}{2c_6}b_1\tilde{\Phi}^2 + \frac{1}{2}b_1\xi^2. \tag{32}$$

With the time derivative of Lyapunov function and application of equation (30), we have

$$\begin{aligned}
\dot{V} &= S\dot{S} + \frac{1}{c_1}\tilde{b}_1\dot{\tilde{b}}_1 + \frac{1}{c_2}\tilde{b}_2\dot{\tilde{b}}_2 + \frac{1}{c_3}\tilde{b}_3\dot{\tilde{b}}_3 + \frac{1}{c_4}b_1\tilde{W}_c\dot{\tilde{W}}_c + \frac{1}{c_5}b_1\tilde{\varepsilon}_c\dot{\tilde{\varepsilon}}_c + \frac{1}{c_6}b_1\tilde{\Phi}\dot{\tilde{\Phi}} + b_1\xi\dot{\xi} \\
&= S\left[-b_1u - b_1\Phi - b_2x_r - b_3\dot{x}_r + \lambda x_s\right] + \frac{1}{c_1}\tilde{b}_1\dot{\tilde{b}}_1 + \frac{1}{c_2}\tilde{b}_2\dot{\tilde{b}}_2 + \frac{1}{c_3}\tilde{b}_3\dot{\tilde{b}}_3 + \frac{1}{c_4}b_1\tilde{W}_c\dot{\tilde{W}}_c + \frac{1}{c_5}b_1\tilde{\varepsilon}_c\dot{\tilde{\varepsilon}}_c + \frac{1}{c_6}b_1\tilde{\Phi}\dot{\tilde{\Phi}} + b_1\xi\dot{\xi} \\
&= S\left[-b_1\left(\eta u_a + (1-\eta)u_{ann} + u_{au} - \hat{\Phi}\right) - b_1\Phi - b_2x_r - b_3\dot{x}_r + \lambda x_s\right] + \frac{1}{c_1}\tilde{b}_1\dot{\tilde{b}}_1 + \frac{1}{c_2}\tilde{b}_2\dot{\tilde{b}}_2 + \frac{1}{c_3}\tilde{b}_3\dot{\tilde{b}}_3 + \frac{1}{c_4}b_1\tilde{W}_c\dot{\tilde{W}}_c + \frac{1}{c_5}b_1\tilde{\varepsilon}_c\dot{\hat{\varepsilon}}_c \\
&\quad + \frac{1}{c_6}b_1\tilde{\Phi}\dot{\tilde{\Phi}} + b_1\xi\dot{\xi},
\end{aligned} \tag{33}$$

$$\dot{V} = \dot{V}_1 + \dot{V}_2 + \dot{V}_3 + \dot{V}_4, \tag{34}$$

where each term on right-hand side of the function is written explicitly as follows.

We apply the inequality $(|\Phi| - |\hat{\Phi}|) \le |\hat{\Phi} - \Phi| = |\tilde{\Phi}|$ and the observer dynamics component equation (31f), $\vartheta = -c_6S$, to V_1:

$$\begin{aligned}
\dot{V}_1 &= Sb_1(\hat{\Phi} - \Phi) + \frac{1}{c_6}b_1\tilde{\Phi}\dot{\tilde{\Phi}} \\
&= b_1\left(S\tilde{\Phi} + \frac{1}{c_6}\tilde{\Phi}\left(-\beta|x|\tilde{\Phi} - \gamma x\left(|\hat{\Phi}| - |\Phi|\right)\right) + \vartheta\right) \\
&\le b_1\left(S\tilde{\Phi} - \frac{1}{c_6}\beta|x|\tilde{\Phi}^2 + \frac{1}{c_6}\gamma|x|\tilde{\Phi}^2 + \frac{1}{c_6}\tilde{\Phi}\vartheta\right) \\
&= -\frac{b_1}{c_6}(\beta - \gamma)|x|\tilde{\Phi}^2,
\end{aligned} \tag{35}$$

where $b_1 > 0$, $\gamma > 0$, $\beta - \gamma > 0$ in Table 1.

$$\begin{aligned}
\dot{V}_2 &= \eta S\left(-b_1 u_a - b_2 x_r - b_3 \dot{x}_r + \lambda x_s\right) + \frac{1}{c_1}\tilde{b}_1\dot{\hat{b}}_1 + \frac{1}{c_2}\tilde{b}_2\dot{\hat{b}}_2 + \frac{1}{c_3}\tilde{b}_3\dot{\hat{b}}_3 \\
&= \eta S\left(-\left(\hat{b}_1 - \tilde{b}_1\right)\frac{1}{\hat{b}_1}\left(kS - \hat{b}_2 x_r - \hat{b}_3\dot{x}_r + \lambda x_s\right) - b_2 x_r - b_3\dot{x}_r + \lambda x_s\right) + \frac{1}{c_1}\tilde{b}_1\dot{\hat{b}}_1 + \frac{1}{c_2}\tilde{b}_2\dot{\hat{b}}_2 + \frac{1}{c_3}\tilde{b}_3\dot{\hat{b}}_3 \\
&= -\eta k S^2 - \eta S x_r\left(\hat{b}_2 - b_2\right) - \eta S\dot{x}_r\left(\hat{b}_3 - b_3\right) - \eta S\frac{\tilde{b}_1}{\hat{b}_1}\left(k_s S - \hat{b}_2 x_r - \hat{b}_3\dot{x}_r + \lambda x_s\right) + \frac{1}{c_1}\tilde{b}_1\dot{\hat{b}}_1 + \frac{1}{c_2}\tilde{b}_2\dot{\hat{b}}_2 + \frac{1}{c_3}\tilde{b}_3\dot{\hat{b}}_3 \\
&= -\eta k S^2 - \tilde{b}_2\left(\eta S x_r - \frac{1}{c_2}\dot{\hat{b}}_2\right) - \tilde{b}_3\left(\eta S\dot{x}_r - \frac{1}{c_3}\dot{\hat{b}}_3\right) - \tilde{b}_1\left(\eta\left(\hat{b}_1\right)^{-1} S\left(k_s S - \hat{b}_2 x_r - \hat{b}_3\dot{x}_r + \lambda x_s\right) + \frac{1}{c_1}\dot{\hat{b}}_1\right) \\
&= -\eta k S^2.
\end{aligned} \tag{36}$$

Adaptive algorithm equations (31d)–31e are applied to $\dot{V}_3$:

$$\begin{aligned}
\dot{V}_3 &= (1-\eta)S\left(-b_1 u_{anm} - b_2 x_r - b_3\dot{x}_r + \lambda x_s\right) + \frac{1}{c_4}\tilde{W}_c\dot{\hat{W}}_c + \frac{1}{c_5}\tilde{\varepsilon}_c\dot{\hat{\varepsilon}}_c \\
&= (1-\eta)S\left(-b_1\left(\hat{W}_c^T Z(z) + \hat{\varepsilon}_c + kS\right) - b_1\left(W_c^T Z(z) + \varepsilon_c\right)\right) + \frac{1}{c_4}b_1\tilde{W}_c\dot{\hat{W}}_c + \frac{1}{c_5}b_1\tilde{\varepsilon}_c\dot{\hat{\varepsilon}}_c \\
&= -(1-\eta)b_1 k S^2 + b_1\left[(1-\eta)S\left(-\hat{W}_c^T Z(z) + \left(W_c^T Z(z)\right)\right] + b_1\left[(1-\eta)S\left(-\hat{\varepsilon}_c + \varepsilon_c\right)\right] + \frac{1}{c_4}b_1\tilde{W}_c\dot{\hat{W}}_c + \frac{1}{c_5}b_1\tilde{\varepsilon}_c\dot{\hat{\varepsilon}}_c \\
&= -(1-\eta)b_1 k S^2 + b_1\tilde{W}_c\left[-(1-\eta)S\,Z(z) + \frac{1}{c_4}\dot{\hat{W}}_c\right] + b_1\tilde{\varepsilon}_c\left[-(1-\eta)S + \frac{1}{c_5}\dot{\hat{\varepsilon}}_c\right] \\
&= -(1-\eta)b_1 k S^2.
\end{aligned} \tag{37}$$

The auxiliary design system equation (29) and Lemma 1 are applied to $\dot{V}_4$:

$$\begin{aligned}
\dot{V}_4 &= -b_1 S u_{au} + \xi\dot{\xi} \\
&= -b_1 S(-\delta - \xi) + b_1\xi\left[-k_\lambda\xi - \frac{S\delta + |\xi|}{\xi} + \tanh\left(\frac{\xi}{\epsilon}\right)\right] \\
&= -b_1 k_\lambda\xi^2 - b_1\left[|\xi| - \xi\tanh\left(\frac{\xi}{\epsilon}\right)\right] \\
&= -b_1 k_\lambda\xi^2 - \kappa,
\end{aligned} \tag{38}$$

where ϵ is a positive constant and $\kappa = 0.2785$, $|\xi| - \xi\tanh(\xi/\epsilon) \le \kappa\epsilon$.

Applying equations (35)–(38), the derivative Lyapunov equation (34) is represented as

Next, adaptive algorithm equations (31a)–31c are applied to $\dot{V}_2$:

$$\dot{V} < -\frac{b_1}{c_6}(\beta - \gamma)|x|\tilde{\Phi}^2 - \eta k S^2 - (1-\eta)b_1 k S^2 - b_1 k_\lambda\xi^2 - \kappa\epsilon < 0. \tag{39}$$

The boundedness of $\tilde{b}_1$, $\tilde{b}_2$, $\tilde{b}_3$, $\tilde{\Phi}$, and S is asymptotic to zero by the Lyapunov stability criterion. Therefore, the closed-loop system is asymptotically stable. Associated with Remark 1, all signals are bounded.

5. Simulations

In this section, the system, combined with the proposed controller, is simulated to reduce the vibration effectiveness. The dynamic system's parameter values are assigned as $m_s = 2.45\,\text{kg}$, $m_u = 1\,\text{kg}$, $k_s = 900\,\text{m}^{-1}$, $c_s = 8\,\text{N s m}^{-1}$, $k_t = 2500\,N\,m^{-1}$, and the parameters of MRE model are shown in Table 1; the initial state $[b_1, b_2, b_3] = [0.01, 0.01, 0.01]$, and $[x_s, \dot{x}_s, x_u, \dot{x}_u] = [0, 0, 0, 0]$. The coefficients of the

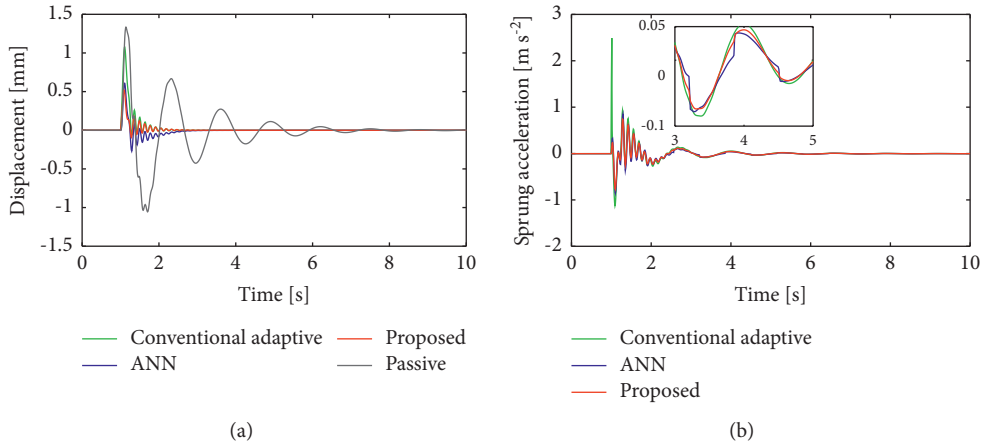


Figure 9: The comparison of sprung mass responses under bump wheel excitation for three different strategies: (a) displacement response and (b) acceleration response.

controller are selected as $c_1 = 2$; $c_2 = 2$; $c_3 = 2$; $c_4 = 4$; $c_5 = 4$; $c_6 = 1$; $\lambda = 2$; and $\delta = 0.1$. Four strategies have been investigated to evaluate controllers: conventional adaptive controller, adaptive neural network controller, the switching adaptive controller (proposed controller), and passive controller. We used the Runge–Kutta 4^{th} order method to solve the differential equation.

5.1. Bump Wheel Excitation. The relative displacement and mass acceleration responses are depicted in Figure 9 for four cases including conventional adaptive controller, ANN controller, proposed controller, and passive system. The relative displacement is reduced significantly using the controllers. The efficiency is the same for the acceleration response, as shown in Figure 9(b). Furthermore, the response is smoother when it reaches a steady state by using the proposed controller.

The switching signal and control forces are shown in Figures 10 and 11, respectively. The switching signal indicates that the adaptive neural network controller takes over in about 1.5 seconds, and then the adaptive controller gradually dominates the controller. Figure 11 presents the control force of the three strategies. From the figure, the proposed controller required a small control effort than the single controller, while the control force jumped to a large value when $\widehat{b}_1$ was near zero using the conventional adaptive controller.

The adaptive parameters and hysteresis state are shown in Figures 12 and 13. From the figures, the parameters achieve a stable state after 2 seconds. The input current calculated by using the proposed algorithm is shown in Figure 14. If the actual force achieves a maximum value, the applied current is set at four amperages. In other cases, the inverse model is used to calculate the required current. These results demonstrate that the proposed controller achieves high efficiency compared to conventional controllers to reduce system vibration.

5.2. Random Road Displacement. To further validate the proposed strategy under random excitation, we choose the road disturbance with an amplitude of 5 mm. From Figure 15, we found that the controllers work well, and the mass displacement responses approach zero quickly. The proposed controller has achieved a positive result while avoiding the drawbacks of the traditional adaptive controllers. The control force, smooth switching, hysteresis state, and applied current are also shown in Figures 16–19. The force generated by the MRE-based isolator is compared for different strategies in Figure 16, where the proposed controller needs a smaller value compared to the other controllers. At the initial time, the hysteresis value is unknown, and the measurement is not achievable. Using an adaptive observer, the value of the hysteresis is estimated and updated based on the hysteresis dynamics equation (31f). Results are depicted in Figure 19 after 0.2 seconds to verify the hysteresis state estimation. The observer has portrayed well the hysteresis properties by using the hysteresis state dynamics.

The efficiency of the proposed controller is based on its adaptability. In the early stages of the control process, the adaptive parameters are in a highly dynamic state, so a singularity phenomenon may occur in this state. The system is dominated temporarily by the ANN controller so that the system works stably and safely. In this stage, the traditional adaptive controller still works as the virtual controller, and the parameters are continuously adapted. When the adaptive values are out of the singularity, the adaptive controller smoothly takes over the system, and the adaptive values update to their true values quickly. A smooth switching algorithm is used to observe the adaptive parameters and decide which controller takes over the system. The smooth switching algorithm has outstanding advantages. When the signals pass through the singularity, the algorithm can suppress this phenomenon, and the controller is temporarily switched to the ANN controller. In this way, the advantages of a single controller are exploited, and their disadvantages are eliminated.

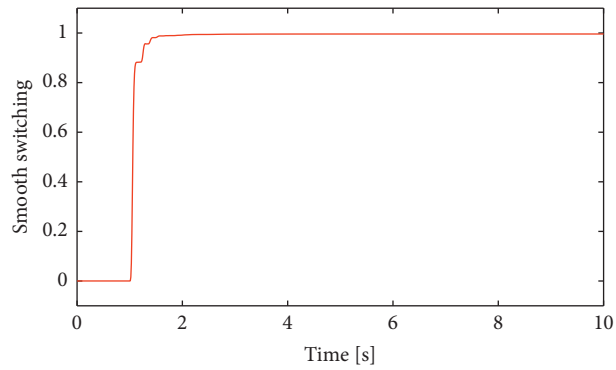


Figure 10: The smooth switching to take over between the controllers under bump wheel excitation.

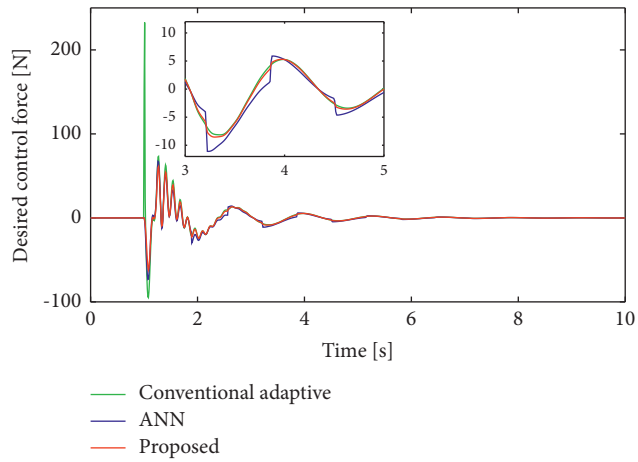


Figure 11: The comparison of control force for the system with the proposed algorithm.

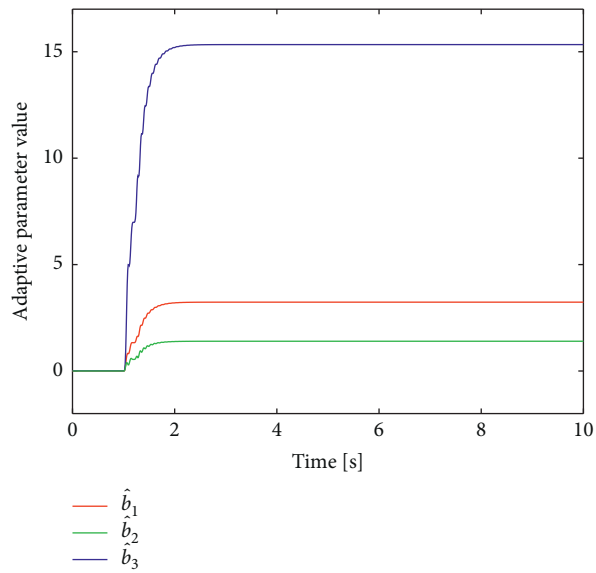


Figure 12: Time history of the adaptive parameters $\hat{b}_1$, $\hat{b}_2$, and $\hat{b}_3$.

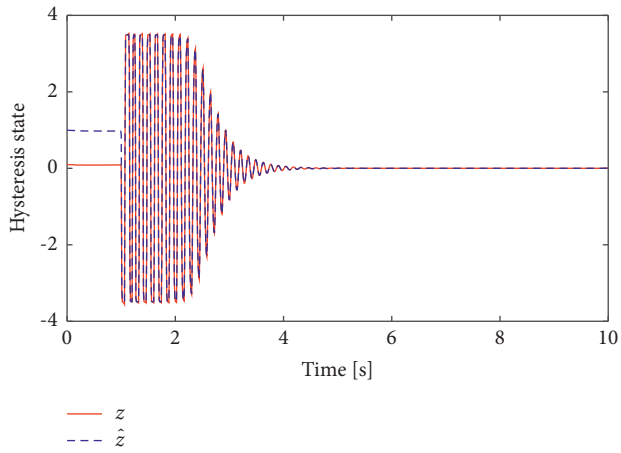


Figure 13: Time history of the hysteresis state.

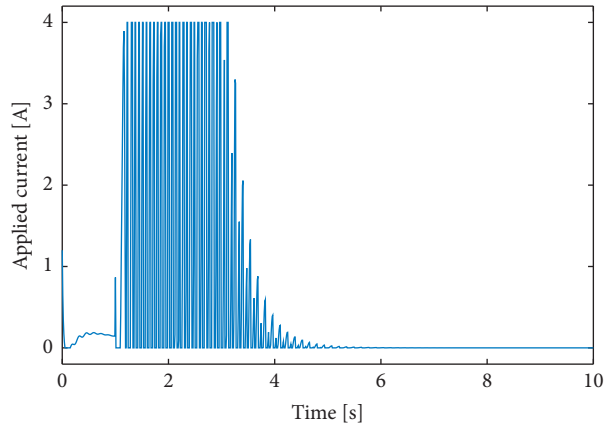


Figure 14: Current input for the system with the proposed algorithm.

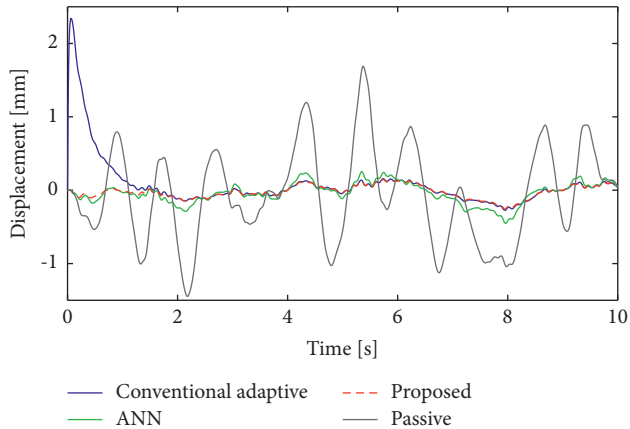


Figure 15: The comparison of sprung mass displacement responses under random excitation for three different strategies.

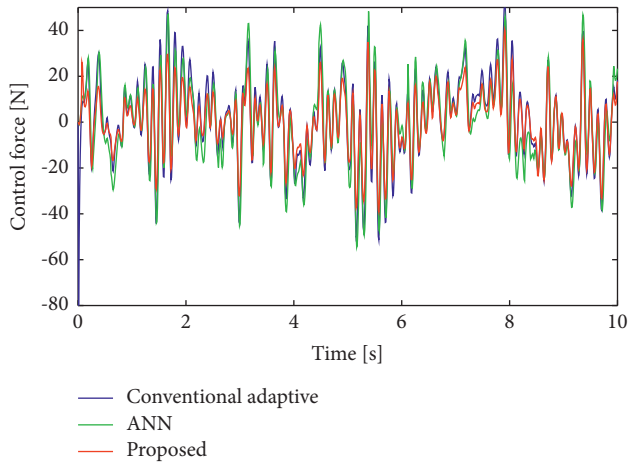


Figure 16: The comparison of control forces for the system for three different strategies.

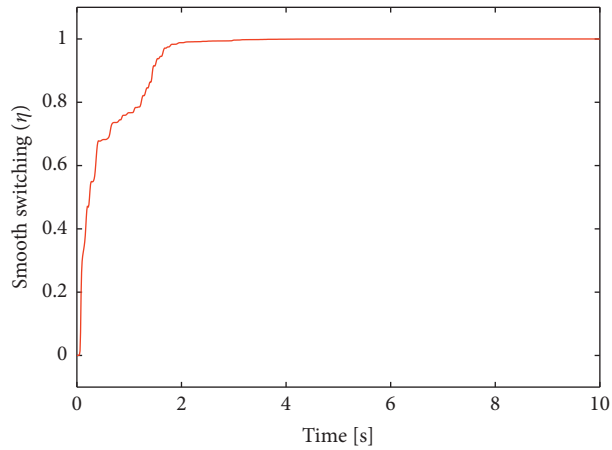


Figure 17: The smooth switching to take over between the controllers under random excitation.

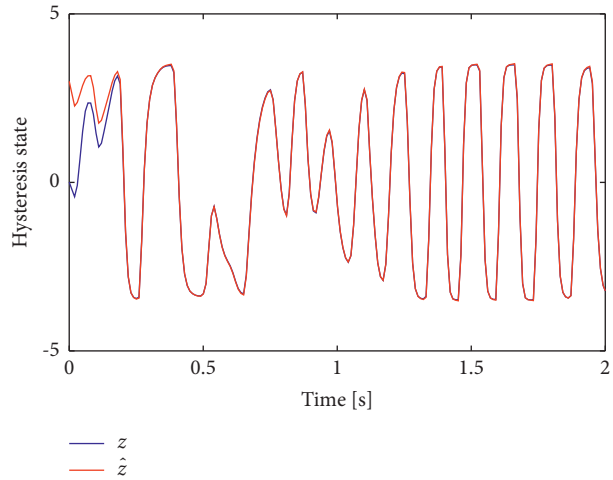


Figure 18: Time history of the hysteresis state.

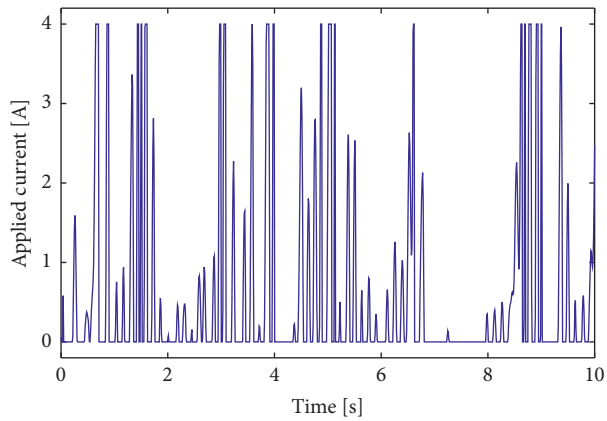


Figure 19: Current input for the system with the proposed algorithm.

6. Conclusions

In this study, the Bouc–Wen model was used to represent the properties of the MRE material. The inverse model was also developed to determine the applied current for the MRE-based isolator. A nonlinear observer was constructed to predict the unknown hysteresis state, and the input constraint was also considered to ensure the system's stability. The control singularity was avoided with such an approach, and the transient behavior improved in the adaptive controller. Based on the unique characteristics of smooth switching equations (20a)–20d, the controller requires less force without loss performance than the single controller. The proposed controller exploited the advantages of adaptive controllers and neural network controllers and eliminated the disadvantages of these controllers with a smooth switching mechanism. Consequently, the denominator part of the adaptive control formula was absorbed near zero to eliminate the singularity problem. The proposed controller overcomes the traditional adaptive controller's disadvantages, including nonsingularity, low control force, and high stability. The simulation results have proved the effectiveness of the proposed control algorithm. The proposed controller significantly improves the vibration system compared to the adaptive controller.

Data Availability

The data used to support the findings of this study are available upon request from the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Y. Li, J. Li, W. Li, and H. Du, "A state-of-the-art review on magnetorheological elastomer device," *Smart Materials and Structures*, vol. 23, no. 12, pp. 1–24, 2014.

[2] J. M. Ginder, M. E. Nichols, L. D. Elie, and J. L. Tardiff, "Magnetorheological elastomers: properties and applications," in *Proceedings of the SPIE 3675, Smart Structures and Materials 1999: Smart Materials Technologies*, pp. 131–138, Newport Beach, USA, July 1999.

[3] X. B. Nguyen, T. Komatsuzaki, and H. T. Truong, "Adaptive parameter identification of Bouc-wen hysteresis model for a vibration system using magnetorheological elastomer," *International Journal of Mechanical Sciences*, vol. 213, Article ID 106848, 2022.

[4] Y. Q. Guo, J. Zhang, D. Q. He, and J. B. Li, "Magnetorheological elastomer precision platform control using OFFO-pid algorithm," *Advances in Materials Science and Engineering*, vol. 2020, Article ID 3025863, 9 pages, 2020.

[5] G. J. Liao, X.-L. Gong, S. H. Xuan, C. J. Kang, and L. H. Zong, "Development of a real-time tunable stiffness and damping vibration isolator based on magnetorheological elastomer," *Journal of Intelligent Material Systems and Structures*, vol. 23, no. 1, pp. 25–33, 2011.

[6] X. B. Nguyen, T. Komatsuzaki, and N. Zhang, "A nonlinear magnetorheological elastomer model based on fractional viscoelasticity, magnetic dipole interactions, and adaptive smooth Coulomb friction," *Mechanical Systems and Signal Processing*, vol. 141, 2020.

[7] Y. Yu, A. N. Hoshyar, H. Li, G. Zhang, and W. Wang, "Nonlinear characterization of magnetorheological elastomer-based smart device for structural seismic mitigation," *International Journal of Social Network Mining*, vol. 139 pages, 2021.

[8] Y. Yu, Y. Li, J. Li, X. Gu, and S. Royel, "Nonlinear characterization of the MRE isolator using binary-coded discrete CSO and ELM," *International Journal of Structural Stability and Dynamics*, vol. 18, no. 08, p. 24, Article ID 1840007, 2018.

[9] Y. Yu, S. Royel, J. Li, Y. Li, and Q. Ha, "Magnetorheological elastomer base isolator for earthquake response mitigation on building structures: modeling and second-order sliding mode control," *Earthquakes and Structures*, vol. 11, no. 6, 2016.

[10] H. Madani, M. Kooshafar, and M. Emadi, "Compressive strength prediction of nanosilica-incorporated cement mixtures using adaptive neuro-fuzzy inference system and artificial neural network models," *Practice Periodical on Structural Design and Construction*, vol. 25, no. 3, August 2020.

[11] X. B. Nguyen, T. Komatsuzaki, Y. Iwata, and H. Asanuma, "Fuzzy semiactive vibration control of structures using magnetorheological elastomer," *Shock and Vibration*, Hindawi, vol. 2017, 15 pages, 2017.

[12] L. M. Jansen and S. J. Dyke, "Semiactive control strategies for MR dampers: comparative study," *Journal of Engineering Mechanics*, vol. 126, no. 8, pp. 795–803, 2000.

[13] Y. Wang and S. Dyke, "Modal-based LQG for smart base isolation system design in seismic response control," *Structural Control and Health Monitoring*, vol. 20, no. 5, pp. 753–768, 2013.

[14] M. D. Symans and S. W. Kelly, "Fuzzy logic control of bridge structures using intelligent semi-active seismic isolation systems," *Earthquake Engineering & Structural Dynamics*, vol. 28, no. 1, pp. 37–60, 1999.

[15] X. B. Nguyen, T. Komatsuzaki, Y. Iwata, and H. Asanuma, "Modeling and semi-active fuzzy control of magneto-rheological elastomer-based isolator for seismic response reduction," *Mechanical Systems and Signal Processing*, vol. 101, pp. 449–466, 2018.

[16] J. Fei and M. Xin, "Robust adaptive sliding mode controller for semi-active vehicle suspension system," *International Journal of Innovative Computing, Information and Control*, vol. 8, no. 1B, pp. 691–700, 2012.

[17] Z. Chen and H. Lu, "Optimal semiactive damping control for a nonlinear energy sink used to stabilize milling," *Shock and Vibration*, vol. 2020, pp. 1–11, Article ID 8837753, 2020.

[18] X. B. Nguyen, T. Komatsuzaki, Y. Iwata, and H. Asanuma, "Robust adaptive controller for semi-active control of uncertain structures using a magnetorheological elastomer-based isolator," *Journal of Sound and Vibration*, ScienceDirect, vol. 434, , pp. 192–212, 2018.

[19] S. Liu, R. Hao, D. Zhao, and Z. Tian, "Adaptive dynamic surface control for active suspension with electro-hydraulic actuator parameter uncertainty and external disturbance," *IEEE Access*, vol. 8, pp. 156645–156653, 2020.

[20] W. Sun, H. Gao, and O. Kaynak, "Vibration isolation for active suspensions with performance constraints and actuator saturation," *IEEE*, vol. 20, no. 2, pp. 675–683, April 2015.

[21] M. U. Saeed, Z. Sun, and S. Elias, "Research developments in adaptive intelligent vibration control of smart civil structures," *Journal of Low Frequency Noise, Vibration and Active Control*, vol. 138 pages, 2021.

[22] Y. Zhang, Y. Liu, and L. Liu, "Minimal learning parameters-based adaptive neural control for vehicle active suspensions with input saturation," *Neurocomputing*, vol. 396, pp. 153–161, 2019.

[23] Z. Rashid, M. Tantray, and E. N. Farsangi, "Acceleration response-based adaptive strategy for vibration control and location optimization of magnetorheological dampers in multistoried structures," *Practice Periodical on Structural Design and Construction*, vol. 27, no. 1, February 2022.

[24] W. He, Y. Dong, and C. Sun, "Adaptive neural impedance control of a robotic manipulator with input saturation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, no. 3, pp. 334–344, 2016.

[25] G. Li and Z. B. Yang, "Modelling and analysis of a magnetorheological damper with nonmagnetized passages in piston and minor losses," *Shock and Vibration*, vol. 2020, Article ID 2052140, 12 pages, 2020.

[26] X. Yuan, T. Tian, H. Ling, T. Qiu, and H. He, "A review on structural development of magnetorheological fluid damper," *Shock and Vibration*, vol. 2019, Article ID 1498962, 33 pages, 2019.

[27] J. Yang, H. Du, W. Li et al., "Experimental study and modeling of a novel magnetorheological elastomer isolator," *Smart Materials and Structures*, vol. 22, no. 11, Article ID 117001, 2013.

[28] X. Gu, J. Li, and Y. Li, "Experimental realisation of the real-time controlled smart magnetorheological elastomer seismic isolation system with shake table," *Structural Control and Health Monitoring*, vol. 27, no. 1, Article ID e2476, 2020.

[29] X. Gu, Y. Yu, J. Li, and Y. Li, "Semi-active control of magnetorheological elastomer base isolation system utilising learning-based inverse model," *Journal of Sound and Vibration*, vol. 406, pp. 346–362, 2017.

[30] M. M. Polycarpou, "Stable adaptive neural control scheme for nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 41, no. 3, pp. 447–451, 1996.

[31] J. Ming and Y. M. Chen, "A smooth switching adaptive controller for linearizable systems with improved transient performance," *International Journal of Adaptive Control and Signal Processing*, vol. 20, no. 9, pp. 431–446, 2006.

[32] X. B. Nguyen, T. Komatsuzaki, and H. T. Truong, "Novel semiactive suspension using a magnetorheological elastomer (MRE)-based absorber and adaptive neural network controller for systems with input constraints," *Mechanical Sciences*, vol. 11, no. 2, pp. 465–479, 2020.

Hindawi
Shock and Vibration
Volume 2021, Article ID 1205473, 8 pages
https://doi.org/10.1155/2021/1205473

Research Article

Intelligent Diagnosis of Rolling Bearing Fault Based on Improved Convolutional Neural Network and LightGBM

Yanwei Xu,[1,2] **Weiwei Cai**,[1] **Liuyang Wang**,[1] **and Tancheng Xie**[1,2]

[1]*School of Mechatronics Engineering, Henan University of Science and Technology, Luoyang 471003, China*
[2]*Intelligent Numerical Control Equipment Engineering Laboratory of Henan Province, Luoyang 471003, China*

Correspondence should be addressed to Yanwei Xu; xuyanweiluoyang@163.com

Received 24 July 2021; Accepted 25 November 2021; Published 9 December 2021

Academic Editor: Claudio Sbarufatti

Aiming at the problems of weak generalization ability and long training time in most fault diagnosis models based on deep learning, such as support vector machines and random forest algorithms, one intelligent diagnosis method of rolling bearing fault based on the improved convolution neural network and light gradient boosting machine is proposed. At first, the convolution layer is used to extract the features of the original signal. Second, the generalization ability of the model is improved by replacing the full connection layer with the global average pooling layer. Then, the extracted features are classified by a light gradient boosting machine. Finally, the verification experiment is carried out, and the experimental result shows that the average training and diagnosis time of the model is only 39.73 s and 0.09 s, respectively, and the average classification accuracy of the model is 99.72% and 95.62%, respectively, on the same and variable load test sets, which indicates that the diagnostic efficiency and classification accuracy of the proposed model are better than those of other comparison models.

1. Introduction

Rolling bearing is one of the most critical components widely used in a modern machine, and it is easy to appear cracks, pitting corrosion, and other local damages or defects on the inner and outer ring raceways and rolling elements of the rolling bearings under the harsh working conditions of high temperature, alternating load, and long-time fatigue. As one key component, once the rolling bearing fails, it will affect the safe operation of mechanical equipment, or even damage the equipment and cause casualties. It is of great significance for the safe operation of the mechanical equipment to avoid the occurrence of catastrophic accidents, if we can accurately, timely, and intelligently identify the faults of the rolling bearing and carry out maintenance as soon as possible.

In recent years, various fields made great achievements in the research of algorithms. In order to overcome the slow convergence speed, poor global search ability, and difficult designing rotation angle of quantum-inspired evolutionary algorithm (QEA), Xing et al. [1] proposed an improved quantum-inspired cooperative coevolutionary algorithm, named MSQCCEA, which is based on combining the strategies of cooperative coevolution, random rotation direction, and Hamming adaptive rotation angle, and the results demonstrate that the proposed MSQCCEA has faster convergence speed and higher convergence accuracy. In order to overcome the low solution efficiency, insufficient diversity in the later search stage, slow convergence speed, and a high search stagnation possibility of differential evolution (DE) algorithm, Deng et al. [2] studied the quantum computing characteristics of quantum evolutionary algorithm (QEA) and, combined with the divide and conquer idea of cooperative evolutionary algorithm (CCEA), proposed an improved differential evolutionary algorithm (HMCFQDE), and the results proved that the proposed HMCFQDE has higher convergence accuracy and stronger stability and a strong ability to optimize high-dimensional complex functions. Deep learning theory has made great progress in the field of fault diagnosis, and as one of the important models of deep learning theory, continuous neural network (CNN) has shown its own value and great potential in the field of bearing fault diagnosis. For example, Verstraete et al. [3] proposed a deep learning-enabled

featureless methodology to automatically learn the features of the data, and the proposed CNN architecture achieved better results. He [4] presented a deep learning-based approach for bearing fault diagnosis and built an optimized deep learning structure LAMSTAR neural network to diagnose the bearing faults, and the approach shows the accurate classification performance on various bearing faults under different working conditions. Hoang and Kang [5] provided a systematic review of deep learning-based bearing fault diagnosis, introduced the three popular deep learning algorithms for bearing fault diagnosis including autoencoder, restricted Boltzmann machine, and convolutional neural network, and reviewed their applications in the area of bearing fault diagnosis. Zhenghong et al. [6] proposed an adaptive deep transfer learning method for bearing fault diagnosis and verified the method with two kinds of datasets, and the results demonstrate the effectiveness and robustness of the proposed method. Sun et al. [7] proposed a novel intelligent diagnosis method for fault identification of rotating machines, which can not only reduce the amount of measured data that contained all the information of faults but also realize the automatic feature extraction in the transform domain, and the proposed method can reduce the need of human labor and expertise and provide a new strategy to more easily handle the massive data. Ding and He [8] proposed a novel energy-fluctuated multiscale feature mining approach based on wavelet packet energy (WPE) image and deep convolutional network (ConvNet) for spindle bearing fault diagnosis, which is quite suitable for spindle bearing fault diagnosis with multiclass classification regardless of the load fluctuation. Chen et al. [9] proposed a rolling bearing fault diagnosis method based on discrete wavelet transform and the convolution neural network so as to achieve the adaptive feature extraction and intelligent diagnosis of rolling bearing faults, and the experimental results showed that the proposed method has the better generalization ability and robustness. Haidong et al. [10] proposed a novel method for intelligent fault diagnosis of rolling bearing based on deep wavelet autoencoder and extreme learning machine, the method is applied to analyze the experimental bearing vibration signals, and the results showed that the method is superior to the traditional methods and standard deep learning methods. Ding and Jia [11] proposed a one-dimensional multiscale convolutional autoencoder fault diagnosis model of rolling bearings based on the standard convolutional autoencoder, and the test results show that the proposed model has a better recognition effect for rolling bearing fault data. These studies have achieved good diagnosis results. Although the convolutional neural network has achieved good results in the field of fault diagnosis, it cannot well separate the feature extraction and classification functions of the model using the softmax layer to classify the features extracted from the convolution layer, and it may lead to poor classification and generalization ability of the model.

Machine learning algorithms play an important role in the field of fault diagnosis. The single machine learning algorithm, such as support vector machine (SVM) [12, 13] and K-nearest neighbor (KNN) [14], and the ensemble learning algorithm, such as random forest algorithm [15] and extreme gradient boosting [16, 17], all have made great achievements in the field of mechanical fault diagnosis. However, it is difficult for these classification algorithms to meet the requirements in terms of efficiency and accuracy in big data and high-dimensional environments. Light gradient boosting machine (LightGBM) [18, 19] is a gradient lifting algorithm based on the decision tree, it optimizes the classification accuracy and computational efficiency based on the boosting algorithm, and it is more suitable for classification in a large sample environment, while there will be a lot of unprocessed redundant signals if the original signals are directly input in LightGBM, and it will consume too much memory space in model training and easily cause overfitting of LightGBM classifier.

In order to solve the above problems, in this study, a bearing fault diagnosis model combined with LightGBM algorithm and the improved convolutional neural network that is optimized by replacing the full connection layer to the global average pooling (GAP) layer is proposed (hereinafter referred to as GCNN). The two kinds of data sets under the same load and variable load conditions are constructed. The improvement effect of the global average pooling layer on the model generalization ability and the effectiveness of the proposed model are proved through the comparative analysis with other models.

2. Improved Convolution Neural Network and LightGBM

The convolution neural network is one kind of feed-forward neural network, which adopts unsupervised or semi-supervised learning mode. It contains convolution calculation and deep structure and can classify the input information according to its hierarchical structure [20]. Figure 1 shows the structure of the convolution neural network [21, 22], and it includes the input layer, convolution layer, pooling layer, full connection layer, and output layer. Convolution layer, pooling layer, and full connection layer constitute the hidden layer.

2.1. Convolution Layer and Pooling Layer. Convolution layer is the most basic structure of the convolution neural network, it is the feature extraction layer, its main function is to extract features from the input data, and it uses the local link, weight sharing, and multiple convolution kernels to extract features from data. The most significant features of convolution layer are local sensing and parameter sharing compared with the general deep learning network structure, which can greatly reduce the model parameters and ensure the sparsity of the network. The convolution layer formula is

$$y^{l(i,j)} = \sum_{j'=0}^{m-1} k_i^{l(j')} x^{l(j+j')} m, \tag{1}$$

where $y^{l(i,j)}$ is convoluted output, $k_i^{l(j')}$ is the j' weight value of the convolution kernel i in the l layer, $x^{l(j+j')}$ is the

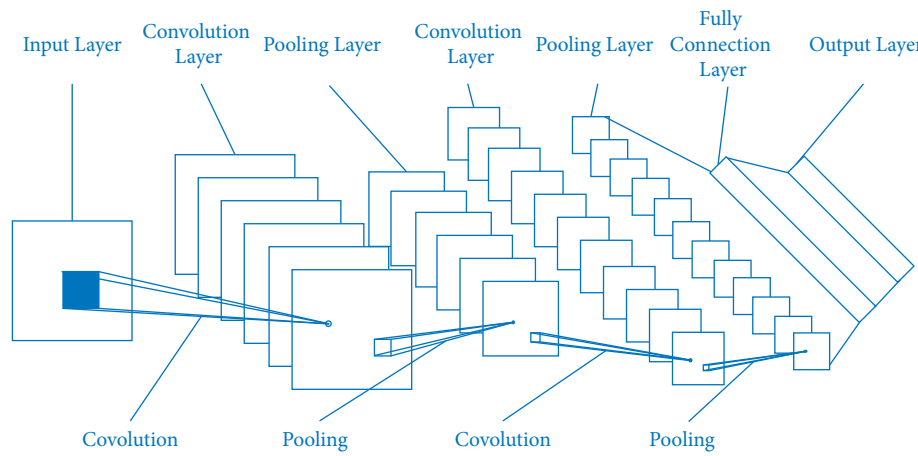


Figure 1: The model of the convolution neural network.

convoluted local region j in the l layer, and m is the width of convolution kernel.

The pooling layer is mainly used to select and filter the feature graph extracted from the convolution layer and replace the results of a single point in the feature graph with the statistics of its adjacent regions so as to reduce the number of nodes in the final fully connected layer. It can reduce the overfitting and improve the fault tolerance of the model. The common pooling methods are maximum pooling and average pooling. Compared with the average pooling method, the maximum pooling can select the most significant features in the region; therefore, in this study, the maximum pooling method is selected to select the maximum value in the region as the pooled value of the region. The expression of the maximum pooling method is

$$p^{l(i,j)} = \max\left\{a^{l(i,j)}\right\};\ (j-1)n+1 \le t \le jn, \tag{2}$$

where p^{lij} is pooled output and n is the width of the pooling area.

The comparison between the full connection layer and the global average pooling layer is shown in Figure 2. It can be seen from Figure 1 that it needs to expand all the features of each feature graph before using the full connection layer, while the global average pooling layer only needs to calculate the average value of each feature graph. It is easy to see from this simple comparison structure diagram that it can greatly reduce the parameter calculation of the classical convolutional neural network using the global average pooling layer to replace the full connection layer.

2.2. Light Gradient Boosting Machine. Gradient boosting decision tree (GBDT) is a long-standing model in machine learning. Its main idea is to use a weak classifier, decision tree, and iterative training to get the optimal model, which has the advantages of good training effect and is not easy to overfit. LightGBM is a framework to realize the GBDT algorithm, it supports efficient parallel training, and it has the advantages of faster training speed, lower memory consumption, better accuracy, support for distribution, and can quickly process massive data. LightGBM uses the negative gradient of the loss function as the residual approximation of the current decision tree to fit the new decision tree. It uses the histogram algorithm, which takes up less memory and reduces the complexity of data separation. It adopts the leaf-wise strategy with depth restriction, and it will find the leaf with the largest splitting gain, the largest amount of data, from all the current leaves every time, and then splits it. In this way, if the splitting times are the same, the leaf-wise strategy can reduce more errors and get better accuracy. LightGBM can skillfully solve the problem that traditional boosting algorithm is very time-consuming in the large sample environment, and the key of LightGBM is to combine two new methods of gradient-based one-side sampling (GOSS) and exclusive feature bundling (EFB). GOSS is a balancing algorithm in reducing the amount of data and ensuring accuracy. GOSS is to reduce the amount of calculation by distinguishing the instances of different gradients, retaining the larger gradient instances, and randomly sampling the smaller gradients, so as to improve the efficiency. EFB is a way to reduce the feature dimension by binding features to improve computing efficiency. Usually, the bundled features are mutually exclusive so that the two features will not lose information.

