

Research Article

The Research on Information Representation of Φ -OTDR Distributed Vibration Signals

Yanzhu Hu,¹ Song Wang,¹ Zhaoyang Wang,² and Yixin Zhang³

¹Automation College, Beijing University of Posts and Telecommunications, Beijing 100876, China

²Automation College, Beijing Institute of Technology, Beijing 100080, China

³Institute of Optical Communication Engineering, Nanjing University, Nanjing 210093, China

Correspondence should be addressed to Song Wang; wongsangwongsang@163.com

Received 25 June 2017; Revised 7 August 2017; Accepted 13 August 2017; Published 18 September 2017

Academic Editor: Taesun You

Copyright © 2017 Yanzhu Hu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper mainly focuses on the representable problem of Φ -OTDR distributed vibration signals. The research included a signal extraction part and a signal representation part. Firstly, in order to extract the better Φ -OTDR signal, the time-domain data should be fully preserved. The 2D-TESP method is used to extract data in this paper. There are 29 characters in the traditional TESP method. The characters' number is reduced from 29 to 13 and the characters' dimension is expanded from 1 to 2 in the 2D-TESP method. Secondly, in order to represent Φ -OTDR signal better, the characteristics of Φ -OTDR data and damped vibration signals are combined in the paper. The EMD method and the NMF method are combined to form the new method in the paper. Some parameters in the proposed method are optimized and adjusted by GA method. After Φ -OTDR data is represented by the proposed method, there is excellent performance both on the length dimension and on the time dimension. Lastly, some experiments are carried out according to the physical truth in this paper. The experiments are carried out in the semianechoic room. The methods of the paper have better performance. The methods are proved to be effective through these experiments.

1. Introduce

Φ -OTDR (phase sensitive optical time-domain reflectometer) was put forward by Taylor and Lee in 1993. It serves as a typical technique for monitoring distributed vibrations. At present, it has a wide range of applications in the field of large building structure, health monitoring [1], perimeter security of important places [2], and so on. However, we hope to grasp its monitoring state more accurately, understand its vibration mode better, and enhance the monitoring efficiency. Therefore, it is of great significance for a standard time-domain analytic expression of the Φ -OTDR data. The Φ -OTDR data is a distributed vibration signal. In order to standardize the time-domain analytic expression, we need to extract and represent useful signals.

In signal extraction, the feature extraction methods in time domain mainly include probability analysis method, time series method, correlation function analysis method, and time-domain waveform feature analysis. Potočnik and Govekar analyzed probabilistic statistics of vibration signals.

They combined probability statistical analysis, feature extraction, and principal component analysis. The combined method is used for evaluating the performance of multiple classifiers [3]. Delpha et al. also combined probability statistical analysis, feature extraction, and principal component analysis. The combined method is used for condition monitoring and fault diagnosis [4]. Ma et al. used the time series model to extract features of vibration signals. The fault signals and nonfault signals are identified effectively by the method [5]. Liu et al. used multiplex detection to eliminate trend correlation analysis. The feature extraction is achieved finally [6]. Alhazza analyzed the waveform of the signal. Alhazza used waveforms to control the system [7]. In a word, the time series, correlation function, and waveform analysis are all based on the condition of obvious signals. But for the condition of poor SNR (signal-to-noise ratio), the probability analysis method has incomparable advantages to other methods because of its unique statistical characteristics. In signal expression, some scholars focused on the NMF (nonnegative matrix factorization) method and

its expansion methods. Gao et al. combined TDF method and NMF method. The combined method is used to diagnose faults [8]. Li et al. combined the S transformation method, NMF method, mutual information method, and multiobjective evolutionary algorithm [9]. Li et al. also, respectively, combined generalized S transform method to NMF method and 2DNMF method [10, 11]. Some other scholars focused on the EMD (empirical mode decomposition) method and its expansion methods. Rai and Upadhyay combined EMD method and K -means method to process the signal [12]. Liu et al. put EMD method into the process of multiplex detection to eliminate trend correlation analysis [6]. In conclusion, NMF and the related methods pay more attention to data model. These researches proceed from the fitting curve of signal separation but ignore the research of vibration mechanism. EMD and the related methods pay more attention to mechanism model. These researches simply proceed from the vibration mechanism and the expression curve but ignore the research of data features.

A summary of the above signal extraction studies is presented. When the SNR is weak, there will be some errors in the extraction of signals. Φ -OTDR technique monitors the vibration of an optical fiber. The shape of optical fiber is a line. There will be weak SNR on the farther point of the optical fiber with the attenuation of signal propagation. Therefore, a signal extraction method which is more suitable for the condition of weak SNR is needed. On the other hand, it is more necessary for data driven and mechanism driven organic combination for the representation of Φ -OTDR distributed vibration signals. The combined method can not only reflect the real situation of data but also reflect the real state of vibration. This paper is divided into two parts. The first part uses the 2D-TESP method to extract the signal. The 2D-TESP method has a more appropriate coding interval and also takes into account the first derivative and the two derivatives. It has good effect on signal extraction under weak SNR. The second part uses the GAEMD-NMF method to express the signal. It combines the EMD method and the NMF method. Some parameters are optimized by GA (genetic algorithm) after the two methods are combined. It can better solve the relative standard time-domain expression problem of distributed vibration signals.

2. Time-Domain Signal Extraction Based on 2D-TESP Algorithm

Signal extraction aims at better signal representation. Signal extraction requires full preservation of signal data prototypes in the time domain. Because of the characteristics of Φ -OTDR technology, it is hoped that effective signals can be extracted on the point with weak SNR. Therefore, the 2D-TESP method is selected to extract the features of the signal.

2.1. TESP Algorithm. TESP (time encoded signal processing) algorithm has two main features [13–16]. The first feature is that the algorithm is based on the time domain which directly processes the signal. The second feature is to convert signals into probabilistic models that contain finite elements. Simply

put, the method is to reencode the signal in the time domain. Typical time-frequency domain analysis algorithms include FFT, WT, and HHT. These algorithms have a long operation time and are also unsuitable for the poor SNR. They are not suited for distributed vibration data characteristics. TESP algorithm has the advantages of intuitive, small computation and simple implementation process. The traditional TESP algorithm is suitable for the feature extraction of speech signals [13]. The TESP method was improved by Wang's team in 2014. They apply the method to feature extraction of sound signals [14]. Wang et al. used it to develop two-dimensional feature extraction. They applied the improved algorithm to signal feature extraction in a complex environment [15, 16]. The specific implementation steps of TESP algorithm are as follows.

(1) The window is divided into time-domain signals on each length node of the optical fiber. In each window, the zero crossing rate of the signal is calculated. The interval between two adjacent zero points is a time period. According to the rule of TESP algorithm, each time period is called meta.

