Research Article

A Rolling Bearing Fault Diagnosis-Optimized Scale-Space Representation for the Empirical Wavelet Transform

Zechao Liu, Jianming Ding ⓘ, Jianhui Lin, and Yan Huang

State Key Laboratory of Traction Power, Southwest Jiaotong University, Chengdu 610031, China

Correspondence should be addressed to Jianming Ding; fdingjianming@126.com

Rolling element bearings have been widely used in mechanical systems, such as electric motors, generators, pumps, gearboxes, railway axles, and turbines, etc. Therefore, the detection of rolling bearing faults has been a hot research topic in engineering practices. Envelope demodulation represents a fundamental method for extracting effective fault information from measured vibration signals. However, the performance of envelope demodulation depends heavily on the selection of the filter band and central frequencies. The empirical wavelet transform (EWT), a new signal decomposition method, provides a framework for arbitrarily segmenting the Fourier spectrum of an analysed signal. Scale-space representation (SSR) can adaptively detect the boundaries of the EWT; however, it has two shortcomings: slow calculation speeds and invalid boundary detection results. Accordingly, an EWT method based on optimized scale-space representation (OSSR), namely, the EWTOSSR, is proposed. The effectiveness of the EWTOSSR is verified by comparisons between the simulation and the experimental signals. The results show that the EWTOSSR can automatically and effectively segment the EWT spectrum to extract fault information. Compared with three well-known methods (the traditional EWT, ensemble empirical mode decomposition (EEMD), and the fast kurtogram), the EWTOSSR exhibits a much better fault detection performance.

1. Introduction

Rolling element bearings have been widely used in mechanical systems, such as electric motors, generators, pumps, gearboxes, railways, and turbines [1]. They are closely associated with the reliability of a mechanical system [2], which is therefore a necessity to monitor their condition. Many methods are used to monitor bearing faults, such as vibration, acoustic, and temperature measurement techniques [3–7]. However, the use of temperature measurements can delay the detection of faults, and acoustic measurements are easily obscured by noise. As a result, vibration monitoring has become a commonly used method.

Bearing defects can evoke a sharp impulse with a wide frequency range, and such evoked impulses modulate other vibration signals. Therefore, to obtain the modulated fault information, a high-frequency resonance technique, also known as envelope analysis, was proposed [8]; the key of envelope analysis is to determine two filter parameters: the bandwidth and the centre frequency. In addition, the

concept of the spectral kurtosis was proposed to select one sub-band with a maximal kurtosis to detect bearing faults [9]. However, the computational cost required for a short-time Fourier transform in spectral kurtosis is high [10]. Consequently, a fast kurtogram algorithm was developed [11]. Unfortunately, the bandwidths of the spectral kurtosis and the fast kurtogram are fixed, being incapable of covering the entire sideband surrounding the resonance frequency. To resolve the fixed bandwidth problem associated with the spectral kurtosis and fast kurtogram approaches, the adaptive spectral kurtosis (ASK) technique was proposed [12]. However, the performance of the ASK method heavily depends on the initial superposition window width [13].

The empirical wavelet transform (EWT) [14], which possesses the fast computational properties of the wavelet transform [15, 16] and the adaptive characteristic of empirical mode decomposition (EMD) [17, 18], provides a framework for arbitrarily determining the bandwidth and centre frequency of a filter. The EWT has been widely applied in the field of rotational mechanical fault detection, and

it obtains excellent fault detection results [19–23]. The key of the EWT is to determine the proper boundaries of spectrum segments associated with arbitrary bandwidths. Many methods for determining these boundaries have been developed, including operational modal analysis [24], order statistics filters [25], sparsity guides [26], Pearson's correlation coefficient [27], and SSR [28]. However, with the exception of SSR, all boundary-determining methods require a priori knowledge, such as the fault characteristic frequency and order information. In contrast, SSR-based EWT is more adaptive and can therefore extract bearing fault information. However, it is computationally expensive, and some overly noisy spectrum segments exist that are subject to the presence of strong noise [29].

To solve the two problems mentioned above, an optimized scale-space representation (OSSR) and an effective scale-space curve (SSC) selection method are proposed in this paper. The calculation cost of the SSR is closely related to the initial scale; the larger the initial scale value, the faster the calculation efficiency of the SSR, and vice versa. However, with an increase in the initial scale value, low-scale information will be lost, resulting in the failure of the boundary detection process. It has been discovered that the minimum number of the SSR conforms to a logarithmic distribution; therefore, a function of this minimum number is used to adaptively adjust the scale value to achieve a faster calculation of the SSR. The differences in the lengths of the SSCs are used to determine the significant

boundaries, and a novel EWT (EWTOSSR) fault detection method based on OSSR is proposed.

The contents followed are organized as follows: the EWT and SSR techniques are introduced in Section 2; Section 3 introduces an OSSR and the boundary selection method; in Section 4, the procedure of the proposed EWTOSSR method is introduced, and the proposed method is verified by a noisy mechanical fault simulation signal in Section 5, and then an experimental verification of the proposed method is conducted in Section 6; finally, the conclusions of this study are drawn in Section 7.

2. Theoretical Background

2.1. Empirical Wavelet Transform. The fundamental principle of EWT, which was introduced by Gilles in 2013 to achieve the adaptive decomposition of signals, is to establish a set of filters covering all resonant frequency bands. The EWT includes a low-pass filter representation approximation and N-1 bandpass filters to describe signals in greater detail. If $(N+1)$ divided boundaries are obtained, the Fourier spectrum of the analysed signal will be divided into N corresponding intervals. The empirical wavelet functions associated with each divided interval are generated using Littlewood–Paley and Meyer's wavelets [14]. As a result, the empirical scale function $\widehat{\phi}_n(\omega)$ and empirical wavelet function $\widehat{\psi}_n(\omega)$ are written as follows [30]:

$$\widehat{\phi}_n(\omega) = \begin{cases} 1 & \text{if } |\omega| \le \omega_n - \tau_n, \\ \cos\left[\dfrac{\pi}{2}\eta\left(\dfrac{1}{2\tau_n}\left(|\omega| - \omega_n + \tau_n\right)\right)\right] & \text{if } \omega_n - \tau_n \le |\omega| \le \omega_n + \tau_n, \\ 0 & \text{otherwise,} \end{cases}$$

$$\widehat{\psi}_n(\omega) = \begin{cases} 1 & \text{if } \omega_n + \tau_n \le |\omega| \le \omega_{n+1} - \tau_{n+1}, \\ \cos\left[\dfrac{\pi}{2}\eta\left(\dfrac{1}{2\tau_{n+1}}\left(|\omega| - \omega_{n+1} + \tau_{n+1}\right)\right)\right] & \text{if } \omega_{n+1} - \tau_{n+1} \le |\omega| \le \omega_{n+1} + \tau_{n+1}, \\ \sin\left[\dfrac{\pi}{2}\eta\left(\dfrac{1}{2\tau_n}\left(|\omega - \omega_n + \tau_n\right)\right)\right] & \text{if } \omega_n - \tau_n \le |\omega| \le \omega_n + \tau_n, \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where ω_n denotes a detected boundary; a transition phase T_n with a width $2\tau_n$ is defined for each ω_n. $\eta(x)$ is an arbitrary $C^k([0,1])$ function which satisfies [13]:

$$\eta(x) = \begin{cases} 0 & \text{if } x \le 0 \text{ and } \eta(x) + \eta(1-x) = 1 \ \forall \ x \in [0,1], \\ 1 & \text{if } x \ge 1, \end{cases} \tag{2}$$

and the most used function $\eta(x)$ is

$$\eta(x) = x^4\left(35 - 84x + 70x^2 - 20x^3\right). \tag{3}$$

The transition phase between two adjacent wavelets can be calculated from the linear relationship $\tau_n = \gamma\omega_n$, where $0 < \gamma < 1$. An empirical wavelet with a divided pattern is shown in Figure 1.

Therefore, a key step of the EWT is to determine the divided boundaries. To effectively obtain these boundaries, as mentioned above, many methods have been proposed,
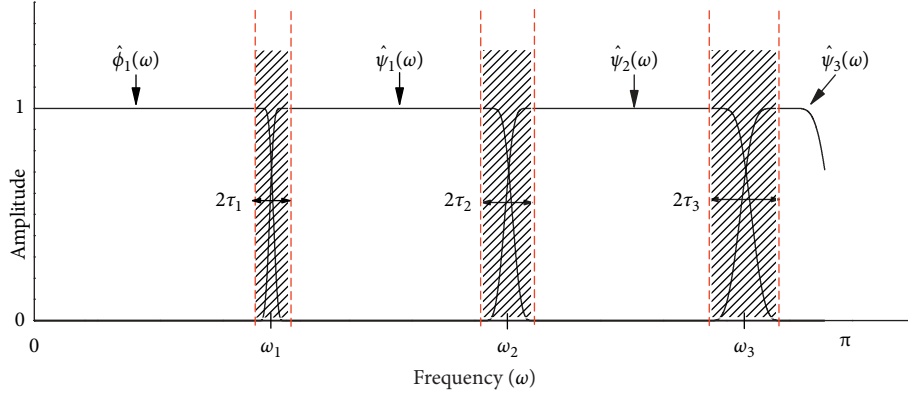
FIGURE 1: An empirical wavelet.

including operational modal analysis [24], order statistics filters [25], sparsity guides [26], Pearson's correlation coefficient [27], and SSR [28].

2.2. Scale-Space Representation.

The SSR $L(f, s)$ of the discrete Fourier transform (DFT) $X(f)$ of a discrete signal $x(n)$ can be expressed as [31]:

$$L(f, s) = X(f) * g(f, s) = \sum_{m=-M}^{+M} X(m - f)g(f, s), \quad (4)$$

where $g(f, s) = 1/\sqrt{2\pi s}e^{-(f^2/2s)}$ is a kernel function, s denotes the scale parameter, and M denotes the width of the kernel function. To ensure the approximation error of the kernel function during the discretization, the value of M is set as $M = C\sqrt{s} + 1$, where $3 \leq C \leq 6$ (with an increase in the scale parameter, the width of the kernel function should also increase) [32].

The SSR can be interpreted as a global trend of the Fourier spectrum under different scale parameters. The scale parameter will update by $\sqrt{s} = N\sqrt{s_0}$, where $\sqrt{s_0} = 0.5$ is the initial scale parameter corresponding to half of the distance between two Fourier spectrum points, and $N = 1, \ldots N_{\max}$ are integers. To perform a finite maximum number of steps, $N_{\max}$ is defined as twice the length of the signal. Detailed local information of characteristic lengths smaller than $\sqrt{s}$ will be removed by the SSR $L(f, s)$. W_N denotes the number of local minima in the N-th scale step; the number of local minima with respect to f of $L(f, s)$ is a decreasing function of the step N (the relationship of N and s has been defined previously), although no new minima will appear as the step N increases. The number of initial local minima of $L(f, s)$ (with an initial scale $\sqrt{s_0}$) will be denoted as $W0$, and each of these local minima defines a scale-space curve Ci ($i \in [1, W0]$) with a length Li ($i \in [1, W0]$), and Li can be regarded as the life span of the i-th minimum [33]. Therefore, Li is defined as

$$L_i = \max \left\{ \frac{N}{\text{the } i\text{-th minimum exists}} \right\}. \quad (5)$$

The supports of meaningful modes in Fourier spectra are achievable by delimiting two long SSCs. Therefore, the selection of the SSC is the fundamental step of the SSR, and it is

TABLE 1: Parameters of the simulated signal.