3. Model of the Improved Convolution Neural Network and LightGBM

3.1. Structure Diagram of the Model. The structure diagram of the improved convolution neural network and LightGBM (GCNN-LightGBM) model is shown in Figure 3, and it is mainly composed of convolution layer, pooling layer, global average pooling layer, and LightGBM classifier. Before the original one-dimensional vibration signal was input into the convolution layer, the random deactivation with a probability of 0.2 was carried out on it, so as to improve the generalization ability of the training model and the stability of fault diagnosis under variable load conditions. There are two convolution layers and two pooling layers. In the first layer, a large convolution kernel is used to obtain more effective information in the low- and medium-frequency bands of the original signal. The feature maps obtained by

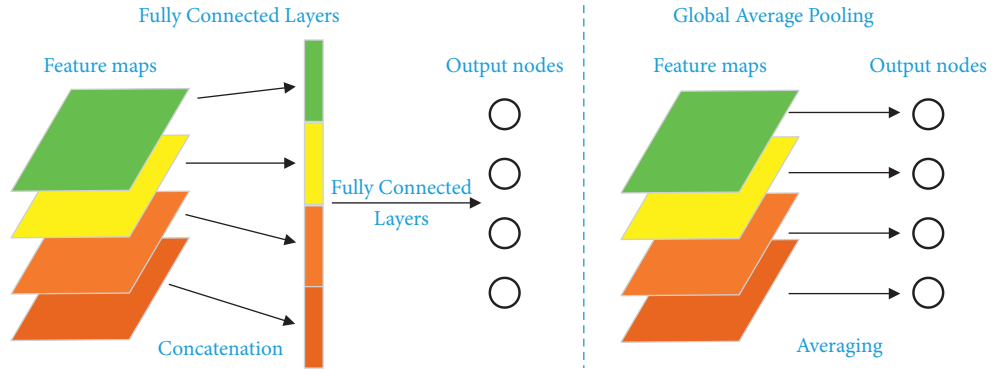


Figure 2: Comparison between the full connection layer and global average pooling layer.

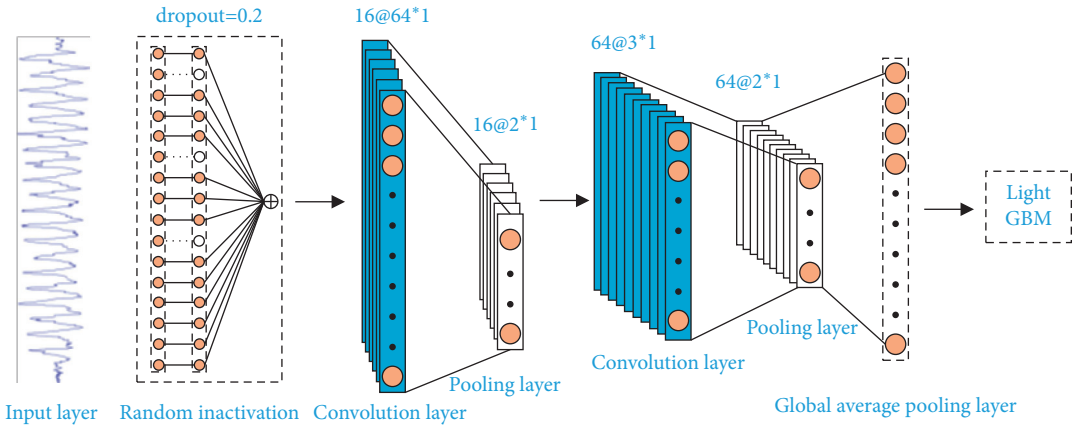


Figure 3: Structure diagram of the model of the improved convolution neural network and LightGBM.

two convolution layers and two pooling layers are input into the global average pooling layer, and the secondary feature extraction and data dimension reduction are realized by averaging each feature map. Finally, the extracted low-dimensional features are input into the LightGBM classifier for classification.

3.2. Parameter Setting of the GCNN-LightGBM Model. The GCNN-LightGBM model uses the improved convolution neural network to extract the adaptive features of the bearing vibration signals. It is very important for the feature extraction effect to select the superparameters of the convolution neural network. Therefore, the parameters of the convolution neural network are trained by softmax classifier (shown as Table 1). After the convolution part is trained, LightGBM is used to replace the softmax layer. The selection of LightGBM parameters is processed by the Bayesian parameter adjustment algorithm. The meanings and values of some important parameters are shown in Table 2.

4. Test and Performance Analysis

4.1. Test Data Set. The test data are selected from the Case Western Reserve University Bearing Data Center. The normal bearing data under different load conditions and the fault bearing data of the inner ring, steel ball, and outer ring

Table 1: Hyperparameter setting of the improved convolution neural network.

Names of the hyper parameters	Parameter value
Input random deactivation rate	0.2
Loss function	Categorical_crossentropy
Optimizer	Adam
Initial learning rate	0.001
Attenuation coefficient	0.5
Attenuation interval	5
Minimum learning rate	0.00001
Number of batches	32
Iterations	20

Table 2: Value of important parameters of LightGBM.

Names of the parameters	Parameter value
n_estimators	500
max_depth	6
num_leaves	34
learning_rate	0.07
bagging_fractin	0.68
bagging_freq	5
feature_fraction	0.6
lambda_l1 , lambda_l2	3.4, 2.1
min_data_in_leaf	33
min_split_gain	0.96
min_sum_hessian_in_leaf	0.003

with the damage diameters of 0.178, 0.356, and 0.534 mm are selected, and the fault bearing is at the drive end. A total of 10 kinds of bearing operation state data are selected. The sampling frequency is set to 12 kHz, and 1,024 data points are collected as a sample each time. The z-sore normalization method is used to preprocess the data before feature extraction so as to accelerate the convergence speed of the convolution neural network. The expression of the z-sore method is

$$x' = \frac{x - u}{\sigma}, \tag{3}$$

where x is the original sample value, u is the mean value of all sample data, σ is the standard deviation of all samples, and x' is the normalized value.

The selected data are divided into three data sets corresponding to the load of 1 HP, 2 HP, and 3 HP. Each data set contains 10,000 samples, there are 10 kinds of bearing operation state, and each bearing state includes 1,000 samples. About 70% of samples are selected as the training set, 20% samples are selected as the verification set, and 10% samples are randomly selected as the test set. The specific data sets are shown in Table 3.

Generally, the distribution of the data set is different due to the amplitude, fluctuation period, and phase inconsistency of the vibration signals under different working conditions. Therefore, it needs the classifier designed has strong generalization ability and robustness. However, it is not realistic to collect and mark enough training samples to make the classifier robust to all the working conditions. In this study, one method using the single load to train the fault diagnosis model and using the test set of the other loads to carry out the fault diagnosis is adopted [23]. For example, it requires the model trained under 1 HP load to not only have high classification accuracy in the 1 HP test set but also in the 2 HP or 3 HP test set, and the variable load adaptive data set constructed to achieve this goal is shown in Table 4.

Table 3: Bearing test data set.

Bearing status	Fault diameter (mm)	Number of samples
Normal	—	700/200/100
Ball failure	0.178	700/200/100
	0.356	700/200/100
	0.534	700/200/100
Inner ring fault	0.178	700/200/100
	0.356	700/200/100
	0.534	700/200/100
Outer ring fault	0.178	700/200/100
	0.356	700/200/100
	0.534	700/200/100

Table 4: Variable load adaptive data set.

Training data set	Training sample	Target test set	Test sample
1 HP	7000	2 HP, 3HP	1000, 1000
2 HP	7000	1 HP, 3HP	1000, 1000
3 HP	7000	1 HP, 2HP	1000, 1000

4.2. Model Validation. In the experiment, the GCNN-LightGBM model uses the deep learning framework Keras in Python language, the classification module directly calls the LightGBM software package, and the established network is used to train and test using different data sets. Because the initialization of input data and neural network weights is random, the average value was calculated after each data set was trained 10 times so as to ensure the reliability of the test results.

In order to verify that the improved convolutional neural network has stronger generalization ability, the contrast model of the classical convolutional neural network and LightGBM is constructed, and the network structure and training parameters of the contrast model are consistent with the GCNN-LightGBM model except for the full connection layer. At the same time, in order to verify that LightGBM has a stronger classification ability than the softmax layer, one contrast model of the improved convolutional neural network and softmax also is constructed, and the feature extraction part of the contrast model is consistent with the model in this study except the softmax classifier.

The recognition accuracy of each model under different load conditions is shown in Figures 4 and 5.

It can be seen from Figures 4 and 5 that the average classification accuracy of the GCNN-LightGBM model is slightly higher than that of the CNN-LightGBM model under the same load condition, while the average classification accuracy of the GCNN-LightGBM model is 2.72% higher than that of the CNN-LightGBM model under variable load condition, and it is verified that the improved convolutional neural network has a better anti-overfitting effect and can improve the generalization ability of the model. The average classification accuracy of the GCNN-LightGBM model is 1.05% and 0.77% higher than that of the GCNN-softmax model, respectively, under the same load and variable load conditions, and it indicates that LightGBM has a stronger classification ability than softmax. LightGBM classifier can achieve good classification results under the same load condition, but the average classification accuracy is less than 68% under the variable load condition, which indicates that it is easy to overfit when LightGBM is directly used to train the original data although it is a very powerful classifier, and it is necessary to extract features from the original data. The classification accuracy between adjacent conditions is high under variable load conditions, which indirectly reflects that the distribution difference of adjacent load data sets is small but the distribution difference of nonadjacent load data sets is large.

4.3. Contrast Test

4.3.1. Accuracy Rate of Fault Diagnosis. Since the classification accuracy of the GCNN-LightGBM model is close to 100% under the same load test set, several deep learning models [24, 25] that have achieved good classification results under the same load conditions are selected to carry out the

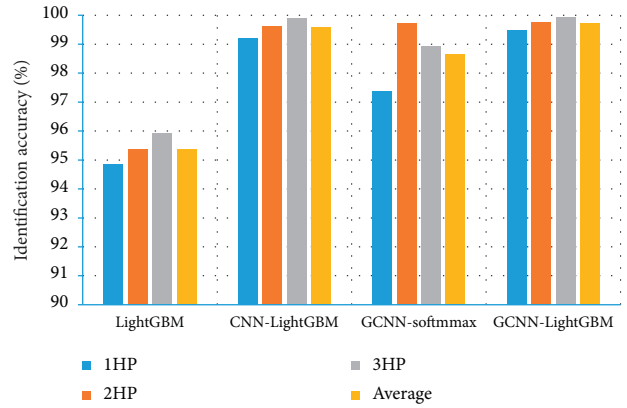


Figure 4: The classification accuracy of each model under the same load condition.

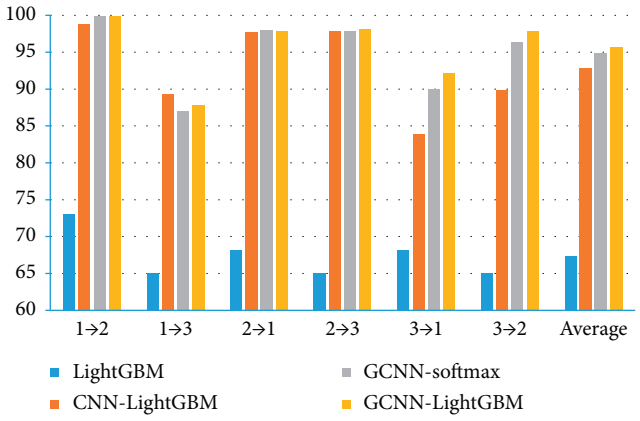


Figure 5: The classification accuracy of each model under different load conditions.

comparative analysis under the variable load test sets, so as to highlight the advantages of the generalization ability and load migration ability of GCNN-LightGBM model. The classification accuracy of different deep learning models under the variable load adaptive data sets is shown in Figure 6.

It can be seen from Figure 6 that the models of CNN-LSTM and WDCNN have strong adaptability when they are trained under the load conditions of 1 HP and 2 HP, and the classification accuracy can reach more than 90% in other variable load test sets. But the load migration ability of the models is not strong when they are trained under the load conditions of 3 HP, and the classification accuracy can only reach about 80% under the load conditions of 1 HP and 2 HP. The classification accuracy of the CNN-SVM model trained under the load condition of 3 HP in other test sets is close to 100%, but the classification accuracy of the CNN-SVM model trained under the load condition of 1 HP in other test sets is even less than 80%. It shows that the overall robustness and load migration ability of the three comparison models are not very strong, although they can achieve good classification results on a variable load test set.

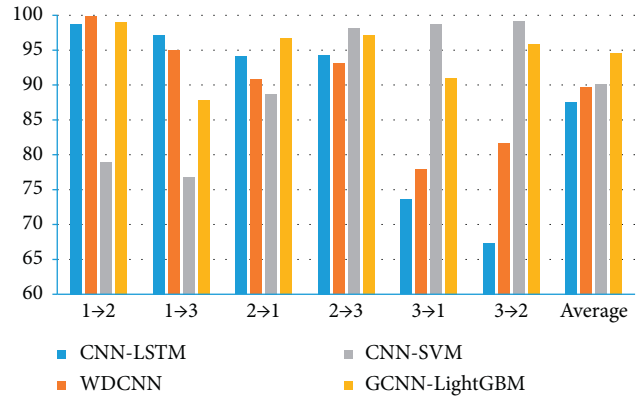


Figure 6: The classification accuracy of each model under the variable load condition.

The lowest classification accuracy of the GCNN-LightGBM model is about 87.89% under the variable load conditions, and it is increased by 20.57%, 9.93%, and 11.10%, respectively, compared with the worst cases of the models of CNN-LSTM, WDCNN, and CNN-SVM. The average classification accuracy of the GCNN-LightGBM model is 94.64%, which is significantly higher than that of other models. It can be seen that the GCNN-LightGBM model has the better overall classification effect under the variable load conditions and also has better generalization ability and load migration ability.

4.3.2. Efficiency of Fault Diagnosis. In order to further highlight the advantages of the GCNN-LightGBM model on the efficiency of rolling bearing fault diagnosis, the training time and diagnosis time of each model, the amount of training parameters of the deep learning model, and the number of required layers of training parameters (excluding pooling layer) are recorded in the process of comparative test, and these parameter values are shown as Table 5. Among them, the average duration is the average value of each model trained for 10 times under different load conditions, and the average value of different loads is calculated again.

It can be seen from Table 5 that the required training parameters and network layers of the GCNN-LightGBM model are the least, especially the amount of the training parameters are several orders of magnitude different from the other three networks, and the required average training and diagnosis time is the smallest among the four models. The training time is shortened by 90.81%, and the fault diagnosis efficiency is improved by 10.44 times compared with the CNN-LSTM model.

These data show that the less the parameters and layers of the model need to be trained, the shorter the time of model training and fault diagnosis. The advantage of the short fault diagnosis time of the GCNN-LightGBM model will be more obvious when there are hundreds or even millions of samples need to be trained or diagnosed, and it can save a lot of time cost.

Table 5: Analysis table of training time and diagnosis time of each model.

Models	Training parameters		Average duration (s)	
	Number of layers	Number	Training	Diagnosis
CNN-LSTM	8	507333	432.14	1.03
WDCNN	7	20755	107.26	0.12
CNN-SVM	4	63518	65.81	0.15
GCNN-LightGBM	2	2413	39.73	0.09

5. Conclusions

In order to solve the problem that weak generalization ability and long training time in most fault diagnosis models based on deep learning, one intelligent diagnosis method of rolling bearing fault based on the improved convolution neural network and light gradient boosting machine is proposed.

(1) First, the random deactivation with a probability of 0.2 was carried out on the original one-dimensional vibration signal, so as to improve the generalization ability of the training model and the stability of fault diagnosis under variable load conditions. Second, the signal was input into the GCNN. In the first layer, a large convolution kernel is used to obtain more effective information in the low- and medium-frequency bands of the original signal. The feature maps obtained by two-layer convolution pooling operation are input into the global average pooling layer, and the secondary feature extraction and data dimension reduction are realized by averaging each feature map. Finally, the extracted low-dimensional features are input into the LightGBM classifier for classification.

(2) The results show that (1) the average classification accuracy of the GCNN-LightGBM is 99.72% for the same load test set and 95.62% for the variable load test set; (2) the GCNN-LightGBM model has the higher average classification accuracy on the variable load test set compared with the models of CNN-LSTM, WDCNN, and CNN-SVM, and it has the stronger generalization ability and load migration ability; (3) the GCNN-LightGBM model only needs two training layers, and the amount of parameter calculation is less than 3,000, the training and fault diagnosis durations are 39.73 s and 0.09 s, respectively, and these data are far lower than other comparison models, which shows that the GCNN-Light GBM model has the advantages of simple structure, less parameter calculation, and high efficiency of training and fault diagnosis.

(3) In this study, the generalization ability of the model is improved according to the change of load. In the future, the robustness of the model will be further improved by adding noise interference to the samples.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors are grateful for the financial support provided by the National Natural Science Foundation of China under grant no. 51805151 and the Key Scientific Research Project of the University of Henan Province of China under grant no. 21B460004.

References

[1] C. Xing, Z. Huimin, S. Shifan et al., "An improved quantum-inspired cooperative co-evolution algorithm with muli-strategy and its application," *Expert Systems with Applications*, vol. 171, no. 1, Article ID 114629, 2021.

[2] W. Deng, S. Shang, C. Xing et al., "Quantum differential evolution with cooperative coevolution framework and hybrid mutation strategy for large scale optimization," *Knowledge-Based Systems*, vol. 224, Article ID 107080, 2021.

[3] D. Verstraete, A. Ferrada, E. Lopez Droguett, V. Meruane, and M. Modarres, "Deep learning enabled fault diagnosis using time-frequency image analysis of rolling element bearings," *Shock and Vibration*, vol. 2017, Article ID 5067651, 2017.

[4] M. He and D. He, "Deep learning based approach for bearing fault diagnosis," *IEEE Transactions on Industry Applications*, vol. 53, no. 3, pp. 3057–3065, 2017.

[5] D. T. Hoang and H. J. Kang, "A survey on Deep Learning based bearing fault diagnosis," *Neurocomputing*, vol. 335, pp. 327–335, 2019.

[6] W. Zhenghong, J. Hongkai, Z. Ke, and X. Li, "An adaptive deep transfer learning method for bearing fault diagnosis," *Measurement*, vol. 151, Article ID 107227, 2020.

[7] J. Sun, C. Yan, and J. Wen, "Intelligent bearing fault diagnosis method combining compressed data acquisition and deep learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 1, pp. 185–195, 2018.

[8] X. Ding and Q. He, "Energy-fluctuated multiscale feature learning with deep ConvNet for intelligent spindle bearing fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 8, pp. 1926–1935, 2017.

[9] R. X. Chen, X. Huang, L. X. Yang, and B. P. Tang, "Rolling bearing fault identification based on convolution neural network and discrete wavelet transform," *Journal of Vibration Engineering*, vol. 31, no. 5, pp. 883–891, 2018, in Chinese.

[10] S. Haidong, J. Hongkai, L. Xingqiu, and W. Shuaipeng, "Intelligent fault diagnosis of rolling bearing using deep wavelet auto-encoder with extreme learning machine," *Knowledge-Based Systems*, vol. 140, pp. 1–14, 2018.

[11] Y. Ding and M. Jia, "A multi-scale convolutional auto-encoder and its application in fault diagnosis of rolling

bearings," *Journal of Southeast University*, vol. 35, no. 4, pp. 417–423, 2019.

[12] Z. Wang, L. Yao, and Y. Cai, "Yongwu. Rolling bearing fault diagnosis using generalized refined composite multiscale sample entropy and optimized support vector machine," *Measurement*, vol. 156, Article ID 107574, 2020.

[13] D. Goyal, A. Choudhary, B. S. Pabla, and S. S. Dhami, "Support vector machines based non-contact fault diagnosis system for bearings," *Journal of Intelligent Manufacturing*, vol. 31, no. 5, pp. 1275–1289, 2020.

[14] S. Dong, T. Luo, L. Zhong, L. Chen, and X. Xu, "Fault diagnosis of bearing based on the kernel principal component analysis and optimized k-nearest neighbour model," *Journal of Low Frequency Noise, Vibration and Active Control*, vol. 36, no. 4, pp. 354–365, 2017.

[15] M. Cerrada, G. Zurita, D. Cabrera, R. V. Sánchez, M. Artés, and C. Li, "Fault diagnosis in spur gears based on genetic algorithm and random forest," *Mechanical Systems and Signal Processing*, vol. 70-71, pp. 87–103, 2016.

[16] Z. Wu, X. Wang, and B. Jiang, "fault diagnosis for wind turbines based on ReliefF and eXtreme gradient boosting," *Applied Sciences*, vol. 10, no. 9, p. 3258, 2020.

[17] J. Tao, C. Qin, W. Li, and C. Liu, "intelligent fault diagnosis of diesel engines via extreme gradient boosting and high-accuracy time-frequency information of vibration signals," *Sensors*, vol. 19, no. 15, p. 3280, 2019.

[18] A. A. Taha and S. J. Malebary, "An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine," *IEEE Access*, vol. 825579 pages, 2020.

[19] P. j. Chun, S. Izumi, and T. Yamane, "Automatic detection method of cracks from concrete surface imagery using two-step light gradient boosting machine," *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 1, pp. 61–72, 2020.

[20] C. Zhang, X. Yao, J. Zhang, and E. Liu, "Tool wear monitoring based on deep learning," *Computer Integrated Manufacturing Systems*, vol. 23, no. 10, pp. 2146–2155, 2017, (in Chinese).

[21] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mechanical Systems and Signal Processing*, vol. 115, pp. 213–237, 2019.

[22] S. Khan and T. Yairi, "A review on the application of deep learning in system health management," *Mechanical Systems and Signal Processing*, vol. 107, pp. 241–265, 2018.

[23] W. Zhang, G. Peng, C. Li, Y. Chen, and Z. Zhang, "A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals," *Sensors*, vol. 17, no. 2, p. 425, 2017.

[24] L. Zhang, W. Xu, and L. Jing, "Study on fault diagnosis methods of rolling bearing based on deep convolutional neural network and support vector machine," *Modular Machine Tool & Automatic Manufacturing Technique*, vol. 2019, no. 7, pp. 69–71, 2019, in Chinese.

[25] S. Hao, F.-X. Ge, Y. Li, and J. Jiang, "Multisensor bearing fault diagnosis based on one-dimensional convolutional long short-term memory networks," *Measurement*, vol. 159, Article ID 107802, 2020.

Hindawi
Shock and Vibration
Volume 2021, Article ID 1069470, 19 pages
https://doi.org/10.1155/2021/1069470

Research Article

Improved Time Domain Substructural Damage Identification Method on Large-Span Spatial Structure

Taoyuan Yang,[1] Kun Liu,[2,3] and Guibo Nie[1]

[1]*Key Laboratory of Earthquake Engineering and Engineering Vibration, Institute of Engineering Mechanics, China Earthquake Administration, Harbin 150000, China*
[2]*Key Lab of Structures Dynamics Behavior and Control of the Ministry of Education, Harbin Institute of Technology, Harbin 150090, China*
[3]*Key Lab of Smart Prevention and Mitigation of Civil Engineering Disasters of the Ministry of Industry and Information Technology, Harbin Institute of Technology, Harbin 150090, China*

Correspondence should be addressed to Guibo Nie; nieguibo0323@163.com

Received 12 August 2021; Revised 1 October 2021; Accepted 6 October 2021; Published 23 October 2021

Academic Editor: Chengwei Fei

The large-span spatial structure is a complex structural type with large number of elements, which makes health monitoring difficult. A time domain global substructural identification method was proposed in this paper to identify the local damage for the large-span spatial structure. The proposed method was an improvement method based on the time domain substructural identification method, which can identify the damage with a reduced structural model, and explicit force identification method, which can give convergent force identification result with incomplete response measurements, and it can assess one or more target substructures without knowing the conditions of the other parts of the structure. The application of the new method to large-span spatial structure was presented, and then an improved global method was proposed to further reduce the computation time and promote it in practice. Two orthogonal spatial square pyramid grid structure models are constructed to validate the time domain global substructural method and the improved global method. The results show that the time domain global substructural damage identification method identifies the small local damage in multiple members with satisfactory accuracy and the improved global method effectively shortens identification time.

1. Introduction

A great quantity of large-span spatial structures, such as stadiums, train stations, airport terminals, and exhibition centers, have been constructed around the world in the past several decades. The large-span spatial structure usually has important functions to undertake activities or transfer passengers, thereby always gathering large numbers of people in it. Once failure or even collapse occurs, it will cause terrible casualties, huge property loss [1], and extremely bad social impact, which makes the structural health monitoring of large-span spatial structure a crucial issue. The small local damage may lead to disastrous damage under strong earthquake, so the small local damage detection becomes a key issue for the structural health monitoring of the large-span spatial structure. The damage detection method based on vibration information [2–5] has been proposed and developed in the past three decades, and it has been applied in various structures (e.g., bridges [6], frame structures, and plane trusses [7, 8]).

Several structural health monitoring methods based on vibration information have been proposed to estimate the local damage of space truss structure in recent years. The square ratio of frequency variation and variation rate of axial strain were adopted to identify local damage of space truss structure [9, 10], and the numerical simulation results show that the single damage location can be detected accurately while the damage quantification needs to be further discussed. Song et al. [11] expanded incomplete mode shape through a dynamic model expansion technique, identified the possible damage members by model strain energy, and then used Least Squares Support Vector

Machine to detect the location and extent of the local damage of a space truss. However, the measurement noise level is a challenge. Another method based on residual modal force and modal strain energy was proposed to detect the local damage of space truss [12], and a space truss was numerically simulated to verify the proposed method. The damage location and quantification can be detected accurately, but the sensitive mode shape is needed in the simulation cases.

The neural networks have been adopted in structural health monitoring widely since last decades. Wu and Zhang [13] presented the damage identification method of grid structures based on BP neural networks, and it has been verified with numerical model of space truss and experimental model of a double-layer cylindrical reticulated shell structure. The erroneous judgment is always unavoidable for some damage locations, and this method cannot give good identification results with missed damaged members. To solve the data explosion problem, Liu et al. [14] introduced the method of substructure, while using probabilistic neural networks (PNN) to improve identification accuracy. The numerical simulation results show that this method is effective, but there are still erroneous judgments. If the damaged members are located in different substructures, the condition assessment will become more complicated, and excessive substructure partitioning will produce some new problems in global detection. He and Yan [15] adopted the method combining wavelet packet with support vector machine to identify the damage of single-layer lattice shell, but the comprehensive correct rate of damage identification was 80%.

Data fusion technique has also been adopted for the damage detection of large-span spatial structure to improve the accuracy of the damage identification. The information of acceleration measurements and strain measurements is used to identify the damage of reticulated shell based on Dempster–Shafer evidence theory [16, 17]. Teng and Yao [18] added wavelet packet analysis on the basis of information fusion. Although both methods improve the accuracy of damage identification, for avoiding data explosion, global structure still needs to be divided into too many parts. Moreover, the damaged members are all located in the same substructure; the scenarios with damaged members located in different substructures are not considered. Therefore, these methods can be used to assess a small amount of concentrated damage.

A probabilistic substructure identification and health monitoring methodology was proposed [19], which does not require any interface measurements or excitation measurements. This method can be applied widely, because only the stochastic model of the input is required. Then, the boundary force in the substructure is modulated as filtered white noise [20], which can be viewed as a continuity condition. This proposed method does not require stationarity of the response. An identifiability-enhanced Bayesian frequency-domain substructure identification approach is proposed without the requirement of input or boundary force measurements [21], in which extra constraints are imposed to enhance the identifiability of the inverse problem. Substructural identification approaches provide effective methodologies for the identification of large-span spatial structures, because they offer the flexibility to isolate some critical substructures for identification. However, for large-span spatial structure, the common substructural damage identification methods based on vibration information still suffer difficulties because of the characteristics of large-span spatial structure including complex structural composition, intensive frequency distribution, and numerous dynamic degrees of freedom.

The response sensitivity-based method was derived by Jahn [22] in 1940s; usually, its application involves model updating iteration and optimization algorithm [23–25] to enhance the identifiability. Liu et al. [26] proposed a substructural condition assessment method based on response sensitivity, and the accuracy and effectiveness are then validated with simulation studies of a plane truss when damage only exists in one of the substructures. This method, with comparatively lesser substructure partitioning, may be suitable for large complex structures. Since this method shows potential for damage identification of large-span spatial structure, in this paper, the global damage identification method based on this substructural method is presented to detect the location and extent of local damage in different substructures with uncompleted acceleration measurements. An improved global substructural damage identification method is further proposed to reduce calculation duration. The global substructural damage identification method and the improved global substructural damage identification method are numerically verified with simulation studies of two double-layer lattice space structures.

2. The Global Substructural Damage Identification Method

The target substructure can be assessed based on the time domain response sensitivity matrix of the substructural system as is described in Appendix B, and the detailed process has been written in previous paper [26]. However, this original substructural method is effective only when the condition of the rest substructure is known. The interface force depends on the local damage extent of the target substructure and the rest substructure, so the interface force sensitivity is related to the local damaged parameters of both substructures. When the local damaged parameter of the rest substructure is unknown, the substructural response sensitivity considering the interface force effect cannot be obtained. To overcome this problem, this paper will propose a global substructural method based on the interrelation between each substructural dynamic equation.

If the whole structure is divided into n substructures, similar to (B.6), the motion equation of ith substructure among the n substructures can be written as

$$M_i\ddot{x}_i + C_i\dot{x}_i + K_i x_i = L_i P_i, \tag{1}$$

where the subscripts i denote the number of substructures; obviously $1 \le i \le n$. The interface force P_i is related to all the

substructural stiffness matrices from K_1 to K_n. Due to the interconnection between substructural motion equations, the global substructural dynamic equation set can be written as

$$\begin{cases} M_1\ddot{x}_1 + C_1\dot{x}_1 + K_1x_1 = L_1P_1, \\ \quad\vdots \\ M_i\ddot{x}_i + C_i\dot{x}_i + K_ix_i = L_iP_i, \\ \quad\vdots \\ M_n\ddot{x}_n + C_n\dot{x}_n + K_nx_n = L_nP_n. \end{cases} \tag{2}$$

In (2) the DOFs of the interface are repeated once compared with the equation of motion of the full structure.

2.1. The Global Substructural Damage Identification Algorithm. Assume that the local damage extent of the jth substructure is damage index vector α_j, which includes the local damaged parameter of every element in the jth substructure. A change in the global substructural stiffness matrix can be described as

$$\Delta K = \sum_j \alpha_j K_j. \tag{3}$$

Performing differentiation of both sides of (1), the motion equation of ith substructure, with respect to the damage index vector α_j, we have

$$M_i\frac{\partial\ddot{x}_i}{\partial\alpha_j} + C_i\frac{\partial\dot{x}_i}{\partial\alpha_j} + K_i\frac{\partial x_i}{\partial\alpha_j} = -\frac{\partial K_i}{\partial\alpha_j}x_i - a_1\frac{\partial K_i}{\partial\alpha_j}\dot{x}_i + L_i\frac{\partial P_i}{\partial\alpha_j}, \tag{4}$$

where $\partial\ddot{x}_i/\partial\alpha_j$, $\partial\dot{x}_i/\partial\alpha_j$, and $\partial x_i/\partial\alpha_j$ are the substructural response sensitivity matrices which can be obtained by solving (4). The substructural response sensitivity algorithm is presented in (B.9)–(B.16).

Let the ith substructural response sensitivity matrix with respect to the jth substructural damage index vector be represented as $S_{i,j}$. All the substructural sensitivity matrices can be calculated, and then the global substructural sensitivity matrix is assembled as

$$S = \begin{bmatrix} S_{1,1} & \cdots & S_{1,j} & \cdots & S_{1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ S_{i,1} & \cdots & S_{i,j} & \cdots & S_{i,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ S_{n,1} & \cdots & S_{n,j} & \cdots & S_{n,n} \end{bmatrix}. \tag{5}$$

The identification equation for the local damage of all the substructures with the global substructural sensitivity matrix can be represented as

$$S\alpha + o\left(\alpha^2\right) = \ddot{X}^{cs} - \ddot{X}^{ms}, \tag{6}$$

where

$$\begin{aligned} \alpha &= \left[\begin{array}{cccc} \alpha_1^T & \cdots & \alpha_i^T & \cdots & \alpha_n^T \end{array}\right], \\ \ddot{X}^{cs} - \ddot{X}^{ms} &= \left[\begin{array}{cccc} X_1^T & \cdots & X_i^T & \cdots & X_n^T \end{array}\right]. \end{aligned} \tag{7}$$

The measured acceleration response $\ddot{X}^{ms}$, the corresponding calculated acceleration response $\ddot{X}^{cs}$, and the global substructural sensitivity matrix S in (6) are known while the higher order term $o(\alpha^2)$ can be omitted. The unknown damage index vector α can be determined from (6) with an iterative approach based on Gauss elimination method. The substructural sensitivity matrix $S_{i,j}$ and the difference values of the response vector X_i are the basic computing unit in the process of solving. The fractional change increment of the stiffness vector $\Delta\alpha_i$ is the result of each iteration, and the final damage index vector α_i is obtained by summarizing all the iterative results.

The substructural stiffness matrix is only related to its own damage parameter, so (4) can be also written as

$$M_i\frac{\partial\ddot{x}_i}{\partial\alpha_j} + C_i\frac{\partial\dot{x}_i}{\partial\alpha_j} + K_i\frac{\partial x_i}{\partial\alpha_j} = \begin{cases} -\dfrac{\partial K_i}{\partial\alpha_j}x_i - a_1\dfrac{\partial K_i}{\partial\alpha_j}\dot{x}_i + L_i\dfrac{\partial P_i}{\partial\alpha_j}, & i = j, \\ L_i\dfrac{\partial P_i}{\partial\alpha_j}, & i \neq j. \end{cases} \tag{8}$$

In (8), if $i \neq j$, only one term is considered on the right-hand side. Obviously, the substructural sensitivity matrix obtained from (8) or (4), $S_{i,j}$ $(i \neq j)$, has less contribution than $S_{i,i}$ to the global substructural sensitivity matrix. A part of the global substructural sensitivity matrix far from the diagonal can be ignored as zero, and then the global

substructural sensitivity matrix becomes sparse, and the computational processes of Gauss elimination can be simplified.

When the full structure is divided into two substructures, after ignoring low contribution sensitivity matrices, the global substructural sensitivity matrix can be written as

$$S_{\mathrm{II}} = \begin{bmatrix} S_{1,1} & 0 \\ 0 & S_{2,2} \end{bmatrix}. \tag{9}$$

According to the solution by Gauss elimination method, the kth iterative algorithm can be written as

$$\begin{aligned} \Delta\alpha_1^k &= S_{1,1}^+ X_1, \\ \Delta\alpha_2^k &= S_{2,2}^+ X_2. \end{aligned} \tag{10}$$

Similarly, when the full structure is divided into three substructures, the simplified global substructural sensitivity matrix and the kth iterative algorithm can be written, respectively, as

$$S_{\mathrm{III}} = \begin{bmatrix} S_{1,1} & S_{1,2} & 0 \\ S_{2,1} & S_{2,2} & S_{2,3} \\ 0 & S_{3,2} & S_{3,3} \end{bmatrix}, \tag{11}$$

$$\begin{aligned} \Delta\alpha_1^k &= \left(S_{2,1}^+ + S_{1,1}^+ S_{1,2} S_{3,2}^+ S_{3,3} S_{2,3}^+ - S_{1,1}^+ S_{1,2} S_{2,2}^+\right)\left[\left(S_{2,3} S_{3,3}^+ S_{3,2} S_{1,2}^+ - S_{2,2} S_{1,2}^+\right) X_1 + X_2 - S_{2,3} S_{3,3}^+ X_3\right], \\ \Delta\alpha_2^k &= S_{1,2}^+ X_1 - S_{1,2}^+ S_{1,1} \Delta\alpha_1^k, \\ \Delta\alpha_3^k &= S_{3,3}^+ X_3 - S_{3,3}^+ S_{3,2} \Delta\alpha_2^k. \end{aligned} \tag{12}$$

The substructural method by dividing the full structure into several substructures can effectively reduce structural model, and then the size of response sensitivity matrix is reduced significantly. When there is local damage in different substructures, the following procedure based on the global substructural damage identification method can be performed to assess all the substructures' conditions.

Step 1: Conduct the dynamic measurement in the structure, and assemble the measured responses. Then, set $k = 0$.

Step 2: The relationship between the interface forces and the measured responses is constructed, and the interface forces are then identified.

Step 3: Compute the response of the structure ($\ddot{X}^{cs}$) and the substructural sensitivity matrices of response ($S_{i,j}$) with respect to the local damaged parameters of different substructure.

Step 4: Let $k = k + 1$, and identify the kth local change increment of the stiffness with the global substructural sensitivity matrix S in (6).

Step 5: Update the finite element model and repeat Steps 2 to 4 until convergence in (13) is met. The final damage index vector α which is used to represent the change of stiffness can be obtained as $\alpha^{k+1} = \sum_{k+1} \Delta\alpha^{k+1}$.

$$\frac{\left\|\alpha^{k+1} - \alpha^k\right\|}{\left\|\alpha^{k+1}\right\|} < \mathrm{Tol1} \begin{cases} \mathrm{Tol1} = 10^{-5} & \text{for the noise free case,} \\ \mathrm{Tol1} = 10^{-3} & \text{for 5\% noise level case.} \end{cases} \tag{13}$$

The flowchart of the global substructural method is shown in Figure 1.

2.2. The Improved Global Substructural Damage Identification Algorithm. The application of global substructural method to large-span spatial structure can reduce the structural model and the size of sensitivity matrix through dividing the full structure into several substructures, but all the elements need to be considered repeatedly in each iterative computation. When the structural model is too large and complex, it will consume a lot of computation time for damage identification with the global substructural method due to the huge number of elements. To enhance the computational efficiency of damage identification, an improved global substructural method can be proposed by shrinking the scope of assessment in the iterative process.

In this method the damage index is calculated at first identification round, and then a critical value is defined for determining the damaged location, in which those elements with the local damage extent larger than the defined critical value are set as the probable damaged elements. After filtering the probable damaged elements, these elements are independently identified, while regarding the others as intact. Then, all the elements are reviewed to avoid missing the damaged element. Based on the global substructural damage identification method, the improved global substructural damage identification method can be performed as in the following steps:

Step 1: Conduct the dynamic measurement in the structure, and assemble the measured responses. Then, set $k = 0$ and $k_1 = 0$.

Step 2: The relationship between the interface forces and the measured responses is constructed, and the interface forces are then identified. Compute the response of the structure and the substructural sensitivity matrices of responses with respect to the local damaged

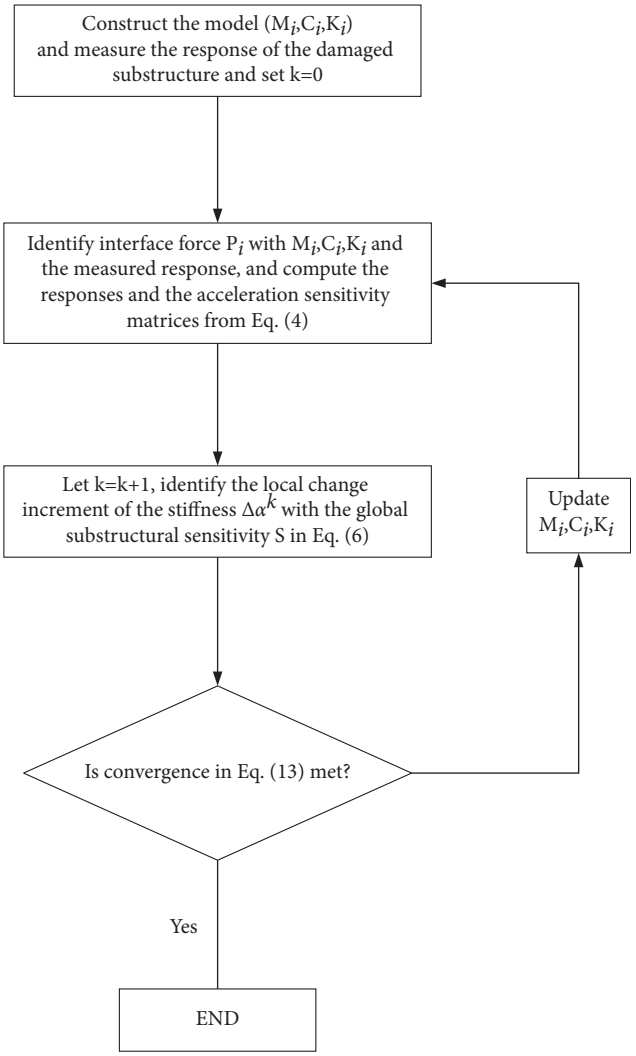


Figure 1: Flowchart of the global substructural method.

parameters of different substructures based on the identified interface forces.

Step 3: Let $k = k + 1$, and identify the kth local change increment of the stiffness with the analytical responses and the corresponding substructural sensitivity matrices in Step 2.

Step 4: Update the finite element model and repeat Steps 2 to 3 until convergence in (14) is met. Define a critical value for determining the damaged location. The elements with the obvious stiffness reduction are selected to be updated in the following update procedures (referred to as "damaged elements"), and the stiffness change of the other elements regarded as intact is set to zero. Set $k_1 = k + 1$; the first-round identification is finished.

$$\frac{\left\|\alpha^{k+1} - \alpha^{k}\right\|}{\left\|\alpha^{k+1}\right\|} < \text{Tol2} \begin{cases} \text{Tol2} = 10^{-2} & \text{for the noise free case,} \\ \text{Tol2} = 10^{-1} & \text{for 5\% noise level case.} \end{cases} \tag{14}$$

Step 5: Let $k = k + 1$, calculate the responses based on the previous updated model of structure and sensitivities of the "damaged elements" in every substructure, and then identify the kth fractional increment of the damaged parameters of the "damaged elements" $\Delta\alpha_d^k$.

Step 6: Update the finite element model and repeat Steps 5 until "damaged elements" convergence in (15) is met.

$$\frac{\left\|\alpha_d^{k+1} - \alpha_d^{k}\right\|}{\left\|\alpha_d^{k+1}\right\|} < \text{Tol3} \begin{cases} \text{Tol3} = 10^{-5} & \text{for the noise free case,} \\ \text{Tol3} = 10^{-3} & \text{for 5\% noise level case.} \end{cases} \tag{15}$$

Step 7: To review the damage identification results, update the finite element model and repeat Steps 2 to 3 until convergence in (14) is met.

The flowchart of the improved global substructural method is shown in Figure 2.

3. Numerical Simulations

3.1. Numerical Simulation of Global Substructural Method. The orthogonal spatial square pyramid grid structure as shown in Figure 3 serves for the following study. This structure is modeled using 240 spatial truss finite elements without internal nodes and 72 hinge nodes giving 150 DOFs. The structure is divided into two substructures as shown in Figure 4. The node number and element number are shown in Figures 3 and 4. Substructure I contains Nodes 1 to 44 and Substructure II contains Nodes 34 to 72. Nodes 34 to 44 are the interface nodes. The structure is hinge-supported at peripheral nodes, and the location of bearing is shown in Figure 3 as solid blue squares. The upper, lower, and web members are all 3.0 meters long. The cross-sectional area of all members is 0.0028 m^2. The plane dimensions of the structure are 18 m × 15 m, and the height of the structure is 2.12 m. The first eight natural frequencies of the structure are 2.724 Hz, 4.621 Hz, 5.438 Hz, 7.029 Hz, 7.030 Hz, 8.057 Hz, 8.462 Hz, and 8.560 Hz, respectively. Rayleigh damping is adopted for the system, and the two damping coefficients are $a_1 = 0.1077$ and $a_2 = 2.1669 \times 10^{-4}$. The mass density and elastic modulus of material are, respectively, 7.85×10^3 kg/m^3 and 2.06 GPa.

Vertical external loads are applied to the structure at Nodes 26 and 47, and they are, respectively, modeled as

$$F_1(t) = 650\sin(20\pi t) + 600\sin(80\pi t) + 550\sin(160\pi t),$$
$$F_2(t) = 500\sin(18\pi t) + 450\sin(70\pi t) + 450\sin(210\pi t), \tag{16}$$

to simulate excitation over a relatively wide range of frequencies.