(2) There are two indexes in the meta. One is the duration; it is usually expressed in D . The other one is the signal form; it is usually expressed in S . At the same time, the following information is obtained according to these two indexes:

- (a) The number of sampling points existing in each element is the duration.
- (b) In each element's duration, the data is derived one by one. The extreme number in each element is obtained.
- (3) The matrix is constructed by using D and S as two dimensions. Each element is coded according to the elements in the matrix.

(4) The probability of each code in the matrix is compiled statistically. Finally, the probability distribution is used as a feature to be inserted into a classifier or cluster.

2.2. Feature Extraction Based on 2D-TESP Algorithm. There are only 29 characters because of the limitations of the encoding principle in traditional TESP algorithms. At the same time, the traditional TESP method achieves higher recognition rate in recognition. This paper focuses on the study of damped vibration. The encoding of 29 characters has far exceeded the requirements of damped vibration in the Φ -OTDR technique. Therefore, the D - S matrix is extended to D - S_1 and D - S_2 matrices in the paper at first. Secondly, the coding principle for each matrix is reduced from 29 to 13. Lastly, the D - S_1 and D - S_2 matrices are encoded together with joint probability distribution statistics to form the A matrix.

2.2.1. The Reduced TESP Symbol Table. In the traditional TESP algorithm, S only represents the first derivative. There may be mutual influence between points in Φ -OTDR technology. Therefore, further analysis of the Φ -OTDR signal is needed. The concept of the inflection point of the two-order derivative is introduced in the traditional method. In this paper, the S matrix that is the extreme point is set as D - S_1 , and the S matrix that is the inflection point is set as D - S_2 .

TABLE 1: The standard 29-symbol encoding table of D -S1 matrix.

D	S					
	1	2	3	4	5	>5
1	1					
2	2	2				
3	3	3	3			
4	4	4	4	4		
5	5	5	5	5	5	
6	6	6	6	6	6	6
...
34	24	25	26	27	28	29
35	24	25	26	27	28	29

TABLE 2: The shrunk 13-symbol encoding table of D -S1 matrix.

D	S			
	1	2	3	>3
1	1			
2	1	1		
3	1	2	2	
4	2	2	3	
5	3	3	3	
6	3	3	4	4
...
34	11	12	13	13
35	11	12	13	13

TABLE 3: The standard 29-symbol encoding table of D -S2 matrix.

D	S					
	1	2	3	4	5	>5
3	1					
4	2	2				
5	3	3	3			
6	4	4	4	4		
7	5	5	5	5	5	
8	6	6	6	6	6	6
...
34	24	25	26	27	28	29
35	24	25	26	27	28	29

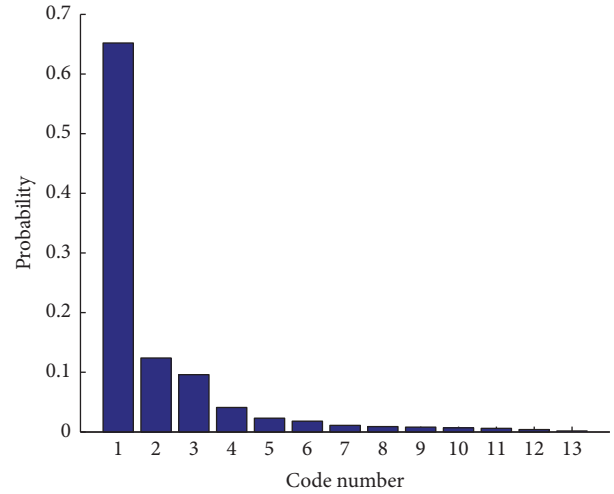
Because of the characteristics of the Φ -OTDR technology, the two indicators named D and S do not require such many codes. The new method employs 13 coded representations. Tables 1 and 2 represent the standard 29 encoded and the shrunk 13 encoded D -S1 matrices, respectively.

From Tables 1 and 2, the longest continuous sampling point is still 35. The statistics of first-order extreme points are reduced from six classes to four classes. And the encoding is reduced from 29 symbols to 13 symbols. Tables 3 and 4 are statistics of the two-order inflection point and the encoding result is reduced to 13 codes too.

By comparing Tables 2–5, it can be seen that the features represented are coarser with the codes' decrease. Actually, the

TABLE 4: The shrunk 13-symbol encoding table of D -S2 matrix.

D	S				
	1	2	3	>3	
3	1				
4	1	1			
5	1	2	2		
6	2	2	3		
7	3	3	3		
8	3	3	4	4	
...
34	11	12	13	13	
35	11	12	13	13	

FIGURE 1: The histogram of D -S1 coding's probability distribution in a signal.

probability distribution of encoded high numbers is almost zero, so it is possible to reduce the encoding to 13 characters.

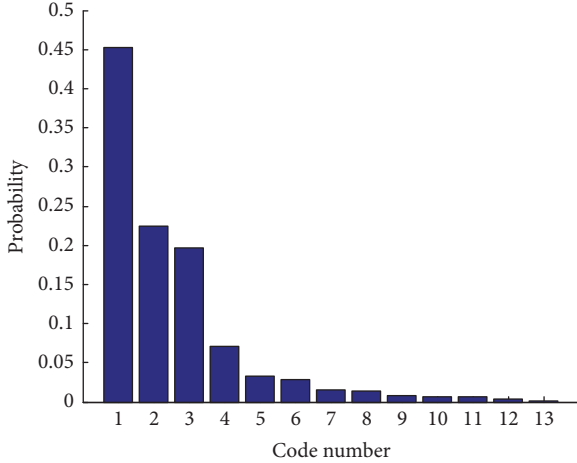
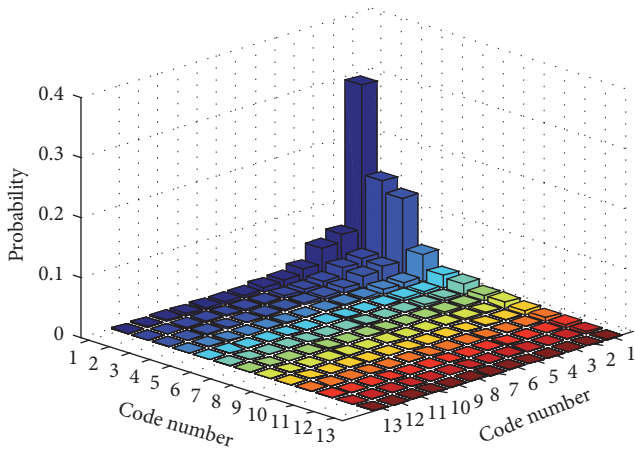
2.2.2. The Improved 2D-TESP Algorithm. After the original algorithm reduces to 13 characters, the feature of the algorithm is further represented as a two-dimensional A matrix. The two dimensions of the A matrix are the 13 codes in the shrunk D -S1 and in the shrunk D -S2 matrix. The values of the A matrices are the joint distribution probabilities of D -S1 and D -S2. We select a window of the signal to be encoded in accordance with the improved 2D-TESP algorithm. Figures 1 and 2 are histograms of probability distributions for D -S1 and D -S2. Figure 3 is a histogram of the A matrix with the joint distribution probability of D -S1 and D -S2. The A matrix is shown in Figure 3.