A_j	$1/T_P$	f_R	β
1	64 Hz	1000 Hz	1500 Ns/m

solved by finding a threshold T such that the SSCs with lengths larger than T are the curves delimiting the supports of meaningful modes. This is a two-class clustering problem. Many methods have been proposed to solve this problem, including k-Means clustering [34] and probabilistic approaches [35] in addition to the mean value method and Otsu's method [36].

2.3. A Discussion of Scale-Space Representation Shortcomings.

A noisy simulated signal is employed to illustrate the shortcomings of SSR when diagnosing a faulty bearing signal. The simulated signal is expressed as follows [37]:

$$s(t) = \sum_{j=0}^{J} A_j e^{-\beta(t - jT_P)} \cos(2\pi f_R(t - jT_P))u(t - jT_P),$$

$$(6)$$

where A_j is the amplitude of the j-th fault impulse, T_P is the time period corresponding to the characteristic fault frequency, f_R denotes the excited resonance frequency, β is the structural damping coefficient, J is the number of impulses, $u(t)$ is a unit step function, and $s(t)$ denotes the simulated signal. The fixed parameters are listed in Table 1.

The sampling frequency is 4096 Hz, and the length of the simulated signal is 8192. The simulated signal and the noised signal (SNR = −14 dB) are shown in Figures 2(a) and 2(b), respectively. The Fourier spectrum of the noised signal and the Fourier spectrum of the squared envelope (i.e., the squared envelope spectrum) of the noised signal are shown in Figures 2(c) and 2(d), respectively.

Because of the large noise in the simulated signal, the characteristic fault frequency cannot be identified from the squared envelope spectrum of the noised signal. The partial discrete SSRs $L(f, s)$ with different scale parameters s are computed using Equation (3) and shown in Figure 3. Obviously, with the s increase, the SSRs become smoother. The SSP and threshold T obtained by different two-class
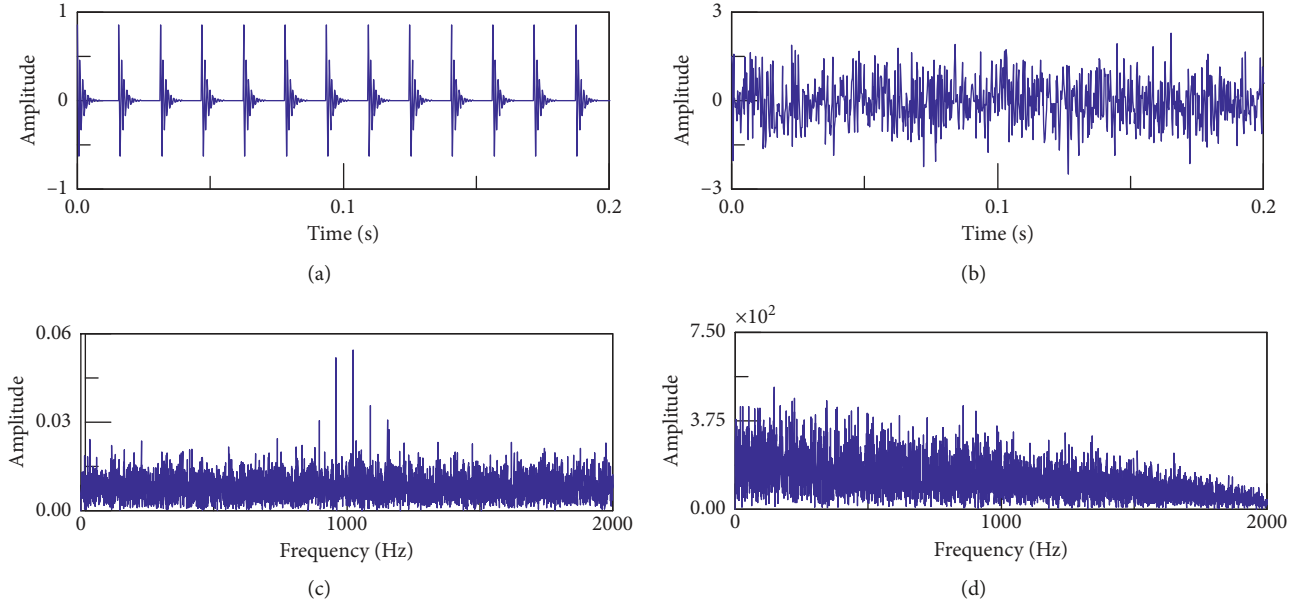
Figure 2: The simulated signal: (a) raw signal, (b) simulated signal with added noise, (c) Fourier spectrum of the noised signal, and (d) squared envelope spectrum of the noised signal.
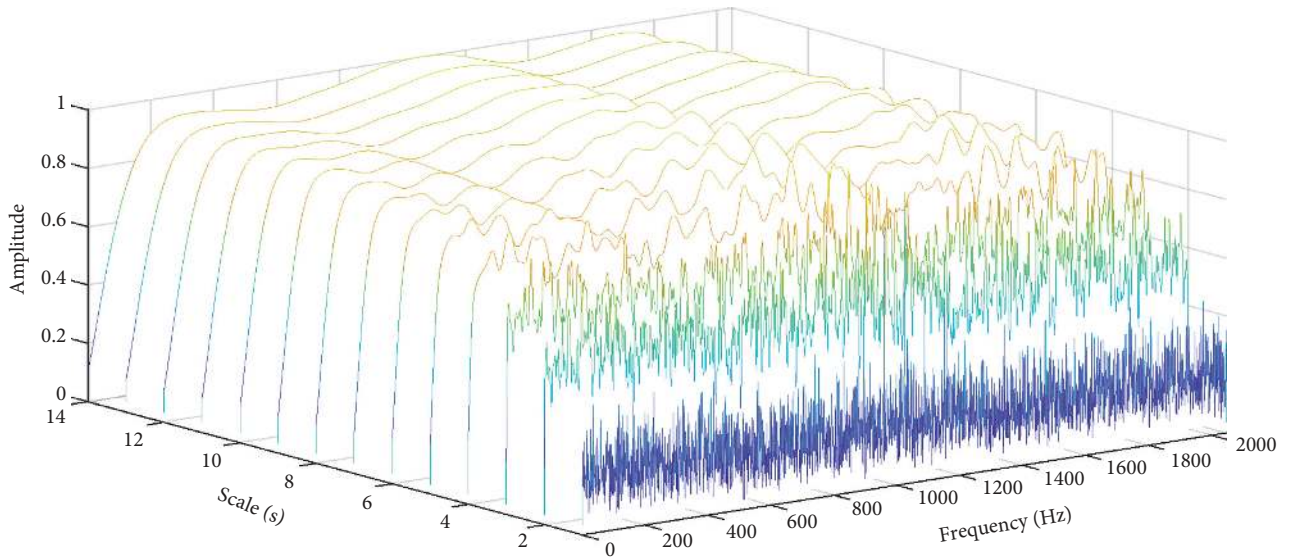

Figure 3: The discrete SSRs L(f, s) with different scale parameters.

clustering methods are shown in Figure 4(a). There is only one resonance frequency band in the simulation signal; hence, two boundaries are expected (i.e., one resonance frequency between two boundaries). The numbers of boundaries determined by the SSR with the different two-class clustering methods (i.e., the Otsu, k-means, mean, and probabilistic methods) are 16, 16, 100, and 48, respectively. The minimum number of boundaries obtained by k-means clustering (16) and the detected boundaries are shown in Figure 4(b), from which too many false boundaries are evident, thereby providing an incorrect determination of the practical resonance frequency band.

The number of minima declines with an increase in the step according to a similar logarithmic law of attenuation, as

shown in Figure 5. There are two attenuation stages with different slopes. However, during the calculation of $L(f, s)$, the scale parameter $\sqrt{s}$ (i.e., the smoothing effect is related to $\sqrt{s}$) increases linearly. As a result, Gaussian functions with an excessive number of steps are used to smooth the Fourier spectra of the analysed signal; consequently, the resulting $L(f, s)$ is so redundant that numerous error boundaries without any information partitions are found. When the step reaches a maximum number, the number of minimum of $L(f, s)$ is still 11, and it is greater than 2 (as expected). Therefore, a fixed number of steps is not suitable for diagnosing fault bearing signals. In addition, the cost of generating the SSR is fairly expensive. Therefore, traditional SSRs have two main shortcomings:
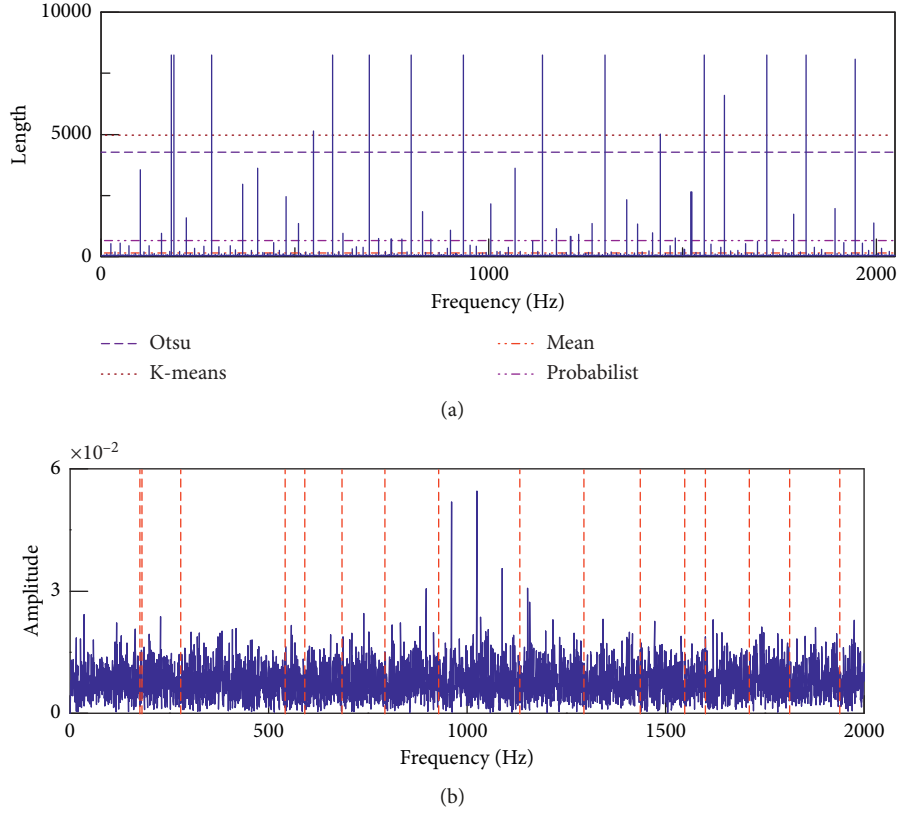
(a)

(b)

FIGURE 4: (a) The SSP and the classification thresholds; (b) the detected boundaries.