The sampling rate is 1000 Hz and the time duration of study is 0.5 s after the load application. The acceleration responses of the structure are calculated using the *Newmark*-

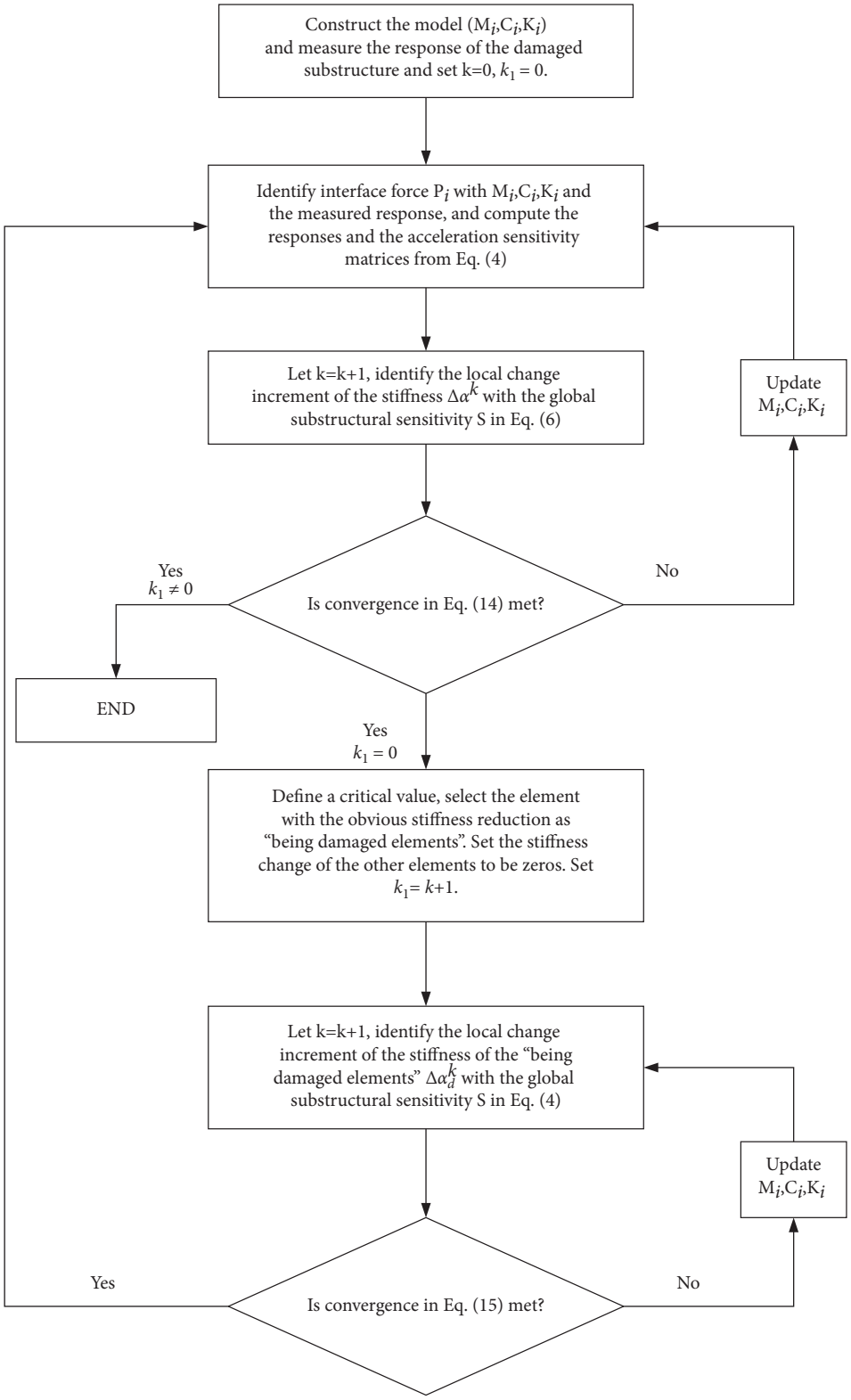


Figure 2: Flowchart of the improved global substructural method.

β method as the "measured" responses. The sensor location is shown in Figure 3 as red hollow circles.

The initial global substructural sensitivity matrix is obtained with the substructural response sensitivity algorithm. The former 120 columns of matrix are related to the local damaged parameters of Substructure I, and the last 120 columns of matrix are related to the local damaged parameters of Substructure II. To compare the contribution

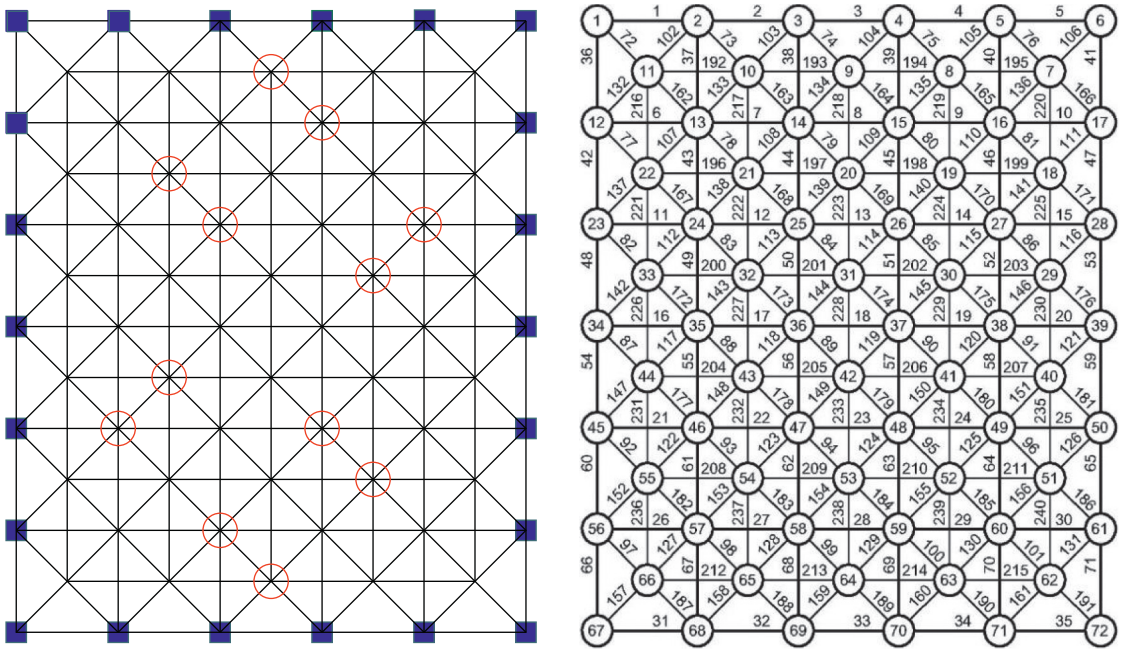


Figure 3: An orthogonal spatial square pyramid grid structure and element number system.

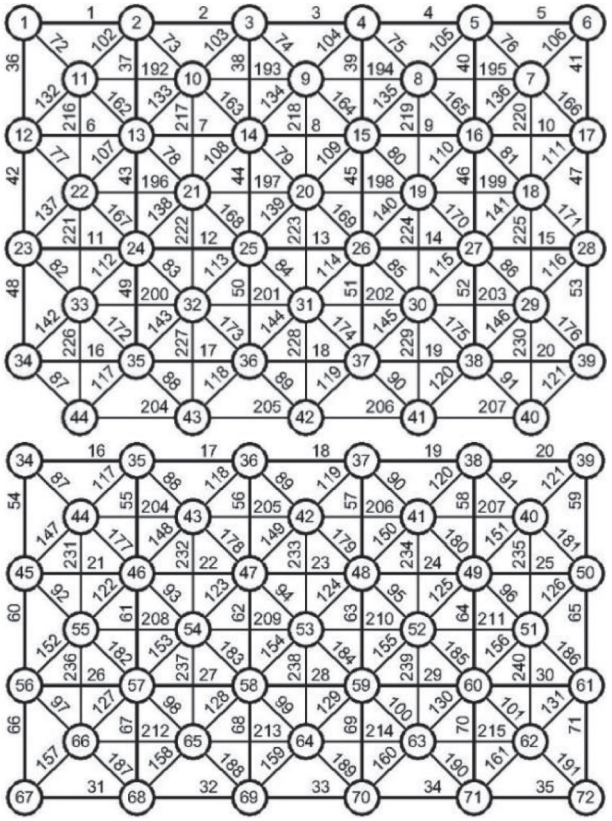


Figure 4: Substructure division.

between different substructural sensitivity matrices, every sensitivity vector norm with respect to single local damaged parameter at the different nodes is calculated. The norm of sensitivity vectors at the sensor locations is shown in Figure 5. The mean of norm of sensitivity vectors is shown in Table 1. It is noted that the norm of sensitivity vectors with respect to the parameters of Substructure I is greater than that of Substructure II at the nodes belonging to Substructure I, but it is the opposite at the nodes belonging to Substructure II. Accordingly, the substructural sensitivity $S_{i,j}$ $(i \neq j)$ has less contribution than $S_{i,j}$ $(i = j)$ to the global substructural sensitivity matrix.

To evaluate the contribution extent of different substructural sensitivity to the global substructural sensitivity matrix, a ratio is defined as

$$\text{Ratio} = \frac{\left\| S_{i,j} \Delta\alpha_j \right\|}{\left\| \sum_j S_{i,j} \Delta\alpha_j \right\|}. \tag{17}$$

Because all the fractional change increments $\Delta\alpha_i$ keep the same level under the small local damage, the defined ratio in (17) can represent the influence of one substructural sensitivity in the identification equation. Small stiffness reduction of 0.1% is simulated in all structural elements, and the initial substructural sensitivity matrices are considered to calculate the ratio. The results of ratio are shown in Table 2, demonstrating that the influence of substructural sensitivity matrix $S_{1,1}$ and $S_{2,2}$ is much larger than the others. In the following numerical simulation, (9) and (10) are applied to assess the condition of structure setting $S_{1,2}$ and $S_{2,1}$ as zero.

Two damage scenarios as shown in Table 3 are studied in this section. 10% stiffness reduction is assumed as local damage of each selected element in each scenario. The damaged members are located in different substructures.

With the global substructural damage identification method, the damage identification results of the two scenarios are shown in Figure 6. The identified results without noise are consistent with the real damaged scenarios, which

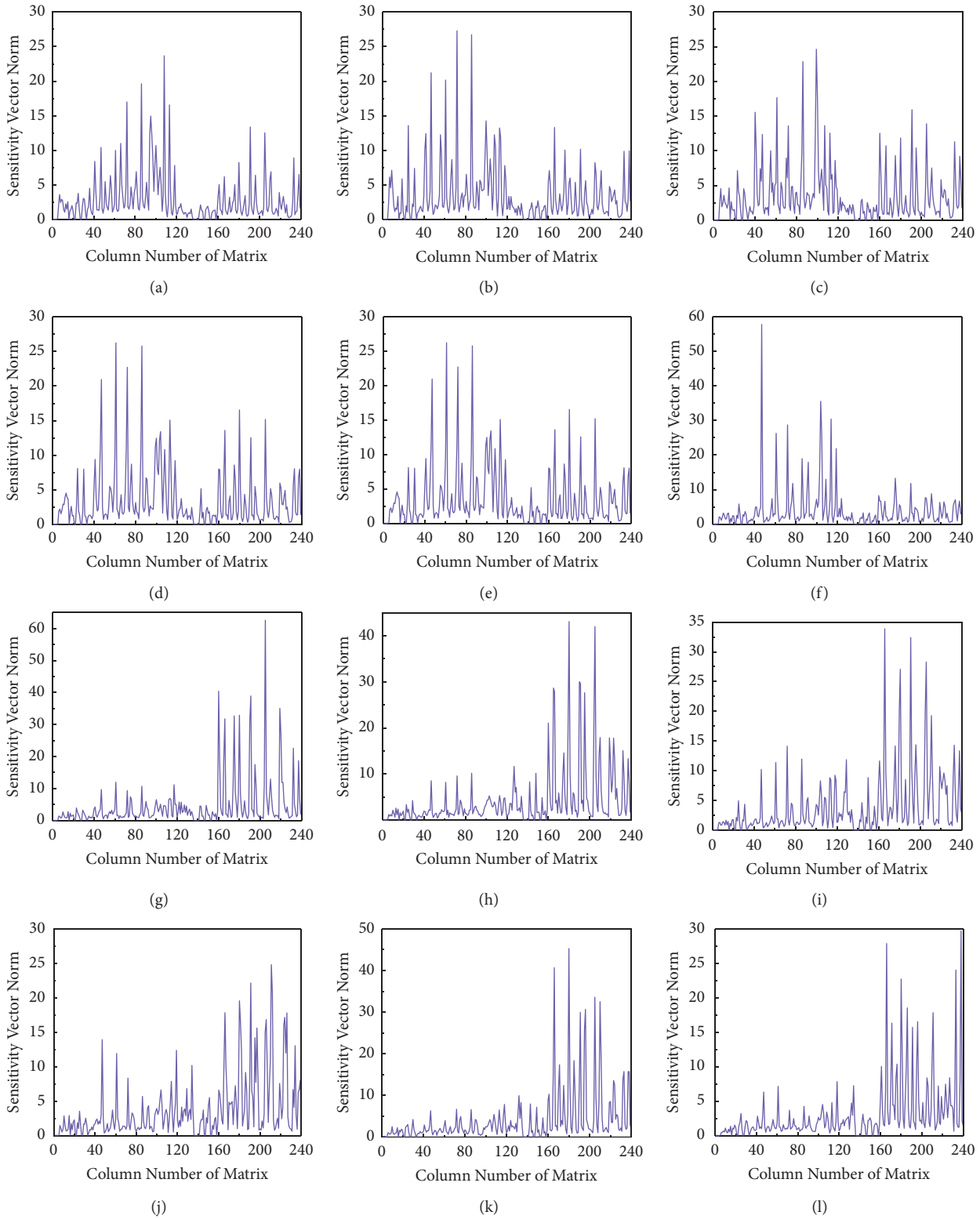


Figure 5: Sensitivity vector norm. (a) At node 9 belonging to Substructure I. (b) At node 15 belonging to Substructure I. (c) At node 21 belonging to Substructure I. (d) At node 25 belonging to Substructure I. (e) At node 37 belonging to Substructure I. (f) At node 30 belonging to Substructure I. (g) At node 43 belonging to Substructure II. (h) At node 46 belonging to Substructure II. (i) At node 48 belonging to Substructure II. (j) At node 52 belonging to Substructure II. (k) At node 58 belonging to Substructure II. (l) At node 64 belonging to Substructure II.

Table 1: Mean value of sensitivity vector norm.

Node number	Mean of sensitivity vector norm	
	Related to parameters of Sub I	Related to parameters of Sub II
9	3.68	1.95
15	4.27	2.44
21	4.18	2.68
25	4.32	2.57
27	4.16	2.11
30	4.84	2.76
All nodes of Sub I	4.24	2.42
43	2.45	6.34
46	2.01	6.24
48	2.22	5.37
52	2.11	5.20
58	1.86	6.09
64	1.43	4.51
All nodes of Sub II	2.02	5.62

Table 2: Contribution extent of substructural sensitivity.

Index of sensitivity matrix	Ratio
$S_{1,1}$	0.88
$S_{1,2}$	0.33
$S_{2,1}$	0.25
$S_{2,2}$	0.92

Table 3: Damage scenarios.

Damage scenario	Damaged element		
	Upper member	Lower member	Web member
Scenario I	12,46	163	206,238
Scenario II	40	84,130	192,240

indicate that the global substructural damage identification algorithm is correct and accurate. There are some small false positives and negatives in the results, because identification of interface force and simplification of sensitivity matrix will generate errors in iterative process. However, the identification results with 5% noise in measurements have larger errors than those in the scenario without noise. The mean values and standard deviations of error are shown in Table 4. The identified error of damaged element 46 in Scenario I reaches about 3.8%, but the damage locations can be identified with satisfactory accuracy in both scenarios.

The size of sensitivity matrices is shown in Table 5, and it is noted that the size of sensitivity matrix is reduced significantly in the identification process. This study shows that the global substructural method can be applied for assessing the double-layer lattice space structure with small damage in multiple members distributed in different substructures by incomplete measured acceleration responses (24% of responses data are measured in this simulation study).

3.2. Numerical Simulation of Improved Global Substructural Method. The orthogonal spatial square pyramid grid structure as shown in Figure 7 serves for the following study. This structure is modeled using 800 spatial truss finite elements without internal nodes and 221 hinge nodes giving 543 DOFs. The structure is divided into three substructures as shown in Figure 8, and the node number and element number are shown in Figure 9. Nodes 64 to 84 and Nodes 138 to 158 are the interface nodes. The structure is hinge-supported at peripheral nodes, and the location of bearing is shown in Figure 9 as solid blue squares. The upper, lower, and web members are all 3.0 meters long. The cross-sectional area of all members is $0.0028\,\mathrm{m}^2$. The plane dimensions of the structure are $30\,\mathrm{m} \times 30\,\mathrm{m}$, and the height of the structure is 2.12 meters. The first eight natural frequencies of the structure are 0.864 Hz, 1.826 Hz, 2.564 Hz, 3.659 Hz, 3.682 Hz, 3.985 Hz, 4.144 Hz, and 4.628 Hz, respectively. Rayleigh damping is adopted for the system, and the two damping coefficients are $a_1 = 0.0368$ and $a_2 = 5.9167 \times 10^{-4}$. The mass density and elastic modulus of material are, respectively, $7.85 \times 10^3\,\mathrm{kg/m}^3$ and 2.06 GPa.

Vertical external loads are applied to the structure at Nodes 67, 71, 141, and 145, and they are, respectively, modeled as

$$\begin{aligned} F_1(t) &= 500\sin(18\pi t) + 450\sin(70\pi t) + 450\sin(210\pi t), \\ F_2(t) &= 650\sin(20\pi t)600\sin(80\pi t) + 550\sin(160\pi t), \\ F_3(t) &= 550\sin(18\pi t) + 500\sin(50\pi t) + 500\sin(160\pi t), \\ F_4(t) &= 600\sin(20\pi t) + 550\sin(30\pi t) + 500\sin(210\pi t), \end{aligned} \tag{18}$$

to simulate excitation over a relatively wide range of frequencies.

The sampling rate is 1000 Hz and the time duration of study is 0.5 s after the load application. The acceleration responses of the structure are calculated using the *Newmark-β* method as the "measured" responses. The sensor location is shown in Figure 7 as red hollow circles. Two damaged scenarios as shown in Table 6 are studied in this section. 10% stiffness reduction is assumed as local damage of selected element in each scenario. Equations (11) and (12) are applied

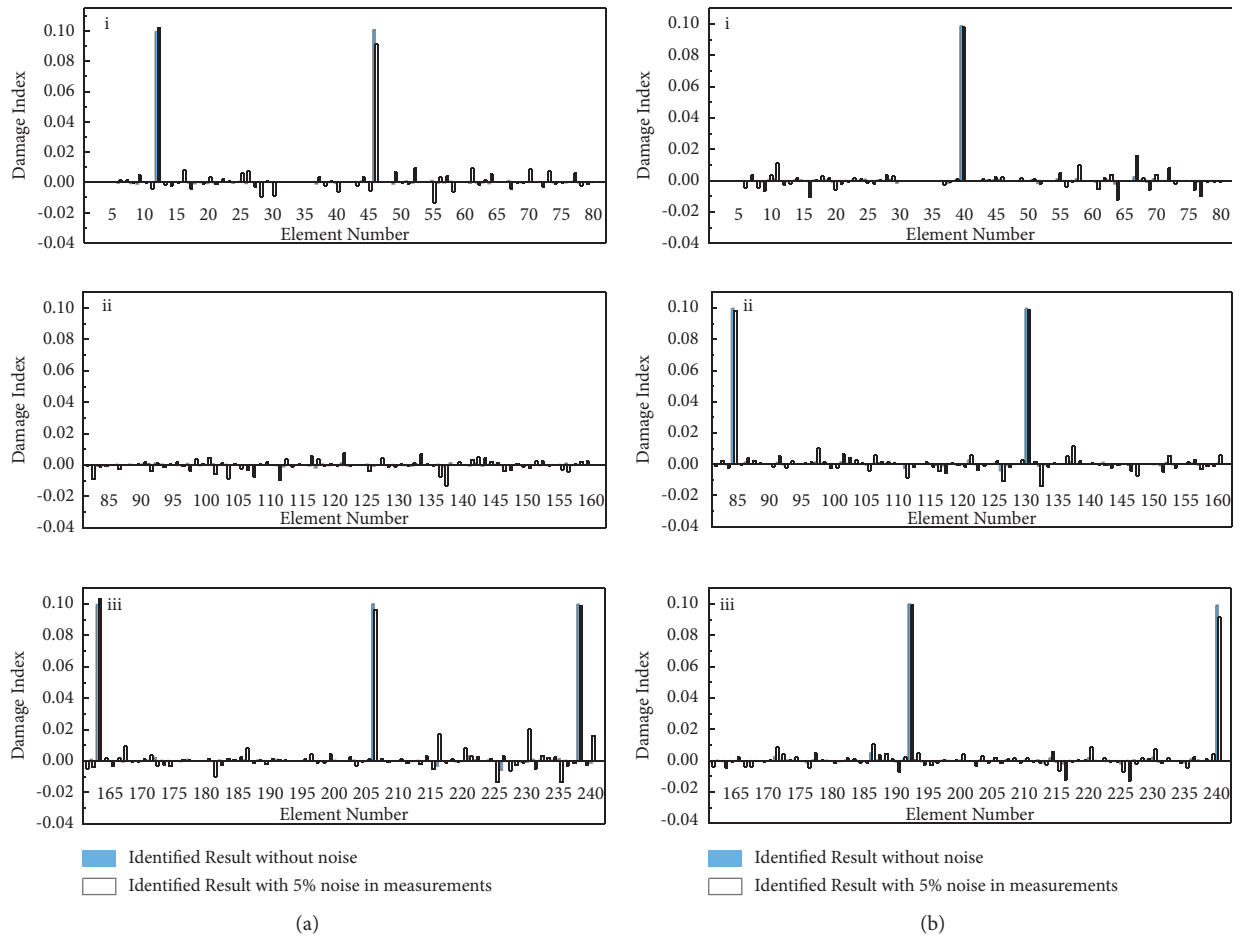


Figure 6: Identified results of global substructural method. (a) Scenario I with (i) part I, (ii) part II, and (iii) part III. (b) Scenario II with (i) part I, (ii) part II, and (iii) part III.

Table 4: Calculated error of the identified results.

Damage scenario	With 5% noise in measurements		Without noise in measurements	
	Mean value (%)	Standard deviation	Mean value (%)	Standard deviation
Scenario I	1.24	1.50	0.18	0.26
Scenario II	1.16	1.34	0.22	0.23

Table 5: Size of sensitivity matrix.

Index of sensitivity matrix	Size of sensitivity matrix
S_{II}	18000×240
S_{11}	9000×120
S_{22}	9000×120

to assess the condition of structure setting $S_{1,3}$ and $S_{3,1}$ as zero.

With the improved global substructural method, the damage identification results of the two scenarios without noise in measurements are shown in Figure 8. The identified results are consistent with the real damage scenarios, and there is almost no error, which indicates that the improved global substructural methodology is correct and accurate. The required computation time on a PC with 3.6 GHz Intel Core i7-4790 CPU and 4 GB memory is shown in Table 7. Compared with that of the global substructural method, the calculation duration of the improved global substructural method is reduced significantly.

In the first-round identification, the number of elements regarded as "damaged elements" is decided by the defined critical value, being related to the efficiency of next step identification. However, if the defined critical value is too large, some damaged elements may be missing. In order to verify that the improved substructural method is still effective even when there are erroneous judgments in the process of filtering, the critical value is defined as 0.09 in this numerical study. After filtering, the damaged element 500 is to be left out

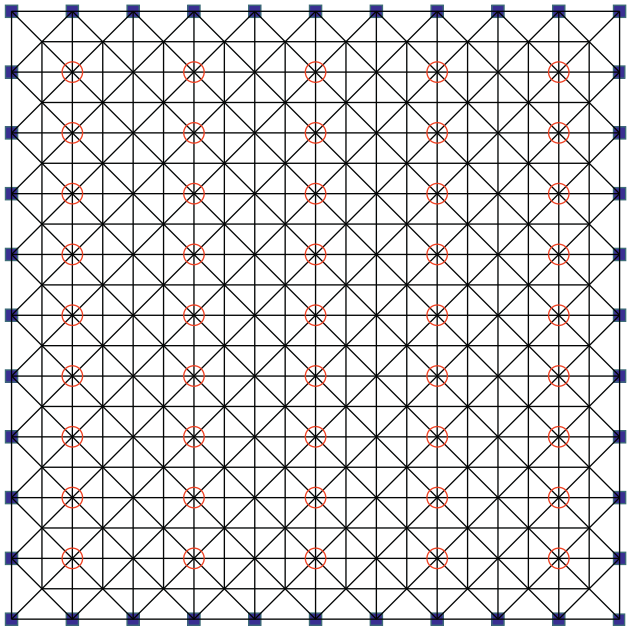


Figure 7: An orthogonal spatial square pyramid grid structure.

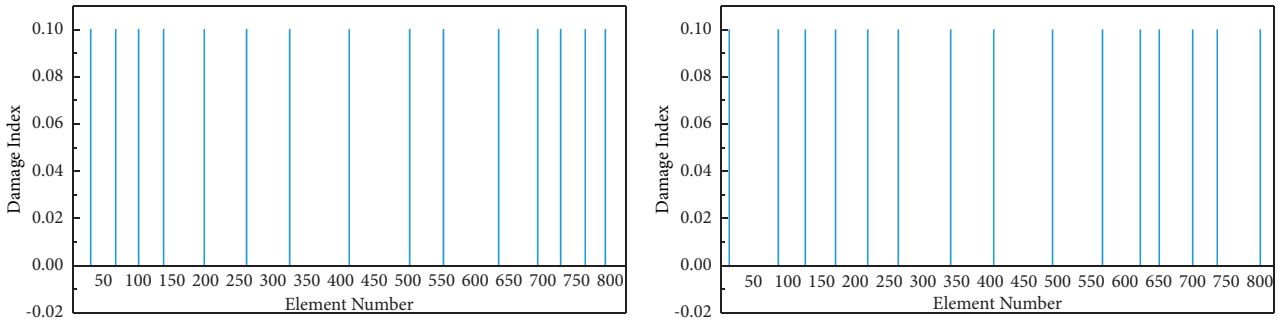


Figure 8: Identified results of improved substructural method. (a) Scenario I. (b) Scenario II.

in Scenario I and the damaged element 736 is to be left out in Scenario II. The identified results of the selected elements of the two scenarios with 5% noise in measurements are shown in Figure 10, and the locations of erroneous judgment are marked by red circles. The final damage identification results after review of the two scenarios are shown in Figure 11. The correct identified results of damaged location indicate that the accuracy of damage identification is not affected.

The mean value and standard deviation of error are shown in Table 8. It is noted that the error between two methods is extremely close, which indicate that the improved global substructural method is stable for error with 5% noise while improving the identification speed. 25% of responses data are obtained for measurement, which is reasonable extent in the practical application. This study shows that the improved global substructural method can be applied for assessing the bigger double-layer lattice space

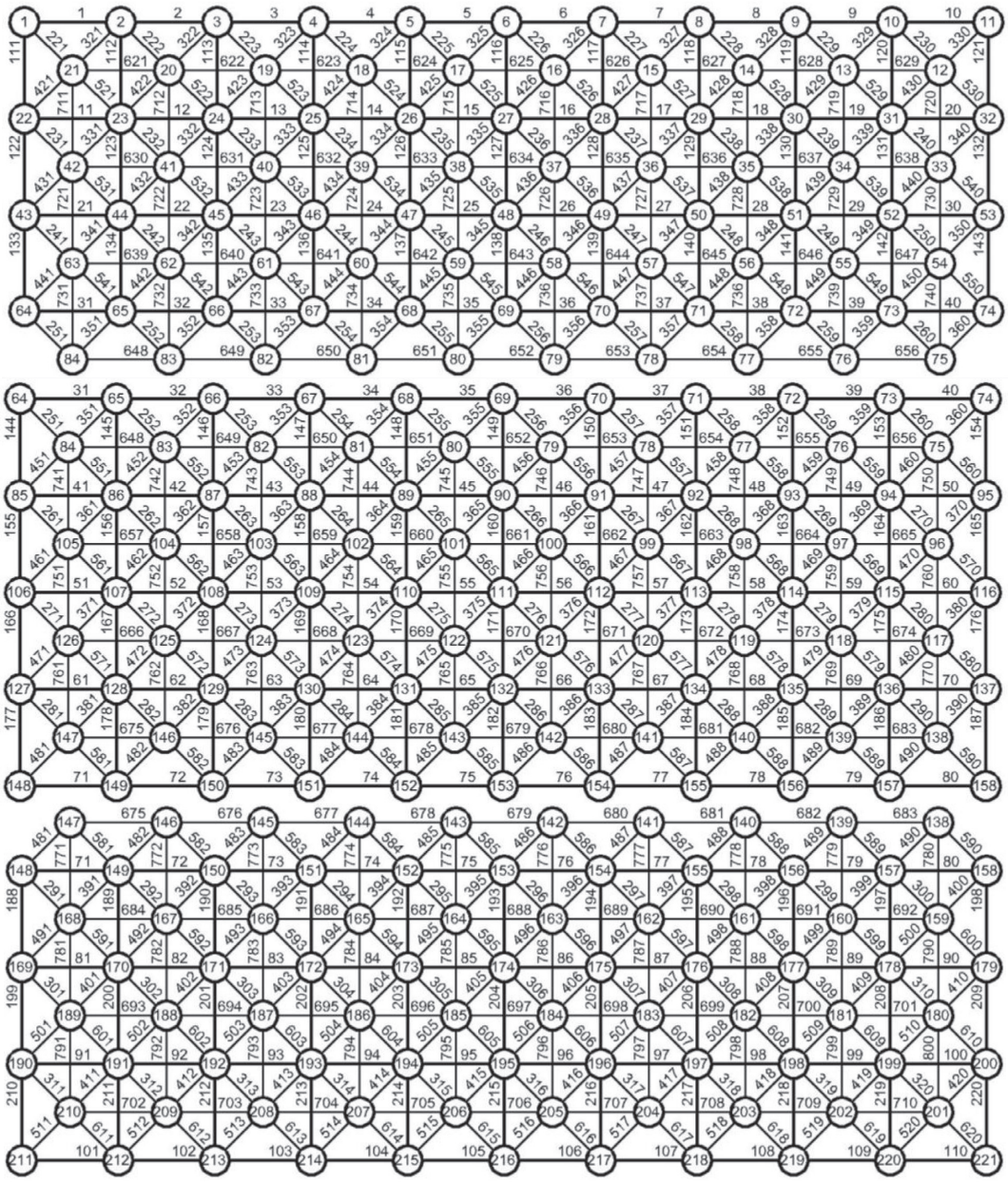


Figure 9: Substructure division and element number system.

Table 6: Damage scenario.

Damage scenario	Damaged element		
	Upper member	Lower member	Web member
Scenario I	27, 64, 98, 135, 195	258, 322, 410, 500, 550	632, 690, 724, 760, 790
Scenario II	12, 85, 125, 170, 218	263, 341, 405, 492, 566	622, 650, 700, 736, 800

Table 7: Required computation time (s) for damage identification.

Damage scenario	Improved global substructural method				Global substructural method
	Filtering	Identifying	Review	Total time	
Scenario I	502	16	407	925	1108
Scenario II	526	21	390	937	1276

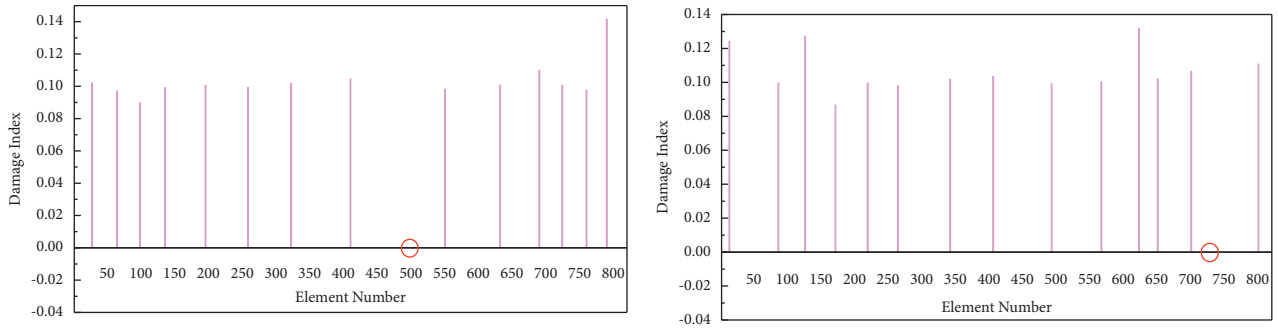


Figure 10: Identified results of selected elements. (a) Scenario I. (b) Scenario II.

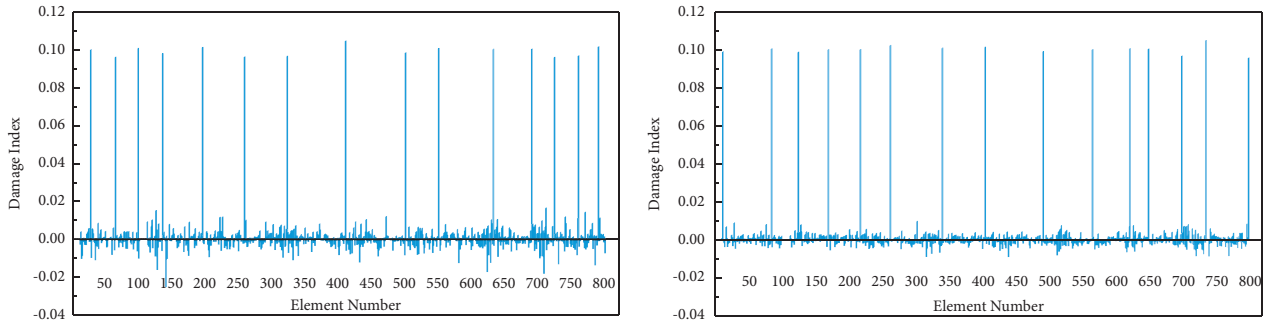


Figure 11: Final identified results after review. (a) Scenario I. (b) Scenario II.

Table 8: Calculated error of the identified result with 5% noise in measurements.

Damage scenario	Improved substructural method		Global substructural method	
	Mean value	Standard deviation	Mean value	Standard deviation
Scenario I	0.61	0.67	0.71	0.79
Scenario II	0.44	0.42	0.50	0.51

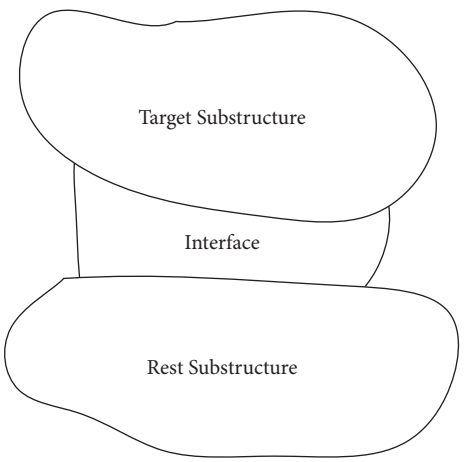


Figure 12: Substructure division.

structure rapidly with dividing substructure to avoid data explosion.

4. Conclusions and Discussions

The proposed global substructural damage identification method can assess all the substructures without knowing the conditions of the other parts of the entire structure, so it is a practical method for the health monitoring of large-span spatial structure. The numerical simulation of a double-layer lattice space structure with local damage distributed in different substructures verifies that the new method is effective in identifying the damage of large-span spatial structure by the incomplete dynamic response information. The small damage in multiple members can be accurately assessed without noise, and the damaged location can be correctly identified with 5% noise in measurements. The divided substructure has less model and smaller sensitivity matrix than the full structure to avoid data explosion, and the appropriate number of sensors ensures the practical feasibility.

The improved global substructural method is proposed to shorten the required computation time, and a larger orthogonal spatial square pyramid grid structure illustrates and verifies this improved global method. The iterative algorithm based on the Gauss elimination of identification is simplified by ignoring a few substructural sensitivities included in the global sensitivity matrix. However, the calculated error will increase with the number of ignored sensitivity matrices growing. The relation between the error and the number of the ignored sensitivity matrices should be further studied. If the number of divided substructures is excessive and the Gauss elimination method suffers difficulty, the Jacobi iteration can be adopted.

Appendix

A. Explicit Interface Force Identification Method

The displacement, velocity and acceleration of the substructure at time t_{i+1} can be obtained from Equations (A.1) to (A.3) similar to a full structure [27] as

$$(x_{\text{sub}})_{i+1} = A_0 L_{\text{sub}} (P_{\text{sub}})_{i+1} + A_d (x_{\text{sub}})_i + A_v (\dot{x}_{\text{sub}})_i + A_a (\ddot{x}_{\text{sub}})_i, \tag{A.1}$$

$$(\dot{x}_{\text{sub}})_{i+1} = B_0 L_{\text{sub}} (P_{\text{sub}})_{i+1} + B_d (x_{\text{sub}})_i + B_v (\dot{x}_{\text{sub}})_i + B_a (\ddot{x}_{\text{sub}})_i, \tag{A.2}$$

$$(\ddot{x}_{\text{sub}})_{i+1} = C_0 L_{\text{sub}} (P_{\text{sub}})_{i+1} + C_d (x_{\text{sub}})_i + C_v (\dot{x}_{\text{sub}})_i + C_a (\ddot{x}_{\text{sub}})_i, \tag{A.3}$$

where

$$\widehat{K}_{\text{sub}} = K_{\text{sub}} + \frac{1}{\beta \Delta t^2} M_{\text{sub}} + \frac{\gamma}{\beta \Delta t} C_{\text{sub}},$$

$$A_0 = K_{\text{sub}}^{-1}; \qquad A_d = \widehat{K}_{\text{sub}}^{-1} \left[\frac{1}{\beta \Delta t^2} M_{\text{sub}} + \frac{\gamma}{\beta \Delta t} C_{\text{sub}} \right],$$

$$A_v = \widehat{K}_{\text{sub}}^{-1} \left[\frac{1}{\beta \Delta t} M_{\text{sub}} + \left(\frac{\gamma}{\beta} - 1 \right) C_{\text{sub}} \right]; \qquad A_a = \widehat{K}_{\text{sub}}^{-1} \left[\left(\frac{1}{2\beta} - 1 \right) M_{\text{sub}} + \frac{\Delta t}{2} \left(\frac{\gamma}{\beta} - 2 \right) C_{\text{sub}} \right],$$

$$B_0 = \frac{\gamma}{\beta \Delta t} \widehat{K}_{\text{sub}}^{-1}; \qquad B_d = -\frac{\gamma}{\beta \Delta t} \widehat{K}_{\text{sub}}^{-1} K_{\text{sub}}, \tag{A.4}$$

$$B_v = \frac{\gamma}{\beta \Delta t} \widehat{K}_{\text{sub}}^{-1} \left[\left(\frac{\beta \Delta t}{\gamma} - \Delta t \right) K_{\text{sub}} + \frac{1}{\gamma \Delta t} M_{\text{sub}} \right]; \quad B_a = \frac{\gamma}{\beta \Delta t} \widehat{K}_{\text{sub}}^{-1} \left[\left(\frac{\beta \Delta t^2}{\gamma} - \frac{\Delta t^2}{2} \right) K_{\text{sub}} + \left(\frac{1}{\gamma} - 1 \right) M_{\text{sub}} \right],$$

$$C_0 = \frac{1}{\beta \Delta t^2} \widehat{K}_{\text{sub}}^{-1}; \qquad C_d = -\frac{1}{\beta \Delta t^2} \widehat{K}_{\text{sub}}^{-1} K_{\text{sub}},$$

$$C_v = \frac{1}{\beta \Delta t^2} \widehat{K}_{\text{sub}}^{-1} (C_{\text{sub}} + \Delta t K_{\text{sub}}); \qquad C_a = \frac{1}{\beta \Delta t^2} \widehat{K}^{-1} \left[(\gamma - 1) \Delta t C_{\text{sub}} - \beta \Delta t^2 \left(\frac{1}{2\beta} - 1 \right) K_{\text{sub}} \right].$$

Combining (A.1) to (A.3), we have the displacement, velocity, and acceleration at $(i+1)$th time instant as function of the responses at the ith time instant as

$$\begin{bmatrix} (x_{\text{sub}})_{i+1} \\ (\dot{x}_{\text{sub}})_{i+1} \\ (\ddot{x}_{\text{sub}})_{i+1} \end{bmatrix} = \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix} L_{\text{sub}} (P_{\text{sub}})_{i+1} + \begin{bmatrix} A_d & A_v & A_a \\ B_d & B_v & B_a \\ C_d & C_v & C_a \end{bmatrix} \begin{bmatrix} (x_{\text{sub}})_i \\ (\dot{x}_{\text{sub}})_i \\ (\ddot{x}_{\text{sub}})_i \end{bmatrix}. \tag{A.5}$$

Equation (A.5) is formulated from the standard *Newmark-β* method for the forward dynamic analysis of a substructure. It can be further rewritten into a general recursive relation as

$$\begin{aligned} \begin{bmatrix} (x_{\text{sub}})_i \\ (\dot{x}_{\text{sub}})_i \\ (\ddot{x}_{\text{sub}})_i \end{bmatrix} &= \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix} L_{\text{sub}} (P_{\text{sub}})_i + \begin{bmatrix} A_d & A_v & A_a \\ B_d & B_v & B_a \\ C_d & C_v & C_a \end{bmatrix} \begin{bmatrix} (x_{\text{sub}})_{i-1} \\ (\dot{x}_{\text{sub}})_{i-1} \\ (\ddot{x}_{\text{sub}})_{i-1} \end{bmatrix}, \\ \begin{bmatrix} (x_{\text{sub}})_{i-1} \\ (\dot{x}_{\text{sub}})_{i-1} \\ (\ddot{x}_{\text{sub}})_{i-1} \end{bmatrix} &= \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix} L_{\text{sub}} (P_{\text{sub}})_{i-1} + \begin{bmatrix} A_d & A_v & A_a \\ B_d & B_v & B_a \\ C_d & C_v & C_a \end{bmatrix} \begin{bmatrix} (x_{\text{sub}})_{i-2} \\ (\dot{x}_{\text{sub}})_{i-2} \\ (\ddot{x}_{\text{sub}})_{i-2} \end{bmatrix}, \\ &\vdots, \\ \begin{bmatrix} (x_{\text{sub}})_1 \\ (\dot{x}_{\text{sub}})_1 \\ (\ddot{x}_{\text{sub}})_1 \end{bmatrix} &= \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix} L_{\text{sub}} (P_{\text{sub}})_1 + \begin{bmatrix} A_d & A_v & A_a \\ B_d & B_v & B_a \\ C_d & C_v & C_a \end{bmatrix} \begin{bmatrix} (x_{\text{sub}})_0 \\ (\dot{x}_{\text{sub}})_0 \\ (\ddot{x}_{\text{sub}})_0 \end{bmatrix}, \end{aligned} \tag{A.6}$$

and the response at time t_i can be written as

$$\begin{bmatrix} (x_{\text{sub}})_i \\ (\dot{x}_{\text{sub}})_i \\ (\ddot{x}_{\text{sub}})_i \end{bmatrix} = \sum_{j=0}^{i-1} \begin{bmatrix} A_d & A_v & A_a \\ B_d & B_v & B_a \\ C_d & C_v & C_a \end{bmatrix}^j \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix} L_{\text{sub}} (P_{\text{sub}})_{i-j} + \begin{bmatrix} A_d & A_v & A_a \\ B_d & B_v & B_a \\ C_d & C_v & C_a \end{bmatrix}^i \begin{bmatrix} (x_{\text{sub}})_0 \\ (\dot{x}_{\text{sub}})_0 \\ (\ddot{x}_{\text{sub}})_0 \end{bmatrix}, \tag{A.7}$$

where both indices i and j denote a power of the corresponding matrices.

The measurement matrix y_{sub} represents the output of the structural system and it can be assembled from the measurements with

$$y_{\text{sub}} = R_a \ddot{x}_{\text{sub}} + R_v \dot{x}_{\text{sub}} + R_d x_{\text{sub}}, \tag{A.8}$$

where R_a, R_v and $R_d \in R^{ns \times N}$ are the output influence matrices for the measured acceleration, velocity and displacement respectively, ns is the number of the measured responses and N is the number of DOFs of the selected substructure.

Vector y_{sub} can be rewritten as

$$y_{\text{sub}} = \begin{bmatrix} R_d & R_v & R_a \end{bmatrix} \begin{Bmatrix} x_{\text{sub}} \\ \dot{x}_{\text{sub}} \\ \ddot{x}_{\text{sub}} \end{Bmatrix}. \tag{A.9}$$

Let $R = \begin{bmatrix} R_d & R_v & R_a \end{bmatrix}$ and substituting Equations (A.7) and (A.9) can be rewritten into the following discrete equation as

$$y_{\text{sub}} (t_i) = \sum_{j=0}^{i-1} R \begin{bmatrix} A_d & A_v & A_a \\ B_d & B_v & B_a \\ C_d & C_v & C_a \end{bmatrix}^j \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix} L_{\text{sub}} (P_{\text{sub}})_{i-j} + R \begin{bmatrix} A_d & A_v & A_a \\ B_d & B_v & B_a \\ C_d & C_v & C_a \end{bmatrix}^i \begin{bmatrix} (x_{\text{sub}})_0 \\ (\dot{x}_{\text{sub}})_0 \\ (\ddot{x}_{\text{sub}})_0 \end{bmatrix}. \tag{A.10}$$

Assuming zero initial response of the structure and let,

$$H_k = R\begin{bmatrix} A_d & A_v & A_a \\ B_d & B_v & B_a \\ C_d & C_v & C_a \end{bmatrix}^k \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix}. \quad (A.11)$$

Equation (A.10) can be rewritten into a matrix convolution relation from t_1 to t_{nt} time instants as

$$Y_{\text{sub}} = (H_{\text{sub}})_L P_{\text{sub}}, \quad (A.12)$$

where

$$Y_{\text{sub}} = \begin{Bmatrix} y_{\text{sub}}(t_1) \\ y_{\text{sub}}(t_2) \\ \vdots \\ y_{\text{sub}}(t_{nt}) \end{Bmatrix},$$

$$(H_{\text{sub}})_L = \begin{bmatrix} H_0 L_{\text{sub}} & 0 & \cdots & 0 \\ H_1 L_{\text{sub}} & H_0 L_{\text{sub}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ H_{nt-1} L_{\text{sub}} & H_{nt-2} L_{\text{sub}} & \cdots & H_0 L_{\text{sub}} \end{bmatrix}, \quad (A.13)$$

$$P_{\text{sub}} = \begin{Bmatrix} P_{\text{sub}}(t_1) \\ P_{\text{sub}}(t_2) \\ \vdots \\ P_{sub}(t_{nt}) \end{Bmatrix}.$$

The inverse problem in Equation (A.12) is ill-posed, and the interface forces are therefore identified through the Tikhonov regularization technique by minimizing the following objective function.

$$J(P_{\text{sub}}, \lambda) = \left\| (H_{\text{sub}})_L P_{\text{sub}} - Y_{\text{sub}} \right\|^2 + \lambda \left\| P_{\text{sub}} \right\|^2, \quad (A.14)$$

where λ is the regularization parameter obtained using the *L*-curve method.