2.3. The Analysis of the Algorithm. In this paper, the improved 2D-TESP algorithm is mainly considered as follows.

(1) Because the damping vibration signal is encoded by TESP, the higher the coding number is, the more the probability drops. In this paper, the original 29-character encoding is reduced to 13-character encoding.

TABLE 5: Comparison of correct rates under different SNRs and different coding methods.

Number	Name	Poor SNR			Good SNR		
		<i>K</i> -means	Hierarchical clustering	Spectral clustering	<i>K</i> -means	Hierarchical clustering	Spectral clustering
(1)	<i>D</i> -S1(29)	84.76%	85.97%	86.72%	94.16%	95.97%	96.32%
(2)	<i>D</i> -S1(13)	84.67%	85.21%	86.34%	93.87%	95.21%	96.14%
(3)	<i>D</i> -S2(29)	82.31%	83.41%	84.91%	94.01%	95.41%	96.41%
(4)	<i>D</i> -S2(13)	81.74%	82.92%	83.34%	93.74%	95.12%	96.34%
(5)	<i>A</i>	93.12%	94.22%	96.12%	94.72%	96.12%	96.82%

FIGURE 2: The histogram of *D*-S2 coding's probability distribution in a signal.FIGURE 3: The histogram of *A* coding's probability distribution in a signal.

(2) The Φ -OTDR technology has its own characteristics. Therefore, the signal is expressed not only by extreme points but also by inflection points.

(3) After the reduction of characters, the 13-character coding must be slightly rougher than the 29-character coding. Some experiments are carried out in *K*-means, hierarchical clustering, and spectral clustering. The effect of 13-character coding has not declined.

3. Time-Domain Signal Representation Based on GAEMD-NMF

Because of the presence of 3D data in the Φ -OTDR technique, the signal representations of “time dimensional” and “length dimensional” need to be required. In the representation of the time axis and length axis, it is necessary to consider the signal's physical model and actual data characteristics in a rational representation. Therefore, the paper combines two aspects of knowledge about the prior vibration model and the actual signal data decomposition. The new method is combined by EMD and NMF. The combined method is optimized by GA method. The parameters of the method are effectively corrected and finally good results are achieved.

3.1. *The EMD of Damped Vibration.* In the time amplitude plane, a damping vibration model is established because the data conforms to the law of damping vibration. The model is shown in

$$f_1(t) = A_1 e^{-t/\tau} \sin(\omega_1 t + \varphi_1) + B_1. \quad (1)$$

In the length amplitude plane, a SA model is established because the data conforms to the law of SA signal. The model is shown in

$$f_2(l) = A_2 \frac{\sin(\omega_2 l + \varphi_2)}{\omega_2 l + \varphi_2} + B_2. \quad (2)$$

After the superposition of signals, the model is shown in

$$F(t, l) = f_1(t) \cdot f_2(l). \quad (3)$$

According to the two models above, the original signal corresponds to the model to determine the various parameters. The concrete steps are as follows.

(1) The absolute maximum value of the signal amplitude on the time axis is A_1 . A_2 is identified as 1.

(2) The peak value points are fitted to a curve. The damping attenuation coefficient is obtained according to the fitted curve. The damping attenuation coefficient is recorded as τ .

(3) ω_1 is determined according to the number of zero crossings on the time axis.

(4) When the signal starts, it must start from zero. The parameter φ_1 is determined as 0 or 1.

(5) The parameters B_1 and B_2 are determined as 0 according to the actual data.

(6) The parameter w_2 is determined by the number of zero crossings on the length axis.

(7) When the signal starts, it must be the maximum. The parameter φ_2 is determined as 0.

3.2. NMF. NMF (nonnegative matrix factorization) was proposed by Lee and Seung in 1999. It is a new matrix decomposition method published in Nature magazine [17]. It is widely used in image analysis, data mining, speech processing, and other fields. In recent years, scholars have made some modifications on the basis of the original NMF algorithm and achieved many results [18, 19]. The advantages of NMF algorithm are two main points. Firstly, its decomposition form and results are very explanatory because it requires matrix elements to be nonnegative. Secondly, the result of matrix decomposition usually has natural sparsity. Its results not only are easy to express, but also reduce the space occupation. Since the value of the vibration signal is positive, the vibration data are processed by using nonnegative matrix factorization to meet the requirements of the algorithm.

The nonnegative matrix factorization algorithm is defined as follows. A nonnegative matrix X_+ is set. It is decomposed into the product of two nonnegative matrices named W_{+d} and H_{+d} :

$$X_+ \approx W_+ \cdot H_+. \quad (4)$$

Here, subscript “+” stands for nonnegative constraints. The parameter d is a low dimensional spatial dimension that approximately describes the raw data. d should meet the requirement “ $(m+n)d < mn$.” For the solution of formula (4), the product of matrices W and H is approached gradually by the original matrix X . Euclidean distance is usually used to represent the error between WH and X . The error function is shown in

$$\begin{aligned} E(X \| WH) &= 0.5 \cdot \|X - W \cdot H\|_F^2 \\ &= 0.5 \cdot \sum_{ij} (X_{ij} - W_i \cdot H_j)^2. \end{aligned} \quad (5)$$

Here, the matrices X , W , and H are nonnegative matrices. When the minimum of formula (5) is obtained, the error between matrix X and matrix WH is minimum. Lee and Seung gave the iterative rules of the corresponding upper formula such as

$$\begin{aligned} W_{ik} &\leftarrow W_{ik} \cdot \frac{X \cdot H^T}{W_{ik} \cdot H \cdot H^T}, \\ H_{ik} &\leftarrow H_{ik} \cdot \frac{W^T \cdot X}{W \cdot W^T \cdot H_{ik}}. \end{aligned} \quad (6)$$

W and H are iterated over the two formulas above. The iteration ends when WH converges. NMF method is completed. However, the traditional NMF method cannot meet the needs of signal representation. The new iterative method needs to be carried out according to the actual situation.

3.3. GAEMD-NMF. The accuracy of the traditional NMF method is not satisfactory. Therefore, this paper uses the combination of EMD and NMF algorithm for signal representation.