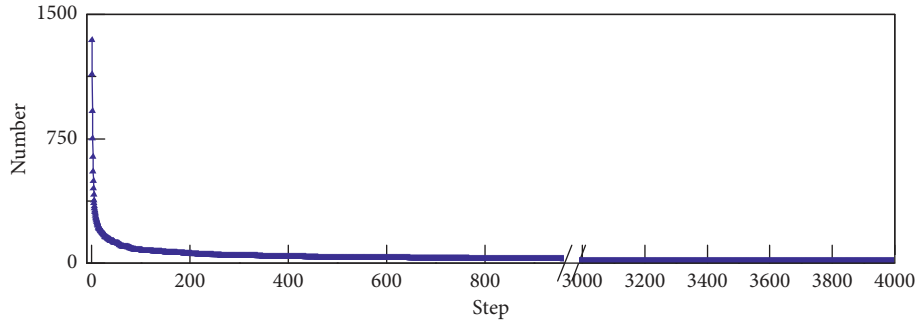FIGURE 5: The number of minima at different numbers of scale steps.

(1) Too many false boundaries are found, preventing a determination of the actual boundaries

(2) The cost of generating the SSR is fairly high

3. The Proposed OSSR Method for Boundaries Selection

3.1. Optimized Scale-Space Representation. To overcome the abovementioned shortcomings, a novel scale-updating mechanism is proposed as follows:

$$\sqrt{s} = \begin{cases} N\sqrt{s_0} & \text{if } W_{N-1} \leq W_T, \\ N^2\sqrt{s_0} & \text{if } W_{N-1} > W_T, \end{cases} \tag{7}$$

where W_{N-1} denotes the number of local minima of $L(f,s)$ at the N-1 th step, and W_T is the minimum threshold

number set as ten percent of the initial minimum number (W_0) according to experimental analysis. When the number of local minima is less than the set threshold W_T, the scale parameter is updated linearly ($\sqrt{s}=N\sqrt{s_0}$); when the number of minima is greater than the set threshold W_T, it is updated quadratically ($\sqrt{s} = N^2\sqrt{s_0}$). Therefore, upon reaching the stage of minimum rapid change (Figure 5), the novel updating mechanism captures the crucial information for the boundary partitions. Meanwhile, in the slow stage, it not only eliminates the scale redundancy but also minimizes the calculation cost with fewer steps.

Meanwhile, the traditional maximum number of steps $N_{\max}$ is replaced by N_e corresponding to the no minimum in $L(f, N_e^2\sqrt{s_0})$. Therefore, the dynamic determination of the maximum step can expand the scale values to perform effectively with relatively few boundary divisions.

To illustrate the advantages of this approach, the optimized scale-space plane (OSSP) of a simulated signal is shown in Figure 6. The time required to calculate the OSSP using the OSSR on a laptop computer (i5-4460 processor, 3.20 GHz) is shown in Figure 7.

From Figure 7, it is obvious that the computational cost of the OSSR is much less than that of the SSR. Through a comparison between Figures 6 and 4(a), the number of false boundaries decreases obviously, and the SSCs (B1 and B2) that delineate the practical resonance band are clearer. Therefore, the OSSR exhibits not only a better performance with regard to the determination of boundaries but also a faster calculation speed than the SSR approach.

3.2. The SSC Selection Method. In the OSSR approach, as the length of the SSCs increase, the probability of becoming a valid boundary increases. Therefore, the problem of finding a meaningful mode can be solved by observing the change in the lengths of SSCs. The lengths L_i of SSCs are sorted in descending order ($L_i^d > L_{i-1}^d > \ldots > L_0^d$). Then, the difference ΔL_i^d in the lengths L_i^d of the SSCs are calculated and expressed as

$$\Delta L_i^d = L_i^d - L_{i-1}^d, \tag{8}$$

The sequences of L_i^d and ΔL_i^d of the simulated signal are depicted in Figure 8. Obviously, the first peak of ΔL_i^d appears at an index of 2. Consequently, the two longest SSCs are selected to delimit the resonance frequency band, and the determined boundaries are shown in Figure 9.

Compared with Figure 5, the OSSR approach and the proposed boundary selection algorithm can more accurately determine the boundaries used to delineate the resonance band containing fruitful fault information, as shown in Figure 9.

4. Procedure of the Proposed Method

The procedure of the proposed EWTOSSR method is summarized in this section, and it is shown in Figure 10.

(1) Compute the Fourier spectrum of the analysed signal

(2) Calculate the OSSP

(3) Detect the boundaries using the differences in the lengths of scale-space curves

(4) Decompose the analysed signals using the EWT

(5) Extract the fault information from the squared envelope spectra of the decomposed signals.

5. Simulation Verification

To verify the effectiveness of the proposed method, the signal simulated using Equation (6) in Section 2 is analysed by the EWTOSSR technique. According the EWTOSSR procedure, the OSSP is shown in Figure 6. The differences in the length of the SSCs are shown in Figure 8. The peak of the difference appears at an index of 2. Consequently, the two longest SSCs are selected to delimit the resonance frequency band. The
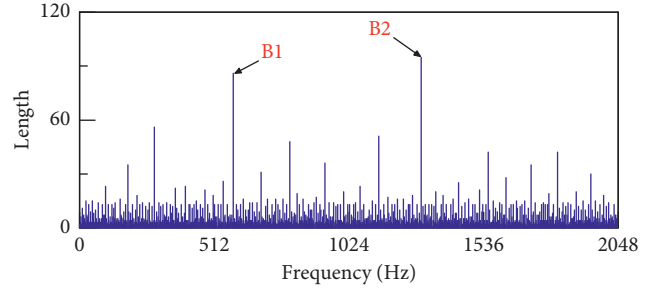

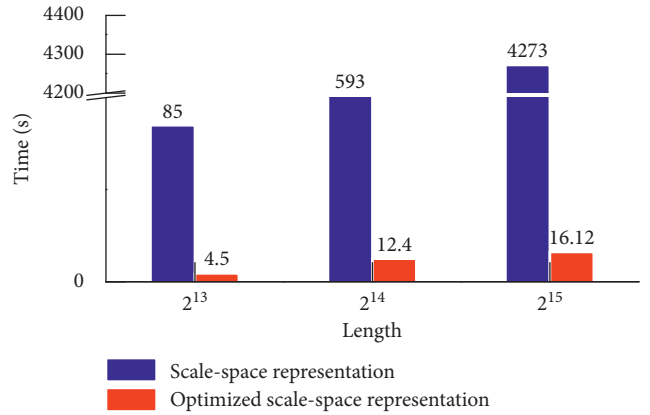FIGURE 6: The OSSP obtained by the OSSR approach.
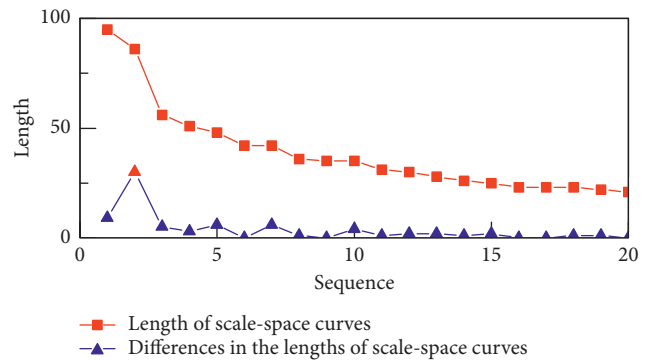

FIGURE 7: A comparison of the calculation costs.

FIGURE 8: The lengths of scale-space curves and the differences in the lengths of scale-space curves.
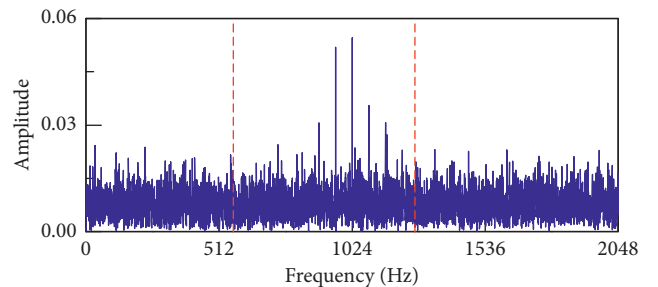

FIGURE 9: The boundaries obtained by the OSSR approach and the differences in the lengths of scale-space curves.
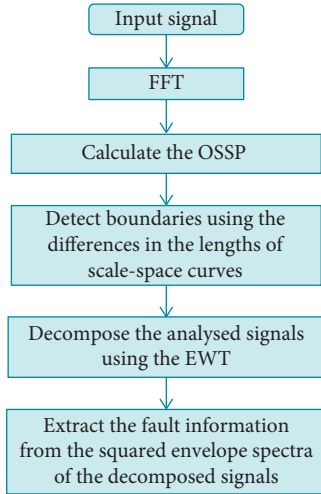
FIGURE 10: The procedure of the proposed method.

determined boundaries (548 Hz and 1299 Hz) are shown in Figure 9. Finally, three sub-band signals are obtained, and their squared envelope spectra are calculated and shown in Figure 11. Consequently, the characteristic fault frequency $1/T_p$ and its first three harmonics can be observed clearly in Figure 11(e). Therefore, the bearing fault has been identified by the proposed EWTOSSR.

To illustrate the advantages of the proposed EWTOSSR technique, three other well-known methods, namely, the traditional EWT [10], ensemble empirical mode decomposition (EEMD; the standard deviation set to 0.3, and the number of ensembles is set to 100) [38], and the fast kurtogram [6], are used to analyse the same signal simulated in Section 2. By the traditional EWT method, 17 subband signals are obtained. However, only the sub-band signals 7–10 approaching the resonant frequency band are depicted in Figure 12 because they have the greatest probability of extracting characteristic fault frequency from their squared envelope spectrum. The EEMD decomposes the noised signal into 13 intrinsic mode functions (IMFs). Because of the IMF5-13 in the low-frequency domain may not be modulated by fault characteristic frequency, only IMF1-4 and their resulting squared envelope spectra are shown in Figure 13. The fast kurtogram are shown in Figure 14(a), and the squared envelope and squared envelope spectrum with the largest kurtosis have been shown in Figures 14(b) and 14(c). The characteristic fault frequency $1/T_p$ obtained by the traditional EWT can be observed only in Figure 12(g); no more harmonics could be detected. The EEMD and fast kurtogram are unable to extract any useful fault information; instead, the fault characteristic frequency and its harmonics are still submerged in strong noise.

Through comparisons of Figures 12–14 with Figure 11, the EWTOSSR fault-detecting performance is superior to those of the three other methods. This is because the OSSR is capable of detecting relatively reasonable and effective frequency band boundaries containing maximal fault information and minimal noise, thereby highlighting the fault features.

6. Experimental Validation

To validate the effectiveness of the proposed EWTOSSR method, a testing bench is constructed, as shown in Figure 15(a). The test bench consists of a motor, a drive wheel, a loading device, a wheelset, and an axle box. The power is conveyed to the driving wheel through rubber belts. The traction power of the driving wheel is then transmitted to the wheelset. Two types of artificial faults on the outer race and rollers are shown in Figures 15(b) and 15(c), respectively [39]. The parameters of the tested wheelset bearing are listed in Table 2. The fault experiments are conducted at a rotation speed of 50 km/h, and the vibration signals are collected at a sampling frequency of 10 kHz.