B. Substructural Damage Identification Method

The equation of motion of a damped linear structure with multiple DOFs can be written as

$$M\ddot{x} + C\dot{x} + Kx = LP(t), \quad (B.1)$$

where M, C and K are the mass, damping and stiffness matrices of the structural system, respectively. $P(t)$ is the vector of external forces on the structure and L is the mapping matrix for the external forces. $\ddot{x}$, $\dot{x}$ and x are the vectors of acceleration, velocity and displacement responses, respectively. The structure is assumed to exhibit Rayleigh damping for discussion as

$$C = a_1 M + a_2 K, \quad (B.2)$$

where a_1 and a_2 are the damping coefficients.

A structure can be divided into several substructures and a target substructure can be selected for the assessment. Based on the sub-division of a structure as shown in Figure 12, Equation (B.1) can be rewritten as

$$\begin{bmatrix} M_{rr} & M_{ri} & 0 \\ M_{ir} & M_{ii} & M_{is} \\ 0 & M_{si} & M_{ss} \end{bmatrix} \{\ddot{x}_r \ddot{x}_i \ddot{x}_s\} + \begin{bmatrix} C_{rr} & C_{ri} & 0 \\ C_{ir} & C_{ii} & C_{is} \\ 0 & C_{si} & C_{ss} \end{bmatrix} \begin{Bmatrix} \dot{x}_r \\ \dot{x}_i \\ \dot{x}_s \end{Bmatrix} + \begin{bmatrix} K_{rr} & K_{ri} & 0 \\ K_{ir} & K_{ii} & K_{is} \\ 0 & K_{si} & K_{ss} \end{bmatrix} \begin{Bmatrix} x_r \\ x_i \\ x_s \end{Bmatrix} = \begin{Bmatrix} L_r P_r \\ L_i P_i \\ L_s P_s \end{Bmatrix}, \quad (B.3)$$

where the subscripts *r*, *i* and *s* denote the DOFs of the target substructure, the interface DOFs between the substructures and the DOFs of the rest substructure.

The following equation on DOFs *r* and *i* of the structure can be extracted from (B.3) as

$$\begin{bmatrix} M_{rr} & M_{ri} \\ M_{ir} & M_{ii} \end{bmatrix} \{\ddot{x}_r \ddot{x}_i\} + \begin{bmatrix} C_{rr} & C_{ri} \\ C_{ir} & C_{ii} \end{bmatrix} \begin{Bmatrix} \dot{x}_r \\ \dot{x}_i \end{Bmatrix} + \begin{bmatrix} K_{rr} & K_{ri} \\ K_{ir} & K_{ii} \end{bmatrix} \begin{Bmatrix} x_r \\ x_i \end{Bmatrix} = \begin{Bmatrix} L_r P_r \\ L_i P_i + P'_s \end{Bmatrix}, \quad (B.4)$$

where $P'_s = -M_{is}\ddot{x}_s - C_{is}\dot{x}_s - K_{is}x_s$ is the set of interface forces. The presence of other parts of the structure is represented by the interfacing forces acting on the target substructure. Accurate knowledge of these forces is therefore a requirement for a successful assessment.

The target substructure is selected for study as an individual structure, and the interface forces can be considered as a set of external forces acting on the target substructure. The subscript '*sub*' is adopted to represent the substructure. Let,

$$
\begin{aligned}
M_{\text{sub}} &= \begin{bmatrix} M_{rr} & M_{ri} \\ M_{ir} & M_{ii} \end{bmatrix}, \\
C_{\text{sub}} &= \begin{bmatrix} C_{rr} & C_{ri} \\ C_{ir} & C_{ii} \end{bmatrix}, \\
K_{\text{sub}} &= \begin{bmatrix} K_{rr} & K_{ri} \\ K_{ir} & K_{ii} \end{bmatrix}, \\
\ddot{x}_{\text{sub}} &= \{ \ddot{x}_r \ \ddot{x}_i \}, \\
\dot{x}_{\text{sub}} &= \begin{Bmatrix} \dot{x}_r \\ \dot{x}_i \end{Bmatrix}, \\
x_{\text{sub}} &= \begin{Bmatrix} x_r \\ x_i \end{Bmatrix}, \\
L_{\text{sub}} &= \begin{bmatrix} L_r & 0 \\ 0 & I \end{bmatrix}, \\
P_{\text{sub}} &= \begin{Bmatrix} P_r \\ P_s' \end{Bmatrix},
\end{aligned}
\tag{B.5}
$$

then the equation of motion of the substructure in Equation (B.4) can be written as

$$
M_{\text{sub}}\ddot{x}_{\text{sub}} + C_{\text{sub}}\dot{x}_{\text{sub}} + K_{\text{sub}}x_{\text{sub}} = L_{\text{sub}}P_{\text{sub}}. \tag{B.6}
$$

Since matrices M_{sub}, C_{sub} and K_{sub} are positive semi-definite similar to the mass, damping and stiffness matrices of the whole structure, the substructural force identification can also be performed similar to that for a full structure, and the external forces identification method is listed in AppendixA.

Assuming the local damage extent of the mth element in the target substructure as α_m, the change of the stiffness matrix can be described as

$$
\Delta K_{\text{sub}} = \sum_m \alpha_m (K_{\text{sub}})_m, \tag{B.7}
$$

where $(K_{\text{sub}})_m$ is the stiffness matrix of the mth element in the target substructure. Performing differentiation to both sides of Equation (B.6) with respect to the parameter α_m, we have

$$
M_{\text{sub}}\frac{\partial \ddot{x}_{\text{sub}}}{\partial \alpha_m} + C_{sub}\frac{\partial \dot{x}_{\text{sub}}}{\partial \alpha_m} + K_{\text{sub}}\frac{\partial x_{\text{sub}}}{\partial \alpha_m} = -\frac{\partial K_{\text{sub}}}{\partial \alpha_m}x_{\text{sub}} - a_1\frac{\partial K_{\text{sub}}}{\partial \alpha_m}\dot{x}_{\text{sub}} + L_{\text{sub}}\frac{\partial P_{\text{sub}}}{\partial \alpha_m}. \tag{B.8}
$$

Equation (B.8) is of the same form as similar equations for the response sensitivity of a full structure [8] except one extra term (the third term) on the right-hand side. When only the first two terms on the right are included, the obtained response sensitivity is for the substructure with an non-varying interface force. Equation (B.8) is solved again by keeping the third term on the right to get the response sensitivity derived from the interface force sensitivity alone. The response sensitivity derived from the interface force sensitivity is noted to be significant [26] and it cannot be ignored in the calculation of the response sensitivity of the substructure.

Vector $y_{\text{sub}}^{\alpha m}$ represents the response sensitivity vector of the substructural system and similar to Equation (A.8) it can be assembled as

$$
y_{\text{sub}}^{\alpha m} = R_a\frac{\partial \ddot{x}_{\text{sub}}}{\partial \alpha_m} + R_v\frac{\partial \dot{x}_{\text{sub}}}{\partial \alpha_m} + R_d\frac{\partial x_{\text{sub}}}{\partial \alpha_m}. \tag{B.9}
$$

Let,

$$
L_{\text{sub}}^{\alpha m} = \begin{bmatrix} -\dfrac{\partial K_{\text{sub}}}{\partial \alpha_m} & -a_1\dfrac{\partial K_{\text{sub}}}{\partial \alpha_m} & 0 \end{bmatrix}. \tag{B.10}
$$

Combining Equations (B.8) and (B.9) and assuming zero initial conditions, similar to Equations (A.1) to (A.10) the following solution of the sensitivity can be obtained.

$$
y_{\text{sub}}^{\alpha m}(t_i) = \sum_{j=0}^{i-1} R \begin{bmatrix} A_d & A_v & A_a \\ B_d & B_v & B_a \\ C_d & C_v & C_a \end{bmatrix}^j \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix} \left(L_{\text{sub}}^{\alpha m} \begin{Bmatrix} x_{\text{sub}} \\ \dot{x}_{\text{sub}} \\ \ddot{x}_{\text{sub}} \end{Bmatrix}_{i-j} + L_{\text{sub}} \left(\frac{\partial P_{\text{sub}}}{\partial \alpha_m} \right)_{i-j} \right), \tag{B.11}
$$

Based on Equation (A.12), the sensitivity of the interface force can be defined as

$$
\frac{\partial P_{\text{sub}}}{\partial \alpha_m} = \left[(H_{\text{sub}})_L \right]^{-1} \frac{\partial (H_{\text{sub}})_L}{\partial \alpha_m} \left[(H_{\text{sub}})_L \right]^{-1} Y_{\text{sub}}, \tag{B.12}
$$

and identified based on the explicit *Newmark- β* method [27],

$$
\begin{Bmatrix} x_{\text{sub}} \\ \dot{x}_{\text{sub}} \\ \ddot{x}_{\text{sub}} \end{Bmatrix}_i = \sum_{j=0}^{i-1} \begin{bmatrix} A_d & A_v & A_a \\ B_d & B_v & B_a \\ C_d & C_v & C_a \end{bmatrix}^j \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix} L_{\text{sub}} (P_{\text{sub}})_{i-j}. \tag{B.13}
$$

Similar to Equations (A.12) and (A.13), let

$$H_k^w = \begin{bmatrix} A_d & A_v & A_a \\ B_d & B_v & B_a \\ C_d & C_v & C_a \end{bmatrix}^k \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix},$$

$$(H_{\text{sub}}^w)_L = \begin{bmatrix} H_0^w L_{\text{sub}} & 0 & \cdots & 0 \\ H_1^w L_{\text{sub}} & H_0^w L_{\text{sub}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ H_{nt-1}^w L_{\text{sub}} & H_{nt-2}^w L_{\text{sub}} & \cdots & H_0^w L_{\text{sub}} \end{bmatrix}, \quad (B.14)$$

$$(H_{\text{sub}}^{\alpha m})_L = \begin{bmatrix} H_0 L_{\text{sub}}^{\alpha m} & 0 & \cdots & 0 \\ H_1 L_{\text{sub}}^{\alpha m} & H_0 L_{\text{sub}}^{\alpha m} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ H_{nt-1} L_{\text{sub}}^{\alpha m} & H_{nt-2} L_{\text{sub}}^{\alpha m} & \cdots & H_0 L_{\text{sub}}^{\alpha m} \end{bmatrix}.$$

The response sensitivity considering the interface force effect can be obtained as,

$$Y_{\text{sub}}^{\alpha m} = \left((H_{\text{sub}}^{\alpha m})_L (H_{\text{sub}}^w)_L + \frac{\partial (H_{\text{sub}})_L}{\partial \alpha_m} \right) [(H_{\text{sub}})_L]^{-1} Y_{\text{sub}}, \quad (B.15)$$

and the sensitivity matrix $\mathbf{S}_{sub}$ for the substructure with NE elements can be defined as

$$S_{\text{sub}} = \begin{bmatrix} Y_{\text{sub}}^{\alpha 1} & Y_{\text{sub}}^{\alpha 2} & \cdots & Y_{\text{sub}}^{\alpha NE} \end{bmatrix}. \quad (B.16)$$

Data Availability

Any reader interested in the data can contact the corresponding author through e-mail.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This study was jointly sponsored by National Key R&D Programs of China (Grant nos. 2017YFC1500601 and 2017YFC1500602); Scientific Research Fund of Institute of Engineering Mechanics, China Earthquake Administration (Grant Nos. 2021B01 and 2021EEEVL0308); and Program for Innovative Research Team in China Earthquake Administration.

References

[1] G. B. Nie, J. W. Dai, C. X. Zhang, and X. D. Zhi, "Failure patterns of large span space structures in Lushan earthquake and numerical simulation," *China Civil Engineering Journal*, vol. 48, no. 4, pp. 1–6, 2015.

[2] O. S. Salawu, "Detection of structural damage through changes in frequency: a review," *Engineering Structures*, vol. 19, no. 9, pp. 718–723, 1997.

[3] O. Vera and N. Chouw, "Damage identification of unreinforced masonry panels using vibration-based techniques," *Shock and Vibration*, vol. 2017, Article ID 9161025, 14 pages, 2017.

[4] S. W. Doebling, C. R. Farrar, and M. B. Prime, "A summary review of vibration-based damage identification methods," *The Shock and Vibration Digest*, vol. 30, no. 2, pp. 91–105, 1998.

[5] X. Liu, N. A. J. Lieven, and P. J. Escamilla-Ambrosio, "Frequency response function shape-based methods for structural damage localization," *Mechanical Systems and Signal Processing*, vol. 23, no. 2009, pp. 1243–1259, 2009.

[6] O. Huth, G. Feltrin, J. Maeck, and K. Nedim, "Damage identification using modal data: experiences on a prestressed concrete bridge," *Journal of Structural Engineering*, vol. 131, no. 12, pp. 1898–1910, 2005.

[7] X. Wang, K. Liu, H. Liu, and Y. He, "Damping identificationwith acceleration measurements based on sensitivity en-hancement method," *Shock and Vibration*, vol. 2018, Article ID 6476783, 16 pages, 2018.

[8] K. Liu, S. S. Law, Y. Xia, and X. Q. Zhu, "Singular spectrum analysis for enhancing the sensitivity in structural damage detection," *Journal of Sound and Vibration*, vol. 333, no. 2014, pp. 392–417, 2014.

[9] W. Ji, Y. Song, and B. Liang, "Damage diagnosis in truss by modal analysis," *Journal of Shenyang Jianzhu University (Natural Science)*, vol. 26, no. 2, pp. 232–237, 2010.

[10] W. Ji, Y. Song, B. Liang, and Y. Song, "Numerical simulation research of for damage diagnosis of truss structure by ANSYS," *Journal of Wuhan University of Technology*, vol. 30, no. 2, pp. 285–288, 2006.

[11] Y. Song, L. Zhang, and Z. Liu, "Damage identification of space truss with incomplete measured data," *China Civil Engineering Journal*, vol. 42, no. 1, pp. 10–14, 2009.

[12] W. Gao, W. Liu, and C. Qian, "Damage detection of space truss using residual modal force and modal strain energy," *Engineering Mechanics*, vol. 24, no. 5, pp. 93–100, 2007.

[13] J. Wu and Y. Zhang, "Damage identification of grid structures based on BP neural networks and experimental study," *World Earthquake Engineering*, vol. 21, no. 4, pp. 71–76, 2005.

[14] C. Liu, Y. Zhang, and J. Wu, "Substructure-oriented damage localization method of the spatial lattice structure," *Building Structure*, vol. 44, no. 6, pp. 79–84, 2014.

[15] H. He and W. Yan, "Structural damage detection with wavelet support vector machine:Introduction and applications," *Structural Control and Health Monitoring*, vol. 14, no. 1, pp. 162–176, 2007.

[16] J. Teng and W. Lu, "Structural damage identification method based on multi-type sensors," *Journal of Southeast University (Natural Science Edition)*, vol. 40, no. 3, pp. 538–542, 2010.

[17] J. Teng and W. Lu, "Damage identification method based on structural dynamic characteristics and strain measurements," *Health Monitoring of Structural and Biological Systems*, vol. 729520, no. 9, 2009.

[18] J. Teng and S. Yao, "A method of earthquake-induced latticed shell damage detection by wavelet packet analysis," *Engineering Mechanics*, vol. 28, no. 9, pp. 183–188, 2011.

[19] K. Yuen, "Substructure identification and health monitoring using noisy response measurements only," *Computer-Aided Civil and Infrastructure Engineering*, vol. 21, no. 4, pp. 280–291, 2006.

[20] K. Yuen and K. Huang, "Real-time substructural identification by boundary force modelling," *Structural Control and Health Monitoring*, vol. 25, no. 4, pp. 1–18, 2018.

[21] K. Yuen and K. Huang, "Identifiability-enhanced Bayesian frequency-domain substructure identification," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 9, pp. 800–812, 2018.

[22] H. A. Jahn, "Improvement of an approximate set of latent roots and modal columns of a matrix by methods akin to those of classical perturbation theory," *Quarterly Journal of Mechanics & Applied Mathematics*, vol. 1, no. 1, pp. 132–144, 1948.

[23] W. Deng, J. Xu, H. Zhao, and Y. Song, "A novel gate resource allocation method using improved PSO-based QEA," *IEEE Transactions on Intelligent Transportation Systems*, vol. 2020, pp. 1–9, 2020.

[24] W. Deng, S. Shang, X. Cai et al., "Quantum differential evolution with cooperative coevolution framework and hybrid mutation strategy for large scale optimization," *Knowledge-Based Systems*, vol. 224, no. 2021, pp. 1–14, Article ID 107080, 2021.

[25] W. Deng, J. Xu, X.-Z. Gao, and H. Zhao, "An enhanced MSIQDE algorithm with novel multiple strategies for global optimization problems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 2020, pp. 1–10, 2020.

[26] K. Liu, S. S. Law, and X. Q. Zhu, "Substructural condition assessment based on force identification and interface force sensitivity," *International Journal of Structural Stability and Dynamics*, vol. 15, no. 2, Article ID 1450046, 2015.

[27] K. Liu, S. S. Law, X. Q. Zhu, and Y. Xia, "Explicit form of an implicit method for inverse force identification," *Journal of Sound and Vibration*, vol. 333, no. 3, pp. 730–744, 2014.

Hindawi
Shock and Vibration
Volume 2021, Article ID 6998232, 14 pages
https://doi.org/10.1155/2021/6998232

Research Article

An Adaptive Noise Reduction Method Based on Improved Dislocation Superposition Method for Abnormal Noise Fault Component of Automotive Engine

Jiaoyi Hou, Pengwei Guo, Aoyu Xu, Dayong Ning, Shengtao Chen, Zengmeng Zhang, Yongjun Gong, Yinglong Chen, Hao Tian, and Hongwei Du

National Center for International Research of Subsea Engineering Technology and Equipment, Dalian Maritime University, Dalian, China

Correspondence should be addressed to Dayong Ning; ningdayongning@163.com

Received 30 June 2021; Revised 28 September 2021; Accepted 30 September 2021; Published 21 October 2021

Academic Editor: Thanh-Phong Dao

The acoustic signal generated by mechanical motion contains the information of its motion state, but when the signal-to-noise ratio (SNR) is low, the accuracy of real-time monitoring mechanical motion state by the acoustic signal is low. This study proposes an adaptive noise reduction method based on the dislocation superposition method (DSM), which can realize the adaptive noise reduction and the extraction of fault a component from the automobile engine abnormal noise signal of low SNR. Firstly, the wavelet coefficients of engine abnormal noise signal are obtained by continuous wavelet transform (CWT), and the fault feature points of the abnormal noise signal in each period are extracted by setting hard threshold function, window function, and feature points extraction algorithm. Then, the signal segments containing fault components are obtained by using the position of feature points to extend the length of the fault component forward and backward, respectively, and Pearson's correlation is calculated by traversal to determine the starting superposition point of each signal segment containing fault components. Finally, the signal segments of the odd group and even group are selected for superposition calculation. When the superposition stop condition is not satisfied, the number of superpositions increased until the stop condition is satisfied, and the superposition signal can be used as a fault component. The experimental results show that, compared with the improved DSM, this method has a good effect on the noise reduction and extraction of fault components of automobile engine cylinder knocking fault, and the effectiveness of this method is verified. This method is used to reduce the noise and extract the fault components of automobile engine cylinder missing fault and knock fault, and good results are obtained.

1. Introduction

At present, cars play an important role in human life and are a necessary condition for human travel. The engine is an important part of the automobile, and its structure is complex [1]. In the working process of the engine, due to wear, fatigue, aging, and other factors, the engine failure may cause serious economic losses and even casualties [2, 3]. When the engine failure produces an abnormal noise signal, the higher SNR of the abnormal noise signal will make engine fault diagnosis easier and more accurate, so it is necessary to improve the SNR of the engine abnormal noise signal [4].

The traditional noise reduction methods mainly include Wiener filtering, spectral subtraction, and minimum mean square deviation, but the effectiveness of noise reduction for the nonstationary signal and short-term transient signal is significantly reduced [5, 6]. The commonly used noise reduction methods mainly include wavelet transform (WT) [7], empirical mode decomposition (EMD) [8], local mean decomposition (LMD) [9], DSM [10], etc.

As a widely used signal processing tool, WT has strong multiresolution analysis ability in the time domain and frequency domain. The wavelet denoising generally uses setting the wavelet coefficient threshold to eliminate the noise information and then uses the inverse WT to

reconstruct the signal from the threshold wavelet coefficient [11]. Li [12] used the wavelet decomposition and reconstruction algorithm to decompose and reduce noise and reconstruct and analyze the spectrum of bearing vibration signal. Experiment shows that the wavelet noise reduction method is very suitable for fault frequency detection of weak vibration signal of the rolling bearing in low SNR cases. Moshrefi et al. [13] used WT and adaptive filtering to denoise the knock signal of the internal combustion engine. This method is applied to real knock signal, which showed superiority compare to previous works and led to a 13.2% improvement in detection accuracy. Kai et al. [14] proposed a wavelet denoising method based on improved threshold function by studying soft threshold function and hard threshold function. By this method, the feature of fault signal is highlighted and the fault diagnosis effect was improved. Wen and You [15] used the wavelet decomposition and reconstruction algorithm to denoise the fault signal of high-speed rolling bearing, and the experimental data showed that most of the noise has been filtered out compared with the original data. Although wavelet denoising has good processing ability for nonstationary signals, the selection of wavelet basis function is very difficult. Different wavelet basis functions have different denoising effects. To solve this problem, Huang et al. [16] proposed the EMD, which can efficiently decompose nonlinear and nonstationary signals without any set of basis functions. Sun et al. [17] used the wavelet threshold noise reduction algorithm based on EMD to solve the problem of complex centrifugal pump vibration signals. The experimental results show that the wavelet soft threshold denoising algorithm based on EMD decomposition has a better noise reduction effect when the centrifugal pump vibration signal is used as the noise reduction object. Ren and Liu [18] proposed an adaptive reduction noise and feature extraction algorithm based on improved EMD and verified the effectiveness and feasibility of the method by simulation signals and examples. Although EMD has been successfully applied in the field of signal noise reduction, it also has its limitations, such as the frequent occurrence of modal mixing. In view of the shortcomings of EMD, Wu and Huang [19] proposed ensemble empirical mode decomposition (EEMD) in 2009. EEMD adds finite amplitude Gauss white noise to the signal, and then the signal with the white noise is decomposed as a whole, to effectively suppress the mode mixing of EMD. However, this method can only reduce modal mixing to some extent.

In recent years, scholars have also proposed some other signal noise reduction methods. Gao et al. [20] proposed a new method based on LMD and wavelet denoising to analyze the signals of bearing outer ring, inner ring, and ball. However, LMD has the phenomenon of modal mixing, which reduces the accuracy of signal decomposition and affects the accuracy of the noise reduction signal. Dayong et al. [21] proposed DSM based on random decrement technique in 2015. Compared with other methods, DSM only calculates in time domain. In the superposition process, the method does not destroy the correlation components in the mixed signal and effectively avoids the modal mixing. In 2019, Dayong et al. [10] used improved DSM to automatically extract engine fault components. Although the improved DSM can automatically extract the fault components of the automotive engine quasiperiodic signal, this method needs to use the pulse number of the encoder to determine the starting superposition point of each quasi-periodic fault signal. In practical applications, automobile engines are not suitable for installing encoders, and encoders are prone to failure when subjected to severe impact, so the improved DSM is limited in some practical applications.

To overcome the above problems, this paper proposes an adaptive noise reduction and extraction method of engine abnormal noise signal fault components based on improved DSM. The essence of this method is the superposition calculation in the time domain, which avoids the modal mixing and can better deal with the acoustic signal with low SNR. The method can adaptively select the starting superposition points, superposition length, and superposition number. By changing the superposition number, the noise reduction degree of fault components can be changed. Compared with the improved DSM, it not only reduces the use of encoders and improves the practicability of DSM but also makes it more convenient to extract fault components.

2. DSM Review

The mathematical expression of the traditional DSM is as follows:

$$\widehat{S}(n) = \frac{1}{K+1} \sum_{K=0}^{K} S(n + KL), \tag{1}$$

where $S(n)$ is the original signal, $\hat{S}(n)$ is the signal processed by DSM (named superposition signal), K is the number of superpositions ($K = 0, 1, 2, \ldots$), and L is the superposition step length (the period of the signal to be processed). Figure 1 shows the graphical description of DSM processing results.

In Figure 1, S is the target signal; N is the interference signal; SN is a mixed signal of S and N; SN_1, SN_2, and SN_3 are superposition signals obtained by superposition of 5, 15, and 20 times, respectively, using equation (1); The superposition step length L is the period of signal S. Compared with signal SN, the component proportion of target signal in signals SN_1, SN_2, and SN_3 increases with the increase of superposition times. On the contrary, the proportion of interference signal decreases with the increase of superposition times [22]. Generally, Pearson's correlation coefficient is used to compare the similarity between the target and the superimposed signals to test the DSM processing effect [10].

3. Adaptive Noise Reduction and Extraction Method of Engine Abnormal Noise Fault Component

3.1. Influence Factors of DSM. Due to the system error of the engine, the actual speed of the engine is slightly changed, which causes the acoustic signal of the engine to be a quasiperiodic signal. Figure 2 shows the schematic of the quasi-periodic signal.

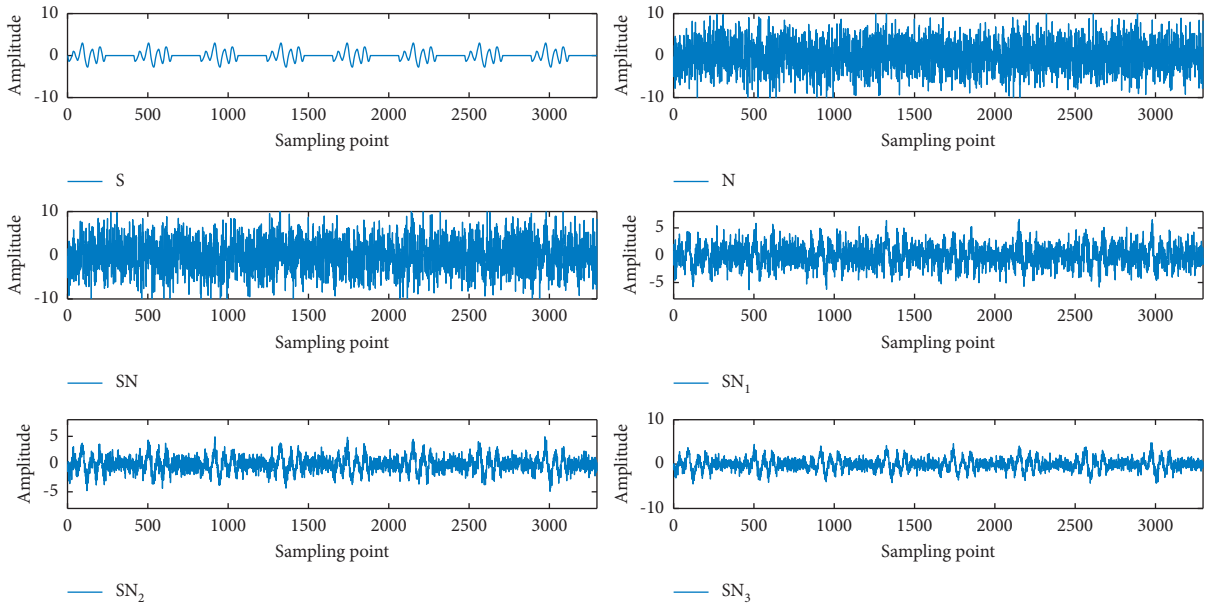


Figure 1: Diagram of DSM. S is the target signal; N is the interference signal; SN is a mixed signal of S and N; SN_1, SN_2, and SN_3 are the superposition signals obtained by superposition of 5, 15, and 20 times, respectively, using equation (1). The superposition step length L is the period of signal S.

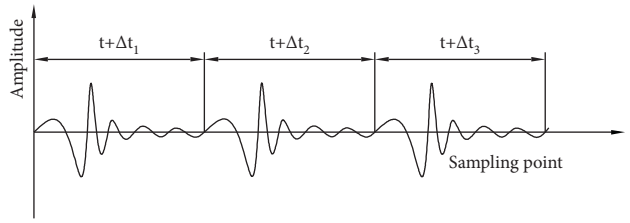


Figure 2: Schematic diagram of the quasiperiodic signal. t is the average period and Δt_1, Δt_2, and Δt_3 are the period deviations, respectively.

When the engine has an abnormal noise impact failure, the fault signal is quasi-periodic, resulting in different superposition step lengths. In the dislocation superposition, it may cause the phase deviation of the impact fault component and even lead to the destruction of the fault component. According to the characteristics of impact failure, it is artificially divided into the "intense change region" and the "stable region," as shown in Figure 3. When the engine impact fault acoustic signal is noise reduction by DSM, due to the less fault energy and information contained in the stable region, to improve the computational efficiency, the intense change region containing more energy and information is separated as the impact fault component [10, 23]. The noise reduction degree of the fault component also depends on the number of superpositions. Theoretically, the more the number of superpositions is, the better the noise reduction effect is. However, the more the number of superpositions is, the longer the consumed time is. When the number of superpositions reaches a certain amount, the noise reduction effect shows a stable trend. Therefore, the adaptive noise reduction process of engine impact fault acoustic signal is a process of automatically finding the starting superposition points of impact fault acoustic

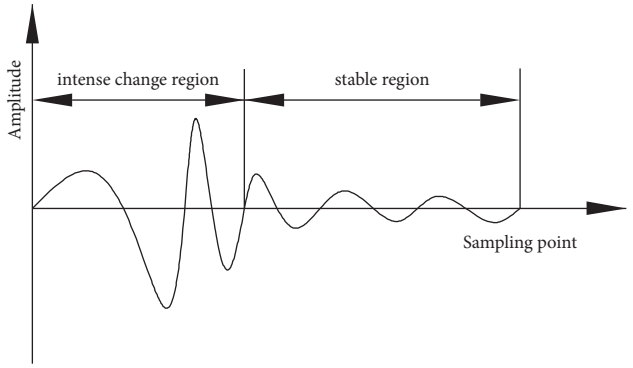


Figure 3: The "intense change region" and the "stable region" of impact fault signal.

components, the length of the intense change region, and the number of superpositions.

3.2. Adaptive Noise Reduction and Extraction Method Based on Improved DSM and CWT. Based on CWT and improved DSM, an adaptive noise reduction method for automobile engine fault acoustic signal is proposed. Figure 4 shows the flowchart of the method. The details are as follows.

3.2.1. Extraction of Feature Point Location of Each Period Impact Fault Component. The "intense change region" has the characteristics of transient, periodic, and large energy which is not easily submerged by background noise [23]. The mother wavelet which is similar to the waveform of the impact fault component is selected for CWT to obtain wavelet coefficients, and the wavelet coefficients are hard thresholding processes. Finally, the window function and feature points extraction algorithm are used to extract the

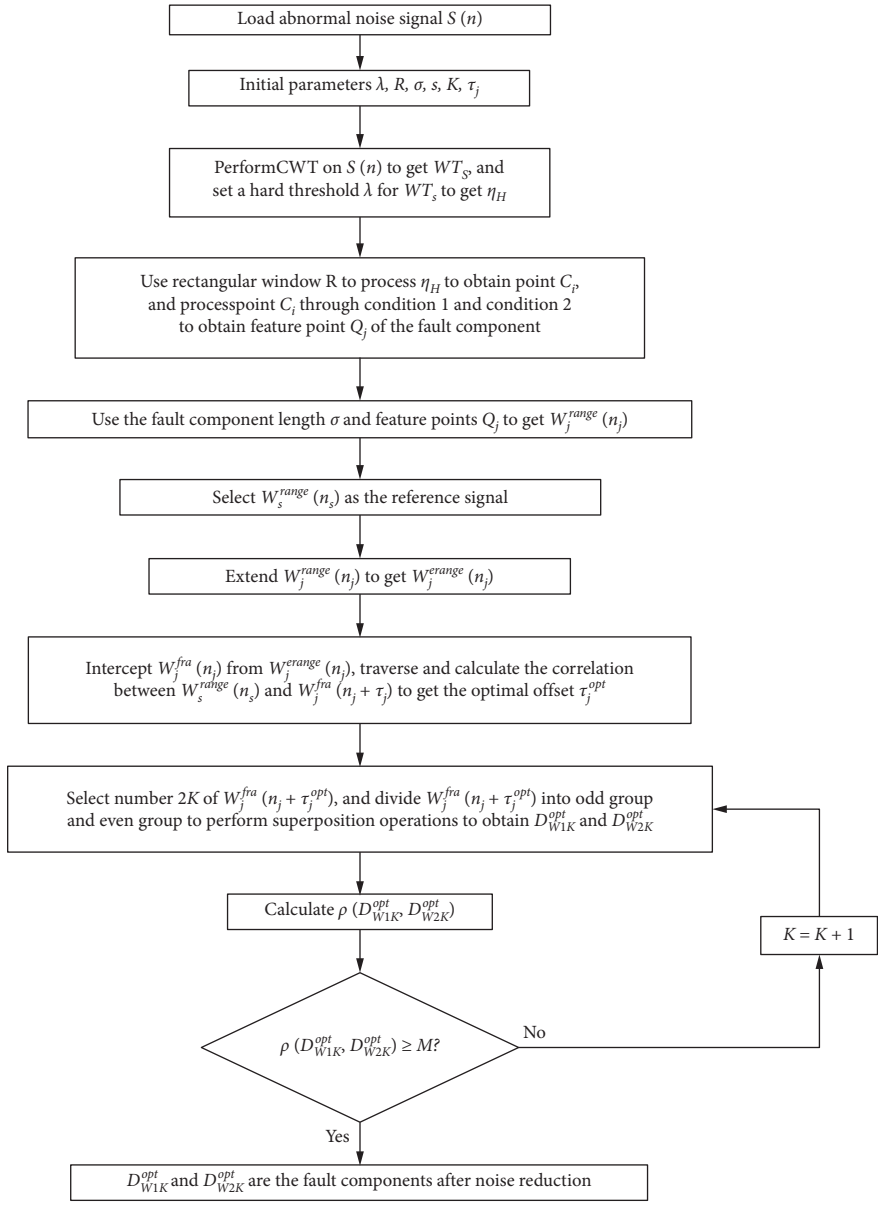


Figure 4: The flowchart of the adaptive noise reduction method of automobile engine acoustic signal based on improved DSM and CWT.

feature points location of the impact fault component by using the threshold wavelet coefficient. The specific steps are as follows:

(1) Load the original fault signal $S(n)$, and select the mother wavelet similar to the intense change region of the fault signal to perform CWT on the signal $S(n)$ to obtain the wavelet coefficients WT_S.

(2) Eliminate the smaller wavelet coefficients containing interference components in the fault signal through the hard threshold function and retain the larger wavelet coefficients containing the impact fault components. The mathematical model of the hard threshold function is

$$\eta_H(\mathrm{WT}_S, \lambda) = \begin{cases} \mathrm{WT}_S, & |(\mathrm{WT}_S)| \geq \lambda, \\ 0, & |(\mathrm{WT}_S)| < \lambda, \end{cases} \quad (2)$$

where $\eta_H(\mathrm{WT}_S, \lambda)$ is the processed wavelet coefficient by hard threshold function, denoted as η_H; WT_S are wavelet coefficients; and λ is the size of the set hard threshold.

(3) Extract the feature point location of each period fault component.

Two diagonally paired time-frequency points (t_1, f_1), (t_2, f_2) are selected to form a rectangular window R, and a time-frequency block is selected from the frame of the thresholding wavelet time-frequency diagram by using the rectangular window R, which is denoted as $\eta_H(a_0, b_0)$, where $t_1 < t_2$, $f_1 < f_2$, $a_0 = t_1$, $t_1 + 1/f_s, \ldots, t_2$, $b_0 = f_1, f_1 + f_s/2a, \ldots, f_2$, f_s is the sampling frequency, and a is the scale factor of CWT. The maximum point of $|\eta_H|$ in time-frequency block $\eta_H(a_0, b_0)$ is taken, which is denoted as $C_0(t_0^{\mathrm{spe}}, f_0^{\mathrm{spe}})$. Keep the frequency $[f_1, f_1 + f_s/2a, \ldots, f_2]$ unchanged, make the

rectangular window R move one time every t_2 along the time axis direction, and constantly frame time-frequency block $\eta_H(a_i, b_i)$ one after another, and the maximum point of $|\eta_H|$ in the time-frequency block $\eta_H(a_i, b_i)$ is denoted as $C_i(t_i^{spe}, f_i^{spe})$. Ifthe selected signal length is not an integer multiple of the rectangular windowlength, the maximum value of i is

$$i = \left[\frac{n_{tot}}{f_s \cdot (t_2 - t_1)}\right], \tag{3}$$

where $[\cdot]$ is rounding function and n_{tot} is the total number of sampling points of the signal $S(n)$.

Let the wavelet coefficient corresponding to the point $C_i(t_i^{spe}, f_i^{spe})$ be η_{HC_i}, and two condition algorithms are set for η_{HC_i} to extract feature points. Condition 1: if $\eta_{HC_i} = 0$, $|\eta_{HC_{i+1}}| > 0$, $\eta_{HC_{i+2}} = 0$, then $C_{i+1}(t_{i+1}^{spe}, f_{i+1}^{spe})$ is called the feature point of the impact fault component. Condition 2: if $\eta_{HC_i} = 0, |\eta_{HC_{i+1}}| > 0, |\eta_{HC_{i+2}}| > 0, \ldots, |\eta_{HC_{i+n}}| > 0, |\eta_{HC_{i+n+1}}| = 0$. Take the point C corresponding to max $\{|\eta_{HC_{i+1}}|, |\eta_{HC_{i+2}}|, \ldots, |\eta_{HC_{i+n}}|\}$ as the feature point of the impact fault component. Finally, all the feature points of the impact fault components are assigned to $Q_j(t_{jj}, f_{jj})$, so the point set of each of the periodic feature points of the impact fault component is $[Q_1, Q_2, \ldots, Q_j]$. Using each periodic feature point of the fault signal to estimate the period of the fault acoustic signal, the period of the fault signal is approximately $T_1 = t_{22} - t_{11}$, $T_2 = t_{33} - t_{22}, \ldots, T_{j-1} = t_{jj} - t_{(jj-1)\ (jj-1)}$.

3.2.2. Extraction of Impact Fault Component by Superimposing Noise Reduction. Since the feature point of the impact fault component is determined by the intense change region of the impact fault signal, the feature point of the impact fault component is located in a certain location of the intense change region of the impact fault signal. According to the feature point $Q_j(t_{jj}, f_{jj})$ of the impact fault component, the sampling point corresponding to the feature point is

$$n_{jj} = t_{jj} \cdot f_s, \tag{4}$$

where n_{jj} is the number of sampling points corresponding to the fault feature point, t_{jj} is the time of the feature point of the impact fault component, and f_s is the sampling frequency.

The range of the impact fault component of each period is as follows: $(\sigma - 1)$ sampling points are extended forward and backward from the n_{jj}th sampling point, and the signal segment is denoted as $W_i^{range}(n_j)$, so $W_i^{range}(n_j)$ includes fault components. $n_j \in [n_{jj} - \sigma + 1, n_{jj} + \sigma - 1]$; σ is the length of the intense change region. Due to the interference of background noise, the feature point location of per period is different in the intense change region. Therefore, the intense change regions in $W_i^{range}(n_j)$ may have location offset of different degrees in different periods. As shown in Figure 5, there is offset τ of different degrees in the intense change regions of different periodic $W_i^{range}(n_j)$, which will affect the dislocation superposition effect and lead to inaccurate superposition results.

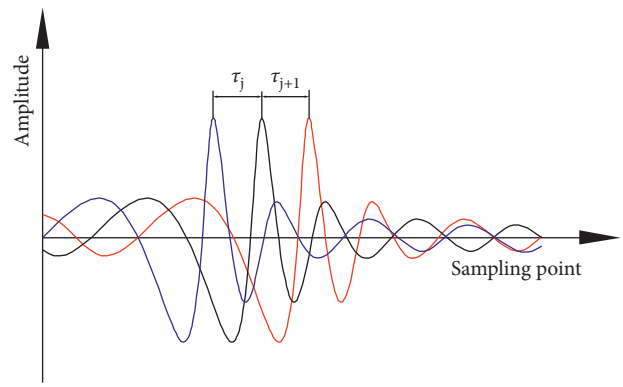


Figure 5: The offset τ of the intense change region in different periods $W_i^{range}(n_j)$.

To solve the problem of the offset of the intense change region, the specific method is as follows.

A segment of $W_i^{range}(n_j)$ containing the fault component is taken as the reference signal, denoted as $W_s^{range}(n_s)$. $W_i^{range}(n_j)$ is extended forward and backward by b sampling points, respectively, to obtain the extended signal segment $W_i^{erange}(n_j)$, $n_j \in [n_{jj} - \sigma + 1 - b, n_{jj} + \sigma - 1 + b]$. The continuous intercept length of $(2\sigma - 1)$ sampling points from the first sampling point in $W_i^{erange}(n_j)$ is denoted as $W_i^{fra}(n_j)$. Set τ_j to the offset of $W_i^{fra}(n_j)$ on $W_i^{erange}(n_j)$, and record the offset signal of $W_i^{fra}(n_j)$ as $W_i^{fra}(n_j + \tau_j)$, as shown in Figure 6. The subset of τ_j is denoted as L, and L is set to $\{0: \sigma - 1 + b\}$. Find the optimal offset τ_j^{opt} of the offset τ_j in $W_j^{fra}(n_j + \tau_j)$; that is, also find the optimal starting superposition point, where the superposition length is $(2\sigma - 1)$ sampling points.

Traverse all the τ_j values in L, and calculate Pearson's correlation coefficient $\rho_{sj}(W_s^{range}(n_s), W_i^{fra}(n_j + \tau_j))$ of the reference signal segments $W_s^{range}(n_s)$ and $W_i^{fra}(n_j + \tau_j)$, respectively; the offset corresponding to the maximum value of $\rho_{sj}(W_s^{range}(n_s), W_i^{fra}(n_j + \tau_j))$ is the optimal offset τ_i^{opt}. The equation of the proposed method is as follows:

$$\left(\tau_1^{opt}, \tau_2^{opt}, \ldots, \tau_j^{opt}\right) = \underset{\tau_j \in L, j=1,2,3,\ldots}{\arg\max} \left[\rho_{sj}\left(W_s^{range}(n_s), W_j^{erange}\left(n_j + \tau_j\right)\right)\right], \tag{5}$$

where τ_j^{opt} is the best offset; argmax $[\cdot]$ is arguments of the maxima; $W_s^{range}(n_s)$ is the reference signal segment; $W_i^{fra}(n_j + \tau_j)$ is the offset signal segment; and $\rho_{sj}(W_s^{range}(n_s), W_i^{fra}(n_j + \tau_j))$ is to calculate the correlation between $W_s^{range}(n_s)$ and $W_i^{fra}(n_j + \tau_j)$.

$2K$ consecutive signals $W_i^{fra}(n_j + \tau_i^{opt})$ are selected, which are divided into two groups according to the parity of the sequence number j, and each group is superimposed with different times according to the sequence number from low to high. The equation of the proposed method is as follows:

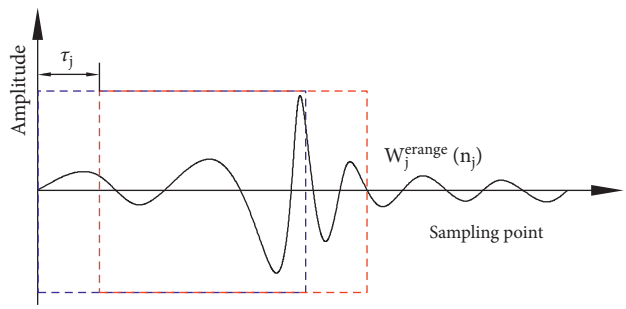


FIGURE 6: $W_i^{fra}(n_j+\tau_j)$ signal segment schematic diagram. The part selected by the blue frame is $W_i^{fra}(n_j)$, and the part selected by the red frame is $W_i^{fra}(n_j+\tau_j)$.

$$\begin{cases} D_{W1K}^{opt} = \dfrac{1}{K+1}\left[W_1^{fra}\left(n_1+\tau_1^{opt}\right)+W_3^{fra}\left(n_3+\tau_{j+3}^{opt}\right)+\cdots+W_{2k-1}^{fra}\left(n_{2k-1}+\tau_{2k-1}^{opt}\right)\right], \\ D_{W2K}^{opt} = \dfrac{1}{K+1}\left[W_2^{fra}\left(n_2+\tau_2^{opt}\right)+W_4^{fra}\left(n_4+\tau_4^{opt}\right)+\cdots+W_{2k}^{fra}\left(n_{2k}+\tau_{2k}^{opt}\right)\right], \\ \rho^{opt}(K) = \rho\left(D_{W1K}^{opt}, D_{W2K}^{opt}\right), \end{cases} \tag{6}$$

where D_{W1K}^{opt} is the signal after the superposition of the odd group; D_{W2K}^{opt} is the signal after the superposition of the even group; K is the number of superpositions; $W_{2K-1}^{fra}(n_{2K-1}+\tau_{2K-1}^{opt})$ is the signal segment of the odd group with fault components; $W_{2K}^{fra}(n_{2K}+\tau_{2K}^{opt})$ is the signal segment of the even group containing the fault component; $\rho^{opt}(K)$ is the correlation between the signal after the superposition of the odd group and the signal after the superposition of the even group.