The objective function is redefined as follows:

$$\begin{aligned} E_1(\widehat{X}_t \| f(t)) &= 0.5 \cdot \|X_t - f(t)\|_F^2 \\ &= 0.5 \cdot \sum_{ij} (X_{ij} - f(t))^2, \\ E_2(\widehat{X}_l \| f(l)) &= 0.5 \cdot \|X_l - f(l)\|_F^2 \\ &= 0.5 \cdot \sum_{ij} (X_{ij} - f(l))^2. \end{aligned} \quad (7)$$

In order to facilitate the convergence calculation, the rotation matrix is defined by the four-element number. It is as shown in

$$R = \begin{bmatrix} 1 - (q_2^2 + q_3^2) & 2(q_1q_2 - q_3q_0) & 2(q_1q_3 + q_2q_0) \\ 2(q_1q_2 + q_3q_0) & 1 - (q_1^2 + q_3^2) & 2(q_2q_3 - q_1q_0) \\ 2(q_1q_3 - q_2q_0) & 2(q_2q_3 + q_1q_0) & 1 - (q_1^2 + q_2^2) \end{bmatrix}. \quad (8)$$

Here, $Q = q_0 + q_1 \cdot i + q_2 \cdot j + q_3 \cdot k$. Q represents quaternion. By adjusting an expression, formula (9) is as follows:

$$\widehat{X} = \begin{bmatrix} f_\infty(t) \\ f_\infty(l) \\ F_\infty(t, l) \\ 1 \end{bmatrix}^T = \begin{bmatrix} f(t) \\ f(l) \\ F(t, l) \\ 1 \end{bmatrix}^T \cdot \begin{bmatrix} & \Delta f(t) \\ R & \Delta f(l) \\ & \Delta F(t, l) \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (9)$$

The GA method is used to optimize the processing of the EMD-NMF method.

(1) Each row or column of the signals X and \widehat{X} is regarded as a population.

(2) The entire matrix is normalized.

(3) The variation convergence step of $\Delta f(t)$, $\Delta f(l)$, and $\Delta F(t, l)$ is 0.01. The convergence step of q_0 , q_1 , q_2 , and q_3 is 0.1.

(4) According to the sample input, we establish the fitness function as shown in

$$J = 0.5E_1 + 0.5E_2, \quad (10)$$

where J is fitness.

(5) If the fitness is reduced, it will inherit the parameter results. If the fitness is enhanced, it will change the parameter results.

4. Experiments

The Φ -OTDR instrument in this paper is the NBX-S3000 instrument from the Japanese Nebrex company. The instrument is shown in Figure 4.

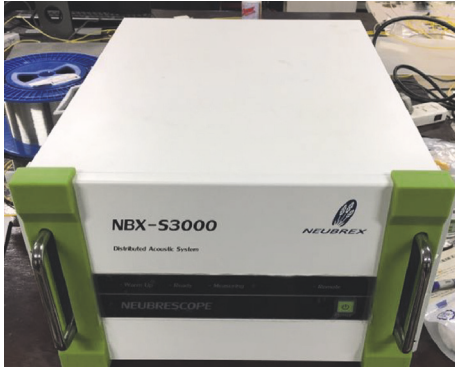


FIGURE 4: Physical picture of Φ -OTDR instrument.

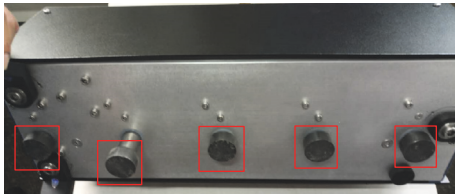


FIGURE 5: A physical picture of five hammers on a knocker.

The actual parameters of the device are as follows:

- (1) Sampling rate: 4000 Hz
- (2) Recording time: 10 s
- (3) Spatial resolution: 0.1 m
- (4) Recording distance: 10 m

The vibration source uses standard vibratory equipment.

Figure 5 is a physical picture of 5 percussion hammers on the knocker. Figure 6 is a physical picture of the knocker.

In order to restrain the environmental noise and ensure the good vibration effect, the anechoic chamber is selected to test. The specific experimental environment is shown in Figures 7 and 8. In Figure 7, the fiber clings to the ground. The instrument is placed on the damping table in Figure 8.

The effects in each observation position on the fiber are shown in Figure 9.

4.1. Signal Extraction Experiment. After encoding and extracting features by the 2D-TESP method, the results are used in the popular clustering algorithms. Popular clustering algorithms are K -means clustering, hierarchical clustering, and spectral clustering. The main work of the first half of this paper is signal extraction. Therefore, clustering targets are limited to two cases: "signal" and "no signal." We compare the effects of the 5 cases. These five situations are "extreme value 29 characters," "extreme value 13 characters," "inflection point 29 characters," "inflection point 13 characters," and "extreme point + inflection point matrix." The clustering accuracy of each case is shown in Table 5. Part 1 in Table 5 is the average accuracy of the signals from 4.4 to 4.6 meters and from 5.9 to 6.1 meters on the fiber. Part 2 in Table 5 is the signal on the fiber from 4.7 meters to 5.8 meters.

Some conclusions can be drawn from the accuracy of Table 5. The 2D-TESP algorithm is the best of the five



FIGURE 6: A physical picture of a knocker.



FIGURE 7: Scene of laying fiber.

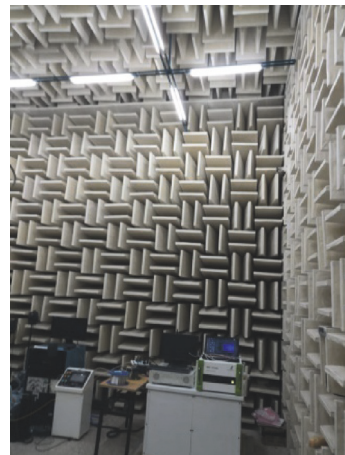


FIGURE 8: Scene of lab.