6.1. Outer-Race Fault Detection. The signals measured when the wheelset bearing in Figure 15(b) with three outer-race faults is installed on the tested bench shown in Figure 15(a) are shown in Figure 16(a). The EWTOSSR method is used to analyse the measured signal. The OSSP obtained according to the EWTOSSR calculation procedure is shown in Figure 16(b), and the calculated differences in the lengths of the SSCs are shown in Figure 16(c). The peak of the difference appears at an index of 6, and thus, the 6 longest scale-space curves are selected as the boundaries. The six frequencies for determining the boundaries are 811 Hz, 1333 Hz, 1615 Hz, 2168 Hz, 2417 Hz, and 3477 Hz, and the detected boundaries are depicted in Figure 16(d). As a result, the analysed signal in Figure 16(a) is decomposed into seven sub-band signals, and the sub-band signals are shown in the left column of Figure 17, while the resulting squared envelope spectra are shown in the right column of Figure 17. The outer-race fault characteristic frequency is expressed as [39]

$$f_{\text{BPFO}} = \frac{Zf_r}{2}\left(1 - \frac{d}{D}\cos\alpha\right), \tag{9}$$

where f_r is the rotation frequency of the wheelset, d is the roller diameter, D denotes the pitch diameter, Z is the number of rollers, and α denotes the contact angle. The rotation frequency f_r corresponding to a running speed of 50 km/h is 5.1 Hz. According to the parameters in Table 2, the outer-race fault characteristic frequency is 41.6 Hz. The sub-band signal filtered by a wavelet transform with Fourier support (811 Hz, 1333 Hz) can be used to extract the fault information. The outer-race characteristic fault frequency f_{BPFO} and its first five harmonics are clearly displayed in Figure 17(i). Consequently, it is determined that the bearing has suffered an outer-race fault.

Similarly, the three other methods are employed to analyse the same signal shown in Figure 16(a) for comparison. The results obtained by the traditional EWT are shown in Figure 18. The SSP and the threshold obtained by k-means are shown in Figure 18(a), and the resulting boundaries are displayed in Figure 18(d). According to the detected boundaries, the signal is decomposed into 25 subband signals. As the frequency band is too narrow, most of them are unable to extract fault information, and only the fifth and sixth sub-band signals contain feature information
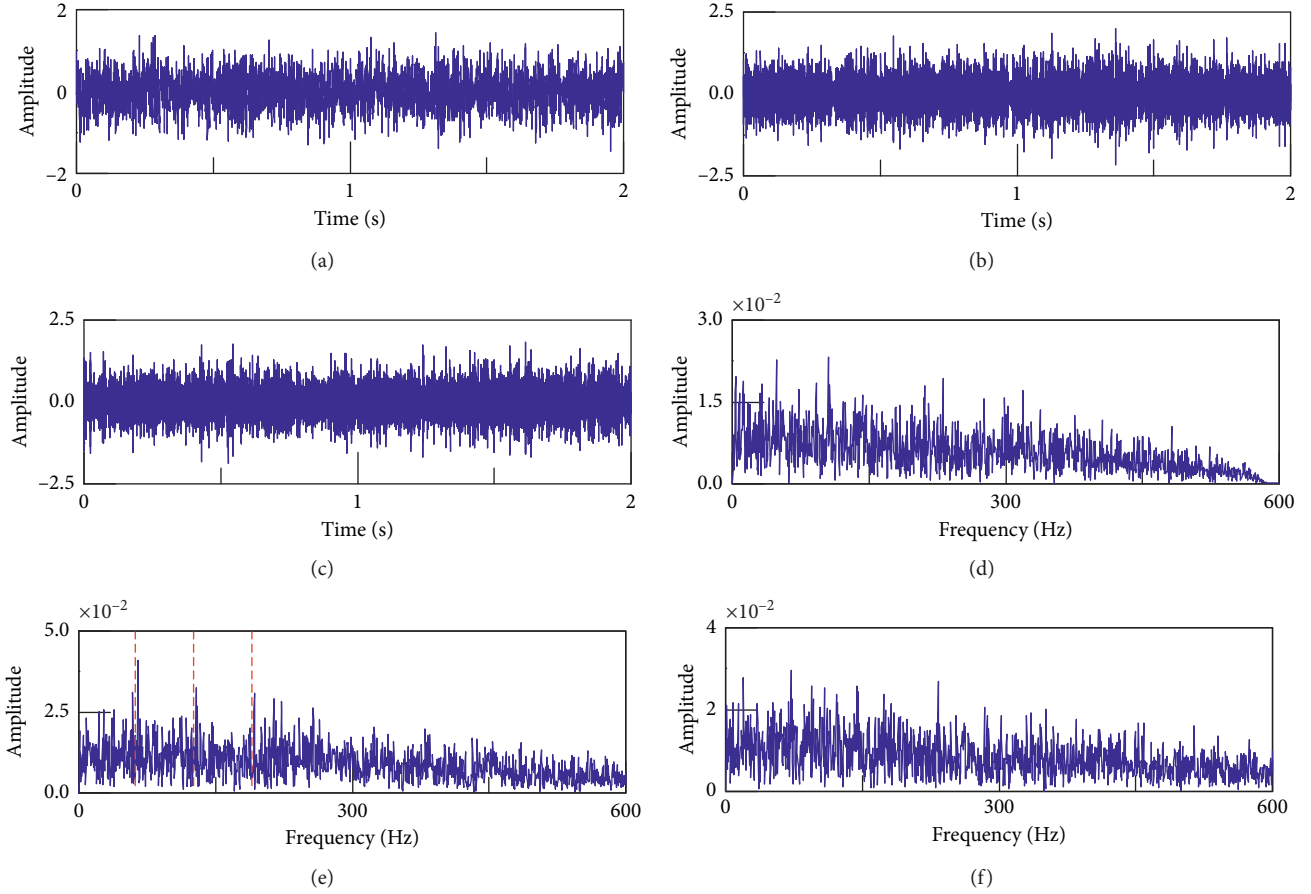
FIGURE 11: The proposed EWTOSSR method for diagnosing a simulated signal: (a–c) sub-band signals; (d–f) squared envelope spectra (red dashed lines: the first three harmonics of the characteristic fault frequency $1/T_p$).

about outer-race faults. The outer-race fault characteristic frequency f_{BPFO} and its previous three harmonics are extracted and shown in Figures 18(e) and 18(f). However, both the amplitude and the number of its harmonics are less than those in Figure 17(i). This is because the second frequency band of EWTOSSR (from 811 Hz to 1333 Hz) in Figure 16(d) is split into two bandwidths determined by the traditional EWT (the first band is from 809 Hz to 992 Hz and the second band is from 992 Hz to 1211 Hz). As a result, the width of the two sub-band signals is too narrow to contain more useful fault information. Then the signal is analysed by the EEMD, the signal is decomposed into 14 IMFs. Because of the IMF5-14 in the low-frequency domain may not be able to modulate the fault characteristic frequency, only IMF1-4 and their resulting squared envelope spectra are shown in Figure 19. The results obtained by fast kurtogram are shown in Figure 20. Obviously, the fault characteristic frequency and its harmonics cannot be identified from Figures 19 and 20. Thus, the EEMD and fast kurtogram could not extract the fault-related repetitive transients from the signal.

By comparison, only the EWT and the proposed EWTOSSR have successfully extracted the fault impulses; however, the fault features obtained by the proposed method are more significant. Thus, the fault-detecting performance of the EWTOSSR is superior to the three comparison methods.

6.2. Rolling Fault Detection. The measured signal of the rolling fault shown in Figure 15(c) is shown in Figure 21(a) under the test speed. Similarly, the OSSP is obtained and shown in Figure 21(f), and the differences in the lengths of the SSCs are shown in Figure 21(b). The peak of the difference appears at an index of 2, and thus, the 2 longest SSCs are selected as the boundaries. The two frequencies for determining the boundaries are 978 Hz and 2524 Hz, and the resulting boundaries are displayed in Figure 21(g). Therefore, the measured signal is decomposed into three sub-band signals, and the sub-band signals and their squared envelope spectra are shown in Figure 21. The rolling fault characteristic frequency f_{BSF} is expressed as

$$f_{BSF} = \frac{Df_r}{2d}\left(1 - \left(\frac{d}{D}\cos\alpha\right)^2\right). \tag{10}$$

According to the rotation frequency of the wheelset and the wheelset bearing parameters listed in Table 2, f_{BSF} is equal to 33.9 Hz. In both Figures 21(i) and 21(j), the rolling fault characteristic frequency f_{BSF} and its harmonics are clearly extracted.