Set the threshold M, when $\rho^{opt}(K) \geq M$, stop the iteration and output the superposition results.

3.2.3. Parameters Setting

(1) Hard Threshold λ Setting. The hard threshold λ is mainly to remove the smaller wavelet coefficients that only contain background noise and retains the larger wavelet coefficients that contain fault components. When the value of λ is too small, the smaller wavelet coefficients generated by background noise may be incompletely removed, resulting in unobvious fault periodicity. When the value of λ is appropriate, the smaller wavelet coefficients are completely removed, only the larger wavelet coefficients containing fault components are retained, and the fault periodicity is obvious. When the value of λ is too large, not only the smaller wavelet coefficients are completely removed, but also part of the larger wavelet coefficients containing fault components will be removed, resulting in a jump in the period of the fault signal. However, the location of the fault component can be found without affecting the superimposing effect. Therefore, the hard threshold λ should be as large as possible.

(2) Rectangular Window R Setting. The rectangular window R is composed of two diagonally paired time-frequency points (t_1, f_1), (t_2, f_2); then, the length of the rectangular window R is l, $l = t_2 - t_1$ and the width $w = f_2 - f_1$. Considering the frequency distribution of the impact fault signal and the length σ of the impact fault component, the time of the intense change region is

$$t_\sigma = \frac{\sigma}{f_s}. \tag{7}$$

The rectangular window starts from $t = 0$ s, so $t_1 = 0$ s. According to the characteristic frequency of the impact fault signal, which is generally distributed in the middle and low frequency, f_1 can be set between 0 Hz and 500 Hz, and $f_2 \leq 5000$ Hz. If t_2 is too large, it will cause the rectangular window to select multiple fault signal periods, which is not conducive to extracting the fault signal period. If t_2 is too small, it will increase the amount of calculation. Therefore, the general value range of l is $2t_\sigma \leq l \leq 4t_\sigma$, which corresponds to 2σ to 4σ sampling points.

(3) The Length of the Intense Change Region σ. The length σ of the intense change region is generally the length of the high-amplitude zone of artificially selected fault component. Literature [10] puts forward the general empirical formula of σ as

$$\sigma = \frac{f_s}{44100} \times 300, \tag{8}$$

where f_s is the sampling frequency of the acoustic signal.

(4) Parameter b Setting. According to the offset of the feature point in the fault component, $b = \sigma - 1$ sampling points are selected to minimize the calculation amount finding the optimal offset τ_j^{opt} more accurately and finding the optimal starting point.

(5) Threshold M Setting. With the increase of superposition number K, the correlation coefficient ρ will not always

increase. When ρ generally increases to 0.9, it will tend to be stable, and there is an unobvious increase trend [10]. To achieve a better noise reduction effect for the fault component, M should be set to a value between 0.75 and 0.85, where $M = 0.80$.

4. Experiment Condition

Figure 7 is the physical diagram of the automobile engine fault detection test bench, which consists of an automobile engine, sound sensor, data acquisition card, encoder, and computer. The engine model is EA211 with the detailed parameters shown in Table 1. The sound sensor with the frequency range of 20 Hz to 20 kHz is placed above the cylinder to receive the abnormal sound signal generated by the engine. The data acquisition card uses the USB-6341 produced by National Instruments Company. When collecting the sound signal, the sampling frequency is set to 44100 Hz. The encoder model is ZSP3806GC, and the resolution is 100P/R. The encoder can synchronously revolve with the crankshaft by fixing it on the front end of the crankshaft through a coupling.

5. Collection and Processing of Experimental Data

5.1. Fault Component Extraction of Engine Knocking Cylinder. Figure 8(a) shows the knocking cylinder abnormal noise signal and encoder pulse signal collected under the condition of engine running at 1800 r/min, Figure 8(b) uses db5 wavelet as the mother wavelet to perform CWT on the knocking cylinder abnormal noise signal to obtain the wavelet time-frequency diagram. The two diagonally paired time-frequency points of the rectangular window R are (0, 0), (0.0227, 4000), respectively. The hard threshold of the wavelet coefficient is $\lambda = 0.94$, and the length of the intense change region is calculated by equation (8) to obtain $\sigma = 300$, $b = 299$. The subset L of offset τ_j is set to $\{0:598\}$.

Table 2 is the sampling point corresponding to the fault feature point n_{jj} of each periodic obtained by data processing using the above method. Table 3 shows the number of sampling points for each period of impact fault signal calculated by using the feature points of each periodic failure component. It can be seen from Table 3 that the number of sampling points in the 6^{th}–8^{th} periods is 8708, which is quite different from the number of sampling points in other periods. This is due to the excessive setting of the hard threshold λ, which leads to the zeroing of the wavelet coefficients corresponding to the feature points of the fault signal components in the 7^{th} and 8^{th} periods after the hard threshold processing, so the number of sampling points in the 6^{th}, 7^{th}, and 8^{th} periods cannot be obtained, but it does not affect the superposition effect.

Selecting $s = 1$, 299 sampling points are extended forward and backward from the first feature point 3530 of the fault component to obtain the reference signal segment $W_1^{range}(n_1)$, as shown in Figure 9. Table 4 shows the optimal offset τ_j^{opt} of each $W_j^{fra}(n_j + \tau_j)$ by equation (5), and $W_j^{fra}(n_j + \tau_j^{opt})$ is obtained according to τ_j^{opt}. The continuous $2K$ signal segments $W_j^{fra}(n_j + \tau_j^{opt})$ including fault components are selected and divided into two groups according to the parity of the sequence number j for superposition operation, and the correlation coefficient $\rho^{opt}(K)$ of the parity group is calculated. Figure 10 shows the variation curve of $\rho^{opt}(K)$ and the number of superpositions.

When $K = 14$, $\rho^{opt}(14) = 0.8030$, which is greater than the threshold M, stopping superimposing. Figure 11 shows the fault component signals D_{W114}^{opt} and D_{W214}^{opt} after noise reduction.

Figure 12 shows the time-domain diagram of the odd group and even group superimposed signals obtained by superimposing 10 times using the improved DSM. The optimal starting point of superposition is the sampling point of the acoustic signal corresponding to the 28^{th} pulse of each cycle encoder signal. The superposition length is set to 600 sampling points, the odd group and even group are superimposed 10 times to obtain D_{W110}^{opt} and D_{W210}^{opt}, and the correlation reaches 0.8020.

According to Figures 11 and 12, the length $\sigma = 150$ of fault component is accurately extracted. Figure 13 shows the fault components W_1, W_2, W_3, and W_4 extracted from D_{W114}^{opt}, D_{W214}^{opt}, D_{W110}^{opt}, and D_{W210}^{opt}, respectively. The correlation between W_1 and W_3 is 0.9029, and the correlation between W_2 and W_4 is 0.9290, which further verifies that the fault component noise reduction method achieves the same effect compared with the improved DSM.

5.2. Fault Component Extraction of Engine Lacking Cylinder. Figure 14 shows the physical picture of the first cylinder lacking of the automobile engine by pulling out the cylinder line of the first cylinder. Figure 15 shows that the engine produces abnormal noise signal for the first cylinder lacking. The db5 wavelet is selected as the mother wavelet to perform CWT on the abnormal noise signal of the lack of cylinder. The time-frequency points of the two diagonally paired of the rectangular window R are set as (0, 500), (0.0227, 4000), respectively. The hard threshold of the wavelet coefficient $\lambda = 1$, $\sigma = 300$, $b = 299$, and the subset L of the offset τ_j is set as $\{0:598\}$.

14 15.

Table 5 shows the sampling points corresponding to the feature points of each period of fault component obtained by the above method. Selecting $s = 1$, the reference signal $W_1^{range}(n_1)$ is obtained by extending 299 sampling points forward and backward from the 1794^{th} sampling point, as shown in Figure 16. Table 6 shows the optimal offset τ_j^{opt} of each $W_j^{fra}(n_j + \tau_j)$ signal by equation (5), and $W_j^{fra}(n_j + \tau_j^{opt})$ is obtained according to τ_j^{opt}. Figure 17 shows the variation curve of the correlation coefficient $\rho^{opt}(K)$ of the superposition signals of the odd group and even group with the number of superpositions K. When $K = 6$, $\rho^{opt}(6) = 0.8048$, which is greater than the threshold M and stopping superimposition, D_{W16}^{opt} and D_{W26}^{opt} of the fault component after noise reduction are got, as shown in Figure 18.

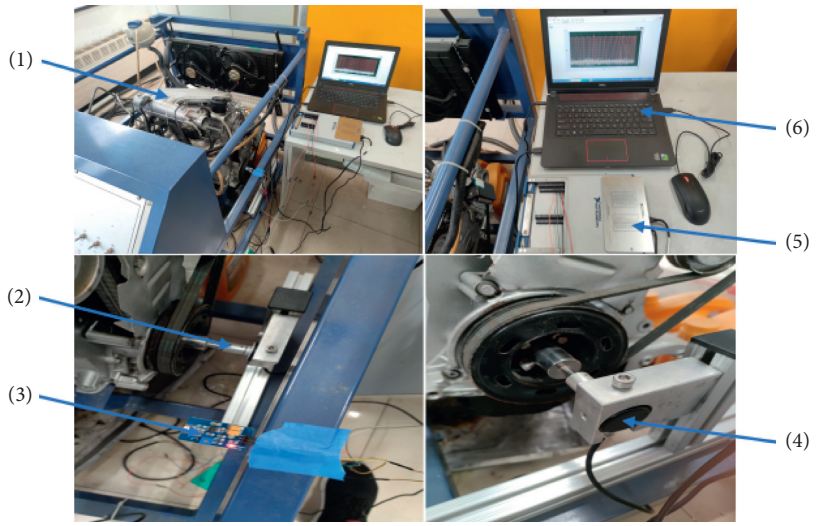


Figure 7: Automobile engine fault detection test bench. (1) Engine; (2) coupling; (3) sound sensor; (4) encoder; (5) data acquisition card; (6) computer.

Table 1: Engine parameters.

Engine type	EA211
Cylinder	4
Displacement	1.6 L
Maximum power	66 kW
Maximum power revolution	5500 rpm
Maximum torque	132 Nm
Maximum torque speed	3800 rpm

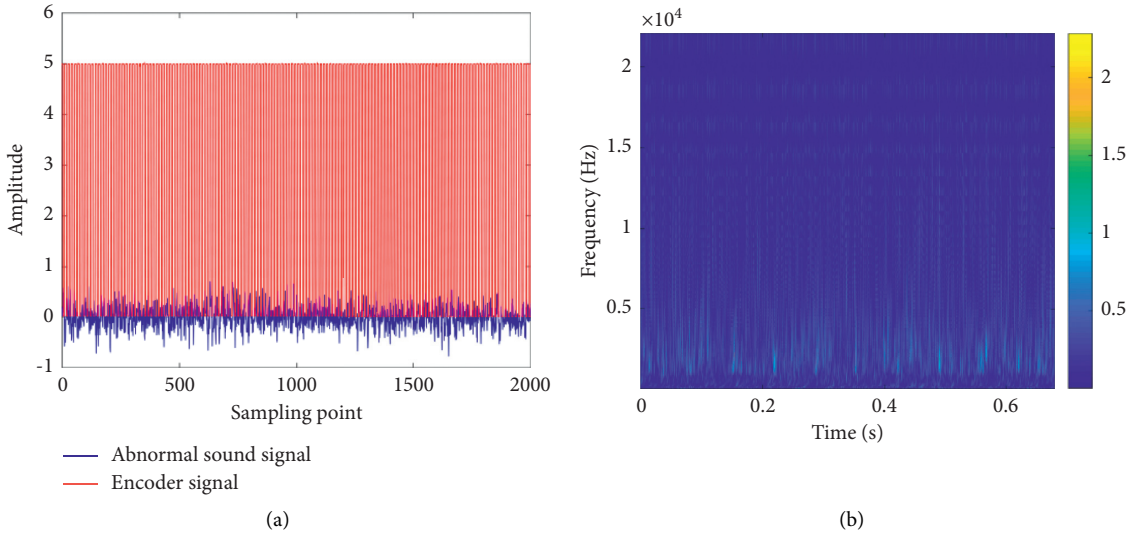


Figure 8: (a) Knocking cylinder and encoder signal collected at 1800 r/min. (b) Wavelet time-frequency diagram.

Table 2: The sampling point corresponding to the fault feature points of each period.

j	1	2	3	4	5	6	7	8	9	10	11	12
Sampling point (n_{jj})	3530	6734	9623	12681	15538	18699	27407	30568	33465	36673	39590	42700

Table 3: Error of cylinder knock fault signal period.

j	1	2	3	4	5	6–8	9	10	11	12
Number of sampling points of each period (n_c)	3204	2889	3058	2857	3161	8708	3161	2897	3208	2917
Actual number of sampling points of each period (n_a)	3153	2953	3025	2846	3174	8680	3166	2917	3190	2944
Error (Δ)	50	−64	33	11	−13	28	−5	−20	18	−27

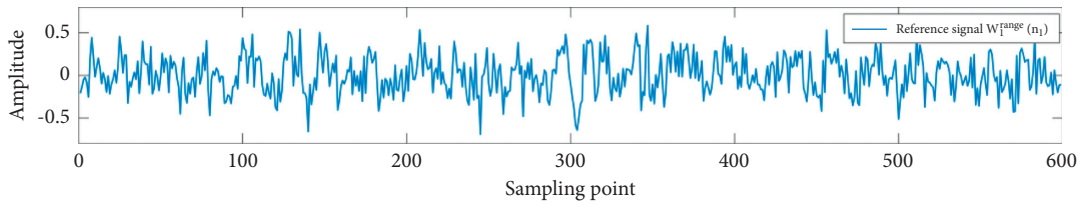


Figure 9: Reference signal segment W_1^{range} (n_1).

Table 4: The optimal offset τ_j^{opt} of W_j^{fra} ($n_j + \tau_j$).

j	1	2	3	4	5	6	7	8	9	10	11	12
The optimal offset (τ_j^{opt})	299	312	348	336	310	347	348	301	337	301	313	300

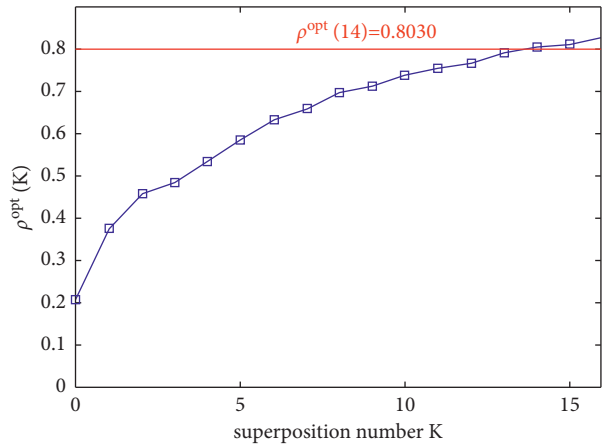


Figure 10: The variation curve of ρ^{opt} (K) with the increase of the superimposition number (K).

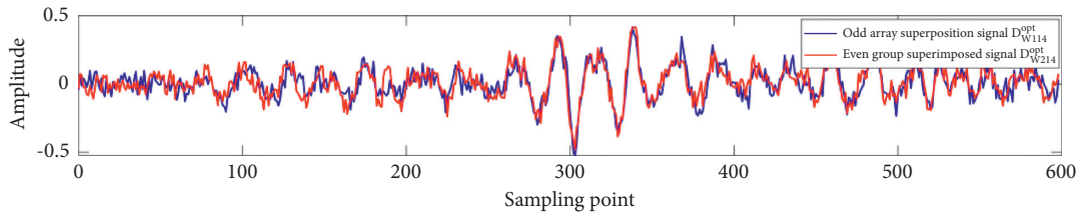


Figure 11: Odd and even groups superposition signals after 14-time superposition.

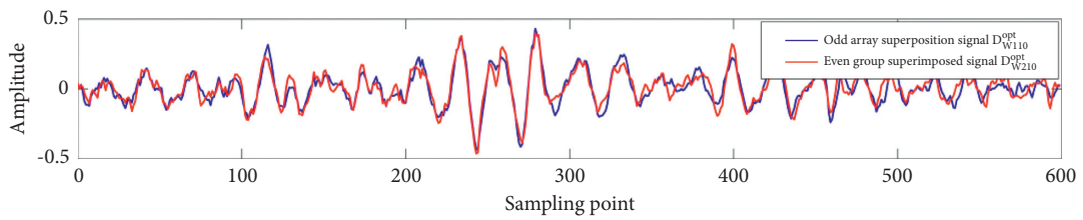


Figure 12: The superimposed signals of the odd group and even group after superimposing 10 times with the improved DSM.

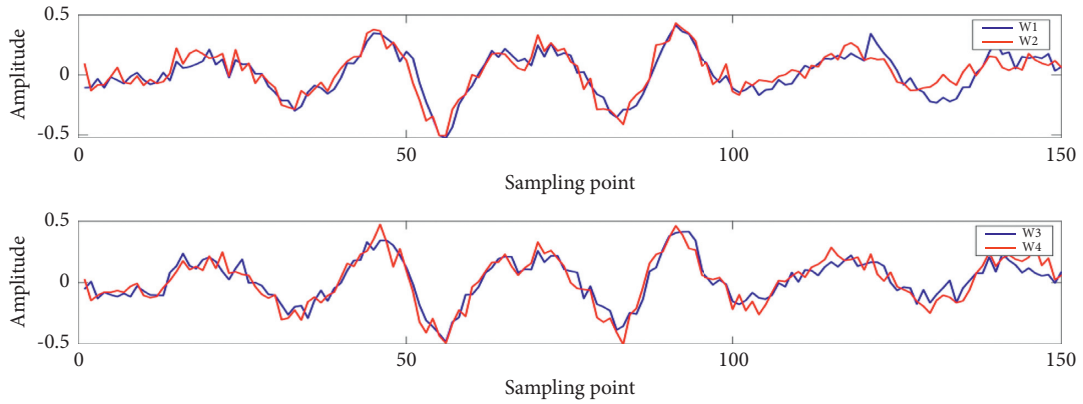


Figure 13: Fault components extracted accurately. W_1, W_2, W_3, and W_4 are fault components accurately extracted from D^{opt}_{W114}, D^{opt}_{W214}, D^{opt}_{W110}, and D^{opt}_{W210}, respectively.

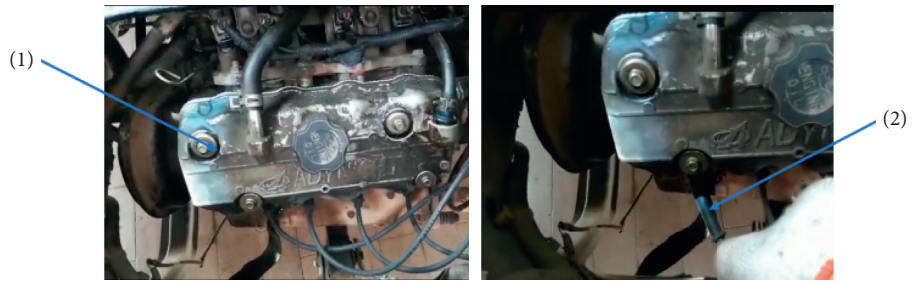


Figure 14: The physical map of the first cylinder lacking of the car engine. (1) Automotive engine; (2) the first cylinder line.

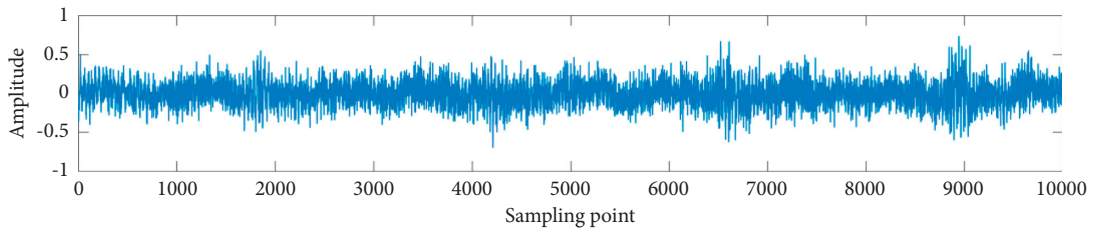


Figure 15: Engine abnormal noise signal of the first cylinder lacking.

Table 5: The sampling point corresponding to the fault feature points of each period.

j	1	2	3	4	5	6	7	8	9	10	11	12
Sampling point (n_{jj})	1794	4204	6567	8885	11161	13606	16002	18301	20798	23331	25749	28247

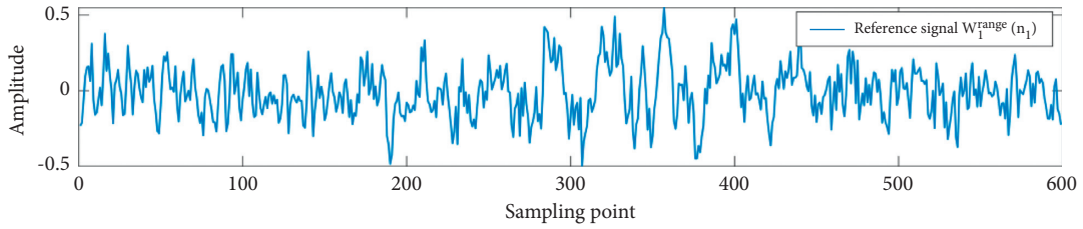


Figure 16: Reference signal W_1^{range} (n_1).

Table 6: The optimal offset τ_j^{opt} of W_j^{fra} $(n_j + \tau_j)$.

j	1	2	3	4	5	6	7	8	9	10	11	12
The optimal offset (τ_j^{opt})	299	268	298	300	298	196	191	266	266	265	301	356

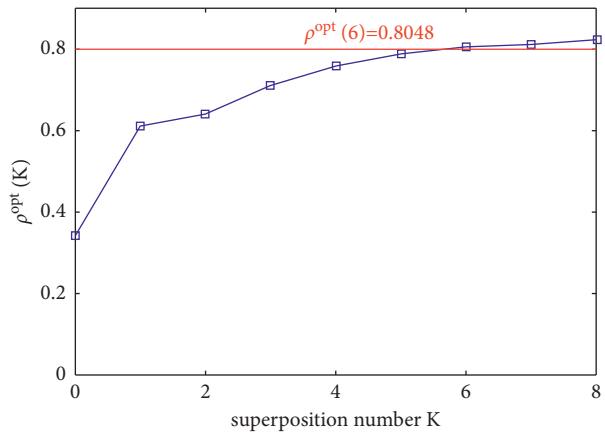


Figure 17: The variation curve of $\rho^{opt}(K)$ with the increase of the superimposition number K.

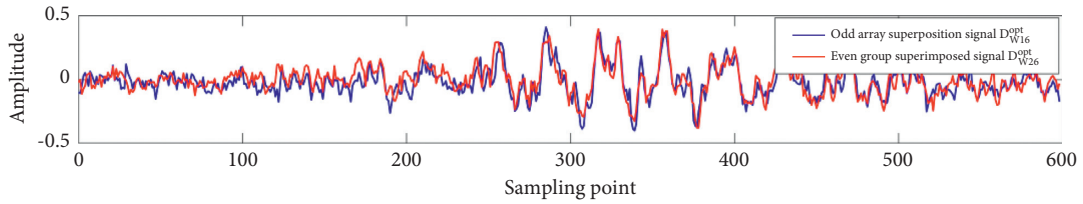


Figure 18: Odd group and even group superposition signals after 6-time superposition.

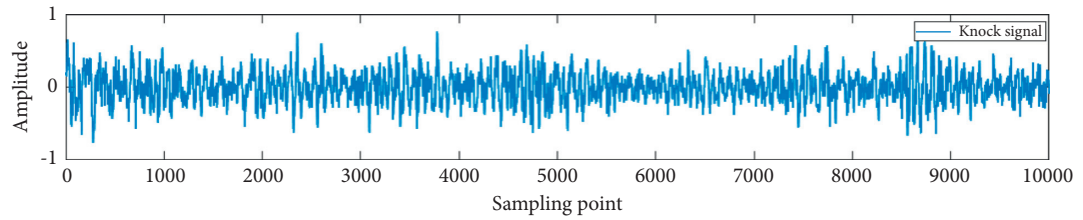


Figure 19: Engine knocking signal.

Table 7: The sampling point corresponding to the fault feature points of each period.

j	1	2	3	4	5	6	7	8	9	10	11	12
Sampling point (n_{jj})	8617	21164	25307	29327	34003	45681	49581	53832	57872	61933	70294	78110

Table 8: The optimal offset τ_j^{opt} of $W_j^{fra}(n_j+\tau_j)$.

j	1	2	3	4	5	6	7	8	9	10	11	12
The optimal offset (τ_j^{opt})	299	228	17	389	167	244	442	236	234	234	447	232

5.3. Fault Component Extraction of Engine Knocking. Figure 19 shows the engine knock signal due to the low gasoline label. The db5 wavelet is selected as the mother wavelet to perform CWT on the knocking signal. The time-frequency points of the two diagonally paired of the rectangular window R are set as (0, 0), (0.0227, 4000), respectively. The hard threshold of the wavelet coefficient $\lambda = 2.8$, $\sigma = 300$, $b = 299$, and the subset L of the offset τ_j is set as $\{0:598\}$.

Table 7 shows the sampling points corresponding to the fault feature points. Due to the large value of hard threshold λ of wavelet coefficient, the feature points of fault components do not show periodic feature, but the interval of sampling points between each two feature points is approximately 4000 times. Table 8 shows the optimal offset τ_j^{opt} of each $W_j^{fra}(n_j+\tau_j)$, and $W_j^{fra}(n_j+\tau_j^{opt})$ is obtained according to τ_j^{opt}. Figure 20 shows the variation curve of the

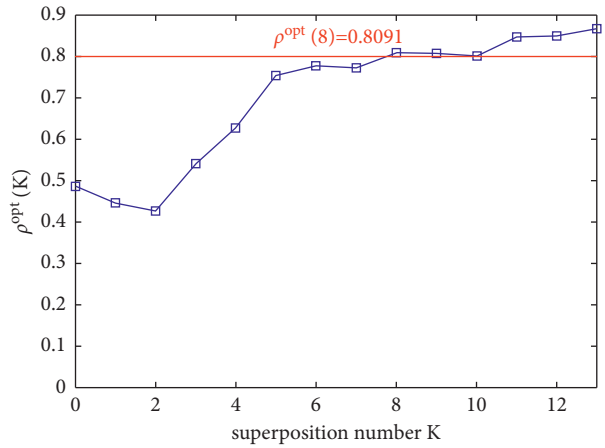


Figure 20: The variation curve of $\rho^{opt}(K)$ with the increase of the superimposition number K.

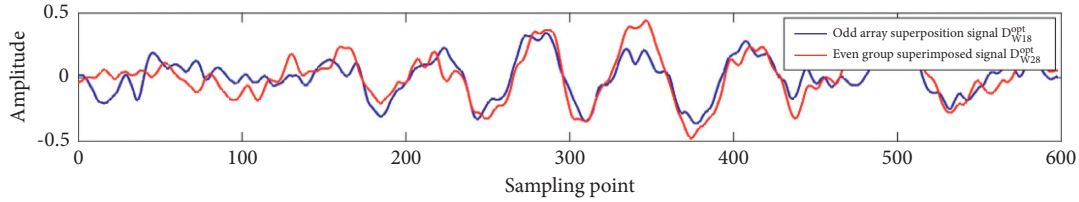


Figure 21: Odd group and even group superposition signals after 6-time superposition.

correlation coefficient $\rho^{opt}(K)$ of the superposition signals of the odd group and even group with the number of superpositions K. When $K=6$, $\rho^{opt}(8)=0.8091$, which is greater than the threshold M, and stopping superimposition, D^{opt}_{W18} and D^{opt}_{W28} of the fault components after noise reduction is got, as shown in Figure 21.

6. Experimental Results Analysis

In order to further verify the effectiveness of this method, the experimental results of this method and the improved DSM are compared by using the knocking cylinder experiment of automobile engine, and then the effectiveness of this method is verified. In this method, odd group signals and even group of engine knocking cylinder signals are superimposed 14 times, and the correlation between the noise reduction signals D^{opt}_{W114} and D^{opt}_{W214} is 0.8030. Then, the noise reduction processing and extraction of engine knocking cylinder failure component are carried out by using the improved DSM. When odd group signals and even group signals are superposed 10 times, the correlation between the noise reduction signals D^{opt}_{W110} and D^{opt}_{W210} is 0.8020. Finally, the fault components W_1, W_2, W_3, and W_4 are accurately extracted from noise reduction signals of D^{opt}_{W114}, D^{opt}_{W214}, D^{opt}_{W110}, and D^{opt}_{W210}, where the correlation between W_1 and W_3 is 0.9029, and the correlation between W_2 and W_4 is 0.9290. The results show that this method can further replace the use of encoders and achieve the same noise reduction effect as the improved DSM.

In the engine cylinder lacking experiment, this method is used to reduce the noise of fault components. The correlation between the noise reduction signal D^{opt}_{W16} and D^{opt}_{W26} obtained by superposition of odd group signals and even group signals for 6 times is 0.8048. In the engine knock experiment, due to the relatively small setting of hard threshold λ, some fault feature points are not selected accurately, resulting in the correlation of superposition signals of odd and even groups after superposition for 2 and 3 times gradually decreasing, as shown in Figure 20. However, with the gradual increase in the number of superpositions, the correlation between D^{opt}_{W18} and D^{opt}_{W28} obtained by superposition of odd group and even group for 8 times is 0.8091.

From the analysis results, the method can replace the improved DSM, effectively reduce the noise, and extract the fault component of the engine abnormal noise signal. By increasing the threshold M and superimposition number K, the accuracy of the extracted fault components can be improved. The method can also use the extracted fault components to establish a database of automobile engine faults, which lays a foundation for the later diagnosis and classification of automobile engine faults.

7. Conclusion

Althoughthe improved DSM can automatically find the starting superposition point andsuperposition length of the periodic signal, it needs to use the encoder toassist in finding the starting superposition point. In practical applications, many occasions are not suitable for the use of encoders, which limits the applicability of DSM. Therefore, based on the improved DSM, this paper proposes an adaptive noise reduction method of automobile engine abnormal noise

signal, which cancels the use of the encoder. Using this method, the fault components of knocking cylinder, cylinder lacking, and knocking are extracted from the abnormal sound signal of automobile engine, and the effectiveness of the method is verified. The method can adaptively obtain the starting superposition point and superposition length from the engine abnormal sound signal and determine the number of superpositions, which replaces the encoder in the improved DSM to determine the starting position of the automobile engine impact fault and greatly improves the applicability of DSM. However, when the hard threshold λ is set to a small value, the selection of fault feature points will be inaccurate, and then the selected fault components will be biased, resulting in inaccurate noise reduction signals. The specific problem of hard threshold setting will be further studied. When the starting superposition point is selected, the accuracy is reduced compared with the improved DSM, resulting in more superposition times than the improved DSM. However, according to the superposition results, the increase of superposition times is within an acceptable range.

In the future, the hard threshold λ setting problem will be further improved, and the corresponding fault component database will be established. The fault type will be determined by comparing the extracted fault component with the fault component in the database.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

Acknowledgments

This work was funded in part by the National Key Research and Development Program under Grants 2018YFC0810406 and 2018YFC0810404, in part by the Fundamental Research Funds for the Central Universities under Grant 3132019118, and in part by the National Natural Science Foundation of China under Grant U1908228.

References

[1] K. Jafarian, M. Darjani, and Z. Honarkar, "Vibration analysis for fault detection of automobile engine using PCA technique," in *Proceedings of the 2016 4th International Conference on Control, Instrumentation, and Automation (ICCIA)*, pp. 372–376, Qazvin, Iran, January 2016.

[2] J. Mohammadpour, M. Franchek, and K. Grigoriadis, "A survey on diagnostic methods for automotive engines," *International Journal of Engine Research*, vol. 13, no. 1, pp. 41–64, 2012.

[3] X. L. Liang, T. Chen, and X. D. Yang, "Application of fuzzy fault tree analysis in the fault diagnosis for automobile engine system," in *Proceedings of the 2017 9th International Conference on Modelling, Identification and Control (ICMIC)*, pp. 711–716, Kunming, China, July 2017.

[4] L. Zhang, Y. Shi, and L. Ren, "Humanoid extraction of abnormal engine sounds by using ICA-R and VANC," in *Proceedings of the 2012 International Conference on Systems and Informatics (ICSAI2012)*, pp. 1687–1692, Yantai, China, May 2012.

[5] J. Xie, J. G. Colonna, and J. L. Zhang, "Bioacoustic signal denoising: a review," *Artificial Intelligence Review*, vol. 54, no. 3, pp. 1–23, 2020.

[6] A. Azarang and N. Kehtarnavaz, "A review of multi-objective deep learning speech denoising methods," *Speech Communication*, vol. 122, pp. 1–10, 2020.

[7] H. Wang and Z. Ma, "Research of acoustic signal de-noising using wavelet transform," in *Proceedings of the 2009 2nd International Congress on Image and Signal Processing*, pp. 1–3, Tianjin, China, October 2009.

[8] S. V. P. S. Nidadavolu, S. K. Yadav, and P. K. Kalra, "Condition monitoring of internal combustion engines using empirical mode decomposition and morlet wavelet," in *Proceedings of the 2009 Proceedings of 6th International Symposium on Image and Signal Processing and Analysis*, pp. 65–72, Salzburg, Austria, September 2009.

[9] X. Chen and Y. Yang, "De-noising for vibration signals based on local mean decomposition," in *Proceedings of the IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society*, pp. 3298–3303, Beijing, November 2017.

[10] N. Dayong, J. Yuhua, S. Hongyu et al., "Separation method of impulsive fault component for gasoline engine based on acoustic signal analysis," *Shock and Vibration*, vol. 2019, Article ID 8573479, 15 pages, 2019.

[11] H. Qiu, J. Lee, J. Lin, and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics," *Journal of Sound and Vibration*, vol. 289, no. 4-5, pp. 1066–1090, 2006.

[12] M. Li, "Weak fault signal detection of rolling bearing," in *Proceedings of the 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE)*, pp. 540–543, Changchun, China, December 2011.

[13] A. Moshrefi, H. Aghababa, and O. Shoaei, "Knock signal denoising employing a new time domain method," in *Proceedings of the 2017 15th IEEE International New Circuits and Systems Conference (NEWCAS)*, pp. 37–40, Strasbourg, France, June 2017.

[14] W. Kai, D. Shaobo, Y. Zilin, and S. Shijie, "Application of wavelet threshold denoising on bearing fault diagnosis," in *Proceedings of the 2019 Chinese Control and Decision Conference (CCDC)*, pp. 1980–1985, Nanchang, China, June 2019.

[15] X. Wen and L. You, "A novel rolling bearing fault detection method based on wavelet transform and empirical mode decomposition," in *Proceedings of the 2019 Chinese Control Conference (CCC)*, pp. 5024–5027, Guangzhou, China, July 2019.

[16] 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.

[17] B. Sun, S. Zhou, and C. Wang, "Application of wavelet soft threshold denoising algorithm based on EMD decomposition in vibration signals," in *Proceedings of the 2019 6th International Conference on Systems and Informatics (ICSAI)*, pp. 7–11, Shanghai, China, November 2019.

[18] G. Ren and Z. Liu, "An improved EMD adaptive denoising and feature extraction algorithm," in *Proceedings of the 2019 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, pp. 1–4, Dalian, China, September 2019.

[19] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assi-sted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2009.

[20] X. Gao, H. Wen, and Pu Wang, "A new method of bearing fault diagnosis based on LMD and wavelet denoising," in *Proceedings of the 2017 29th Chinese Control and Decision Conference (CCDC)*, pp. 4155–4162, Chongqing, 2017.

[21] N. Dayong, S. Changle, G. Yongjun, Z. Zengmeng, and H. Jiaoyi, "Extraction of fault component from abnormal sound in diesel engines using acoustic signals," *Mechanical Systems and Signal Processing*, vol. 75, pp. 544–555, 2016.

[22] N. Dayong, S. Hongyu, X. Aoyu, G. Yongjun, D. Hongwei, and H. Jiaoyi, "Adaptive noise reduction method of synchronous hydraulic motor acoustic signal based on improved dislocation superposition method," *IEEE Access*, vol. 8, pp. 37161–37172, 2020.

[23] J. Yuhua, *Research and Experiment on Extraction of Engine Fault Components Based on Abnormal Sound Analysis*, Dalian Maritime University, Dalian, China, 2019.

Hindawi
Shock and Vibration
Volume 2021, Article ID 6985661, 13 pages
https://doi.org/10.1155/2021/6985661

Research Article

PZT Actuators' Effect on Vibration Control of the PRRRP 2-DOF Flexible Parallel Manipulator

Amin Valizadeh [1] **and Morteza Shariatee** [2]

[1]*Department of Mechanical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran*
[2]*Department of Mechanical Engineering, Iowa State University, Ames, IA, USA*

Correspondence should be addressed to Amin Valizadeh; amin.valizadeh@mail.um.ac.ir

Received 14 April 2021; Revised 21 June 2021; Accepted 1 July 2021; Published 10 July 2021

Academic Editor: Claudio Sbarufatti

Thanks to their advantages over rigid ones, interest for lightweight parallel manipulator was increased. Besides, structural flexibility effects at high operational speeds are more significant. Thus, developing an appropriate model for the assessment of the dynamic properties of flexible mechanisms and linkages to gain effective vibration control will raise high demand. Therefore, this paper represents the dynamic and kinematic modeling using the assumed mode method and first-type Lagrange equations of the 2-DOF planar parallel manipulator with two flexible links. To truly predict vibrations of the manipulator without any major simplifying assumptions, nonlinear dynamic modeling, which thoroughly attempts to represent the flexible behavior of the links, is considered. As a result, an active damping approach is being studied with PZT actuators. The results show that this approach is effective in damping the vibrations of the links that give accurate trajectory control.

1. Introduction

With regard to parallel manipulators with a lightweight structure, a planar parallel manipulator with lightweight linkages provides a high-speed alternative positioning mechanism for manipulators of serial architecture. These robots are used in a wide range of applications, from simple selection and location of robotic systems for industrial applications to microsurgical applications, maintenance of nuclear power plants, or space robotics [1]. The interest in research into flexible connection manipulators and mechanisms was significantly increased to make full use of the potential offered by flexible manipulators. It is, nevertheless, particularly challenging to control flexible manipulators so that precise positioning can be maintained. For a two-link flexible manipulator, the problem becomes more complex. The dynamics are highly nonlinear and complex due to the flexibility of the system [2, 3]. Although lightweight links are more likely to meet high-speed and high-acceleration requirements, the inertia and forces from the actuators are more likely than ever to deflect and vibrate [4]. The structural flexibility effects at high end-effector speeds are much more significant. Manipulators and mechanisms with flexible links are systems with a variety of degrees of freedom. They are described by coupled nonlinear partial differential equations of motion. The dynamic model formulation of manipulators with flexible links and mechanisms was based on different discretization ways of flexible links to devise and apply a real-time controller for joint movements and vibration removal. The most popular approaches are the finite element method (FEM) [5, 6] and the assumed mode method (AMM) [7, 8]. It has been commonly established to model a flexible single link manipulator. Various approaches were developed, mainly divisible into two categories: the approach to numerical analysis and the assumed method mode (AMM) [9, 10]. AMM examines approximate models by solving a partial differential equation that characterizes the system's dynamic behavior.

Previous studies have been reported using this approach to model a flexible single-link manipulator [11, 12]. Zhou

et al. recently developed dynamic equations for a flexible three-PRS manipulator with regard to vibrational analyses using the FEM method [13], taking into consideration link flexibility. Kang and Mills introduced a dynamic flexible-link 3-PRR planar parallel manipulator by employing AMM [14]. The existing parallel 2-DOF manipulator with solid links for an optimal design was studied [7]. In high-speed pick-and-place applications, this manipulator is also productive [15]. Nevertheless, no research has been conducted on the dynamic modeling of the mechanism, taking into account the flexibility in which industrial operations are inevitable. This paper considers a method of active damping using piezoelectric materials. Deformation of the flexible links produces shear stress that PZT materials can counteract due to the voltage control applied. PZT can achieve better performance in vibration damping than other transducer materials as PZT has higher stress constant [16–19].

Due to the promising results, the parallel kinematic machine (PKM) is the greatest increasing need of the machining and pick and place industry. Due to their high structural stiffness and rigidity, PKMs' absolute positioning error is reduced. The 2-DOF PRRRP PKM machine tool proves to achieve competing accuracies in the end tool [15]. However, the heavy and bulky links used to give adequate stiffness and accuracy significantly increase equipment costs, motor torques (power), and energy consumption. To overcome this issue, as a real-world requirement, research on the use of lightweight robot links is carried out in this paper. To overcome the positioning error due to the flexibility of the links, an active vibration control system based on PZT actuators is implemented.

In the present paper, an AMM modeling of the flexible links following Lagrangian method and a PD feedback control with linear velocity feedback (L-type) is used to correctly attenuate vibration due to trajectory tracking. This is followed by a proper PD trajectory control. The proposed active vibration damping approach was verified by simulations for flexible linkage manipulators.

2. Kinematic Modeling

Figure 1 illustrates the planar 2-DOF parallel manipulator with two flexible links. The manipulator architecture is 2-PRRRP, while R and P represent revolute and prismatic joints, respectively. In a plane that works properly for pick and place tasks, the end-effector offers high precision 2-DOF translational motion. The end-effector position vector, two active prism joints, and two passive revolute joints are presented as follows, respectively, about the fixed X-y reference framework displayed in Figure 2:

$$\begin{aligned} \overline{X}_e &= [\, x \;\; y \,]^T, \\ \overline{q} &= [\, q_1 \;\; q_2 \,]^T, \\ \overline{\beta} &= [\, \beta_1 \;\; \beta_2 \,]^T. \end{aligned} \tag{1}$$

As a deformation assumption and design criteria of the flexible links, the influence of transverse, shear, and rotary inertia has not been taken into account since the beam is long and slender. Links only vibrate horizontally, and the torsion and vertical bending are not considered. Besides, the beam properties variation can be neglected across the whole body and cross section [20].

Consequently, Euler–Bernoulli beam theory can be employed to simulate the manipulator's elastic behavior. The product of position-and time-dependent functions, i.e., AMM, expresses the deflection of the link, w_i, as

$$w_i(x,t) = \sum_{j=1}^{r} \eta_{ij}(t)\varphi_j(\xi). \tag{2}$$

In the equation above, $\xi = x/l$, j and r denote the jth vibration mode and a finite number of assumed modes, respectively.

To select the boundary conditions, one can take many different approaches in the AMM. The optimum set is found closest to the system's natural modes among the hypothesized modes in ideal situations. Thus, no assumption can be made about the employed set of hypothesized modes since several structural factors of the manipulator determine the natural modes [21]. In this study, pin-free modeling, implemented in a flexible PKM [16, 19], is considered for boundary conditions, which causes a significant deflection in the flexible links and remarkably investigates the effect of PZT actuators on damping the vibrations of flexible links. The alternative in the future needs an evaluation based on the robot's actual structure to improve outcomes in the results.

Given the boundary conditions of the flexible links on S_i and the end-effector, the selected normalized shape function that satisfies a pin-free boundary condition is as follows:

$$\varphi_j(\xi) = \frac{1}{2\,\sin(\gamma_j)} \times \left[\sin(\gamma_j \xi) + \frac{\sin(\gamma_j)}{\sinh(\gamma_j)} \sinh(\gamma_j \xi) \right], \tag{3}$$

where

$$\begin{aligned} &0 \le \xi \le 1, \\ &\gamma_j = (j + 0.25)\pi. \end{aligned} \tag{4}$$

The shape functions in the first three mode shapes of the flexible link are shown in Figure 3. The inverse kinematics problem is solvable by expanding the following restrictive equation, as shown in Figure 2:

$$OE = OA_i + q_i + b_i + w_i(l), \quad i = 1, 2, \tag{5}$$

which yields

$$\begin{cases} q_1 = \pm\sqrt{L^2 - (x-d)^2 + w_1^2(l)} + y, \\ q_2 = \pm\sqrt{L^2 - (x+d)^2 + w_2^2(l)} + y. \end{cases} \tag{6}$$

Equation (6) states that four solutions are available for the inverse kinematics of the mechanism. The four alternatives are consistent with four types of mechanism work modes. The deflection term, i.e., w_i, should be drawn from dynamic modeling to approach the solutions.