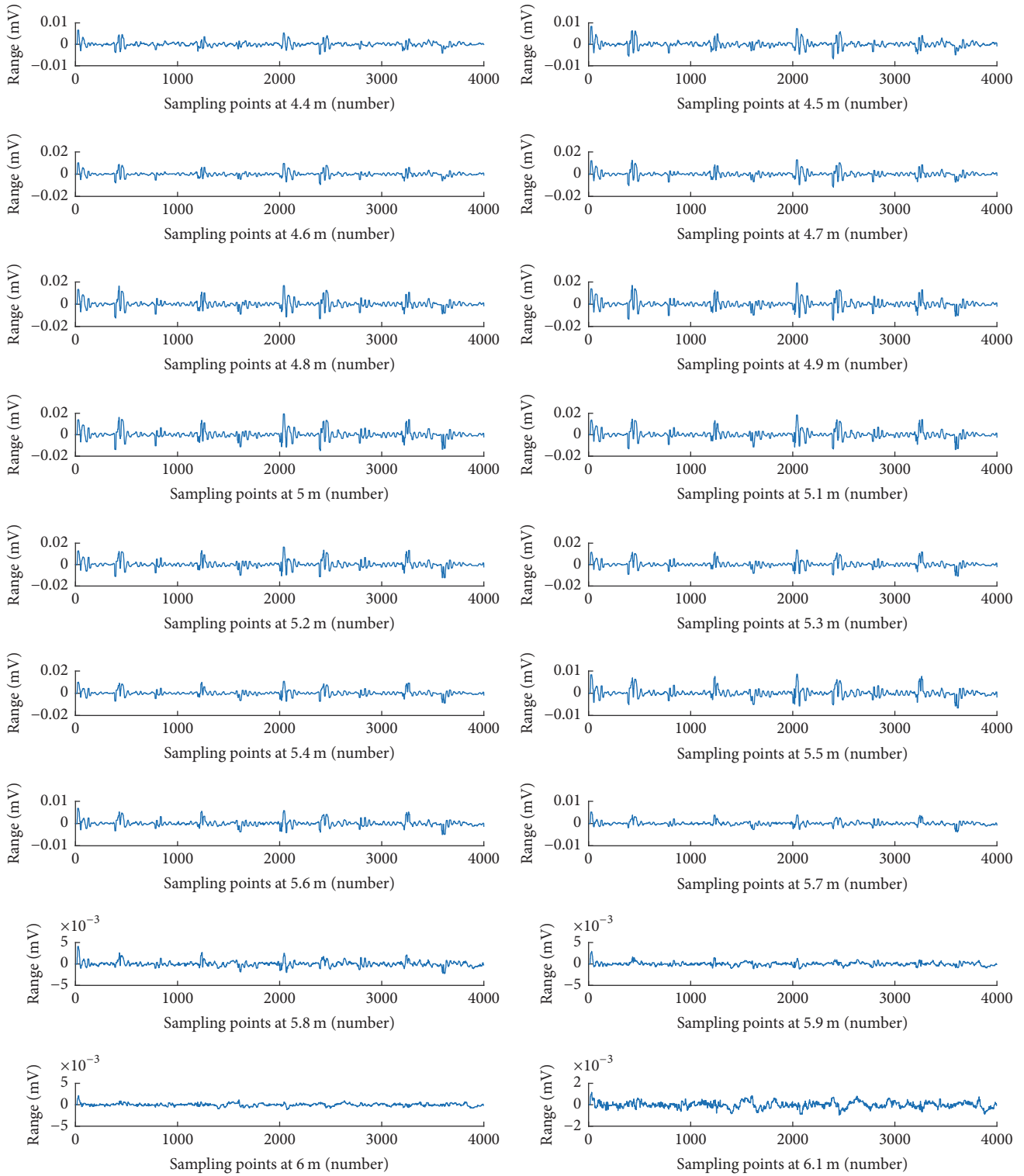


FIGURE 9: Time-domain vibration signals of each point in the effective range.

encoding methods. Spectral clustering is the best of the three clustering methods. In the case of good SNR, the accuracy of the 2D-TESP algorithm is about 0.5% higher than that of the traditional TESP algorithm. But in the case of poor SNR, the accuracy of the 2D-TESP algorithm is about 8% higher

than that of the traditional TESP algorithm. In the case of strong and weak SNR, the accuracies of 2D-TESP algorithm are relatively close. They differ by about 1 percentage point.

In this section, what is the best sort of coding in the improved TESP? The results obtained are shown in Figure 10.

TABLE 6: Comparison of correct rates under different SNRs and different methods.

Number	Name	Poor SNR			Good SNR		
		K-means	Hierarchical clustering	Spectral clustering	K-means	Hierarchical clustering	Spectral clustering
(1)	2D-TESP	93.12%	94.22%	96.12%	94.72%	96.12%	96.82%
(2)	FFT	87.21%	87.95%	88.45%	92.14%	92.71%	93.25%
(3)	WT	89.37%	89.79%	90.12%	92.72%	92.99%	93.65%
(4)	HHT	90.25%	90.97%	91.85%	92.84%	93.12%	93.89%

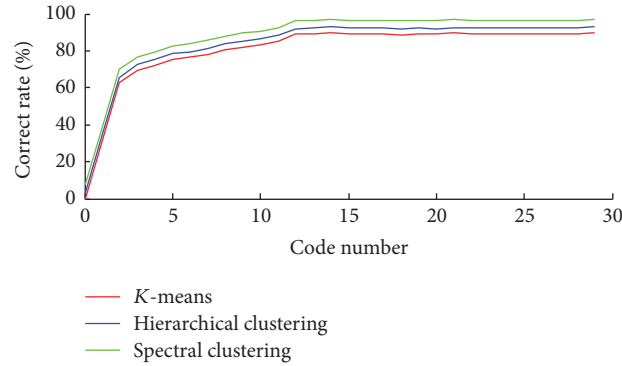


FIGURE 10: Comparison of the number of coded characters.

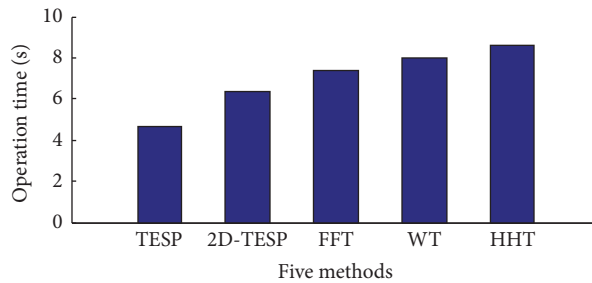


FIGURE 11: Operation time comparison among different methods.

From Figure 10, the correct rate between 13 and 29 characters is almost the same. But in 13 characters, it will have a certain effect on the accuracy. The accuracy rate tends to decrease linearly within 3 characters.

In order to verify the quality of this method, the method and another three methods are compared in the paper. Comparison is also made between three clustering methods and two different SNR situations. The results are shown in Table 6.

Some conclusions can be drawn from the accuracy of Table 6. The 2D-TESP algorithm is the best of the four methods. Spectral clustering is still the best of the three clustering methods. In the case of good SNR, the accuracy of the 2D-TESP algorithm is about 1.5% higher than that of the other algorithms. But in the case of poor SNR, the accuracy of the 2D-TESP algorithm is about 4% higher than that of the other algorithms. The method not only has better accuracy, but also has shorter operation time. The same set of signals is made in the experiments. The signal's size is 1 MB. The PC has 1.6 GHz * 4 clocked CPU and 8 G memory. Matlab software is used on this PC. The results are shown in Figure 11.

As shown in Figure 11, the method in this paper is superior to several other methods. Because the method is only computing in the time domain, its operation time is about 1 s higher than FFT, about 1.2 s higher than WT, and about 1.5 s higher than HHT.