To further illustrate the advantages of the EWTOSSR method, the three other methods are used to analyse the same measured signal in Figure 21(a) for comparison.
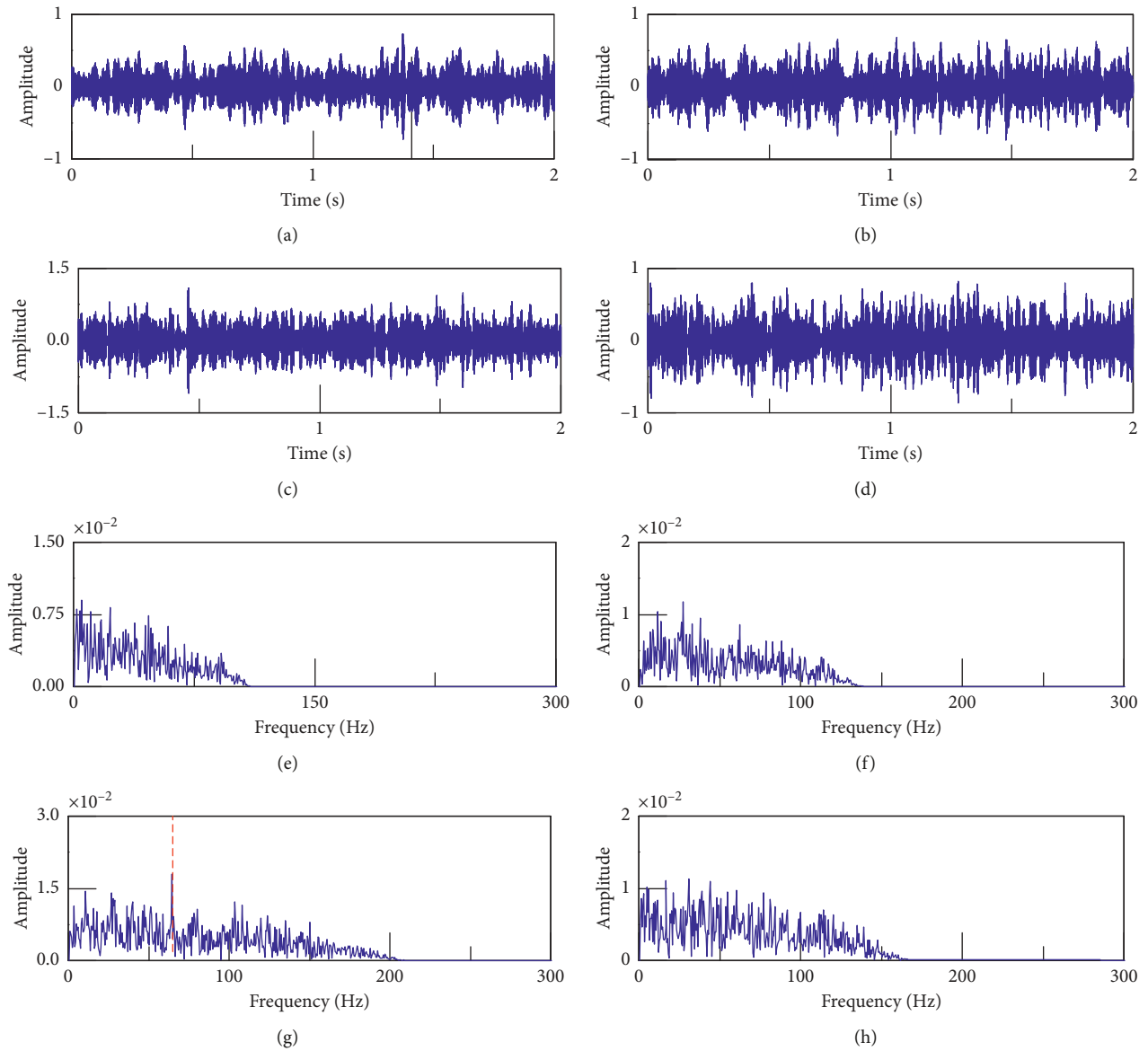
Figure 12: Results obtained by the traditional EWT and SSR: (a–d) sub-band signals; (e–h) squared envelope spectra (red dashed line: characteristic fault frequency $1/T_p$).
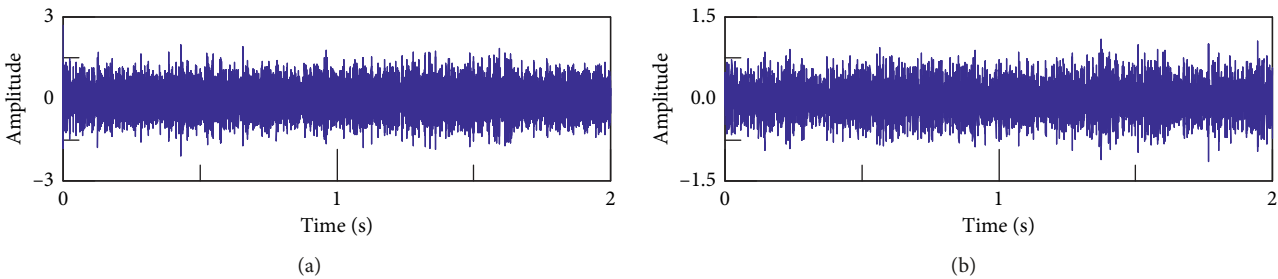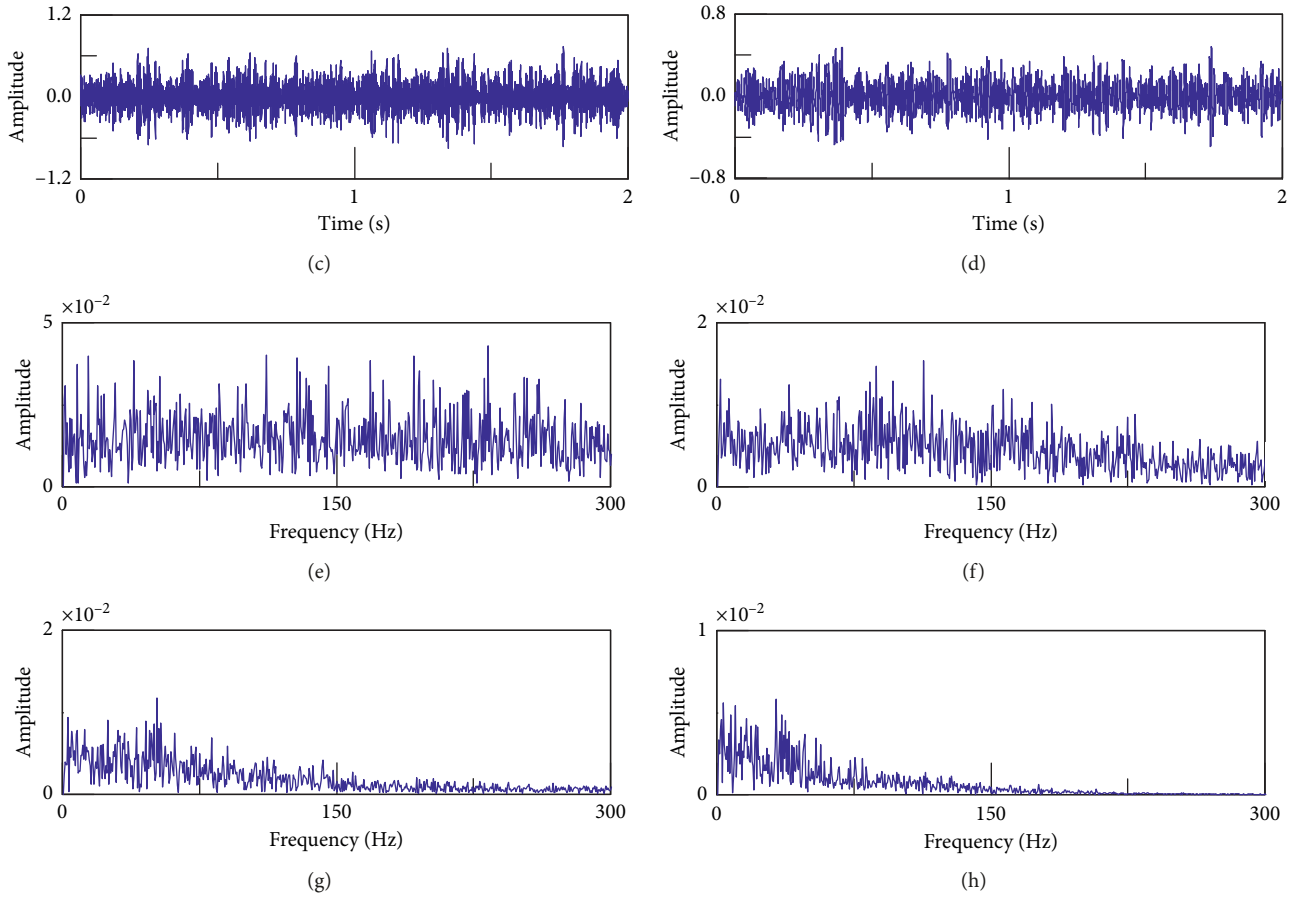

Figure 13: Continued.

FIGURE 13: Results obtained by EEMD: (a–d) IMFs; (e–h) squared envelope spectra.
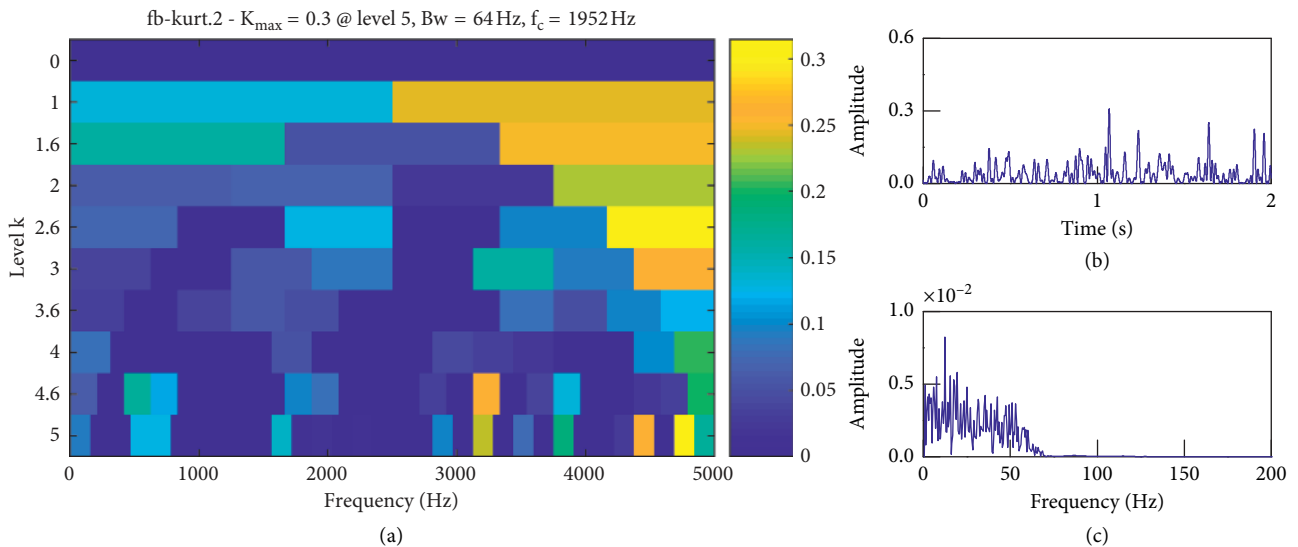

FIGURE 14: Results obtained by the fast kurtogram: (a) kurtogram; (b) squared envelope of the node with the largest kurtosis value; (c) squared envelope spectrum of (b).

According to the process of the EWT, the SSP and threshold obtained by k-means are firstly calculated and shown in Figure 22(a), and the corresponding boundaries are depicted in Figure 22(b). Therefore, there are 16 sub-band signals obtained by EWT, and the sixth and seventh sub-band signals have the most obvious fault characteristic frequency in the squared envelope spectra. Thus, only the sixth and seventh sub-band signals and their squared envelope
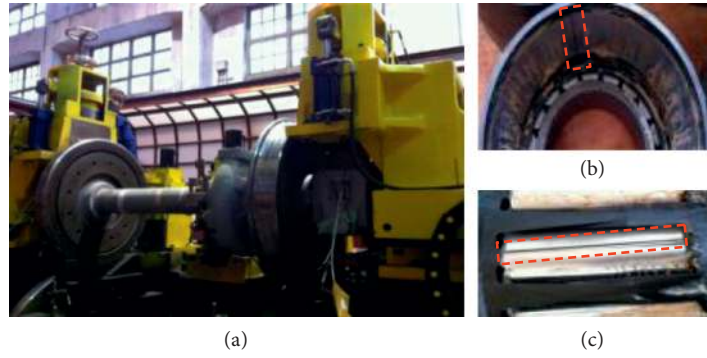
FIGURE 15: The test bench for investigating wheelset bearing faults and bearing defects: (a) bench; (b) outer-race defect; (c) roller defect.

TABLE 2: The parameters of the tested wheelset bearing.