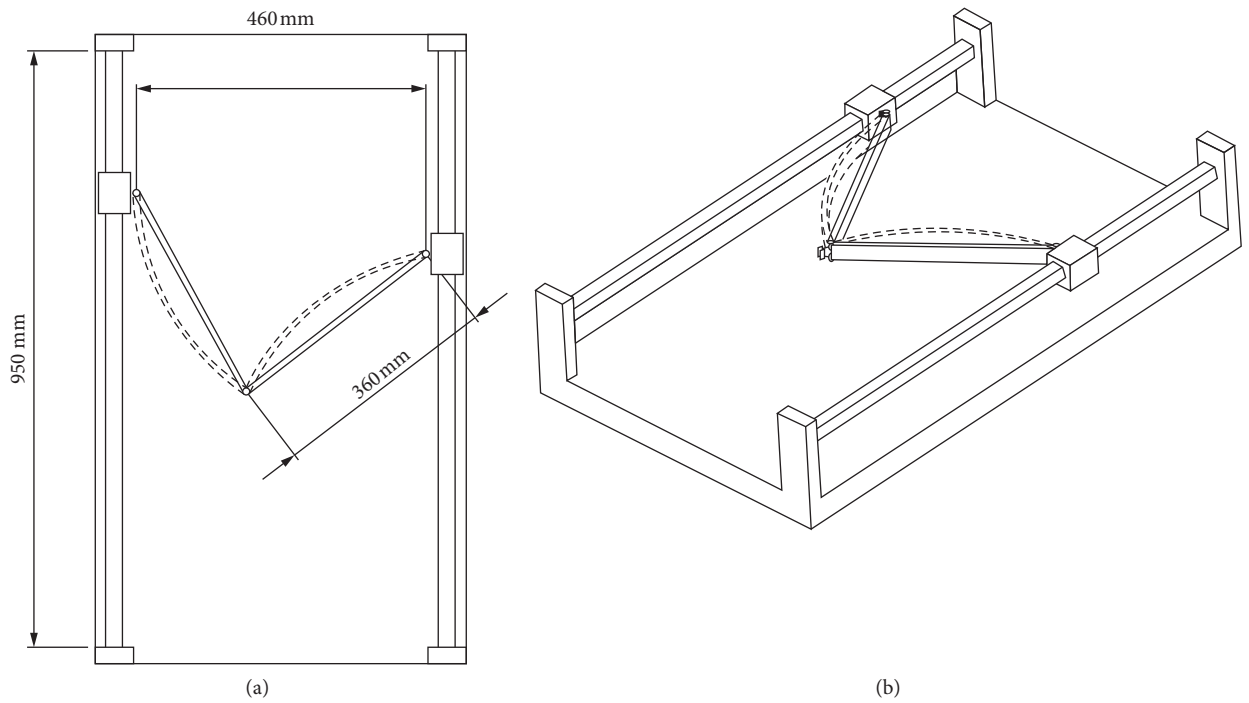


Figure 1: Architecture of 2-PRRRP with flexible links (dashed lines denote the deflection of the links).

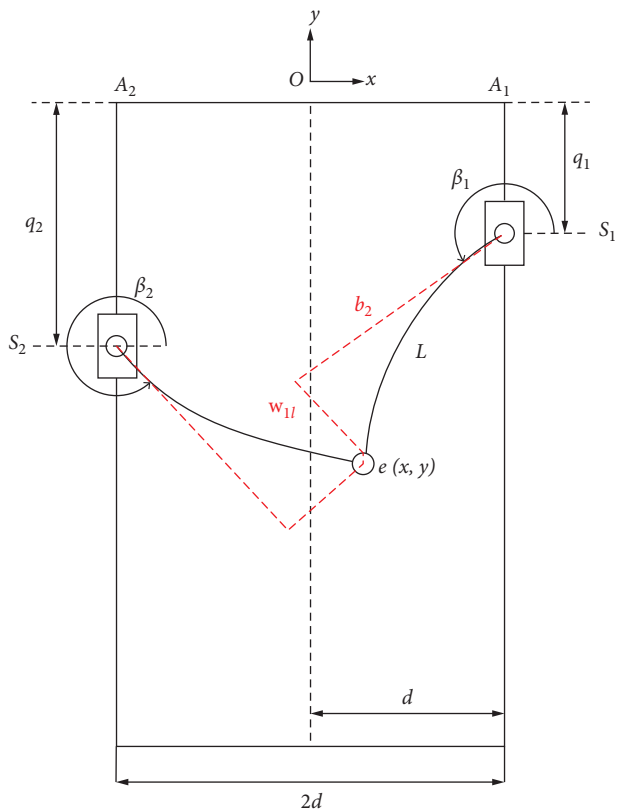


Figure 2: Coordinate system of 2-PRRRP with flexible links.

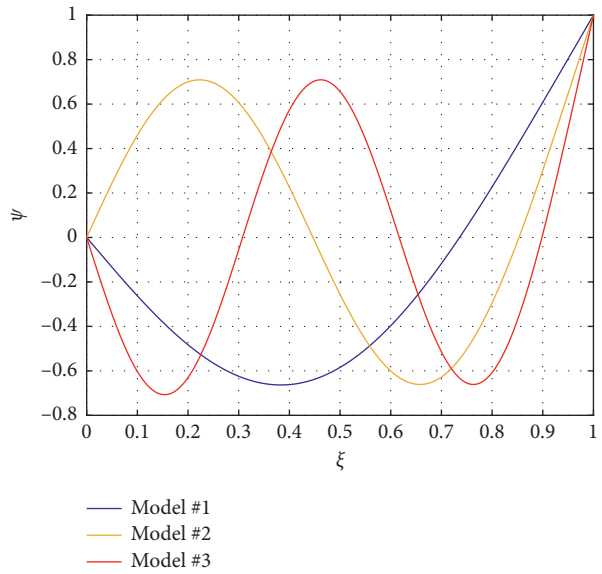


Figure 3: First three mode shapes of the flexible link.

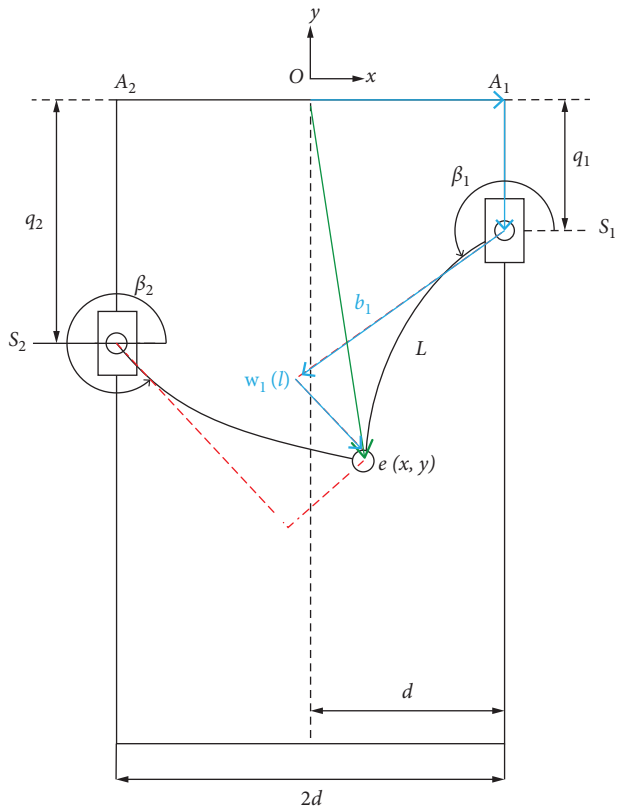


Figure 4: The closed kinematic chain of 2-PRRRP with flexible links.

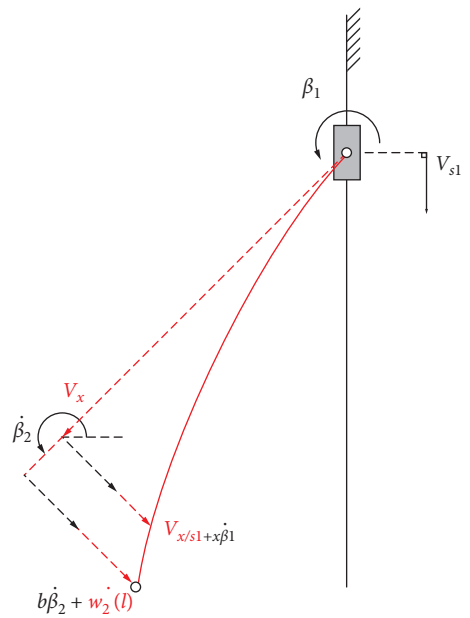


Figure 5: The kinetic energy of the flexible link.

Table 1: Dynamic parameters.

Flexible Link	Density (kg/m^3)	2770
	Young's modulus (GPa)	73
	Dimension (mm)	$360 * 30 * 5$
PZT actuator	Young's modulus (GPa)	69
	Dimension (mm)	$50 * 25 * 0.75$
Slider	Mass (kg)	0.3

Table 2: Feedback control gains.

K_P	1000 (N/m)
K_D	600 (N-s/m)
K_I	3000 (Volts-s/m)

3. Dynamic Modeling

In terms of link flexibility, the generalized coordinates are taken as follows:

$$X_{\text{flex}} = \left[\, \overline{q} \;\; \overline{\beta} \;\; \overline{X}_e \;\; \overline{\eta} \,\right]^T, \tag{7}$$

where

$$\overline{\eta} = \left[\;\; \eta_{11} \;\; \eta_{12} \;\; \eta_{13} \;\; \eta_{21} \;\; \eta_{22} \;\; \eta_{23} \,\right]^T_{6\times 1}. \tag{8}$$

Figure 4 shows the manipulator's closed kinematic chain with the flexible link deflection. Expanding equation (5), as mentioned, leads to four constraint equations, as follows:

$$\begin{cases} x_e = x_{Ai} + L\cos\beta_i - w_i(l)\sin\beta_i, \\ y_e = q_i + L\sin\beta_i + w_i(l)\cos\beta_i. \end{cases} \tag{9}$$

As shown in Figure 5, the kinematic energy is taken into account in expecting the deformation of the links:

$$T = \sum_{i=1}^{2} T_s + T_l + T_e, \tag{10}$$

where

$$\begin{cases} T_e = \dfrac{1}{2} m_e \left(\dot{x}_e^2 + \dot{y}_e^2 \right), \\ T_s = \displaystyle\sum_{i=1}^{2} \frac{1}{2} m_s \dot{q}_i^2, \\ T_l = \displaystyle\sum_{i=1}^{2} \frac{1}{2} \int_0^L \rho_A v_x^2 \, \mathrm{d}x, \end{cases} \tag{11}$$

where ρ_A, m_s, and m_e are mass of the end-effector, mass of the sliders, and mass per length, respectively. Put the kinematic energy and the potential energy induced by link deformation for each coordinate in first-type Lagrangian equations to derive the dynamic modeling yielding the equations of motion for the flexible-link parallel manipulator:

$$\begin{bmatrix} M_{11} & M_{12} & 0 & M_{14} \\ M_{12}^T & M_{22} & 0 & M_{24} \\ 0 & 0 & M_{33} & 0 \\ M_{14}^T & M_{24}^T & 0 & M_{44} \end{bmatrix} \begin{bmatrix} \ddot{\overline{q}} \\ \ddot{\overline{\beta}} \\ \ddot{\overline{X}} \\ \ddot{\overline{\eta}} \end{bmatrix} + \begin{bmatrix} V_1 \\ V_2 \\ 0 \\ V_4 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & K \end{bmatrix} \begin{bmatrix} \overline{q} \\ \overline{\beta} \\ \overline{X} \\ \overline{\eta} \end{bmatrix} = \begin{bmatrix} F_q \\ 0 \\ F_{\text{ext}} \\ 0 \end{bmatrix} + \begin{bmatrix} J_1 \\ J_2 \\ J_3 \\ J_4 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix}. \tag{12}$$

Each of the components has been defined in the Appendix. Equation (12) is a differential-algebraic equation (DAE) that can be simplified by removing the Lagrangian constraint term λ_i as follows:

$$\begin{bmatrix} \overline{M}_{11}^{2\times2} & \overline{M}_{12}^{2\times6} \\ \overline{M}_{21}^{6\times2} & \overline{M}_{22}^{6\times6} \end{bmatrix} \begin{bmatrix} \ddot{\overline{X}}^{2\times1} \\ \ddot{\overline{\eta}}^{6\times1} \end{bmatrix} = \begin{bmatrix} \overline{V}_{11}^{2\times1} \\ \overline{V}_{21}^{6\times1} \end{bmatrix}. \tag{13}$$

4. Vibration Control

Structural flexibility of the links transfers undesirable vibration to the end-effector, which leads to poor tracking efficiency. When the links are flexible for linear actuators, the vibration attenuation is difficult. Thus, an active damping approach with the help of PZT is implemented. Attaching the link surface, the PZT generates a shear force that suppresses the structural vibration of the links along the length. For flexible link manipulators, the aim of the above analysis is to make the rigid mode variable follow the required trajectory or converge to a certain point while suppressing the flexible link modes. The voltage of the PZT actuators will specify the L-type approach as explained in Appendix [22]. The voltage applied to the PZT actuator can be defined as follows:

$$V_i(t) = -k_I\left[\dot{w}_i(a_2, t) - \dot{w}_i(a_1, t)\right], \tag{14}$$

where k_i represents the linear velocity feedback gain for PZT and a_1 and a_2 are the start- and end-point positions of the PZT actuators from the linear actuators, S_i, along the length. $\dot{w}_i(a_i, t)$ denotes the linear velocity of each link at a_i. The virtual work conducted by the i-th PZT actuator can be measured as follows:

$$\delta W_{\text{PZT}} = cV_i(t)\sum_{j=1}^{r}\left[\varphi_f'(a_2) - \varphi_f'(a_1)\right]\delta\eta_{ij}, \tag{15}$$

where c denotes a positive constant implying the bending moment by applying voltage. Fulfilling the PZT actuator's L-type layout is based on the position of the PZT actuator. The PZT actuator is to be installed in a region where the form function and the derivative has a similar variation trend within to achieve a stable control movement: $x \in [a_1, a_2]$.

$$(\varphi(a_2) - \varphi(a_1))(\varphi'(a_2) - \varphi'(a_1)) \geq 0. \tag{16}$$

The application of this case to higher frequency modes is limited as the fulfillment of the equation for higher frequencies is only carried out in small areas on the link.

5. Simulation Results

Two linear actuators are equipped with a simple proportional-derivative-type (PD) feedback controller system as follows:

$$f_i = -k_p(q_{di} - q_i) - k_d(\dot{q}_{di} - \dot{q}_i), \tag{17}$$

where k_p and k_d represent PD feedback gains, respectively. q_{di} and $\dot{q}_{di}$ indicate desired values for linear actuators obtained from (5).

Tables 1 and 2, respectively, include dynamic parameters and feedback control gains. Using a fourth-order, Runge–Kutta method with MATLAB software was integrated into the normal differential equations at 1 msec integration intervals.

The desired trajectory, which accelerates and decelerates smoothly, has a sinusoidal function:

$$x_e = \frac{x_f}{t_f}t - \frac{x_f}{2\pi}\sin\left(\frac{2\pi}{t_f}t\right). \tag{18}$$

The objective is that the end-effector moves 2 mm (x_f) within 10 msec (t_f).

Figure 6 illustrates the end-effector's tracking error following the desired trajectory with and without PZT actuators in the X direction. The activated profile of the PZT actuator, known as "active damping," decreases continuously as a result of the PZT actuator's damping effect, while oscillation at the initial acceleration is significant. The tracking error in active damping mode is rapidly decreased, accordingly. The label "not damping," shown in Figure 6, demonstrates that the PZT actuator is not activated. Thus, it refers to the typical features of the undamped system with flexible connections. The Y-direction movement of the end-effector, which should retain the Y position on 0 m, is also shown in Figure 7. The damping also damped the coordinate oscillations.

Also, to determine the effects of different acceleration values, the proposed path was tested using accelerations two and four times faster than the determined amount, which is exhibited in Figure 8 (i.e., the end-effector moves 4 mm and 8 mm in the same time frame in the second and third scenario). Figure 9 shows the tip deformation of every flexible link, w_i, confirming the prominent role of the PZT actuator when vibrating the links structurally as the structural vibrations are damped thoroughly after 60 msec. As reflected, the increase of acceleration puts more vibration on the manipulator, resulting in more deflection of flexible links, which were fully dampened by the active damping method in all three scenarios.

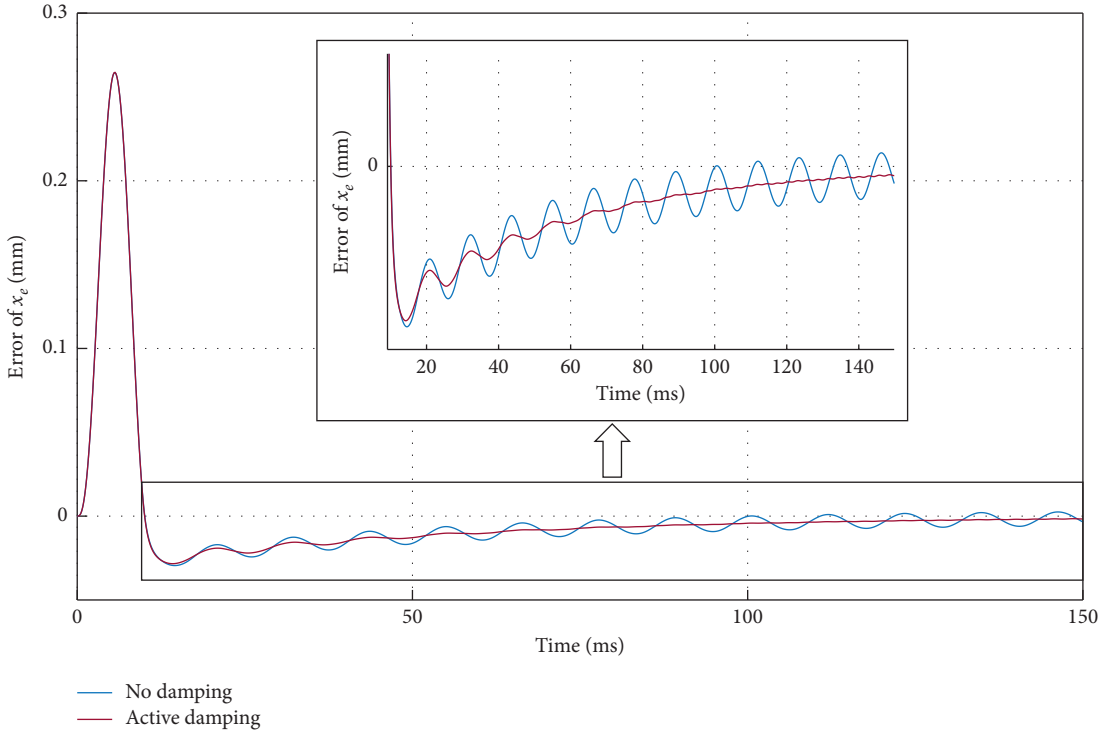


FIGURE 6: End effector's tracking error in X-direction between desired and actual path (blue line refers to the undamped system where red line refers to active PZT damping).

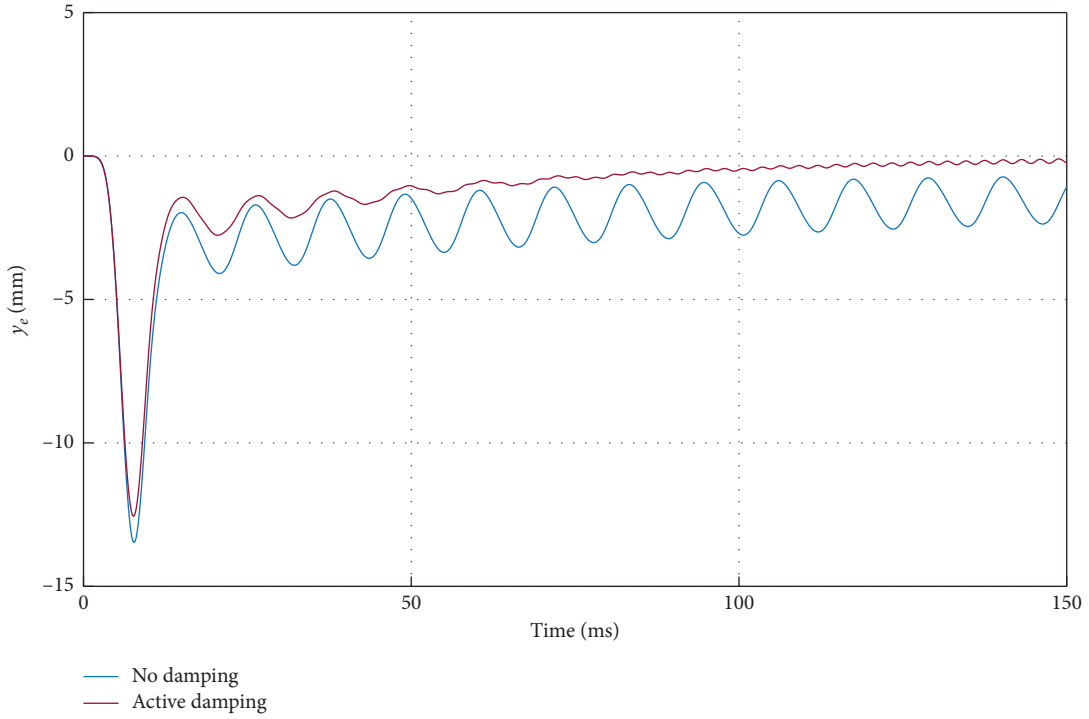


FIGURE 7: Y-directional movement of the end-effector.

As discussed, the dynamic model for the manipulator is studied only by the first three modes. To address this, we performed a power spectral density (PSD) analysis of the manipulator based on frequency-domain for the proposed trajectories with and without active damping; the results plotted in Figures 10–12 determine that the first mode has the

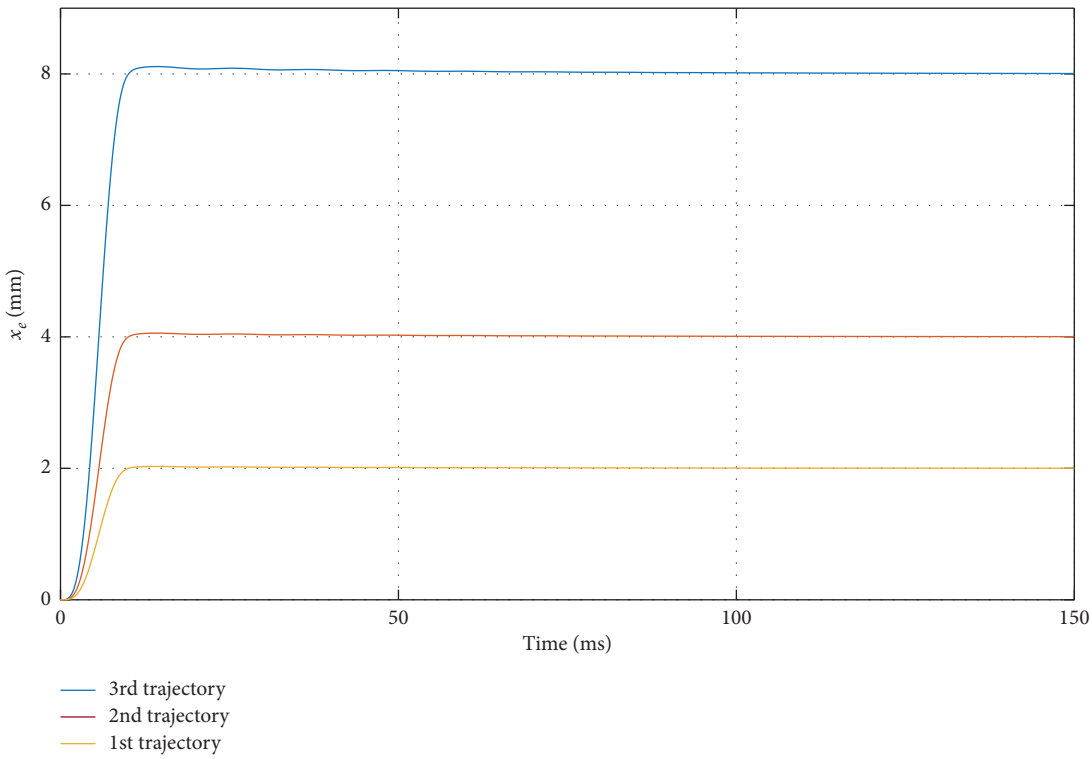
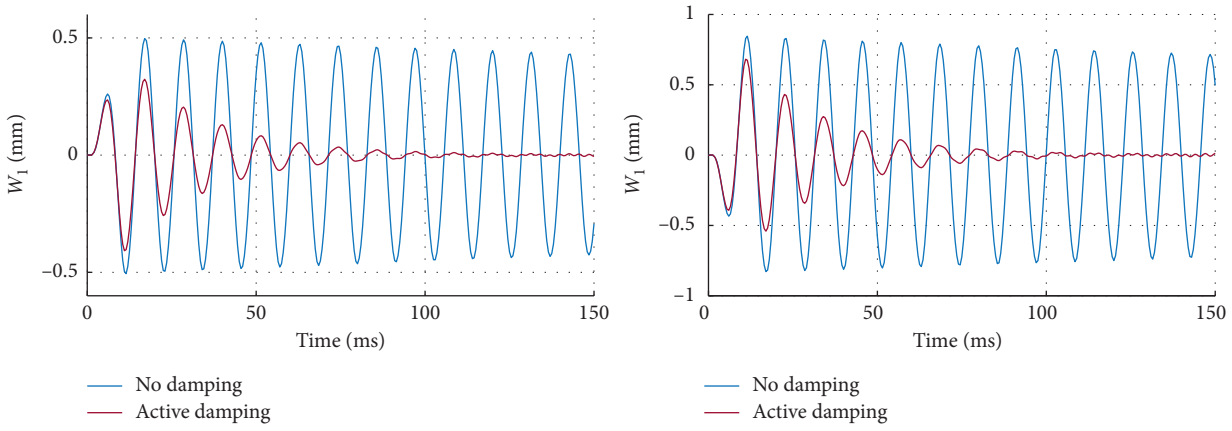


Figure 8: Accelerated trajectories (1st, 2nd, and 3rd trajectories move 2 mm, 4 mm, and 8 mm within 10 msec).



(a)

Figure 9: Continued.

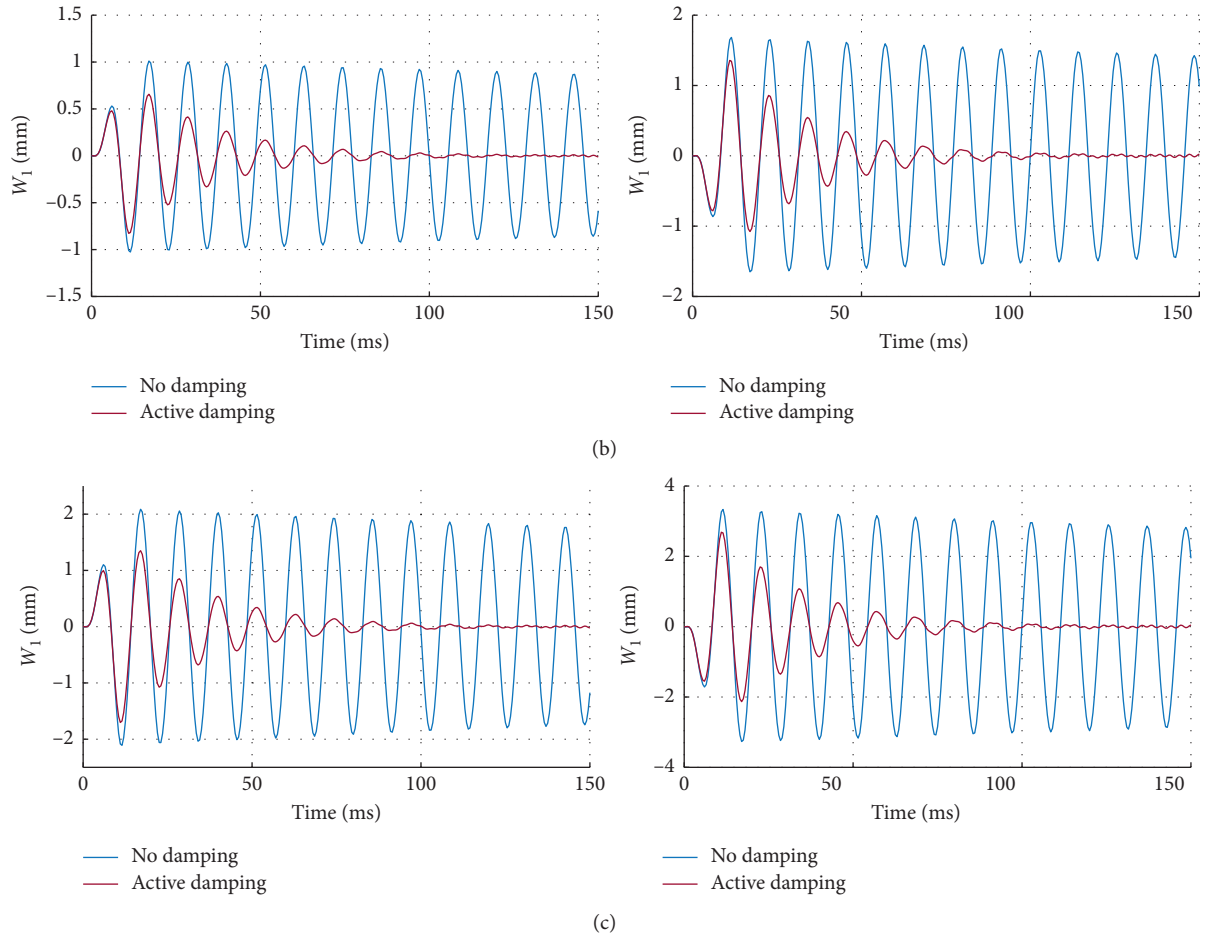


Figure 9: Flexible deformation of each links. (a) 1st trajectory. (b) 2nd trajectory. (c) 3rd trajectory.

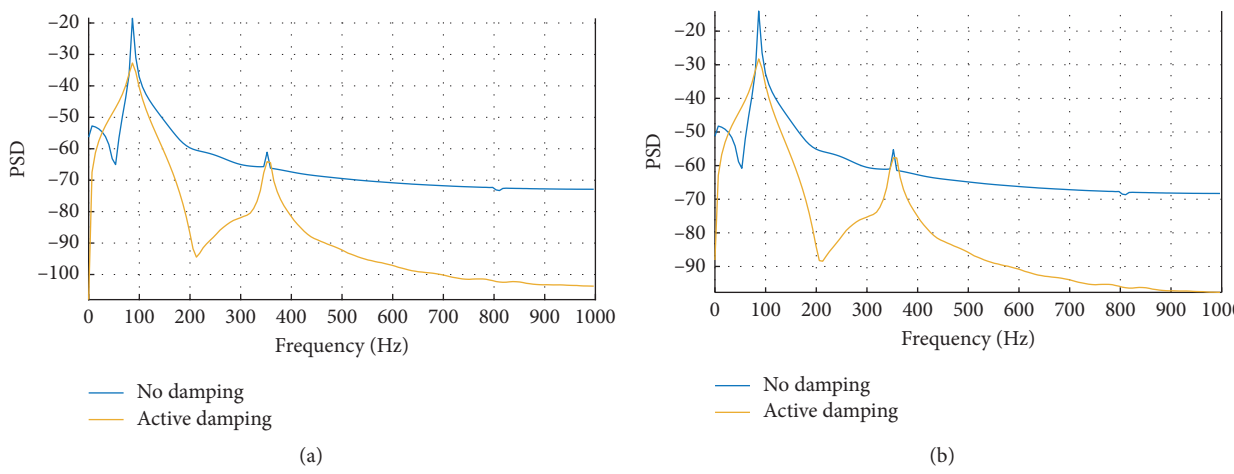


Figure 10: PSD of the first three modes of each link for the 1st trajectory. (a) 1st link. (b) 2nd link.

significant strength of the energy, while the variation at second and third modes are much weaker and had the same measures. Figures 10–12 also prove that the active damping method appropriately reduced the vibration in all given trajectories. This study's results match the research conducted by Zhang et al., in which the vibrations of an experimentally moving platform with flexible links are damped using the PZT actuators as an active control method [23]. Furthermore, well-established researches in this area obtained comprehensive damping performance by considering less than three modes of the manipulator [19, 24] and validated by experimental setups [25], which indicates the accuracy of our results.

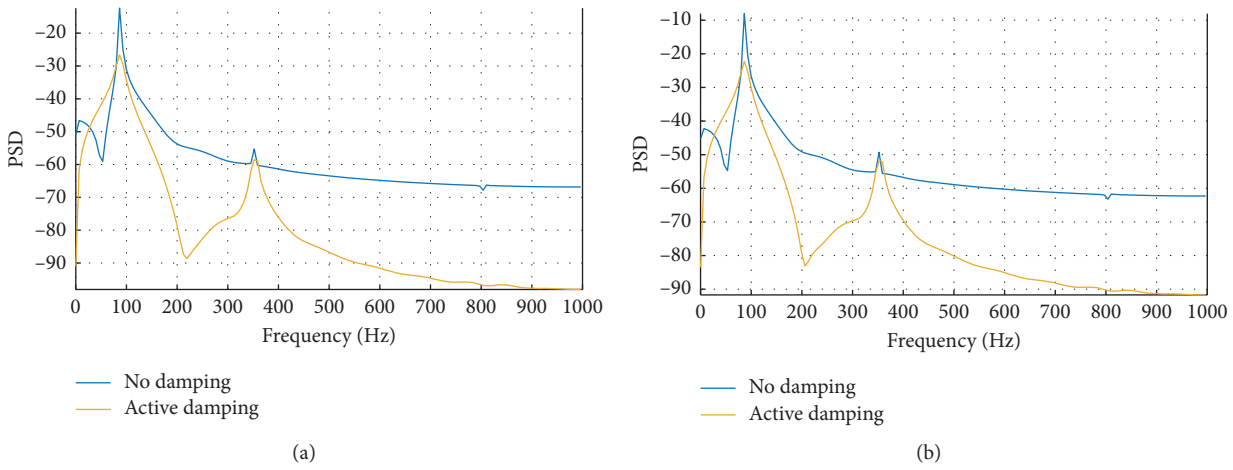


Figure 11: PSD of the first three modes of each link for the 2nd trajectory. (a) 1st link. (b) 2nd link.

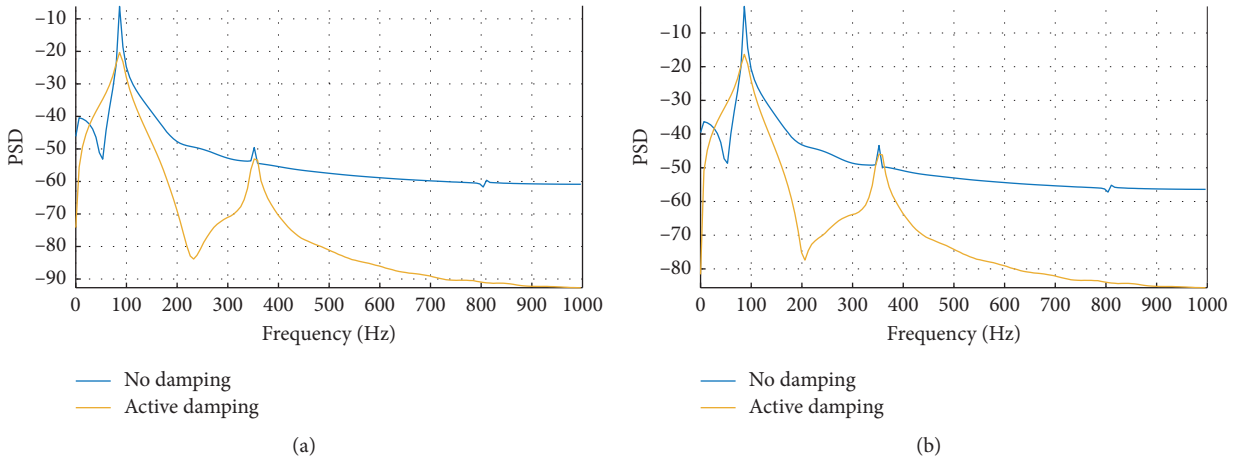


Figure 12: PSD of the first three modes of each link for the 3rd trajectory. (a) 1st link. (b) 2nd link.

6. Conclusions

2-DOF parallel manipulators containing lightweight and flexible materials are among the most discussed subjects by industrial scholars due to their extensive applications, such as pick and place tasks. In addition to the flexibility of robot links, working at high speeds makes it impossible to consider the deformation of the links as rigid. In this study, the deformation of the 2-PRRRP manipulators' links during accelerated trajectories is addressed by approaching the assumed mode method. Using the first-type Lagrangian equations, the dynamic equations of the 2-DOF flexible-link planar parallel manipulator are represented. The active damping method with the PZT actuators is considered to attenuate the structural vibration of flexible links. The piezoelectric materials can achieve a suitable damping performance with an L-type control strategy to counter structural vibration of flexible connections, which substantially reduces the time of setting of the end-effector leading to precise tracking of the trajectory. Relevant results indicate no requirement to use solid heavy links in 2-DOF parallel manipulators and spend less energy by considering them flexible.

Appendix

First-type Lagrangian equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_i}\right) - \frac{\partial (T - V)}{\partial q_i} = Q_i + \sum_{k=1}^{m} \lambda_k \frac{\partial \Gamma_k}{\partial q_i}. \tag{A.1}$$

Potential energy induced by link deformation:

$$U = \frac{1}{2} \sum_{i=1}^{2} E_i I_i \int_0^l \left(\frac{\partial^2 w_i(x)}{\partial x^2}\right)^2 \mathrm{d}x. \tag{A.2}$$

Equation of motion components:

$$M_{11} = (m_s + m_l)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$M_{12} = \left(\frac{m_l l}{2}\right)\begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix},$$

$$M_{22} = \left(\frac{m_l l^2}{3}\right),$$

$$M_{33} = m_p\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$M_{14} = \rho_A\begin{bmatrix} c_1\int\varphi_1 d\xi & c_1\int\varphi_2 d\xi & c_1\int\varphi_3 d\xi & 0 & 0 & 0 \\ 0 & 0 & 0 & c_2\int\varphi_1 d\xi & c_2\int\varphi_2 d\xi & c_2\int\varphi_3 d\xi \end{bmatrix},$$

$$M_{24} = \rho_A\begin{bmatrix} \int\varphi_1\xi d\xi & \int\varphi_2\xi d\xi & \int\varphi_3\xi d\xi & 0 & 0 & 0 \\ 0 & 0 & 0 & \int\varphi_1\xi d\xi & \int\varphi_2\xi d\xi & \int\varphi_3\xi d\xi \end{bmatrix},$$

$$M_{44} = \rho_A\begin{bmatrix} \widehat{M} & 0 \\ 0 & \widehat{M} \end{bmatrix},$$

$$\widehat{M} = \begin{bmatrix} \int\varphi_1^2 d\xi & 0 & 0 \\ 0 & \int\varphi_2^2 d\xi & 0 \\ 0 & 0 & \int\varphi_3^2 d\xi \end{bmatrix},$$

$$K = -EI\begin{bmatrix} \widehat{K} & 0 \\ 0 & \widehat{K} \end{bmatrix},$$

$$\widehat{K} = \begin{bmatrix} \int\varphi_1''^2 d\xi & 0 & 0 \\ 0 & \int\varphi_2''^2 d\xi & 0 \\ 0 & 0 & \int\varphi_3''^2 d\xi \end{bmatrix},$$

$$V_1 = \begin{bmatrix} -\rho_A\dot{\beta}_1 s_1\sum_{j=1}^{3}\dot{\eta}_{1j}\int\varphi_j d\xi - 0.5 m_l l\dot{\beta}_1^2 s_1 \\ -\rho_A\dot{\beta}_2 s_2\sum_{j=1}^{3}\dot{\eta}_{2j}\int\varphi_j d\xi - 0.5 m_l l\dot{\beta}_2^2 s_2 \end{bmatrix},$$

$$V_2 = \begin{bmatrix} \rho_A\dot{q}_1 s_1^2\sum_{j=1}^{3}\dot{\eta}_{1j}\int\varphi_j d\xi \\ \rho_A\dot{q}_2 s_2^2\sum_{j=1}^{3}\dot{\eta}_{2j}\int\varphi_j d\xi \end{bmatrix},$$

$$
V_4 = \begin{bmatrix}
-\rho_A \dot{q}_1 \dot{\beta}_1 s_1 \int \varphi_1 \mathrm{d}\xi \\
-\rho_A \dot{q}_1 \dot{\beta}_1 s_1 \int \varphi_2 \mathrm{d}\xi \\
-\rho_A \dot{q}_1 \dot{\beta}_1 s_1 \int \varphi_3 \mathrm{d}\xi \\
-\rho_A \dot{q}_2 \dot{\beta}_2 s_2 \int \varphi_1 \mathrm{d}\xi \\
-\rho_A \dot{q}_2 \dot{\beta}_2 s_2 \int \varphi_2 \mathrm{d}\xi \\
-\rho_A \dot{q}_2 \dot{\beta}_2 s_2 \int \varphi_3 \mathrm{d}\xi
\end{bmatrix},
$$

$$
J_1 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix},
$$

$$
J_2 = \begin{bmatrix} -ls_1 - w_1 c_1 & lc_1 - w_1 s_1 & 0 & 0 \\ 0 & 0 & -ls_2 - w_2 c_2 & lc_2 - w_2 s_2 \end{bmatrix},
$$

$$
J_3 = \begin{bmatrix} -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \end{bmatrix},
$$

$$
J_4 = \begin{bmatrix}
-s_1 \int \varphi_1 \mathrm{d}\xi & c_1 \int \varphi_1 \mathrm{d}\xi & 0 & 0 \\
-s_1 \int \varphi_2 \mathrm{d}\xi & c_1 \int \varphi_2 \mathrm{d}\xi & 0 & 0 \\
-s_1 \int \varphi_3 \mathrm{d}\xi & c_1 \int \varphi_3 \mathrm{d}\xi & 0 & 0 \\
0 & 0 & -s_2 \int \varphi_1 \mathrm{d}\xi & c_2 \int \varphi_1 \mathrm{d}\xi \\
0 & 0 & -s_2 \int \varphi_2 \mathrm{d}\xi & c_2 \int \varphi_2 \mathrm{d}\xi \\
0 & 0 & -s_3 \int \varphi_3 \mathrm{d}\xi & c_2 \int \varphi_3 \mathrm{d}\xi
\end{bmatrix},
$$

$$
F_q = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix},
$$

$$
F_{\text{ext}} = \begin{bmatrix} F_x \\ F_y \end{bmatrix},
\tag{A.3}
$$

where $c_i = \cos(\beta_i), s_i = \sin(\beta_i)$.

Voltage generated by PZT sensor:

$$
\begin{aligned}
V_s &= k_s \varepsilon \\
&= k_s \frac{\sigma}{E} \\
&= k_s \frac{MC/I}{E} \\
&= k_s \left(\frac{\partial^2 w(x,t)}{\partial x^2} \right).
\end{aligned}
\tag{A.4}
$$

Using strain rate feedback,

$$
\begin{aligned}
V_i(t) &= -k' \dot{\varepsilon} \\
&= -k' \frac{\dot{V}_s}{k_s} \\
&= -k_I \dot{w},
\end{aligned}
\tag{A.5}
$$

where $-k'$ is control gain applied to the PZT actuator.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] S. K. Dwivedy and P. Eberhard, "Dynamic analysis of flexible manipulators, a literature review," *Mechanism and Machine Theory*, vol. 41, no. 7, pp. 749–777, 2006.

[2] J. M. San Martins, Z. Mohamed, M. O. Tokhi, J. Sá da Costa, and M. A. Botto, "Approaches for dynamic modelling of flexible manipulator systems," *IEE Proceedings - Control Theory and Applications*, vol. 150, no. 4, pp. 401–411, 2003.

[3] M. Benosman and G. L. Vey, "Control of flexible manipulators: a survey," *Robotica*, vol. 22, no. 5, pp. 533–545, 2004.

[4] W. Z. Gebrehiwot, "Design and control of a five bar linkage parallel manipulator with flexible arms," MSc thesis, Politecnico Di Milano, Milan, Italy, 2009.

[5] H. Asada, Z. D. Ma, and H. Tokumaru, "Inverse dynamics of flexible robot arms: modeling and computation for trajectory control," *Journal of Dynamic Systems, Measurement, and Control*, vol. 112, no. 2, pp. 177–185, 1990.

[6] F. Wang and Y. Gao, *Advanced Studies of Flexible Robotic Manipulators: Modeling, Design, Control and Applications*, World Scientific Publishing, Singapore, 2003.

[7] X. J. Liu, J. Wang, and G. Pritschow, "On the optimal kinematic design of the PRRRP 2-DoF parallel mechanism," *Mechanism and Machine Theory*, vol. 41, no. 9, pp. 1111–1130, 2006.

[8] G. Piras, W. L. Cleghorn, and J. K. Mills, "Dynamic finite-element analysis of a planar high speed, high-precision parallel manipulator with flexible links," *Mechanism and Machine Theory*, vol. 40, no. 7, pp. 849–862, 2005.

[9] M. O. Tokhi, Z. Mohamed, and M. H. Shaheed, "Dynamic characterisation of a flexible manipulator system," *Robotica*, vol. 19, no. 5, pp. 571–580, 2001.

[10] M. O. Tokhi, Z. Mohamed, and A. K. M. Azad, "Finite difference and finite element approaches to dynamic modelling of a flexible manipulator," *Journal of Systems and Control Engineering*, vol. 211, no. 2, pp. 145–156, 1997.

[11] G. G. Hasting and W. J. Book, "A linear dynamic model for flexible robot manipulators," *IEEE Control Systems Magazine*, vol. 7, no. 1, pp. 61–64, 1987.