4.2. Signal Expression Experiment. Since the Φ -OTDR signal is 3D data, the expression of the signal needs to be seen from the length axis and the time axis.

4.2.1. Positioning Problem on Length Axis. Figure 12 shows the state of the fiber with one vibration source. The vertical intersection point between the vibration source and the optical fiber is the strongest point of vibration.

There are five hammers with equal intervals on the knocker. The striking distance between two strikes is 10 cm. The five signals are not the same on the time axis. Five different signals need to be separated. The result is shown in Figure 13.

The spacing of the two hammers on the length axis is fixed as 10 cm. The five hammers have four pitches. The specific values of the four distances under the four methods are

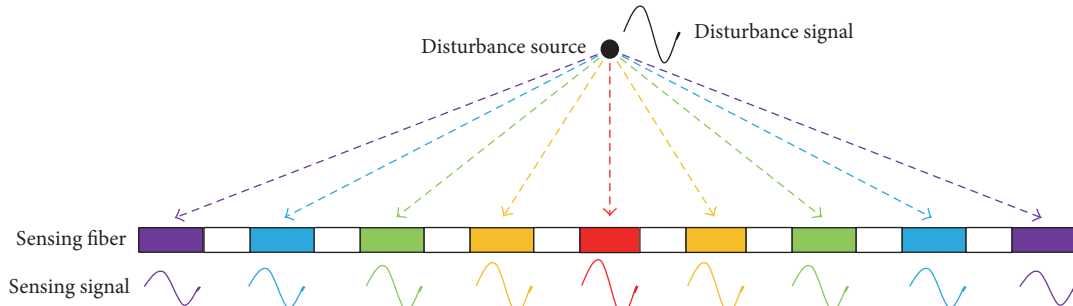


FIGURE 12: Perceptual model on length axis.

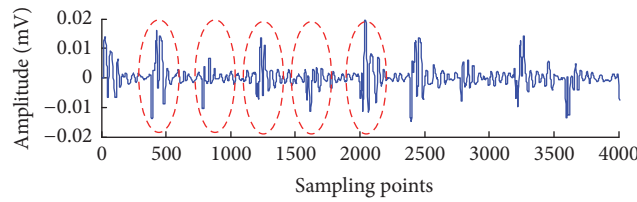


FIGURE 13: Five different signals of hammers.

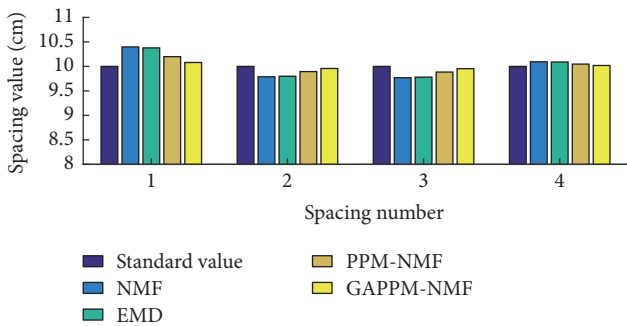


FIGURE 14: Comparison of positioning performance of four different methods on length axis.

compared. The result is shown in Figure 13. The four methods are NMF, EMD, EMD-NMF, and GAEMD-NMF.

As you can see from Figure 14, the GAEMD-NMF method works best. The EMD-NMF method has good performance and the NMF and EMD methods have the worst performance. After 100 sets of trial and error, the error range between the four methods is as shown in Figure 15.

From Figure 15, the GAEMD-NMF method in this paper is the best method for the four spacings. At the same time, the range of error fluctuation is also minimum. The range of error fluctuations in GAEMD-NMF is reduced by half compared with EMD-NMF and by 1/3 compared with NMF and EMD. The error range of the GAEMD-NMF method at 4 distances is from -0.1 cm to 0.1 cm.

4.2.2. Vibration Modes of Signals on Time Axis. The most important thing about the performance of signal on time axis is whether the spectrum analysis is consistent or not. In this

paper, a high-precision vibration sensor is arranged along the fiber. The sensor signal, the original signal, NMF, EMD, EMD-NMF, and GAEMD-NMF are compared in power spectrum. The result is shown in Figure 16. The method which compares the power spectrum similarity between different signals in this paper is as follows.

- Step 1.* Each ten Hz are noted as a frequency segment.
- Step 2.* The accumulated energy per frequency segment is calculated.
- Step 3.* The total energy of signal is calculated.
- Step 4.* Each energy is divided by the total energy. The corresponding percentage of each segment is got.
- Step 5.* The probability fluctuation of each segment between two signals is calculated.
- Step 6.* These probability fluctuations are summed by absolute values.
- Step 7.* The sum of the fluctuation probability is reduced by 1. The result is probability value of the similarity.

From Figure 16, several experimental results are as follows.

- (1) There is large high-frequency noise in the original signal. The same performance is in the treated signal by NMF method and EMD method. This shows that the signal-to-noise separation is poor.
- (2) There is no high-frequency noise in the EMD-NMF and GAEMD-NMF. This shows that the signal separation is in good condition.