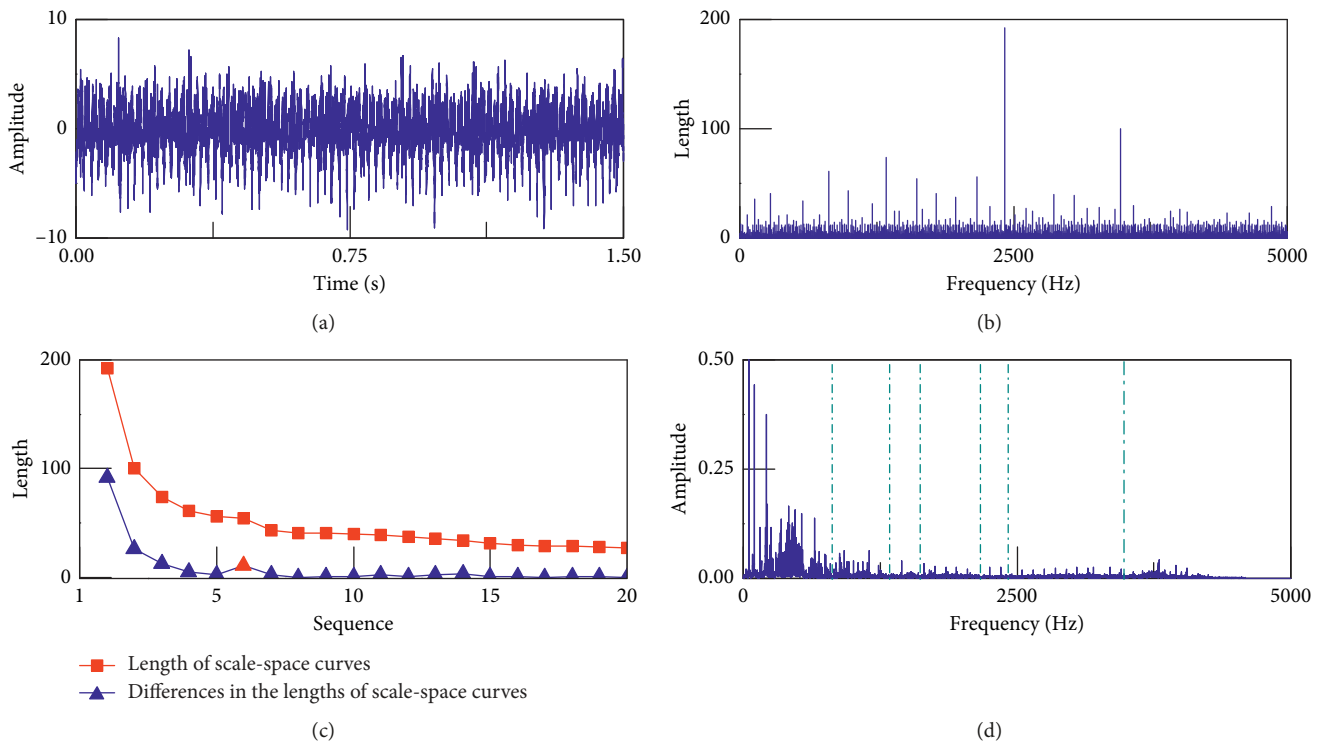| Roller number	Roller diameter (mm)	Pitch diameter (mm)	Contact angle (rad)	Load (kN)
19	26.9	180	0.1571	50

FIGURE 16: Results obtained by the EWTOSSR method: (a) measured signal; (b) OSSP; (c) differences in the lengths of the scale-space curves; (d) detected boundaries (green dash-dotted lines).

spectra have been illustrated in Figure 22. The same signal is also processed by the EEMD, and 14 IMFs are obtained. Because the IMF5-14 are low-frequency narrow-band signals which would not be modulated by defect-related frequency, only the first four IMFs and their squared envelope spectra are illustrated in Figure 23. The result obtained by fast kurtogram is shown in Figure 24. Obviously, the dominant components in Figures 22(e) and 22(f), Figure 23(e),

and Figure 24(c) are fault characteristic frequency and its harmonics.

Although these three methods can extract the characteristic rolling fault frequency f_{BSF} and its harmonics, both the amplitudes of f_{BSF} and the number of its harmonics are less than those extracted by the EWTOSSR method. Therefore, the EWTOSSR method can be used to effectively determine the boundaries containing fruitful
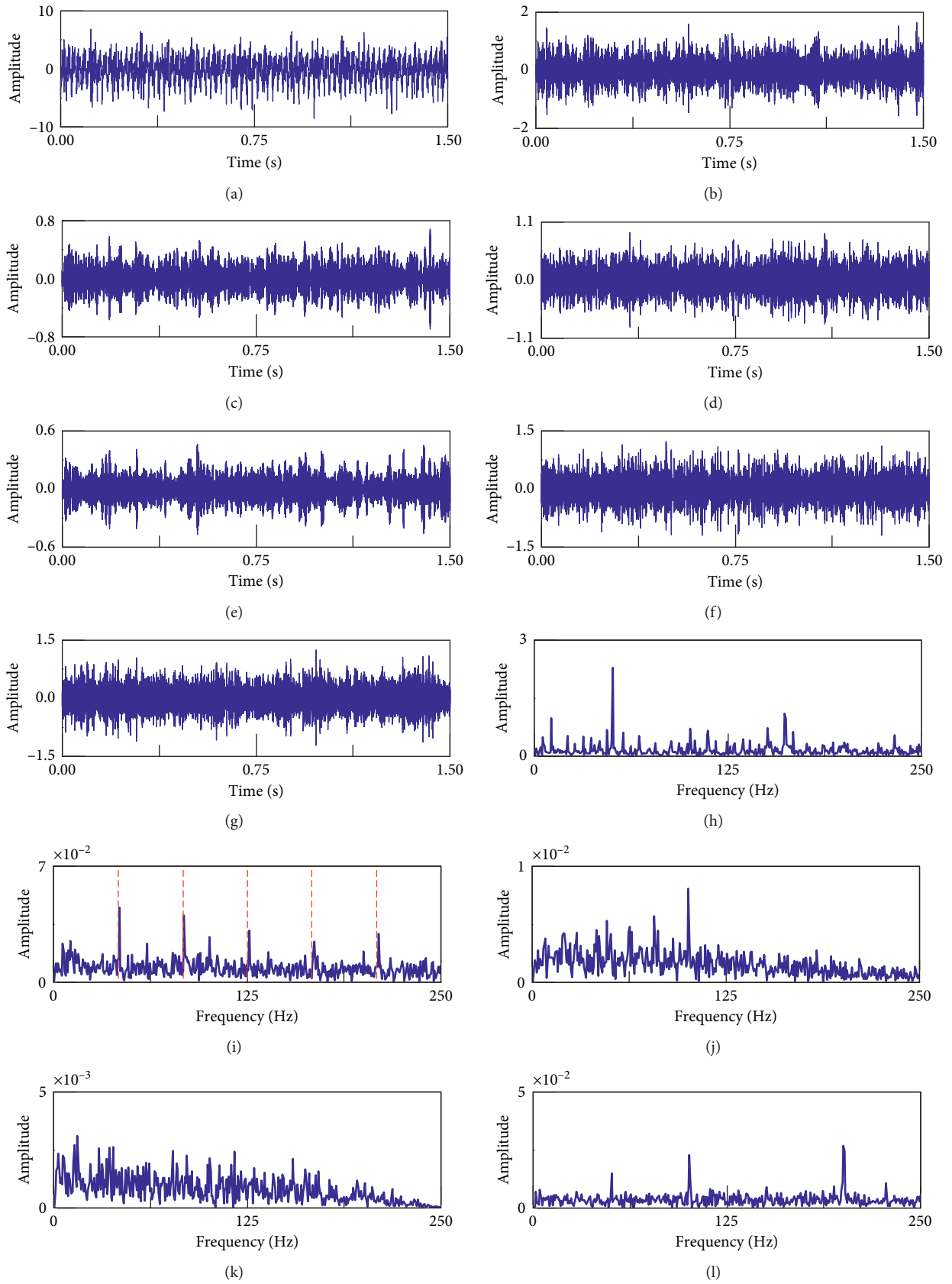
(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

(j)

(k)

(l)

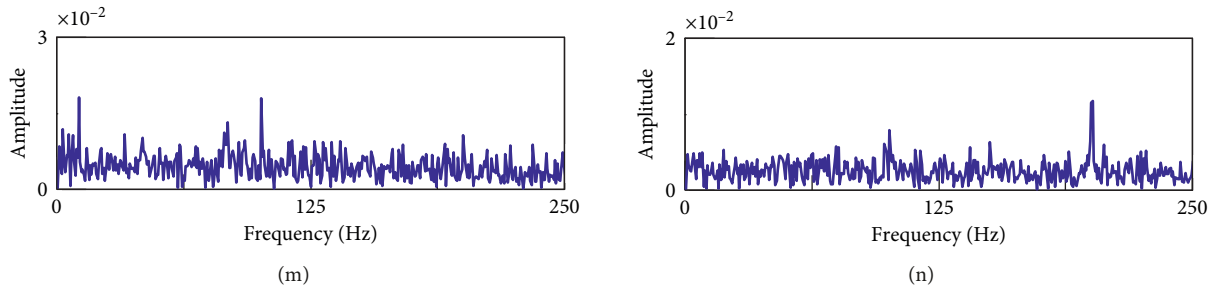FIGURE 17: Continued.

(m)

(n)

FIGURE 17: Sub-band signals and their envelope spectra: the left column shows the sub-band signals; the right column shows the squared envelope spectra corresponding to the left column (red dashed lines: f_{BPFO} and its harmonics).
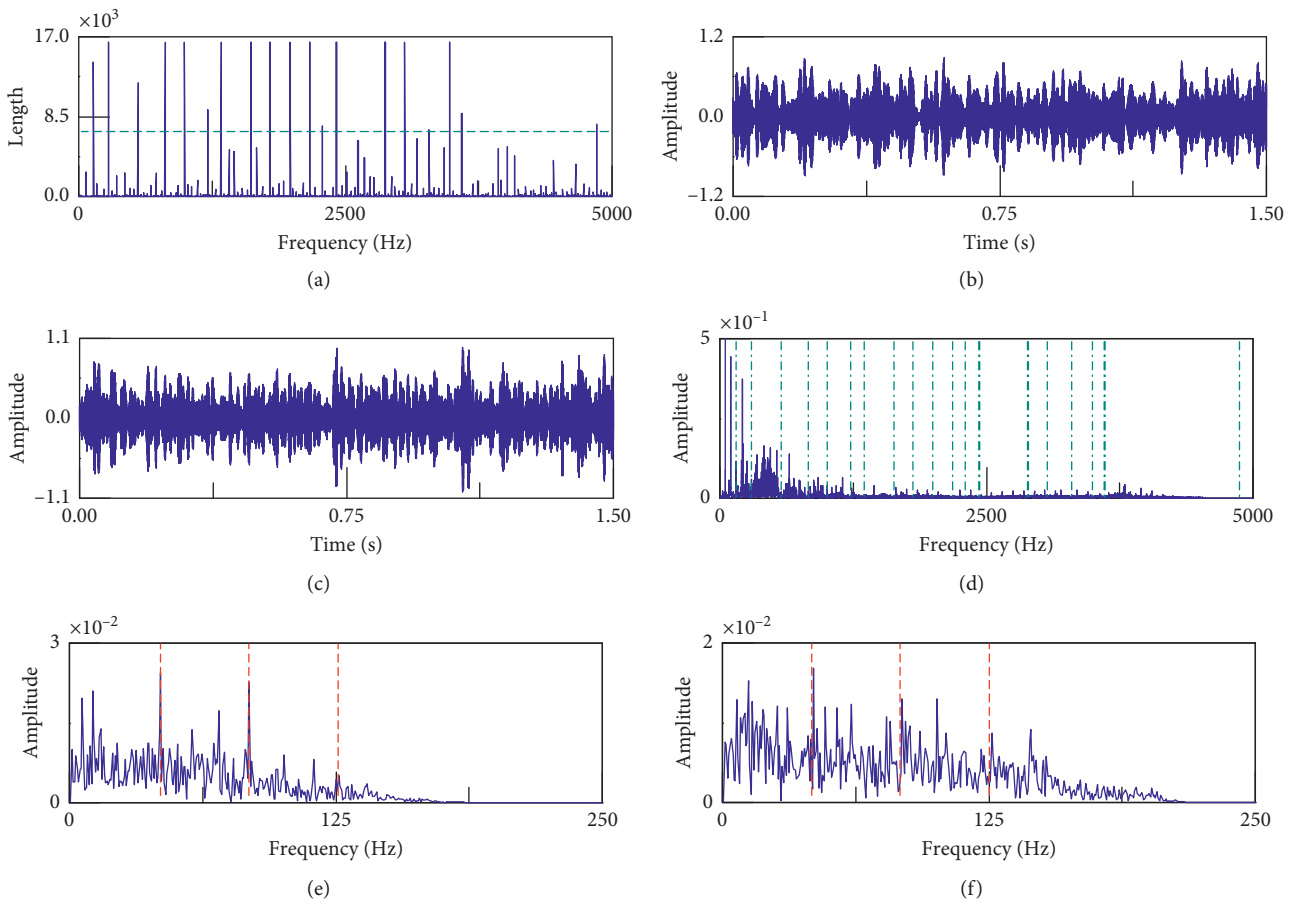