[12] A. K. M. Azad, *Analysis and Design of Control Mechanism for Flexible Manipulator Systems*, PhD Thesis, The University of Sheffield, Sheffield, UK, 1994.

[13] Z. Zhou, J. Xi, and C. K. Mechefske, "Modeling of a fully flexible 3PRS manipulator for vibration analysis," *Journal of Mechanical Design*, vol. 128, no. 2, pp. 403–412, 2006.

[14] X. Zhang, J. K. Mills, and W. L. Cleghorn, "Study on the effect of elastic deformations on rigid body motions of a 3-PRR flexible parallel manipulator," in *IEEE International Conference on Mechatronics and Automation*, pp. 1805–1810, Harbin, China, August 2007.

[15] J. Wu, T. Li, X. Liu, and L. Wang, "Optimal kinematic design of a 2-DOF planar parallel manipulator," *Tsinghua Science and Technology*, vol. 12, no. 3, pp. 269–275, 2007.

[16] B. Kang and J. K. Mills, "Dynamic modeling of structurally-flexible planar parallel manipulator," *Robotica*, vol. 20, no. 3, pp. 329–339, 2002.

[17] K. S. Fu, R. C. Gonzalez, and G. S. G. Lee, *Robotics: Control, Sensing, Vision and Intelligence*, McGraw-Hill, New York, NY, USA, 1987.

[18] M. W. Spong and M. Vidyasagar, *Robot Dynamics and Control*, J. Wiley, Chichester, UK, 1989.

[19] B. Kang and J. K. Mills, "Study on piezoelectric actuators in vibration control of a planar parallel manipulator," in *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pp. 1268–1273, 2003.

[20] M. O. Tokhi and A. K. M. Azad, *Flexible Robot Manipulators: Modelling, Simulation and Control*, IET, London, UK, 2008.

[21] M. H. Korayem and H. N. Rahimi, "Nonlinear dynamic analysis for elastic robotic arms," *Frontiers of Mechanical Engineering*, vol. 6, no. 2, pp. 219–228, 2011.

[22] D. San Sun and J. K. Mills, "PZT actuator placement for structural vibration damping of high speed manufacturing equipment," in *Proceedings of the American Control Conference*, pp. 1107–1111, San Diego, CA, USA, June 1999.

[23] X. Zhang, J. K. Mills, and W. L. Cleghorn, "Experimental implementation on vibration mode control of a moving 3-PRR flexible parallel manipulator with multiple PZT transducers," *Journal of Vibration and Control*, vol. 16, no. 13, pp. 2035–2054, 2010.

[24] X. Zhang, J. K. Mills, and W. L. Cleghorn, "Vibration control of elastodynamic response of a 3-PRR flexible parallel manipulator using PZT transducers," *Robotica*, vol. 26, no. 5, pp. 655–665, 2008.

[25] X. Zhang, J. K. Mills, and W. L. Cleghorn, "Dynamic modeling and experimental validation of a 3-PRR parallel manipulator with flexible intermediate links," *Journal of Intelligent and Robotic Systems*, vol. 50, no. 4, pp. 323–340, 2007.

Hindawi
Shock and Vibration
Volume 2021, Article ID 1483594, 12 pages
https://doi.org/10.1155/2021/1483594

Research Article

A Deep Learning Framework for Damage Assessment of Composite Sandwich Structures

Viviana Meruane,[1,2] **Diego Aichele**,[1] **Rafael Ruiz**,[3] **and Enrique López Droguett**[1]

[1]*Department of Mechanical Engineering, Universidad de Chile, Beauchef 851, Santiago, Chile*
[2]*Millennium Nucleus on Smart Soft Mechanical Metamaterials, Beauchef 851, Santiago, Chile*
[3]*Department of Civil Engineering, Universidad de Chile, Blanco Encalada 2002, Santiago, Chile*

Correspondence should be addressed to Viviana Meruane; vmeruane@ing.uchile.cl

Received 21 April 2021; Revised 10 June 2021; Accepted 21 June 2021; Published 1 July 2021

Academic Editor: Claudio Sbarufatti

The vibrational behavior of composite structures has been demonstrated as a useful feature for identifying debonding damage. The precision of the damage localization can be greatly improved by the addition of more measuring points. Therefore, full-field vibration measurements, such as those obtained using high-speed digital image correlation (DIC) techniques, are particularly useful. In this study, deep learning techniques, which have demonstrated excellent performance in image classification and segmentation, are incorporated into a novel approach for assessing damage in composite structures. This article presents a damage-assessment algorithm for composite sandwich structures that uses full-field vibration mode shapes and deep learning. First, the vibration mode shapes are identified using high-speed 3D DIC measurements. Then, Gaussian process regression is implemented to estimate the mode shape curvatures, and a baseline-free gapped smoothing method is applied to compute the damage images. The damage indices, which are represented as grayscale images, are processed using a convolutional-neural-network-based algorithm to automatically identify damaged regions. The proposed methodology is validated using numerical and experimental data from a composite sandwich panel with different damage configurations.

1. Introduction

The vibration characteristics of composite structures are sensitive to debonding. In particular, mode shape curvatures are extensively used to identify debonding regions in composite materials [1–4], and a greater damage localization accuracy is achieved as the number of measured degrees of freedom (DOFs) increases [1]. However, the number of DOFs that can be acquired simultaneously is largely restricted in conventional vibration measurement techniques. To overcome this limitation, high-speed digital image correlation (DIC) techniques have been implemented for full-field vibration measurements and damage assessment [3–5].

In general, the damage can be identified by examining the changes in the modes of the damaged structure with respect to the undamaged modes. However, modes from the damaged structures cannot always be matched to a corresponding "baseline" mode in the undamaged structure. This has driven the development of baseline-free damage-assessment algorithms, which include gapped smoothing (GS) [2, 6, 7] and wavelet-based [8] methods. In wavelet-based methods, a continuous or discrete wavelet transform is used to detect abrupt changes in the mode shape displacements or curvatures, which are related to damage. However, the accuracy of these methods is particularly sensitive to the family and order of the wavelets selected [9]. The GS method was initially proposed by Ratcliffe and Bagaria [6], who assumed that the undamaged mode shapes can be estimated using a smoothed version of the damaged mode shapes. Then, the damage indices are computed from the difference between the shapes of the undamaged and damaged curvature modes. This method has proven to be useful in different damage detection and localization applications, such as damage identification in

beams [10], beam-like structures [11], and plate-like structures [2, 4, 7, 12]. Yoon et al. [7] implemented the GS method with mode shape curvatures to assess damage in plate-like structures under the assumption that stiffness reductions are related to damage. Their method was successful in identifying delamination in experimental composite panels. Qiao et al. [2] investigated the application of three damage-assessment methodologies in composite laminates: the generalized fractal dimension, strain energy method, and GS. The experimental and numerical data of a composite panel with delamination were used to validate the proposed approach. The experimental panel was excited using lead-zirconate-titanate actuators, and the vibratory response was captured using a scanning laser vibrometer with polyvinylidene fluoride sensors. In this case, the best results were obtained using the GS method. A principle similar to that of the GS method was used by Rucevskis et al. [12] to detect damage in plates. In their implementation, the damage indices were formulated as the difference between the measured mode shape curvatures of the damaged and undamaged panels. The undamaged mode-shaped curvatures were estimated using a smooth polynomial version of the damaged modes. The algorithm was investigated using the simulated data of a panel under different damage scenarios, considering the damage size, measurement noise, and sensor distribution. Then, the experimental data of an aluminum panel with a cut as the damage were employed to validate the approach.

Second-order displacement derivatives required to determine mode shape curvatures are frequently computed using the central difference method, which greatly amplifies the experimental noise. Another approach to obtain derivatives without noise amplification is by using Gaussian process (GP) regression models [13], which are effective nonparametric regression techniques [14]. Meruane et al. [4] combined the GS method with curvature mode shapes estimated through GP regression. They demonstrated that GP regression allows to obtain noise-free mode shape curvatures from mode shape displacements with noise, thus improving the damage identification results compared to those using the conventional GS method.

Previous methods, such as wavelet-based or GS methods, have been implemented to calculate damage indices distributed over the surface of a structure. Given the damage indices, the range of damage index values corresponding to the damaged and undamaged states must be determined. This can be viewed as a semantic segmentation problem, where each pixel must be classified as damaged or undamaged. The most straightforward solution is to use a statistical approach [3] under the assumption that the damage indices in the undamaged regions follow a normal statistical distribution. Therefore, the outliers were considered as damage indices corresponding to statistically significant characteristics, such as damaged elements. Alternatively, automatic thresholding techniques can be implemented, which are frequently used for the automated visual inspection of defects. The valley-emphasis method has been demonstrated to be particularly effective for damage assessment [4]. Unlike the statistical approach, this method automatically performs image segmentation without requiring parameter tuning. Recently, the introduction of deep learning techniques has generated tremendous progress in semantic segmentation. In particular, convolutional neural networks (CNNs) have obtained remarkable results for image segmentation [15], mainly because of their structure. That is, a CNN extracts relevant features from the input images in an incremental manner with no need for domain expertise. This allows the identification of hidden relationships in the images not evident to the naked eye, in many cases exceeding human precision. However, the application of CNNs for the identification of delaminated regions in composite panels has not yet been investigated.

This article presents a novel damage-assessment algorithm for composite sandwich structures based on full-field vibration mode shapes and deep learning. First, the vibration mode shapes were identified from high-speed DIC displacement measurements. Then, the curvature mode shapes were computed using a GP regression, and a baseline-free GS method was applied to compute the damage indices. The damage indices, which are represented as grayscale images, were processed using a CNN-based algorithm to identify the damaged regions automatically. The proposed methodology was validated using numerical and experimental data from a composite sandwich panel under different damage scenarios. Furthermore, to highlight the advantages of our approach over existing methods, the results obtained were compared with those of a similar approach that uses an automatic thresholding technique instead of a CNN for image segmentation [4].

2. Estimation of Curvatures Using GP Regression

The use of the GS technique requires the estimation of the curvatures of the damaged plates. This task is generally realized by applying a finite difference technique on the identified experimental vibration mode, thus making it susceptible to noise in the vibration modes. In this study, the use of a GP to estimate the plate's curvature is motivated by two main reasons: (1) GP can clean the noise from each vibration mode and (2) offer a smooth estimation of the second derivative (with the use of a squared exponential kernel).

Let us define the grid point coordinates using vector $\mathbf{X} = [(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)]$and the measured mode shape displacements as $\phi_r = [\phi_r(x_1, y_1), \phi_r(x_2, y_2), \ldots, \phi_r(x_n, y_n)]$. The root mean square normalization is implemented as follows:

$$\varphi_r\left(x_i, y_j\right) = \phi_r\left(x_i, y_j\right)\sqrt{\frac{N_x N_y}{\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\phi_r^2\left(x_i, y_j\right)}}, \quad (1)$$

where N_x and N_y correspond to the number of grid points in the x and y directions, respectively, and $\varphi_r(x_i, y_j)$is the normalized r^{th} mode shape at points (x_i, y_j). Because the mode shape displacements include experimental noise, they can be expressed as

$$\varphi_r(x_i, y_i) = f(x_i, y_i) + \varepsilon, \tag{2}$$

where ε represents additive Gaussian noise, with a mean of 0 and variance σ_n^2. The mean mode shape displacements at points $(x_*, y_*)$are predicted as

$$\varphi_r^s(x_*, y_*) = \mathbf{k}_*^{\mathrm{T}}\mathbf{K}^{-1}\phi_r, \tag{3}$$

where $\mathbf{k}_*$ contains the kernel values between point $(x_*, y_*)$and grid points $\mathbf{X}$:

$$\mathbf{k}_*[i] = k((x_*, y_*), (x_i, y_i)). \tag{4}$$

Function k is the GP autocorrelation function. In this study, as in [4], a squared exponential kernel with additive noise was used:

$$k\left((x_i, y_i), \left(x_j, y_j\right)\right) = e^{-(1/2)\left(\left((x_i - x_j)^2/s_x^2\right) + \left((y_i - y_j)^2/s_y^2\right)\right)} + \sigma_n^2\delta_{ij}, \tag{5}$$

where $s_x^2$and $s_y^2$are the length scales that define the correlation between grid points, which ultimately drives the smoothness of the mode shape. Parameter σ_n^2accounts for the covariate noise and corresponds to the Kronecker delta. The selection of the squared exponential kernel is motivated by the need to have a GP that could be at least twice differentiable (to allow curvature estimation in damaged plates). $\mathbf{K}$ is the kernel matrix evaluated at the grid points and is defined as

$$\mathbf{K}[i, j] = k\left((x_i, y_i), \left(x_j, y_j\right)\right). \tag{6}$$

The mode shape curvatures are computed as

$$\begin{aligned}\frac{\partial^2\varphi_r^s(x_*, y_*)}{\partial x^2} &= \left[\mathbf{k}_{*xx}\right]^{\mathrm{T}}\mathbf{K}^{-1}\phi_r,\\ \frac{\partial^2\varphi_r^s(x_*, y_*)}{\partial y^2} &= \left[\mathbf{k}_{*yy}\right]^{\mathrm{T}}\mathbf{K}^{-1}\phi_r.\end{aligned} \tag{7}$$

Vectors $\mathbf{k}_{*xx}$and $\mathbf{k}_{*yy}$contain the second-order derivatives of the autocorrelation function evaluated at points $(x_*, y_*)$and grid point $\mathbf{X}$. Finally, the damage-assessment algorithm utilizes the mode shape Laplacian, which is formulated as

$$\nabla^2\varphi_r\left(x_i, y_j\right) = \frac{\partial^2\varphi_r^s\left(x_i, y_j\right)}{\partial x^2} + \frac{\partial^2\varphi_r^s\left(x_i, y_j\right)}{\partial y^2}. \tag{8}$$

3. GS Method

In the GS method, the undamaged mode shape curvatures are calculated using a smoothed version of the damaged mode shape curvatures (Laplacian). The undamaged mode shape curvatures are approximated using first-order base functions, as follows:

$$\nabla^2\varphi_r\left(x_i, y_j\right) = \mathbf{g}_{i,j}^{\mathrm{T}}\boldsymbol{\theta}_{i,j}, \tag{9}$$

where $\mathbf{g}_{i,j}$is a vector of base functions and $\theta_{i,j}$denotes its coefficients:

$$\begin{aligned}\mathbf{g}_{i,j}^{\mathrm{T}} &= \left[1, x_i, y_j\right],\\ \boldsymbol{\theta}_{i,j}^{\mathrm{T}} &= \left[a_0, a_1, a_2\right].\end{aligned} \tag{10}$$

Let us consider the neighboring points of (x_i, y_j); then, (9) can be expressed in the matrix form as follows:

$$\boldsymbol{\lambda}_r\left(x_i, y_j\right) = \mathbf{G}_r^{\mathrm{T}}\left(x_i, y_j\right)\theta_{i,j}, \tag{11}$$

where

$$\begin{aligned}\boldsymbol{\lambda}_r^{\mathrm{T}}\left(x_i, y_j\right) &= \left[\nabla^2\varphi_r\left(x_{i-1}, y_{j-1}\right), \nabla^2\varphi_r\left(x_i, y_{j-1}\right),\right.\\ &\qquad \left.\nabla^2\varphi_r\left(x_{i+1}, y_{j-1}\right), \ldots, \nabla^2\varphi_r\left(x_{i+1}, y_{j+1}\right)\right],\\ \mathbf{G}_r^{\mathrm{T}}\left(x_i, y_j\right) &= \left[\mathbf{g}_{i-1,j-1}, \mathbf{g}_{i,j-1}, \mathbf{g}_{i+1,j-1}, \ldots, \mathbf{g}_{i+1,j+1}\right].\end{aligned} \tag{12}$$

The coefficients are estimated using least squares resulting in

$$\widehat{\theta}_{i,j}^r = \left(\mathbf{G}_r^{\mathrm{T}}\left(x_i, y_j\right)\mathbf{G}_r\left(x_i, y_j\right)\right)^{-1}\mathbf{G}_r^{\mathrm{T}}\left(x_i, y_j\right)\boldsymbol{\lambda}_r\left(x_i, y_j\right). \tag{13}$$

These coefficients are used to calculate the undamaged mode shape curvature as

$$C_r\left(x_i, y_j\right) = \mathbf{g}_{i,j}^{\mathrm{T}}\widehat{\theta}_{i,j}^r. \tag{14}$$

The measure of damage at point $(x_i, y_j)$is estimated by the difference in the curvatures of the undamaged and damaged modes, represented by damage index d_r:

$$d_r\left(x_i, y_j\right) = \left|\nabla^2\varphi_r\left(x_i, y_j\right) - C_r\left(x_i, y_j\right)\right|. \tag{15}$$

Ultimately, this expression is expanded to consider the first m modes:

$$d\left(x_i, y_j\right) = \sum_{r=1}^{m} d_r\left(x_i, y_j\right). \tag{16}$$

4. Deep Learning and Semantic Segmentation

Deep learning models have shown excellent performance in various tasks involving image recognition and computer vision, such as image classification [16, 17], natural language processing [18], and image segmentation [15]. In particular, CNNs have been extensively used for image classification, where the network output to an image is a class label. This is achieved by arranging convolutional, pooling, and fully connected layers, as illustrated in Figure 1.

The convolution operation utilizes weight matrix $\mathbf{K}$, denominated as a filter or kernel, to obtain feature matrix $\mathbf{S}$ from input matrix $\mathbf{A}$ as follows:

$$\mathbf{S} = \mathbf{A} * \mathbf{K}, \quad \text{where } \mathbf{S}(i, j) = \sum_n\sum_m \mathbf{A}(i - m, i - n)\cdot\mathbf{K}(m, n). \tag{17}$$

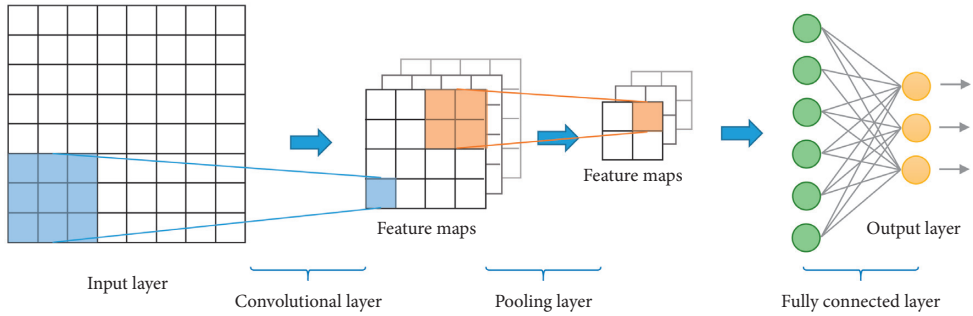


Figure 1: CNN architecture for image classification.

Then, the feature matrix is added to bias matrix **B**, and the activation function is employed to build the feature map **H**:

$$\mathbf{H} = f(\mathbf{A} * \mathbf{K} + \mathbf{B}). \tag{18}$$

The convolution layer applies various kernels and biases to the input matrix, and by operating convolutional layers sequentially, high-level features can be extracted. Pooling layers can then be used to reduce the number of features. For instance, a max-pooling layer provides only the maximum value of the next feature map within a rectangular cell. Finally, a feed-forward neural network located at the end of the CNN delivers the predicted class labels.

Considering the significant quantity of parameters in a CNN, precautions must be taken to prevent overfitting or overadjustment of the CNN to the training data. Overfitting results in an inadequate generalization; therefore, the network is unable to predict unseen cases. Regularization techniques, such as dropout [19] and batch normalization [20], can be implemented to prevent overfitting. In addition, the early stopping strategy, which stops training when the validation error begins to increase, helps prevent overfitting.

Although the most common application of a CNN is image classification, it has also been implemented for image segmentation. In image segmentation, the classification is performed pixel by pixel. Long et al. [21] were the first to introduce a fully CNN for image segmentation, in which the fully connected layers were replaced by convolutional layers. They used interpolation layers to guarantee that the output size equals the input size, which is essential for image segmentation. Ronneberger et al. [22] modified this architecture to allow training with fewer images, and the proposed architecture was named U-Net because of its shape, as illustrated in Figure 2.

The U-Net architecture is characterized by a contraction-expansion configuration. The contraction part is built by arranging convolutional layers using 3×3 kernels, rectified linear activation functions, and pooling layers. For a certain number of convolutional layers, max-pooling with stride 2 was implemented. The combination of a convolutional layer followed by max-pooling is a contraction step. In Figure 2, each contraction step is composed of two convolutional layers and one max-pooling layer; at each step, the number of channels (kernels) is doubled.

In the expansive part, the pooling layers are replaced by upsampling layers, which have the opposite purpose of pooling layers, thereby increasing the size of the input matrix. To increase the localization accuracy, features from the contracting part are joined to the features in the upsampled output. This is represented by the segmented lines in Figure 2. Finally, a 1×1 convolution layer is employed to transform the feature vectors to the required number of classes.

5. Damage-Assessment Methodology

The proposed damage-assessment methodology comprises the following steps:

(1) The experimental mode shapes are identified using a high-speed DIC system, as described in Section 5.3.

(2) The mode shape curvatures are estimated using a Gaussian regression process. The values of the length scale parameters ($s_x^2$and s_y^2) and the noise variance (σ_n^2)are the same as those used in [4] because the application case is the same: $s_x = s_y = 5d_x$and $\sigma_n^2 = 1$. As the GP is used merely to clean the noise from vibration modes and facilitate the curvature estimation, a robust selection of GP hyperparameters (s_x^2, s_y^2, σ_n^2) is considered unnecessary as long as the GP mean corresponds to the observations, i.e., no bias is introduced (which is demonstrated in [4] by studying the residual error). However, a new hyperparameter selection is recommended for new applications.

(3) Damage indices are obtained according to the procedure presented in Section 3.

(4) The damaged regions are identified using a CNN with a customized version of the U-Net architecture, which is presented in Section 6.1. The CNN was trained using a database created using a numerical model of the composite sandwich panel.

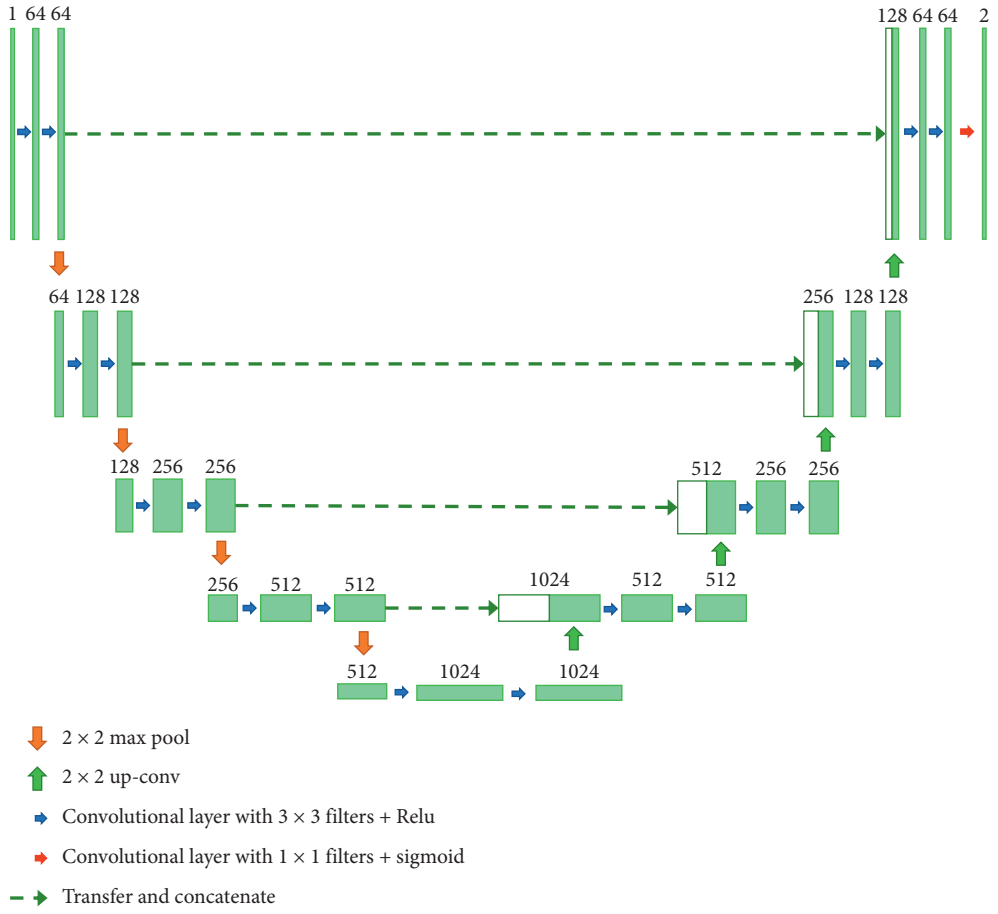


Figure 2: Scheme of the U-Net architecture.

This methodology was applied to identify debonding damage in an aluminum composite sandwich panel using numerical and experimental data. An automatic thresholding technique was used to contrast the image segmentation results obtained with the proposed approach.

The intersection over union (IoU) metric, which is widely used in image segmentation and object detection problems [23], was employed to evaluate the segmentation. In the damage identification problem, we have the true damaged region of the panel and the predicted damaged region, as illustrated in Figure 3. The true positives (TP) are defined as the intersection between both regions; false negatives (FN) correspond to the actual damage that was not detected, whereas the false positives (FP) were incorrectly detected damage. Considering this, the IoU metric is calculated as

$$\text{IoU} = \frac{\text{area of overlap}}{\text{area of union}} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}}. \quad (19)$$

5.1. Application Case. An aluminum honeycomb sandwich panel with dimensions 0.35 m × 0.25 m × 0.021 m was used in our case study. The skins are made of aluminum sheets with 0.8 mm thickness and the properties listed in Table 1,

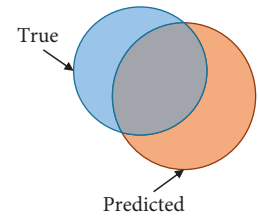


Figure 3: Representation of true and predicted damaged regions.

Table 1: Skin characteristics.

Thickness	0.8 mm
Elastic modulus	6.9×10^{10} Pa
Poisson's ratio	0.33
Density	2700 kg/m^3

Table 2: Core characteristics.

Cell size	19.1 mm
Foil thickness	5×10^{-5} m
Thickness	10 mm
Density	20.8 kg/m^3
Compressive strength	0.448 MPa
Longitudinal shear strength (σ_{xy})	0.345 MPa
Longitudinal shear modulus (G_{xy})	89.63 MPa
Transversal shear strength (σ_{yz})	0.241 MPa
Transversal shear modulus (G_{yz})	41.37 MPa

Figure 4: Experimental panel with speckle pattern.

Table 3: Experimental damage cases.

Case	Normalized damage size		Shape
	Damage 1	Damage 2	
1	0.09	—	Circular
2	0.12	—	Circular
3	0.14	0.07	Square
4	0.11	0.17	Circular

whereas the honeycomb core is made of aluminum with the characteristics presented in Table 2.

Figure 4 shows the experimental panel. For DIC measurements, the face of the panel was painted with a speckled pattern. The aluminum panel was built by bonding the skin to the honeycomb core with an epoxy resin. To introduce debonding damage to one of the skins, a region was intentionally left without an adhesive.

The panels were manufactured with four damage configurations, as listed in Table 3, describing the summary of the damage scenarios with the corresponding attributes, which include circular and square debonding damage shapes and a range of damage sizes. In the third and fourth cases, the panel has two debonded regions, whereas the first and second cases have one debonded region. The normalized damage size, which ranged from 0.07 to 0.17, is defined as the size of the damaged region (diameter or side length) divided by the diagonal length of the panel.

5.2. Numerical Model. The numerical model considers the composite panel as three layers of shell elements connected by linear springs. The exterior shells represent the skin, and the interior shell represents the honeycomb core. The springs act as the epoxy adhesive layer; therefore, the damaged region is represented as a zone with reduced spring stiffness. The model was built using the Structural Dynamics Toolbox (SDT) [24] using MATLAB®, and the layers were modeled with isotropic four-node shell elements (see Figure 5).

Experimental noise is always present in mode shapes identified from experimental data, which is why we decided to introduce noise artificially into the numerical mode shapes to make them similar to the experimental ones. In particular, the noise is introduced by adding a random sample to the mode amplitude at each grid point, where the samples are obtained from a Gaussian distribution with zero mean and standard deviation equal to 10% of the maximum mode amplitude. A database of 3500 panels with a range of damage scenarios was generated to train and evaluate the damage-assessment algorithms. The panels in the database had circular debonded regions with normalized damage sizes ranging between 0 and 0.25, and both the damage location and size were defined randomly.

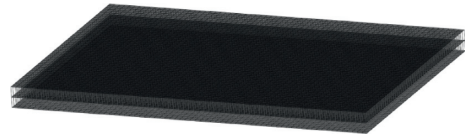


Figure 5: Numerical model representing the sandwich panel.

5.3. Experimental Setup and Measurements. Figure 6 presents the experimental setup, where the panel is suspended by elastic cords while it is excited by an electrodynamic shaker. The panel displacements were captured by two high-speed cameras connected to the DIC software. The DIC system is a Q450 high-speed DIC system manufactured by Dantec Dynamics. The acquisition frequency was 7530 fps and the picture resolution was 1 MP.

The experimental mode shapes are identified according to the following procedure:

(1) First, the natural frequencies of the panel are identified by an impact test

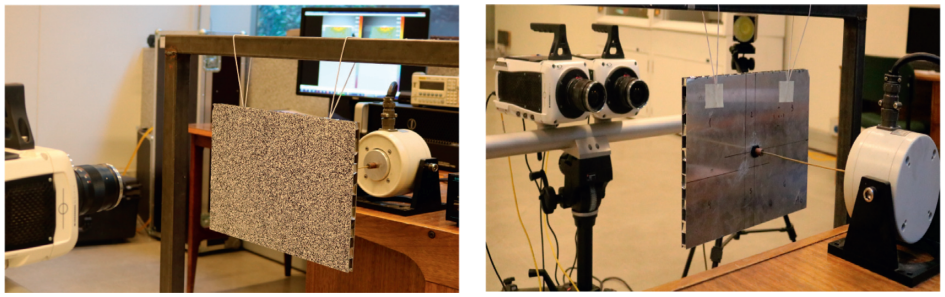
(2) The shaker is configured to vibrate with a sinusoidal signal at the natural frequency

(3) The panel vibration is recorded with the cameras, and the displacements are calculated using the DIC software

(4) The displacements are exported to MATLAB, and the operational mode shapes are identified

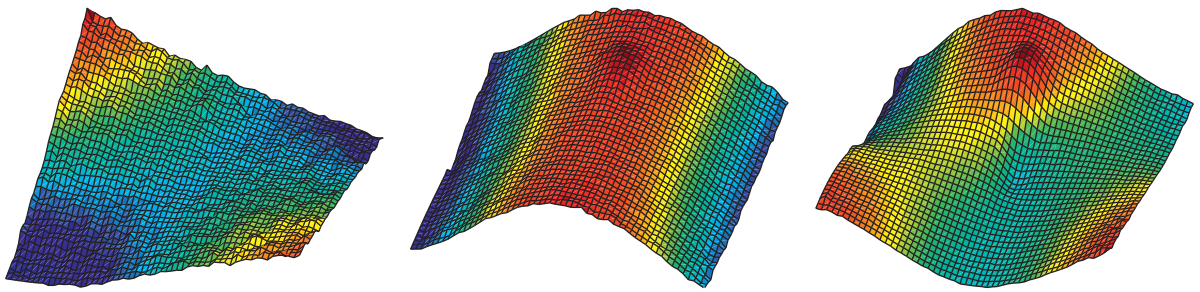
(5) Steps 2 to 4 are repeated for each natural frequency

Mode shapes with frequencies up to 2000 Hz were identified, and the number of experimental modes in this frequency range varied between 6 and 11 for each panel.



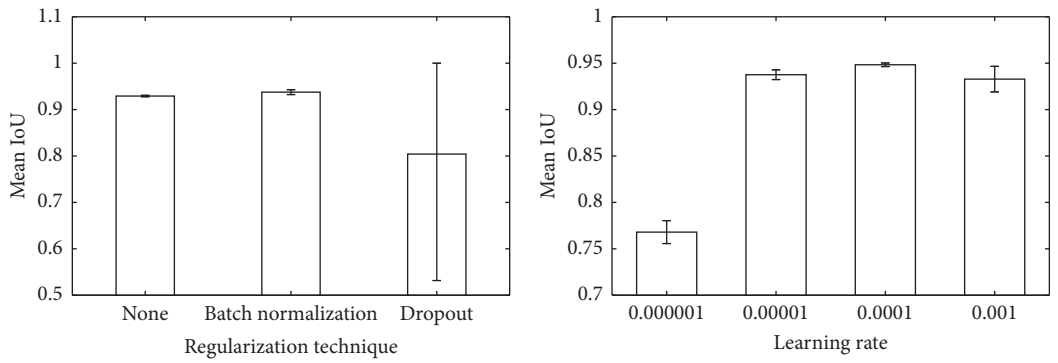
(a) (b)

Figure 6: Experimental setup: (a) panel with the speckle pattern and (b) shaker attachment.



(a) (b) (c)

Figure 7: Example of the first three experimental mode shapes. (a) 488 Hz. (b) 612 Hz. (c) 968 Hz.



(a) (b)

Figure 8: Validation performance with different (a) regularization strategies and (b) learning rates.

Figure 7 illustrates the first three experimental modes obtained for the first panel (Case 1).

6. Results

The numerical database was divided into training, validation, and testing sets. A total of 2800 panels were used for training, 175 for validation, and 525 for testing. The validation set was used during training and to tune the model, whereas the testing set was used to test the final model.

6.1. Optimization of Model Parameters. The first test was performed using the standard U-Net architecture, as described in Section 4. The algorithm was initially trained using the Adam optimizer, and the learning rate was set to 0.00001. The loss function was defined as (1 – IoU). An early stopping strategy was adopted with a validation patience of 50 epochs. Therefore, if the validation loss did not improve after 50 epochs, the training was stopped. To define the best regularization strategy, three cases were evaluated: no regularization, batch normalization, and

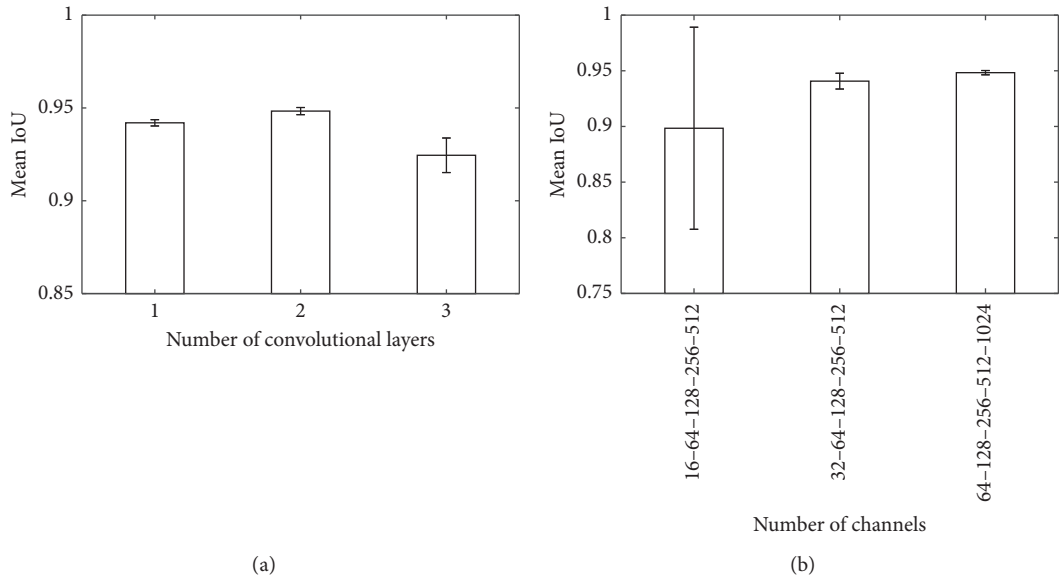


Figure 9: Validation performance with respect to the (a) number of convolutional layers and (b) number of channels.

Table 4: Segmentation model configuration.

Optimizer	Adam
Learning rate	0.0001
Regularization strategy	Batch normalization
Model architecture	U-Net
Number of channels	64–128–256–512–1024
Number of convolutional layers per step	2

dropout (20%). Because variability exists in each training process, each case was trained five times. Figure 8(a)shows the mean validation IoU obtained in each case, and the error bars represent the standard deviation. The best performance was obtained by batch normalization. Therefore, this regularization strategy was adopted. Next, a sensitivity analysis was performed with respect to the learning rate. The results shown in Figure 8(b)indicate that the best learning rate is 0.0001.

Finally, two additional sensitivity analyses were conducted. The first analysis explored the optimal number of convolutional layers at each step of the U-Net algorithm. In the second analysis, the optimal number of channels was investigated. The results are illustrated in Figure 9, and the final configuration of the deep learning segmentation model is summarized in Table 4.

6.2. Numerical Damage Assessment. Figure 10 shows the damage-assessment performance of the testing data as a function of the normalized damage size. The results were compared with the results obtained using an automatic thresholding method, as described in [4]. The results clearly indicate that by using a CNN for segmentation, the damage is identified with a significantly higher exactitude, which allows for the detection of smaller-sized damages.

Some examples of damage identified by both approaches are shown in Figure 11. The damage indices tended to increase at the edges of the panels. This effect is most clearly observed in cases with small damage sizes. Indeed, in cases with small or no damage, larger damage indices are at the edges. This causes the automatic thresholding method to identify damage incorrectly at the edges, but the CNN-based approach is capable of learning that these indices on the edges do not correspond to damage. Furthermore, the CNN is capable of detecting damages as small as a 0.05 normalized size. Damage of this size is not discerned by the human eye in the damage index image or by the thresholding methods.

The automatic thresholding method merely finds regions where the damage indices exceed a certain threshold and identifies those regions as damaged. In contrast, the CNN-based approach can learn different damage index patterns and identify whether an increase in damage indices corresponds to actual damage. This enables the identification of small damages that are not identifiable with other methods and prevents the detection of false damage.

6.3. Experimental Damage Assessment. To validate the approach with the experimental data, four experimental damage scenarios were considered, as listed in Table 1. The damage identified using the proposed CNN-based methodology is shown in Figure 12, whereas the results of the automatic thresholding method are presented in Figure 13. Table 5 summarizes the IoU obtained using the proposed approach compared to the IoU obtained using the automatic thresholding method. On average, the CNN approach performs better, although it is not a significant improvement. The main advantage of the CNN-based approach is that it does not detect false damages and can detect small damages. For example, the CNN-based approach correctly identified the smaller-sized damage in Case 3, but the

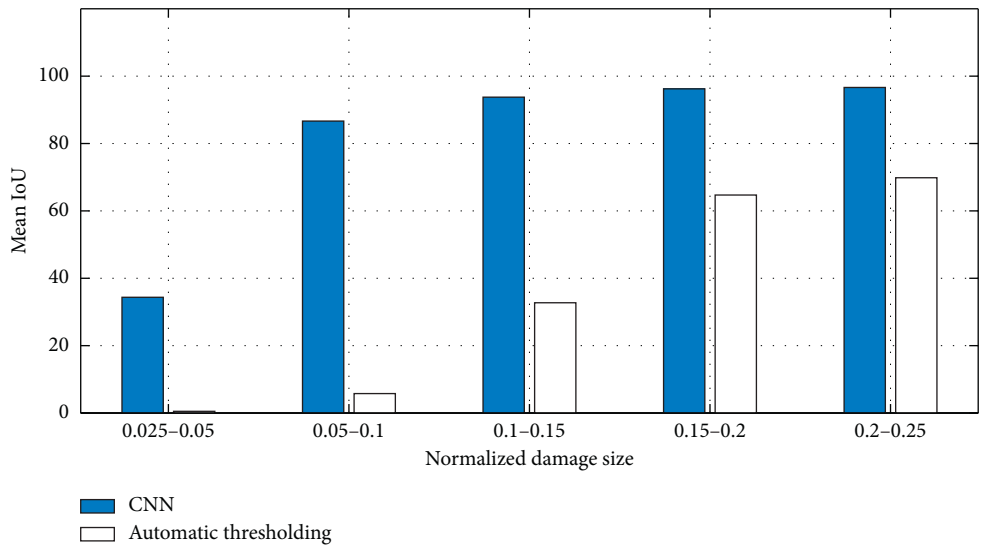


Figure 10: Performance of the damage-assessment methodologies.

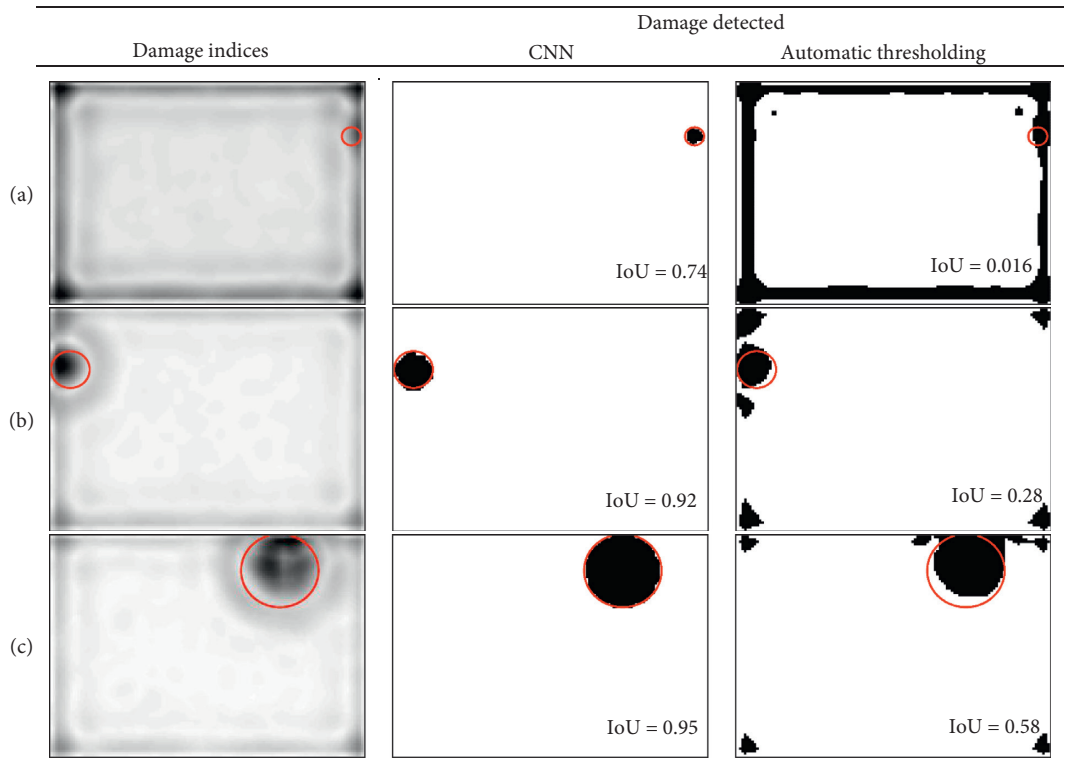


Figure 11: Damage detected using CNN and automatic thresholding method. Normalized damage sizes: (a) 0.05, (b) 0.1, and (c) 0.2. The red circles indicate the true damage.

automatic thresholding method did not. Thus, the automatic thresholding method performed satisfactorily in Case 3 because this model can catch the bigger damages very well, but it misses the small ones completely. Conversely, the CNN-based method is much better at identifying small damages. Furthermore, the experimental investigation indicates that the proposed approach is capable of correctly generalizing the numerical data because it accurately detects

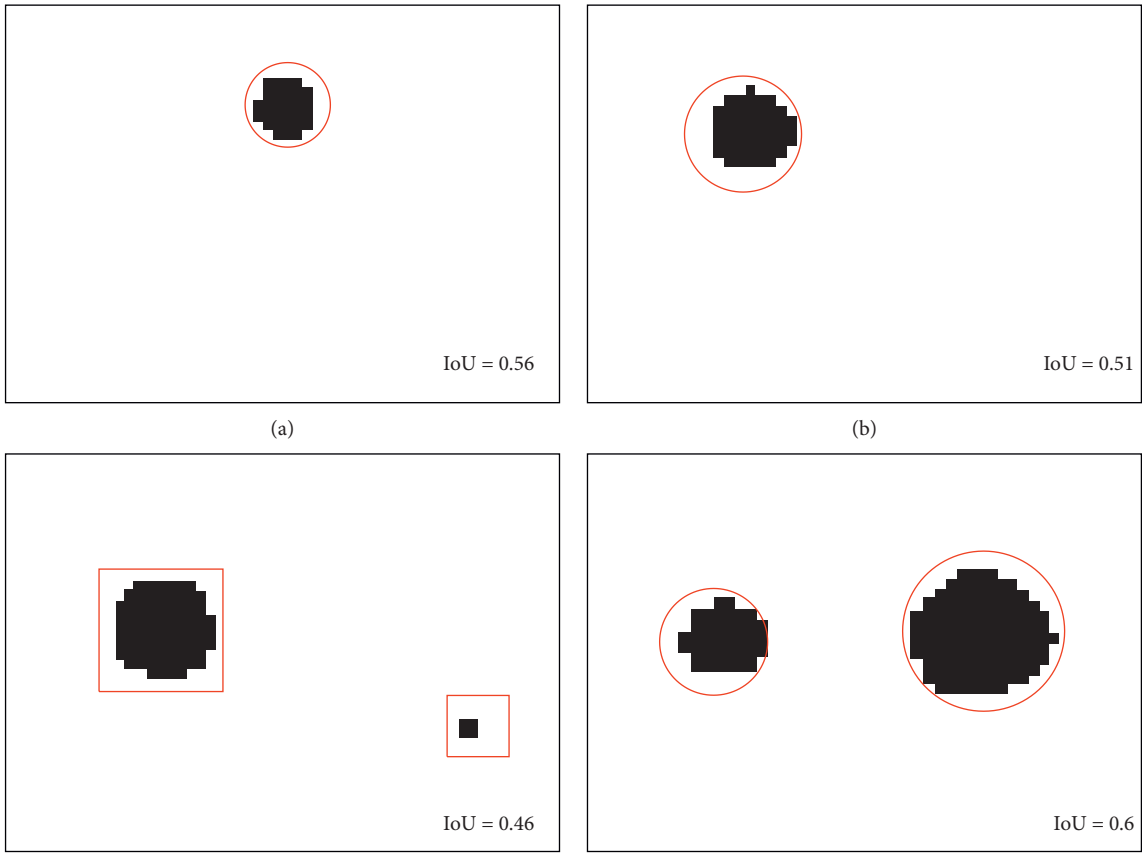


Figure 12: Experimental damage identified with the proposed CNN-based approach. Normalized damage sizes: (a) 0.09 (Case 1), (b) 0.12 (Case 2), (c) 0.14 and 0.07 (Case 3), and (d) 0.11 and 0.17 (Case 4). The red circles/squares denote the true damage region.