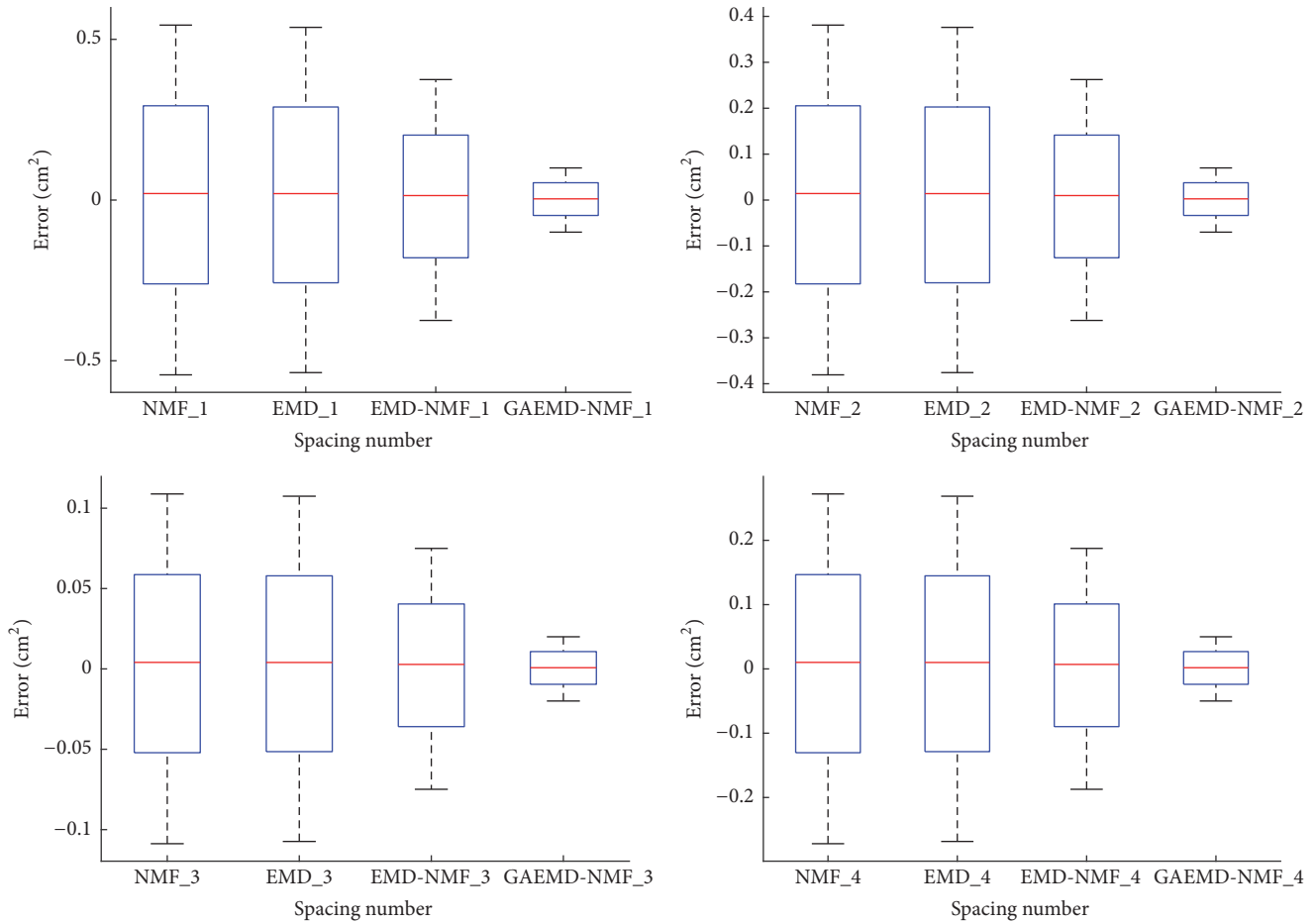


FIGURE 15: Error fluctuation on four distances by four different methods.

(3) The trends of the spectrogram of EMD-NMF, GAEMD-NMF, and sensor signals are consistent. This shows the validity of this method.

(4) The spectrograms of the original signal, the EMD method, and the NMF method are not quite consistent with the other three classes. This shows that the signal expression method is not ideal.

(5) Because the device is different, there may be a difference in the effective value of the power spectrum. Therefore, each frequency segment of the power spectrum is normalized. The similarity between GAPPN-NMF and sensor is 95.63%. The similarity between GAEMD-NMF and sensor is 95.63%. The similarity between EMD-NMF and sensor is 90.74%. This shows that the GAEMD-NMF method is 5.11% higher than the EMD-NMF method.

5. Conclusion

In this paper, we study the information representation of Φ -OTDR distributed vibration signals, and the main conclusions are as follows.

(1) The improved 2D-TESP signal extraction method has good signal extraction effect. The number of coded characters is reduced and the dimension is extended. This will effectively reduce redundancy and increase effectiveness. Particularly, it

has good signal extraction effect on the fiber point with poor SNR.

(2) The improved GAEMD-NMF signal expression method has a good expression effect on the length axis. Because the new method takes into account the characteristics of data model, it is easier to locate the strongest point of the signal in the positioning process. The positioning error effect is effectively suppressed.

(3) The improved GAEMD-NMF signal expression method has good expression effect on the time axis. Because the new method takes into account the characteristics of mechanism model, the vibration model is closer to the theoretical model. The result is compared with the high-precision sensor, and the spectrogram has a great similarity.

Conflicts of Interest

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

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (no. 61627816 and no. 61503034), the National Key Technology R&D Program