(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 18: Results obtained by the traditional EWT: (a) SSP (green dashed line: the detected threshold by k-means clustering), (b, c) sub-band signals; (d) detected boundaries (green dash-dotted lines); (e, f) squared envelope spectra (red dashed lines: f_{BPFO} and its harmonics).
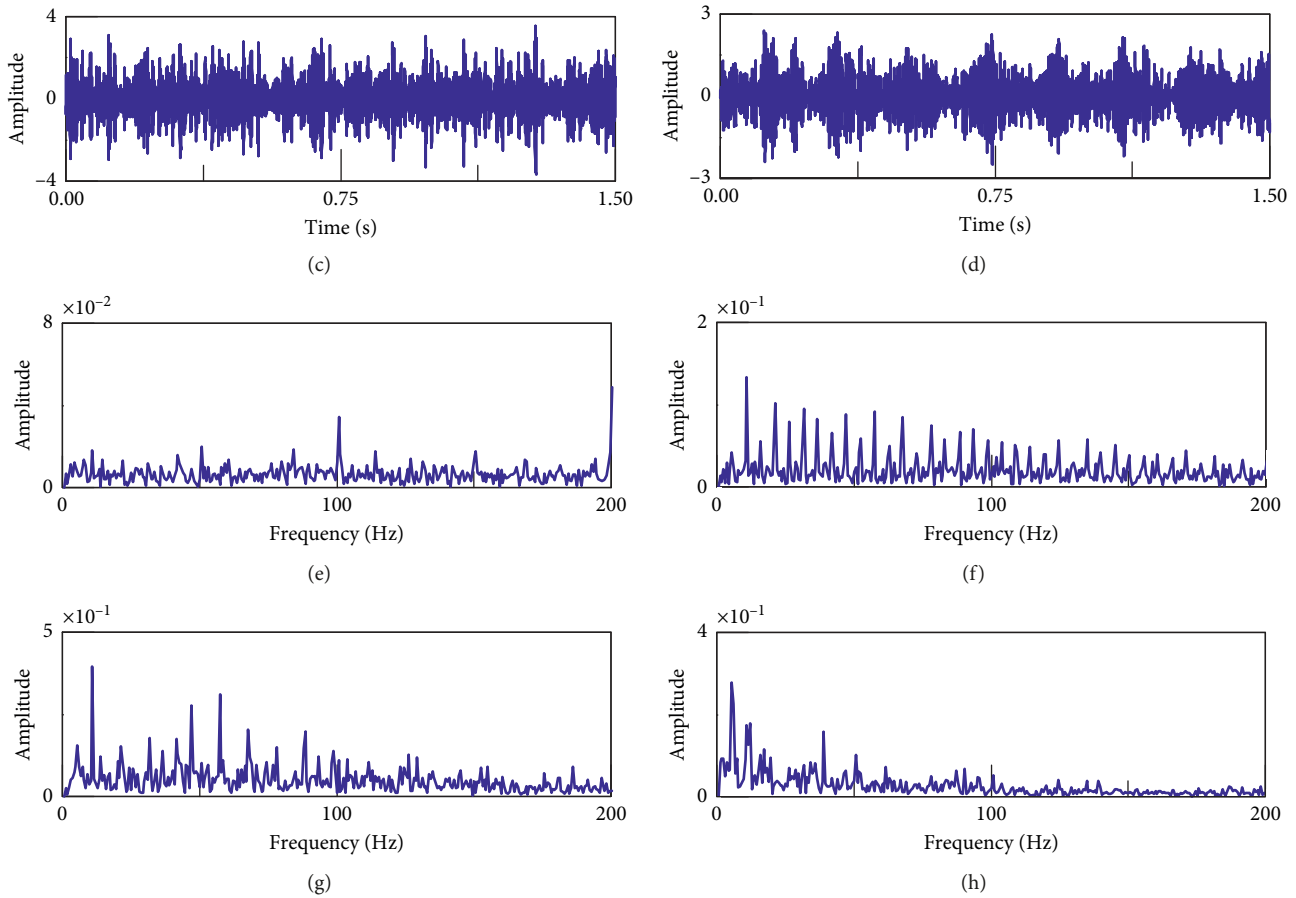

(a)

(b)
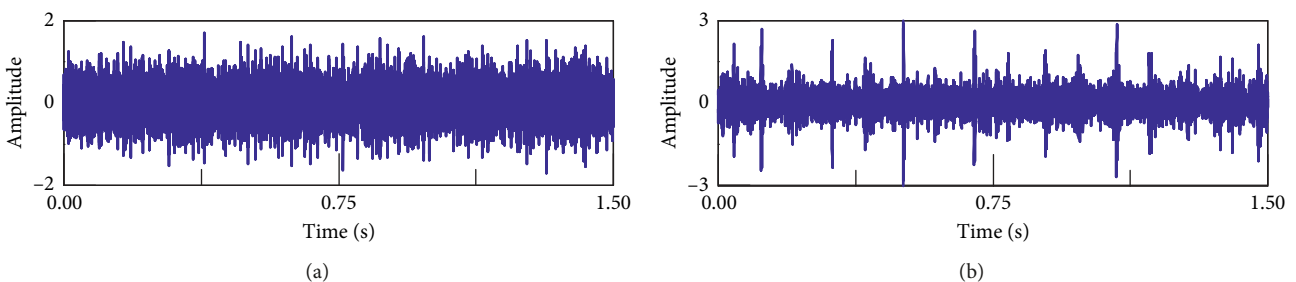
FIGURE 19: Continued.

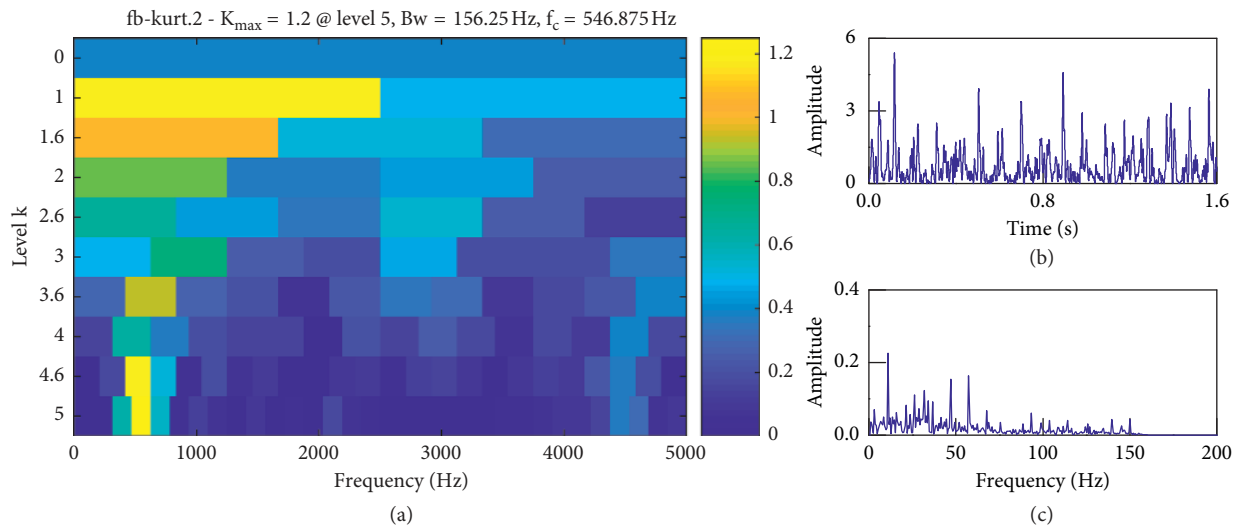FIGURE 19: Results obtained by EEMD: (a–d) IMFs; (e–h) squared envelope spectra.

FIGURE 20: Results obtained by the fast kurtogram: (a) kurtogram; (b) squared envelope of the node with the largest kurtosis value; (c) squared envelope spectrum of (b).

fault information, and it is suitable for extracting weak faults from measured signals.

6.3. Multiple Fault Detection.
The vibration signals measured when there are the multiple outer-race faults (Figure 15(b))

and roller faults (Figure 15(c)) are shown in Figure 25(a). The OSSP and the differences in the lengths of the SSCs are shown in Figures 25(b) and 25(f), respectively. Similarly, the peak of the difference appears at an index of 2, and thus, the 2 longest SSCs are selected as the boundaries. The two frequencies for determining the boundaries are 977 Hz and 2576 Hz, and the
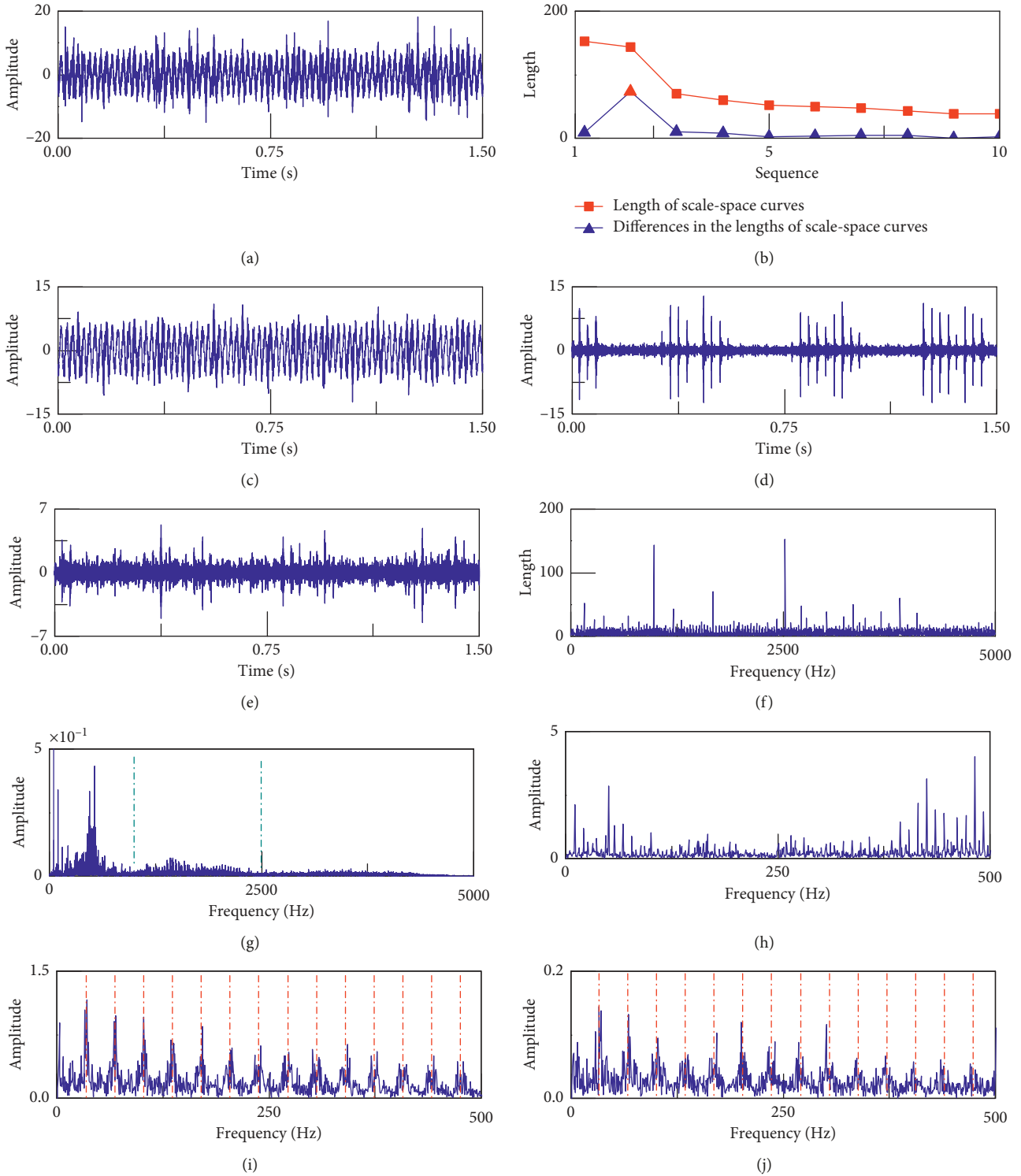
FIGURE 21: Results obtained by the EWTOSSR method: (a) measured signal; (b) differences in the lengths of scale-space curves; (c–e) sub-band signals; (f) OSSP, (g) detected boundaries (green dash-dotted line); (h–j) squared envelope spectra (red dash-dotted lines: f_{BSF} and its harmonics).