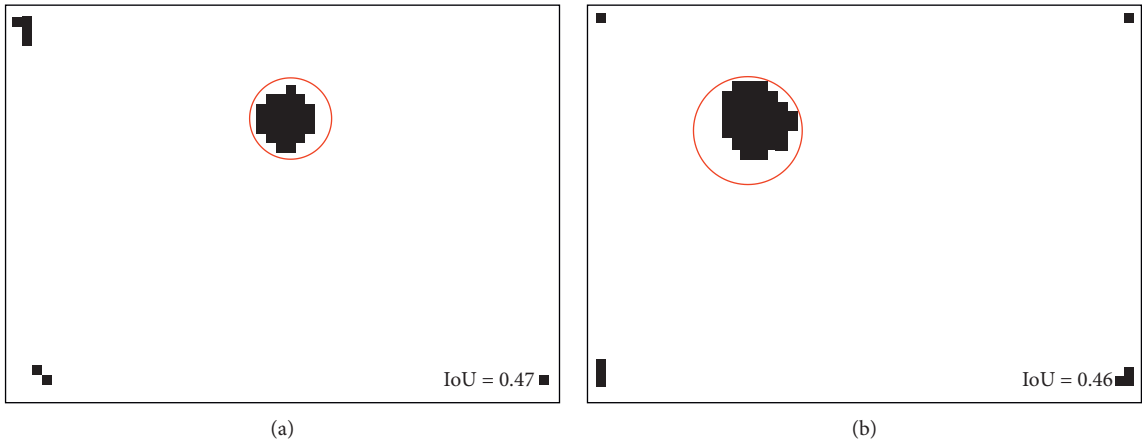
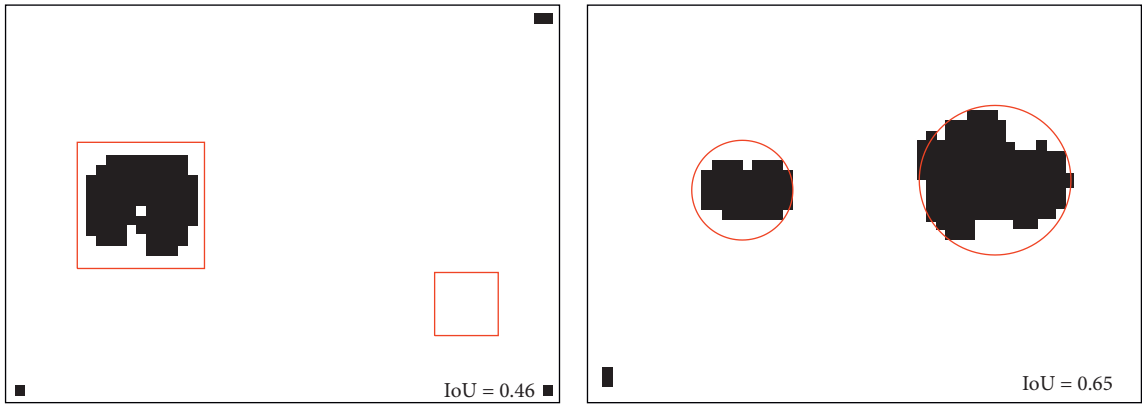


Figure 13: Continued.



(c) (d)

Figure 13: Experimental damage identified with an automatic thresholding method. Normalized damage sizes: (a) 0.09 (Case 1), (b) 0.12 (Case 2), (c) 0.14 and 0.07 (Case 3), and (d) 0.11 and 0.17 (Case 4). The red circles/squares denote the true damage region.

Table 5: Performance of the damage detected using CNN and automatic thresholding method.

Case	IoU	
	CNN	Automatic thresholding
1	0.56	0.47
2	0.51	0.46
3	0.46	0.46
4	0.60	0.65

experimental damage with precision despite having been trained using only data from the numerical model.

7. Conclusions

A damage-assessment methodology using full-field vibration modes and deep learning was developed and implemented to assess the debonding damage in composite sandwich structures. The main novelty of this approach is that the damage indices, represented as grayscale images, are processed using a CNN to automatically identify the damaged regions. The results showed that, compared with automatic thresholding methods, the CNN can better identify damaged regions with respect to IoU. In particular, the CNN enables the identification of smaller damages, significantly improving the results of existing approaches; this is essential from a practical perspective.

The proposed approach can learn different damage index patterns and correctly identify whether an increase in damage indices corresponds to actual damage. This advancement enables the identification of damage that is too small to be identified by other methods and prevents the detection of false damage. The results indicate that the proposed approach can correctly assess damages with normalized sizes greater than 0.05.

Although the experimental results are encouraging, the number of cases studied is not statistically significant, and therefore further experimental analysis is required. In addition, the proposed approach was validated using a simple sandwich plate structure. This structure does not necessarily represent a real structure with geometrical changes and different types of joints and edge conditions. Therefore, applications with more complex and realistic structures will be investigated in the future. In particular, the effects of different boundary conditions and geometrical changes on the damage identified must be analyzed. Because the approach searches for discontinuities in the structure, it is important to discriminate between changes caused by damage and variations due to geometry or boundary conditions.

Data Availability

Experimental and numerical data used in this article are available for download in the following website: http://www.lvmr.cl/des_en.htm.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The authors acknowledge the financial support provided by the Chilean National Fund for Scientific and Technological Development (FONDECYT) under grant nos. 1170535 and 1190720 and the Millennium Science Initiative of the Ministry of Economy, Development, and Tourism, grant "Millennium Nucleus on Smart Soft Mechanical Metamaterials."

References

[1] P. M. García, J. V. A. D. Santos, and H. Lopes, "A new technique to optimize the use of mode shape derivatives to localize damage in laminated composite plates," *Composite Structures*, vol. 108, no. 1, pp. 548–554, 2014.

[2] P. Qiao, K. Lu, W. Lestari, and J. Wang, "Curvature mode shape-based damage detection in composite laminated plates," *Composite Structures*, vol. 80, no. 3, pp. 409–428, 2007.

[3] F. Seguel and V. Meruane, "Damage assessment in a sandwich panel based on full-field vibration measurements," *Journal of Sound and Vibration*, vol. 417, pp. 1–18, 2018.

[4] V. Meruane, I. Fernandez, R. O. Ruiz, G. Petrone, and D. E. Lopez, "Gapped Gaussian smoothing technique for debonding assessment with automatic thresholding," *Structural Control Health Monitoring*, vol. 26, 2019.

[5] V. Meruane, M. Lasen, E. L. Droguett, and A. O Bernardin, "Modal strain energy-based debonding assessment of sandwich panels using a linear approximation with maximum entropy," *Entropy*, vol. 19, no. 11, pp. 619–622, 2017.

[6] C. P. Ratcliffe and W. J. Bagaria, "Vibration technique for locating delamination in a composite beam," *AIAA Journal*, vol. 36, no. 6, pp. 1074–1077, 1998.

[7] M. K. Yoon, D. Heider, J. W. Gillespie, C. P. Ratcliffe, and R. M. Crane, "Local damage detection using the two-dimensional gapped smoothing method," *Journal of Sound and Vibration*, vol. 279, no. 1-2, pp. 119–139, 2005.

[8] C. C. Chang and L. W. Chen, "Damage detection of a rectangular plate by spatial wavelet based approach," *Applied Acoustics*, vol. 65, no. 8, pp. 819–832, 2004.

[9] A. Katunin, "Stone impact damage identification in composite plates using modal data and quincunx wavelet analysis," *Archives of Civil and Mechanical Engineering*, vol. 15, no. 1, pp. 251–261, 2015.

[10] C. P. Ratcliffe, "Damage detection using a modified laplacian operator on mode shape data," *Journal of Sound and Vibration*, vol. 204, no. 3, pp. 505–517, 1997.

[11] A. K. Pandey, M. Biswas, and M. M. Samman, "Damage detection from changes in curvature mode shapes," *Journal of Sound and Vibration*, vol. 145, no. 2, pp. 321–332, 1991.

[12] S. Rucevskis, R. Janeliukstis, P. Akishin, and A. Chate, "Mode shape-based damage detection in plate structure without baseline data," *Structural Control and Health Monitoring*, vol. 23, no. 9, pp. 1180–1193, 2016.

[13] C. E. Rasmussen, "Gaussian processes in machine learning," in *Advanced Lectures on Machine Learning*Springer, Berlin, Germany, 2004.

[14] J. Ko and D. Fox, "GP-BayesFilters: bayesian filtering using Gaussian process prediction and observation models," *Autonomous Robots*, vol. 27, no. 1, pp. 75–90, 2009.

[15] Y. Guo, Y. Liu, T. Georgiou, and M. S. Lew, "A review of semantic segmentation using deep neural networks," *International Journal of Multimedia Information Retrieval*, vol. 7, no. 2, pp. 87–93, 2018.

[16] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: a comprehensive review," *Neural Computation*, vol. 29, no. 9, pp. 2352–2449, 2017.

[17] C. Modarres, N. Astorga, E. L. Droguett, and V. Meruane, "Convolutional neural networks for automated damage recognition and damage type identification," *Structural Control and Health Monitoring*, vol. 25, no. 10, 2018.

[18] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing [review article]," *IEEE Computational Intelligence Magazine*, vol. 13, no. 3, pp. 55–75, 2018.

[19] S. Wager, S. Wang, and P. S. Liang, "Dropout training as adaptive regularization," *Advances in Neural Information Processing Systems*, Stanford University, Stanford, CA, USA, 2013.

[20] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," 2015, https://arxiv.org/abs/1502.03167.

[21] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440, Boston, MA, USA, June 2015.

[22] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Munich, Germany, October 2015.

[23] M. A. Rahman and Y. Wang, "Optimizing intersection-over-union in deep neural networks for image segmentation," in *Proceedings of International Symposium on Visual Computing*, pp. 234–244, Las vegas, NV, USA, December-2016.

[24] E. Balmès, J. P. Bianchi, and J. M. Leclère, *Structural Dynamics Toolbox User's Guide*, SDTools, Paris, France, 2011.

Hindawi
Shock and Vibration
Volume 2021, Article ID 9972119, 7 pages
https://doi.org/10.1155/2021/9972119

Research Article

Lubrication State Recognition Based on Energy Characteristics of Friction Vibration with EEMD and SVM

Hai-jie Yu, Hai-jun Wei, Jing-ming Li, Da-ping Zhou, Li-dui Wei, and Hong Liu

Merchant Marine College, Shanghai Maritime University, Shanghai 201306, China

Correspondence should be addressed to Hai-jun Wei; haijun_welson@163.com

Received 11 March 2021; Revised 10 April 2021; Accepted 19 April 2021; Published 23 April 2021

Academic Editor: Franco Concli

In order to identify different lubrication states, lubrication experiments were carried out on a Bruker UMT-3 tester. The experimental results show that the frequency band energy characteristics of friction vibration signals are different under different lubrication states. Based on this, a lubrication state recognition method with ensemble empirical mode decomposition (EEMD) and support vector machine (SVM) was proposed. The vibration signals were decomposed into a finite number of stationary intrinsic mode functions (IMFs) with the EEMD method. The first six IMF components containing the main friction information were retained to calculate the energy ratio and construct the feature vector. The experimental results show that the mixed lubrication state can be identified by hundred percent, and there is a slight confusion between boundary lubrication and dry friction. The results show that frequency band energy of friction vibration signals is an effective feature to identify different lubrication states, and the proposed method can be used to identify different lubrication states.

1. Introduction

It is well known that friction vibration is caused by the relative movement of friction pairs. Friction vibration can reflect the characteristics and wear state of the friction system. Compared with friction coefficient, wear surface, and wear debris, friction vibration signals can be acquired online and in real-time without affecting the normal operation of the equipment. Real-time monitoring of equipment lubrication status is of great help to the efficient operation and maintenance of machinery and equipment.

As a typical nonlinear signal, the feature extraction of friction vibration has always been a difficulty in research. In recent years, the study mainly focuses on the qualitative analysis of friction vibration, but the quantitative analysis is rarely reported. Sun decomposed the vibration signals of reciprocating sliding friction pair by wavelet packet and analyzed the chaotic characteristics of the friction vibration [1]. Liu et al. analyzed the correlation between the frictional vibration in the normal and tangential directions and pointed out that the friction vibration in different directions has strong correlation [2]. Li et al. made multifractal analysis of the friction vibration signals in the running-in process [3].

In recently years, empirical mode decomposition (EMD) has been widely used to deal with nonlinear signal problems [4–7]. It adaptively decomposes signals into several stationary basic mode components by subtracting the local mean of signals in an iterative way. Because of the defects of the algorithm, false intrinsic mode functions may be generated in the decomposition results. Therefore, Huang proposed the ensemble empirical mode decomposition method (EEMD), which solves the modal aliasing problem of EMD by adding white noise components to the original signals to maintain the continuity of signals in different regions [8]. The most common method of EMD and its improved algorithm in mechanical fault diagnosis is to combine this method with other technologies to extract fault frequency [9–13]. However, there is no obvious periodicity in friction vibration and it is difficult to distinguish the lubrication state by the change of frequency. When the lubrication state changes, the energy in the same frequency band of the friction vibration signal will have a big difference. The signal energy in these frequency bands contains the main friction information, and the change of the signal energy in one or several frequency bands represents the change of the lubrication states. Therefore, the lubrication

states can be distinguished according to the change of frequency band energy.

Support vector machine (SVM) is an intelligent optimization algorithm proposed by Vapnik et al. in 1995 on the basis of statistical learning theory [14, 15]. Based on the principle of structural risk minimization, the SVM can solve practical problems such as small samples, high dimensions, nonlinearity, and local minimum points well. A large number of studies have shown that SVM parameters are the main factors affecting the performance of support vector machines. The differential evolution (DE) algorithm is an optimization algorithm of heuristic parallel random search based on floating-point vector coding proposed by Storn and Price in 1995 [16]. The principle of the algorithm is relatively simple with fewer control parameters. The algorithm has a strong global search ability and robustness and can improve the speed of optimization. In recent years, scholars have done a lot of work to improve the optimization performance of this method. In this paper, in order to improve the classification accuracy of support vector machine, the kernel parameters and penalty parameter of SVM are optimized by using the DE algorithm.

Based on the energy distribution characteristics of friction vibration, a lubrication state recognition method combining EEMD and SVM was proposed in this article. The vibration signals were decomposed by EEMD, and the energy ratio of the high-order intrinsic mode function (IMF) components containing the main friction information was calculated. Then, the standard mode feature vector was constructed with the energy ratio as the element, and the different lubrication states were identified by the SVM.

The innovations and main contributions of this paper are as follows:

(1) Frequency band energy is innovatively used as a characteristic to distinguish different lubrication states

(2) The IMFs are innovatively applied to construct the energy eigenvector of frequency band

(3) A lubrication state identification method based on EEMD and SVM is proposed to effectively monitor the lubrication state of equipment

The remainder of this paper is organized as follows: Section 2 presents the EEMD method, SVM method, DE algorithm, and the recognition method. The design and implementation of operational different lubrication states' tests are provided in Section 3. The results and discussion have been presented in Section 4. Finally, Section 5 would conclude the paper.

2. Method

2.1. EEMD. To solve the modal mixing problem of EMD, EEMD was invented on the basis of EMD algorithm in 2009 [8]. By adding white noise components to the original signal, the signals in different regions are kept continuous and the degree of mode aliasing is reduced. The decomposition steps are as follows:

(1) A Gaussian white noise is added to the original signal to produce a new signal, that is,

$$x_i(t) = x(t) + n_i(t), \tag{1}$$

where $x(t)$ is the original signal and $n_i(t)$ is the Gaussian white noise.

(2) The signal $x_i(t)$ is decomposed by EMD, that is,

$$x_i(t) = \sum_{j=1}^{J} c_{i,j}(t) + r_{i,j}(t), \tag{2}$$

where $c_{i,j}(t)$ is the j^{th} IMF decomposed after adding white noise for the i^{th} time and $r_{i,j}(t)$ is the remainder term.

(3) Repeat steps (1) and (2) for m times, and add white noise signals with different amplitudes to each decomposition to get the IMF set as follows: $\{c_{1,j}(t), c_{2,j}(t), \ldots, c_{m,j}(t)\ j = 1, 2, \ldots, J\}$.

(4) Based on the principle that the statistical mean value of unrelated sequences is zero, the average calculation is carried out for the abovementioned corresponding IMF:

$$c_j(t) = \frac{1}{m} \sum_{i=1}^{m} c_{i,j}(t), \tag{3}$$

where $c_j(t)$ is the j^{th} IMF decomposed by EEMD, $i = 1, 2, \ldots, m$, and $j = 1, 2, \ldots, J$.

2.2. SVM. SVM classification is a machine learning method based on statistical learning theory and structural risk minimization. It is applicable to both linearly separable and linearly nonseparable samples. For more details, please refer to [17].

With the sample (x_i, y_i) $(x_i \in R^d;\ y_i \in \{-1, +1\};\ i = 1, 2, \ldots, n)$, x_i is a feature vector and y_i is class label. If the sample is linearly separable, then SVM transforms the classification problem into a convex quadratic optimization problem, as shown in the following equation:

$$\left\{ \min \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n} \xi_i s.\text{t. } y_i[(\omega \cdot x_i) + b] \geq 1 - \xi_i, \right. \tag{4}$$

where ω is the weight vector, C is the penalty factor, ξ is the relaxation factor, and b is the bias.

The dual description of the above optimization problem is obtained by Lagrange operator. Under the condition that $y_i[(\omega \cdot x_i) + b] = 1$, the classification decision function can be obtained, that is,

$$f(x) = \operatorname{sgn}\left\{ \sum_i \alpha_i y_i (x_i \cdot x) + b^* \right\}, \tag{5}$$

where α_i is the Lagrange coefficient.

If the samples are linearly nonseparable, the samples in the input space can be mapped into the high-dimensional

linearly separable feature space by nonlinear mapping, and the optimal classification decision function in the feature vector can be obtained by kernel function, that is,

$$f(x) = \operatorname{sgn}\left\{\sum_i \alpha_i^* y_i K\left(x_i, x_j\right) + b^*\right\}, \quad (6)$$

where $K(x_i, x_j)$ is the kernel function.

2.3. Differential Evolution Algorithm. The basic idea of DE is to extract the search step and direction information from the current population and add random difference and crossover to improve the diversity of the population. After the above mutation and crossover operation, a temporary population is generated. Then, one-to-one selection of the two populations is carried out based on greedy thought to generate a new generation of population. The population evolves continuously according to the above method until the termination condition of the algorithm is satisfied. The details are described as follows:

(1) Four different individuals are randomly selected from the population to generate a difference vector to mutate the optimal individuals of each generation, which can not only improve the convergence speed of the algorithm but also maintain a high population diversity to a certain extent, that is,

$$v_i^{g+1} = x_{\text{best}}^{g+1} + k\left[\left(x_{s_1}^{g+1} - x_{s_2}^{g+1}\right) + \left(x_{s_3}^{g+1} - x_{s_4}^{g+1}\right)\right], \quad (7)$$

where v_i^{g+1} is the mutant individual obtained from each individual x_i^g of g generation by mutation operation, x_{best}^{g+1} is the best individual of $g+1$ generation, and k is the scaling factor.

(2) In order to improve the diversity of the population, the crossover operation mode is

$$y_i^{g+1} = \begin{cases} v_{i,j}^{g+1}, & \text{rand}(j) \le \text{CR}, \\ x_{i,j}^{g+1}, & \text{rand}(j) > \text{CR}, \end{cases} \quad (8)$$

where $\text{rand}(j)$ is a random value on [0, 1] and the crossover rate (CR) is a specified constant on [0, 1].

(3) If the values of the parameters exceed the corresponding bounds, they will be randomly and uniformly reinitialized within the given range. Then, the target function values of all test vectors are evaluated and selected. If the objective function value of the test vector is less than or equal to the objective function value of the corresponding objective vector, the next generation replaces the objective vector with the test vector. Otherwise, the target vector will be retained for the next generation. The selection operation can be represented as follows:

$$x_i^{g+1} = \begin{cases} y_i^{g+1}, f\left(y_i^{g+1}\right) < f\left(x_i^g\right), \\ y_i^{g+1}, f\left(y_i^{g+1}\right) < f\left(x_i^g\right), \end{cases} \quad (9)$$

where f is the target function.

2.4. Recognition Method. The collisions and breakage of the rough peak between the friction pairs are characterized by microimpact and random distribution on the contact interface, which stimulates the high-frequency dynamic response of the coupling system [18]. Therefore, the lower orders of IMF components should be removed and the higher orders of IMF components should be retained. As shown in Figure 1, the proposed lubrication state recognition method based on energy characteristics with EEMD and SVM is described as follows:

(1) Under the states of mixed lubrication, boundary lubrication, and dry friction, the samples are sampled several times at a certain sampling frequency to obtain enough samples.

(2) The collected vibration signals are decomposed by EEMD, and several IMF components are obtained.

(3) The energy ratio of the first n IMF components is calculated:

$$E_j = \int_{-\infty}^{+\infty} c_j^2(t)\mathrm{d}t \quad j = 1, 2, 3, \ldots, n,$$
$$T_j = \frac{E_j}{\sum_{j=1}^{n} E_j}, \quad (10)$$

where E_j is the energy of the j^{th} IMF component and T_j is the energy proportion of the j^{th} IMF component.

(4) The feature vector $T = [T_1, T_2, \ldots, T_j]$ is constructed.

(5) Lubrication states are classified and identified with SVM.

3. Experiment

3.1. Apparatus. The experiments were conducted on a piece of commercial equipment, Bruker UMT-3 tester, as illustrated in Figure 2. The equipment is able to control the disk rotational speed and the load. Friction force was measured by the sensors of the equipment. A triaxial acceleration sensor, fixed on the pin specimen (model 356B17ICP, PCB Piezotronics Company) with a range of ±5 g and a sensitivity of 1000 mv/g, was used to measure the vibration signal. A data acquisition system (VibPilot, $m+p$ international) was used to collect the data.

3.2. Experimental Method. In order to obtain the standard mode feature vectors of typical states, it is necessary to carry out experiments on the experimental bench to obtain the friction vibration signals under different lubrication states by controlling the amount of lubricating oil to change the lubrication states and according to the friction coefficient to judge the lubrication states. First, the experiment was carried out under the condition of oil lubrication. When the lubrication experiment lasted to the 50$^{\text{th}}$ min, the severe wear experiment without lubrication was carried out, at which dry

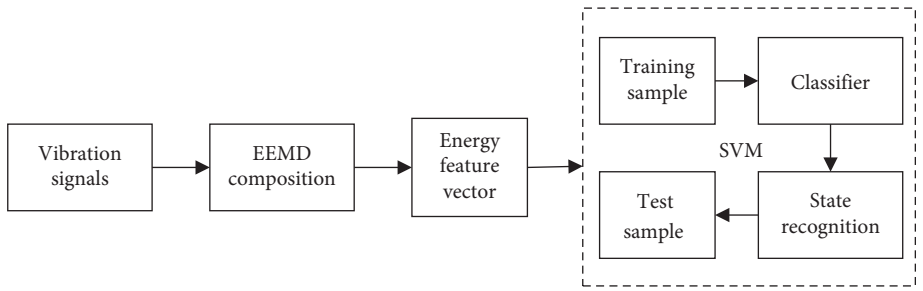


Figure 1: States' recognition flowchart.

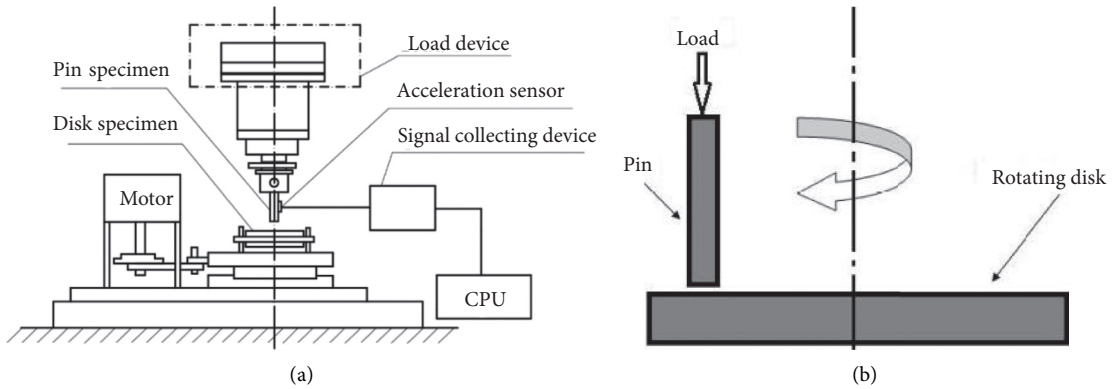


Figure 2: Schematic diagram of the UMT-3 tester. (a) Experiment device and (b) tribological pair.

absorbent cottons were used to clean the oil. Meanwhile, the vibration signals were collected by using the triaxial acceleration sensor with a sampling frequency of 51200 Hz and a sampling duration of 2 s per minute.

4. Results and Discussion

4.1. Different Lubrication States. The whole experimental process is shown in Figure 3. In the actual mechanical operation, the friction pairs are usually in the mixed lubrication state, and the friction coefficient in this state is about 0.1. In the beginning stage, the friction coefficient experienced a short decline and then entered a stable weak fluctuation state, corresponding to the running-in wear stage and the stable wear state, respectively, which were analyzed in detailed in [19]. Until the 50^{th} minute, the oil was sucked away, and there was only a very small amount of adsorbed oil film between the friction pairs. The lubrication state changed from the mixed lubrication to the boundary lubrication with the friction factor varied between 0.25 and 0.4. With the aggravation of wear, the oil films in the friction pairs were destroyed. The lubrication state changed from the boundary lubrication to the dry friction with large fluctuation of friction factor at the 70^{th} minute, under which the friction factor was greater than 0.4.

Figure 4 shows the spectrum diagram of the vibration signals in different lubrication states. As analyzed in the literature [18], the friction vibration can be divided into low-frequency and medium-high-frequency parts. The low-frequency part corresponds to the eigenfrequency and coupling frequency of the system. The medium-high-frequency part is closely related to the surface characteristics of the friction pair. The impact and fracture of the microconvex bodies are characterized by microimpact and randomly distributed on the contact interface, which stimulates the medium-high-frequency dynamic response of the coupling system. In the process of deterioration of lubrication state, the amplitude of high-frequency components increases continuously, which is very useful to state recognition.

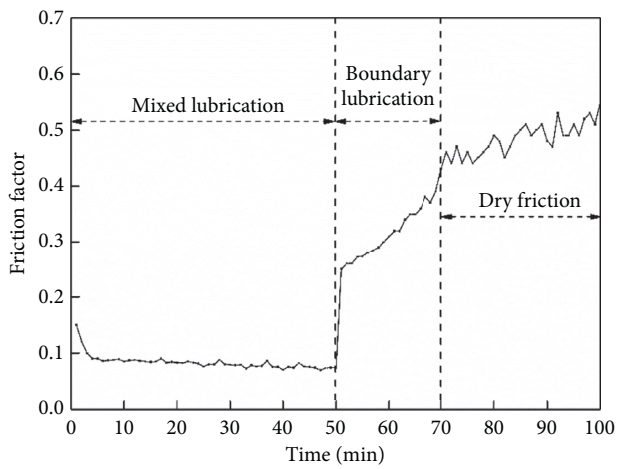


Figure 3: Temporal variation of friction factor in different lubrication states.

4.2. Energy Feature Vectors. The vibration signals collected during the experiment were decomposed by EEMD. The

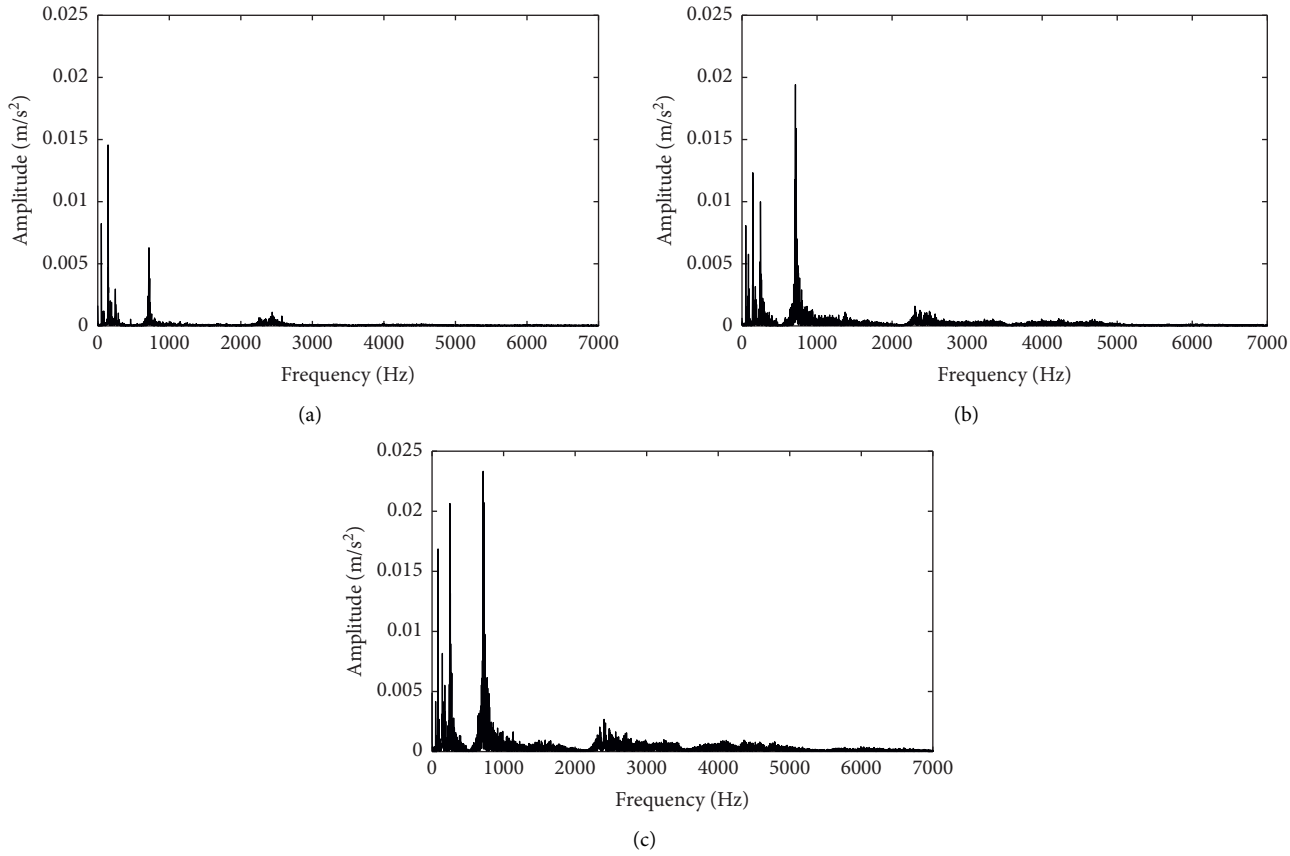


FIGURE 4: Spectrum diagram of the vibration signal in different lubrication states. (a) Mixed lubrication (20^{th} min), (b) boundary lubrication (60^{th} min), and (c) dry friction (80^{th} min).

ensemble average times and the amplitude coefficient of Gaussian white noise were 100 and 0.4, respectively [3]. Figure 5 shows the EEMD decomposition results of friction vibration signals in the dry friction (80^{th} min). The original signal was decomposed into 10 IMF components c_1–c_{10} and a remainder term r_{10}. The IMF components represent frequency band components of the original signal and are arranged in order from high to low frequencies. The remainder term is actually a trend line, that is, the wave with very low frequency (very long period), which can be regarded as the foundation of other IMF components. The characteristics cannot be directly seen from components and remainder term.

In this paper, the first 6 order components were retained to calculate the energy ratio and construct the feature vectors. Figure 6 shows the energy ratio of the friction pair under three states. As expected, the energy distribution in different states was markedly different.

4.3. Lubrication State Classification and Recognition with SVM. In order to illustrate the accuracy of the method proposed in this article, the 100 samples obtained from the experiment were divided into two groups. 50 samples were taken out as training samples, and the remaining 50 samples were taken as test samples. Considering the relatively large number of training samples and the small number of features, polynomial kernel was used as the kernel function of SVM [12]. Also, the differential evolution algorithm was used to obtain the optimal penalty parameter and kernel parameters.

The results of state recognition are shown in the confusion matrix in Figure 7. In this confusion matrix, all 25 samples in the mixed lubrication state were accurately identified. However, 2 out of 10 samples in the boundary lubrication state were incorrectly recognized as dry friction. For the 15 samples in the dry friction state, 3 samples were incorrectly recognized as boundary lubrication. Confusion

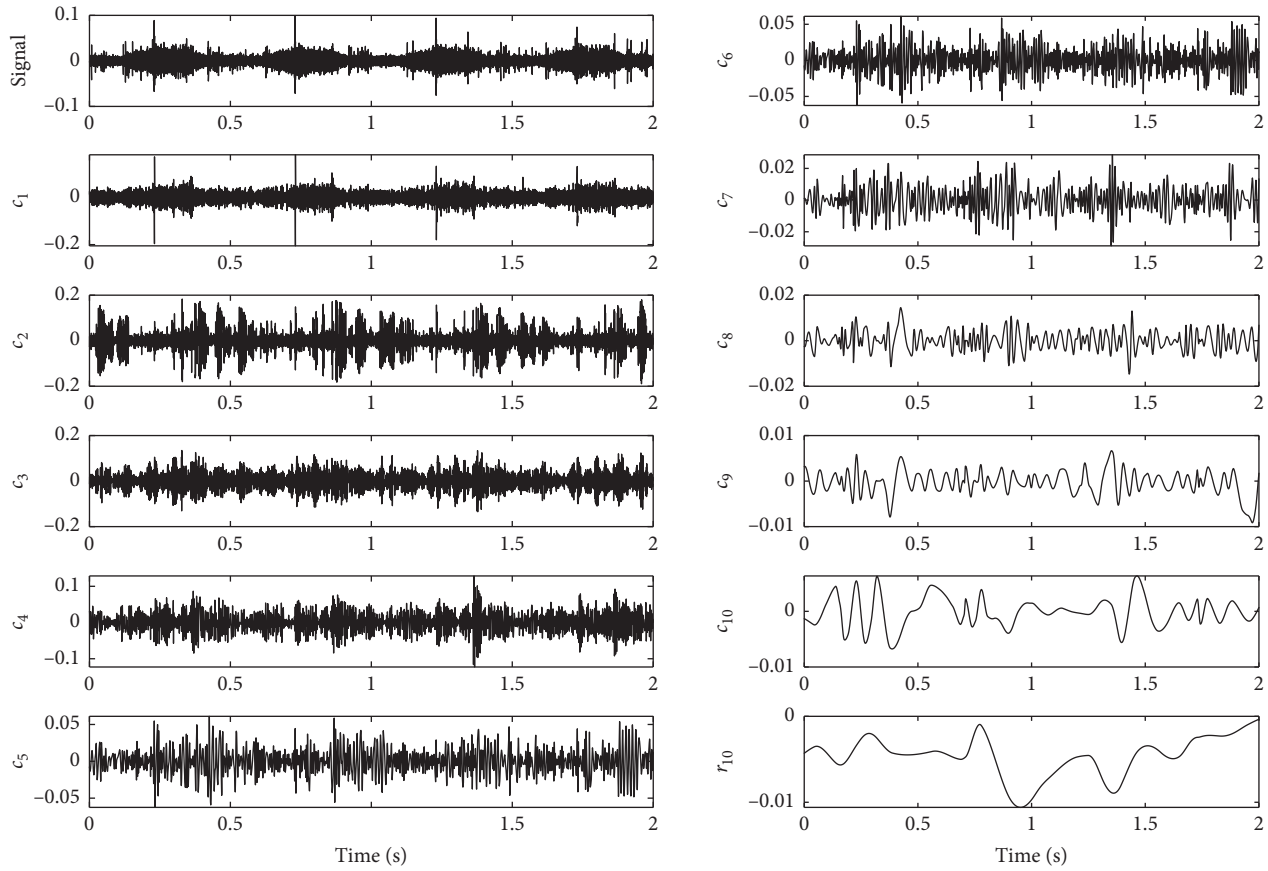


Figure 5: EEMD decomposition results of friction vibration signals in the dry friction.

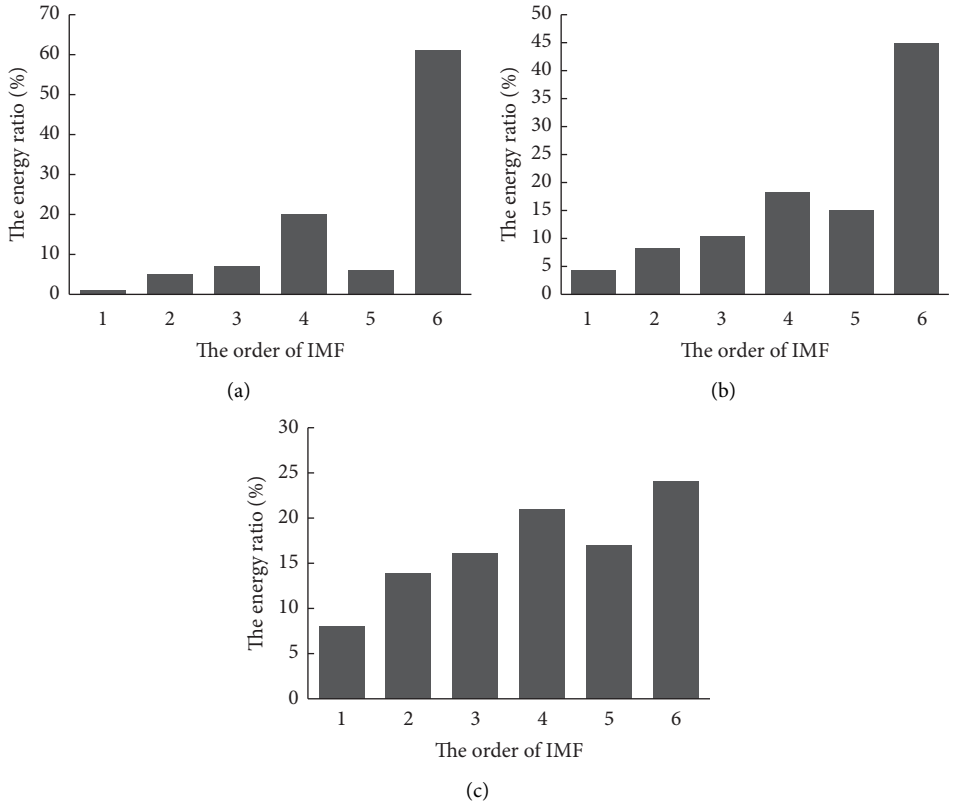


Figure 6: Energy ratio based on EEMD under three states. (a) Mixed lubrication (20^{th} min), (b) boundary lubrication (60^{th} min), and (c) dry friction (80^{th} min).

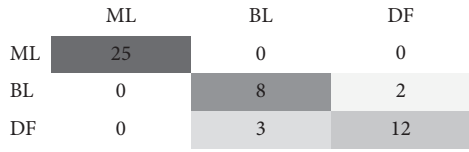


Figure 7: Confusion matrix for the results of state recognition (ML, mixed lubrication; BL, boundary lubrication; DF, dry friction).

between the mixed lubrication state and the dry friction state can partly be explained by high degree of similarity between the two states. The results suggest that the SVM method can identify different lubrication states well.

5. Conclusion

In this article, vibration signals under different lubrication conditions were obtained by using the friction testing machine. The friction vibration signals were decomposed by using the EEMD method, and the medium-high-frequency components containing the main friction information were retained. The energy ratio of IMF components was calculated, and the feature vectors were constructed to provide the basis for lubrication state identification. The SVM was used to recognize the different lubrication states of friction pairs. The results illustrate that the method presented in this work can accurately recognize different lubrication states. The main conclusions are as follows:

(1) The medium-high-frequency components of the friction vibration signal contained the lubrication state information, and the retained IMF components by EEMD decomposition can be used to construct the energy feature vectors under different lubrication states.

(2) Mixed lubrication states can be accurately identified by the SVM. However, there was confusion between boundary lubrication and dry friction, and further research is needed to improve the recognition accuracy.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was funded by the National High Technology Research and Development Program of China (Grant no. 2013AA040203) and the Shanghai Natural Science Foundation (Grant no. 17ZR1412700).

References

[1] D. Sun, G. Li, H. Wei, and H. Liao, "Experimental study on the chaotic attractor evolvement of the friction vibration in a running-in process," *Tribology International*, vol. 88, pp. 290–297, 2015.

[2] T. Liu, G. Li, H. Wei, and D. Sun, "Experimental observation of cross correlation between tangential friction vibration and normal friction vibration in a running-in process," *Tribology International*, vol. 97, pp. 77–88, 2016.

[3] J. Li, H. Wei, L. Fan, and L. Wei, "Multifractal detrended fluctuation analysis of frictional vibration signals in the running-in wear process," *Tribology Letters*, vol. 65, no. 2, 2017.

[4] K. Guo, X. Zhang, H. Li, and G. Meng, "Application of EMD method to friction signal processing," *Mechanical Systems and Signal Processing*, vol. 22, no. 1, pp. 248–259, 2008.

[5] Y. Lei, Z. He, and Y. Zi, "Application of the EEMD method to rotor fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing*, vol. 23, no. 4, pp. 1327–1338, 2009.

[6] Q. Gao, C. Duan, H. Fan, and Q. Meng, "Rotating machine fault diagnosis using empirical mode decomposition," *Mechanical Systems and Signal Processing*, vol. 22, no. 5, pp. 1072–1081, 2008.

[7] F. Wu and L. Qu, "Diagnosis of subharmonic faults of large rotating machinery based on EMD," *Mechanical Systems and Signal Processing*, vol. 23, no. 2, pp. 467–475, 2009.

[8] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2011.

[9] H. Wang, J. Chen, and G. Dong, "Feature extraction of rolling bearing's early weak fault based on EEMD and tunable Q-factor wavelet transform," *Mechanical Systems and Signal Processing*, vol. 48, no. 1-2, pp. 103–119, 2014.

[10] S. Zhao, L. Liang, G. Xu, J. Wang, and W. Zhang, "Quantitative diagnosis of a spall-like fault of a rolling element bearing by empirical mode decomposition and the approximate entropy method," *Mechanical Systems and Signal Processing*, vol. 40, no. 1, pp. 154–177, 2013.

[11] M. Žvokelj, S. Zupan, and I. Prebil, "Multivariate and multiscale monitoring of large-size low-speed bearings using ensemble empirical mode decomposition method combined with principal component analysis," *Mechanical Systems and Signal Processing*, vol. 24, no. 4, pp. 1049–1067, 2010.

[12] M. J. Zhang, K. Chai, J. Huang, and H. Chen, "Combined improved EEMD with SVM in the bearing low dimensional small sample fault diagnosis," *Applied Mechanics and Materials*, vol. 427-429, pp. 354–357, 2013.

[13] M. J. Zhang, J. Huang, K. Chai, and H. Chen, "Bearing binary classification intelligent diagnosis by combined improved EEMD with SVM," *Applied Mechanics and Materials*, vol. 341-342, pp. 1066–1070, 2013.

[14] V. D. Sánchez A, "Advanced support vector machines and kernel methods," *Neurocomputing*, vol. 55, no. 1-2, pp. 5–20, 2003.

[15] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, Berlin, Germany, 1995.

[16] R. Storn, "Differential evolution-a simple and efficient heuristic for global optimization over continuous space," *Journal of Global Optimization*, vol. 11, 1997.

[17] C. Nello, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge University Press, Cambridge, UK, 2000.

[18] M. Di Bartolomeo, G. Lacerra, L. Baillet, E. Chatelet, and F. Massi, "Parametrical experimental and numerical analysis on friction-induced vibrations by a simple frictional system," *Tribology International*, vol. 112, pp. 47–57, 2017.

[19] D. Sun, G. Li, H. Wei, H. Liao, and T. Liu, "Investigation on frictional vibration behavior of tribological pairs under different wear states," *Journal of Tribology*, vol. 137, no. 2, Article ID 021606, 2015.