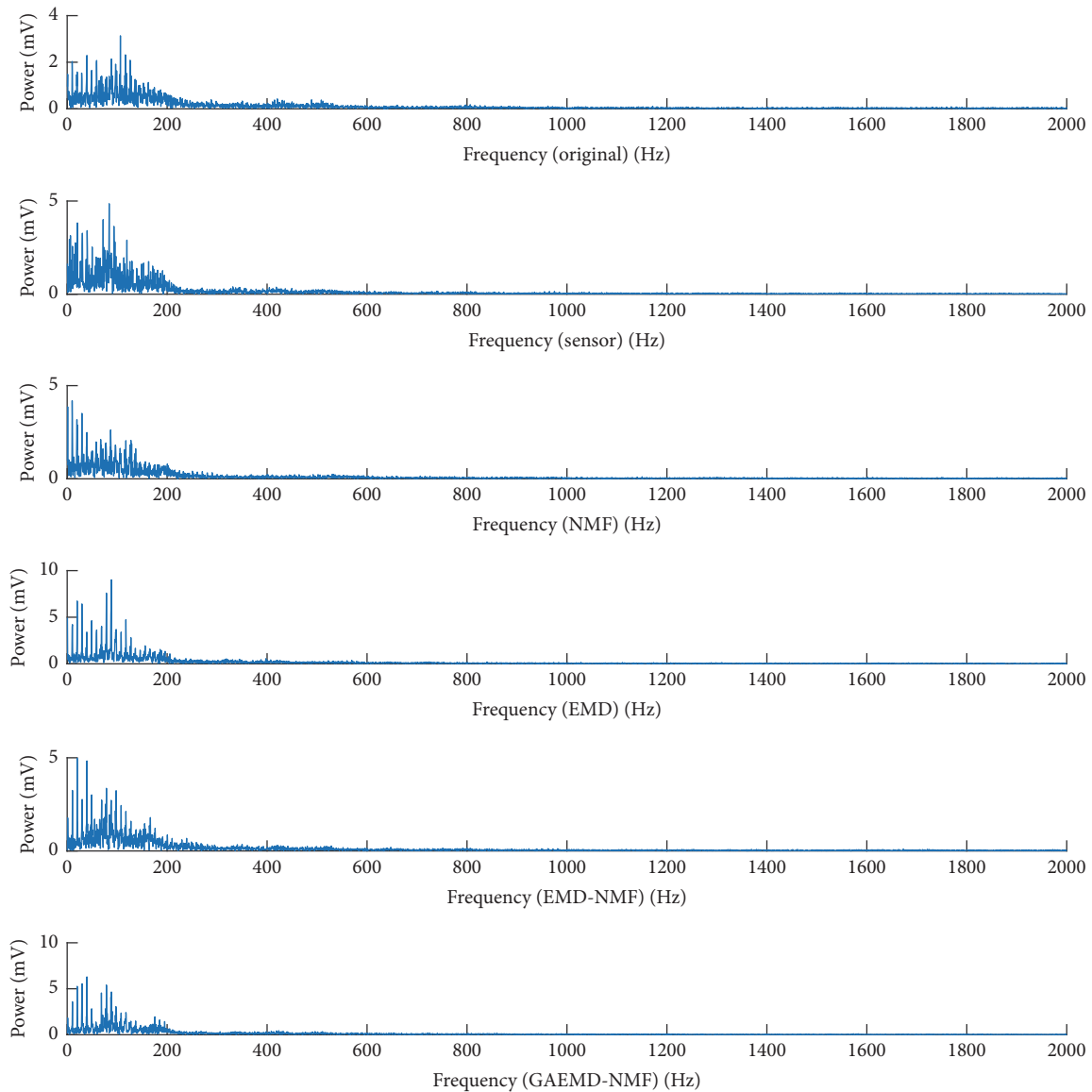


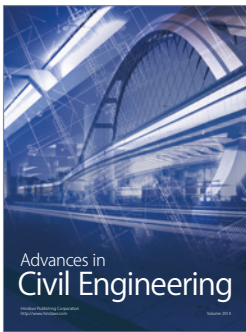
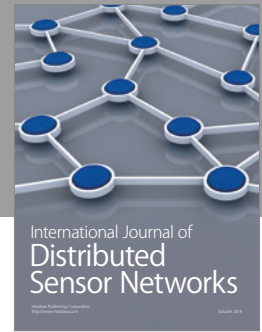
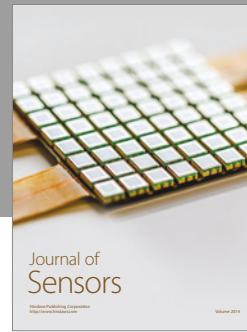
FIGURE 16: Comparison of power spectrum between four different methods, sensor signal, and original signal.

of China (no. 2015BAK40B03 and no. 2016YFC0701309), and Subproject of Key Project of Beijing, China (no. D161100004916002 and no. Z171100002417022).

References

- [1] J. A. Bucaro and E. F. Carome, "Single fiber interferometric acoustic sensor," *Applied Optics*, vol. 17, no. 3, pp. 330–331, 1978.
- [2] J. C. Juarez, E. W. Maier, K. N. Choi, and H. F. Taylor, "Distributed fiber-optic intrusion sensor system," *Journal of Lightwave Technology*, vol. 23, no. 6, pp. 2081–2087, 2005.
- [3] P. Potočnik and E. Govekar, "Semi-supervised vibration-based classification and condition monitoring of compressors," *Mechanical Systems and Signal Processing*, vol. 93, pp. 51–65, 2017.
- [4] C. Delpha, D. Diallo, and A. Youssef, "Kullback-Leibler Divergence for fault estimation and isolation: Application to Gamma distributed data," *Mechanical Systems and Signal Processing*, vol. 93, pp. 118–135, 2017.
- [5] J. Ma, F. Xu, K. Huang, and R. Huang, "GNAR-GARCH model and its application in feature extraction for rolling bearing fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 93, pp. 175–203, 2017.
- [6] H. Liu, J. Zhang, Y. Cheng, and C. Lu, "Fault diagnosis of gearbox using empirical mode decomposition and multi-fractal detrended cross-correlation analysis," *Journal of Sound and Vibration*, vol. 385, pp. 350–371, 2016.
- [7] K. A. Alhazza, "Adjustable maneuvering time wave-form command shaping control with variable hoisting speeds," *Journal of Vibration and Control*, vol. 23, no. 7, pp. 1095–1105, 2017.
- [8] H. Gao, L. Liang, X. Chen, and G. Xu, "Feature extraction and recognition for rolling element bearing fault utilizing short-time fourier transform and non-negative matrix factorization," *Chinese Journal of Mechanical Engineering*, vol. 28, no. 1, pp. 96–105, 2015.

- [9] B. Li, P.-L. Zhang, H. Tian, S.-S. Mi, D.-S. Liu, and G.-Q. Ren, "A new feature extraction and selection scheme for hybrid fault diagnosis of gearbox," *Expert Systems with Applications*, vol. 38, no. 8, pp. 10000–10009, 2011.
- [10] B. Li, P.-L. Zhang, D.-S. Liu, S.-S. Mi, G.-Q. Ren, and H. Tian, "Feature extraction for rolling element bearing fault diagnosis utilizing generalized S transform and two-dimensional non-negative matrix factorization," *Journal of Sound and Vibration*, vol. 330, no. 10, pp. 2388–2399, 2011.
- [11] B. Li, P.-L. Zhang, S.-B. Liang, Y.-T. Zhang, and H.-B. Fan, "Feature extraction for engine fault diagnosis utilizing the generalized S-transform and non-negative tensor factorization," *Proceedings of the Institution of Mechanical Engineers Part C: Journal of Mechanical Engineering Science*, vol. 225, no. 8, pp. 1936–1949, 2011.
- [12] A. Rai and S. H. Upadhyay, "Bearing performance degradation assessment based on a combination of empirical mode decomposition and k-medoids clustering," *Mechanical Systems and Signal Processing*, vol. 93, pp. 16–29, 2017.
- [13] M. Sher, N. Ahmad, and M. Sher, "TESPAR feature based isolated word speaker recognition system," in *Proceedings of the 18th International Conference on Automation and Computing (ICAC '12)*, pp. 318–321, September 2012.
- [14] Y. Wang, X. Wu, X. Li, and J. Zhou, "Vehicle type recognition in sensor networks using improved time encoded signal processing algorithm," *Mathematical Problems in Engineering*, vol. 2014, Article ID 142304, 8 pages, 2014.
- [15] Y. Wang, W. Li, Y. Zhang, W. Zhang, Z. Li, and J. Zhou, "Vehicle sound signal recognition in border surveillance based on improved TESP algorithm," *Journal of Sichuan University*, vol. 46, no. 2, pp. 122–127, 2014.
- [16] Y. Wang and J. Zhou, "Vehicle type recognition in WSN based on ITESP algorithm," in *Proceedings of the 10th IEEE International Conference on Ubiquitous Intelligence and Computing, UIC 2013 and 10th IEEE International Conference on Autonomic and Trusted Computing (ATC '13)*, pp. 668–671, December 2013.
- [17] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.
- [18] S. Shon, S. Mun, D. Han, and H. Ko, "Non-negative matrix factorisation-based subband decomposition for acoustic source localisation," *Electronics Letters*, vol. 51, no. 22, pp. 1723–1724, 2015.
- [19] A. Ozerov and C. Fevotte, "Multichannel nonnegative matrix factorization in convolutive mixtures for audio source separation," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 18, no. 3, pp. 550–563, 2010.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