resulting boundaries are displayed in Figure 25(g). Therefore, the measured signal is decomposed into three sub-band signals. The sub-band signals and their squared envelope spectra are shown in Figure 25. The rolling characteristic fault frequency f_{BSF} and its first five harmonics are clearly

extracted in Figure 25(i). Meanwhile, the outer-race fault characteristic frequency f_{BPFO} and its first eleven harmonics are clearly displayed in Figure 25(j).

To verify the superiority of the proposed EWTOSSR method, the three other methods are used to analyse the same
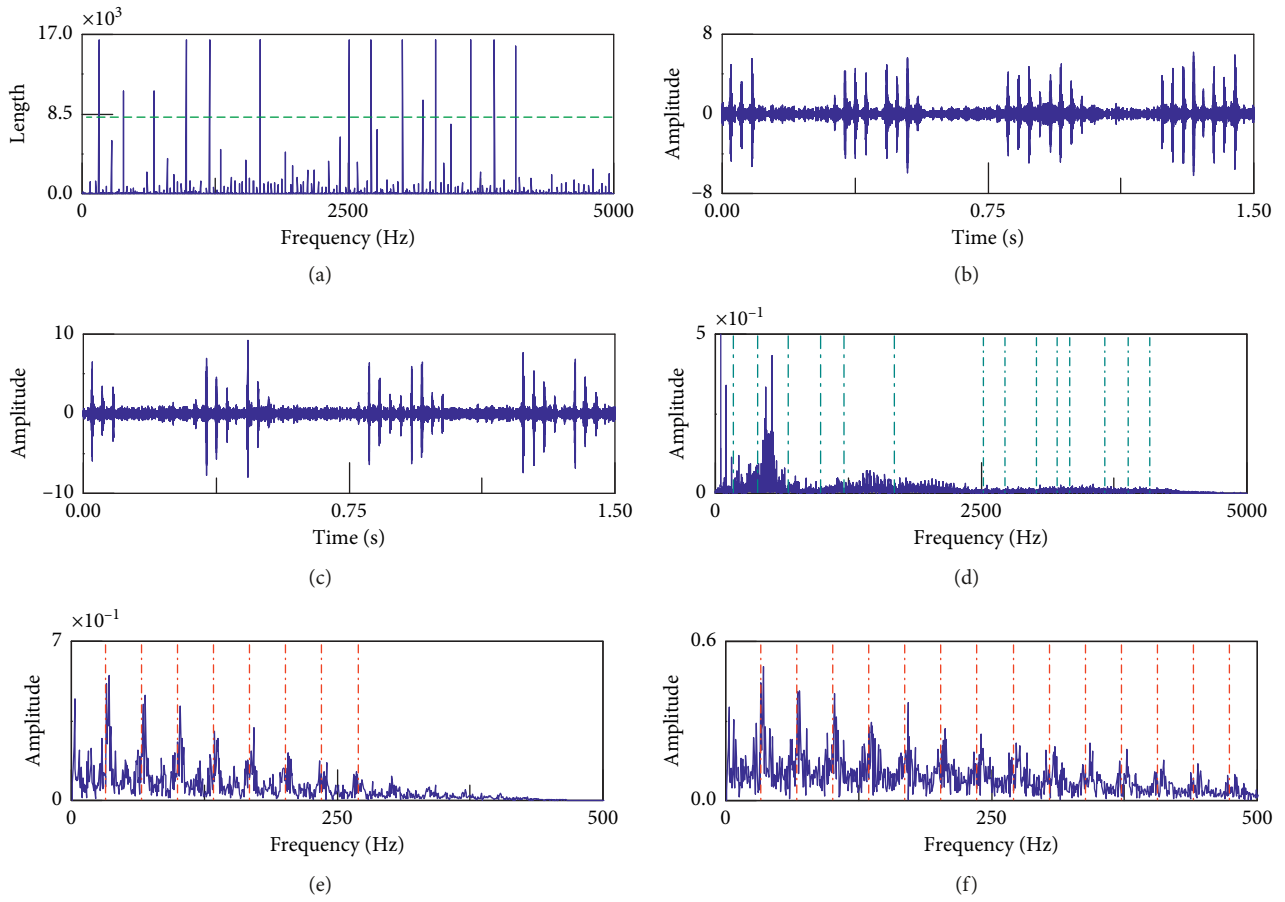
(a)

(b)

(c)

(d)

(e)

(f)

Figure 22: Results obtained by the traditional EWT: (a) SSP (green dashed line: the detected threshold by k-means clustering); (b, c) sub-band signals; (d) detected boundaries (green dash-dotted lines), (e, f) squared envelope spectra (red dash-dotted lines: f_{BSF} and its harmonics).
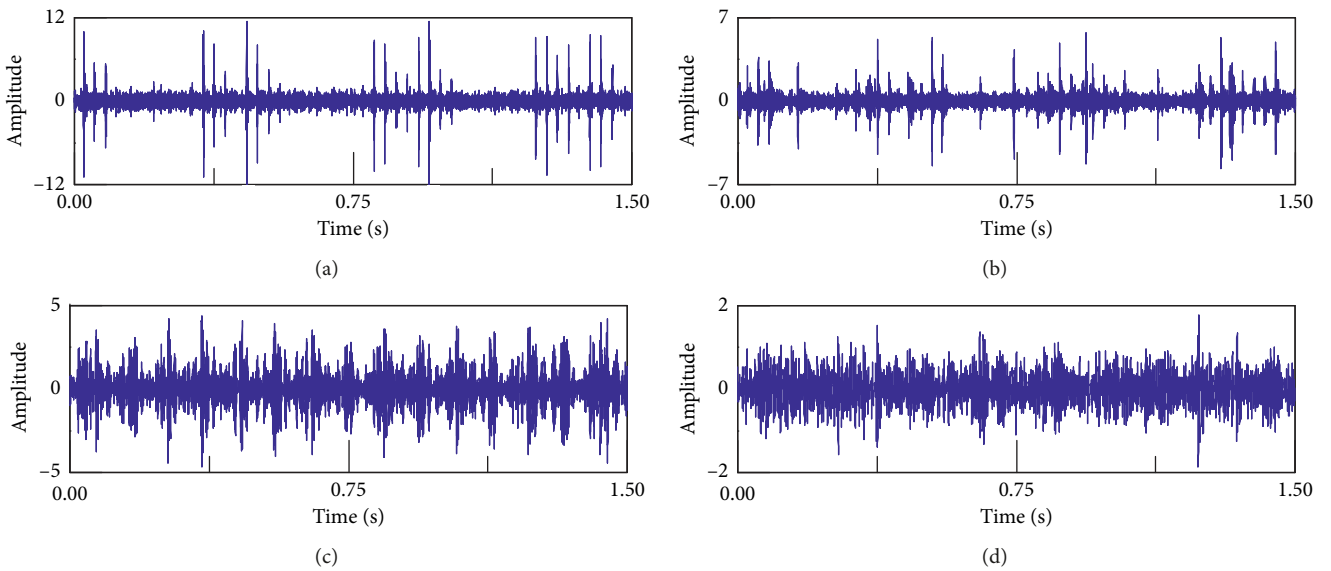
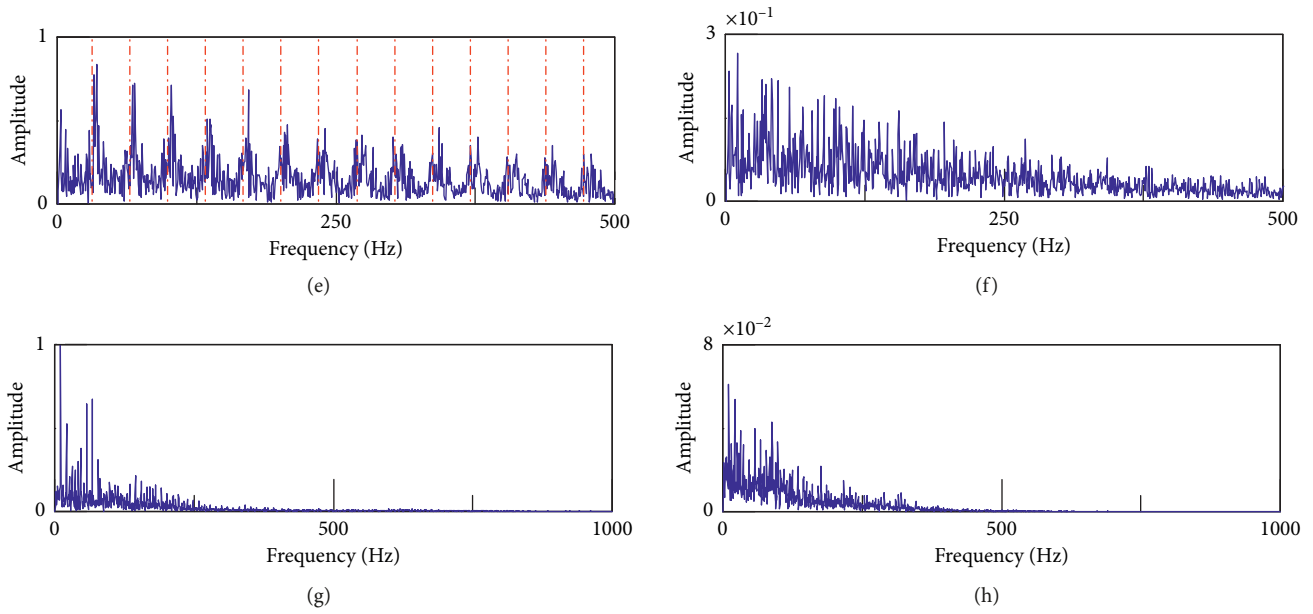

(a)

(b)

(c)

(d)

Figure 23: Continued.

(e)

(f)

(g)

(h)

FIGURE 23: Results obtained by EEMD: (a–d) IMFs; (e–h) squared envelope spectra (red dash-dotted lines: f_{BSF} and its harmonics).
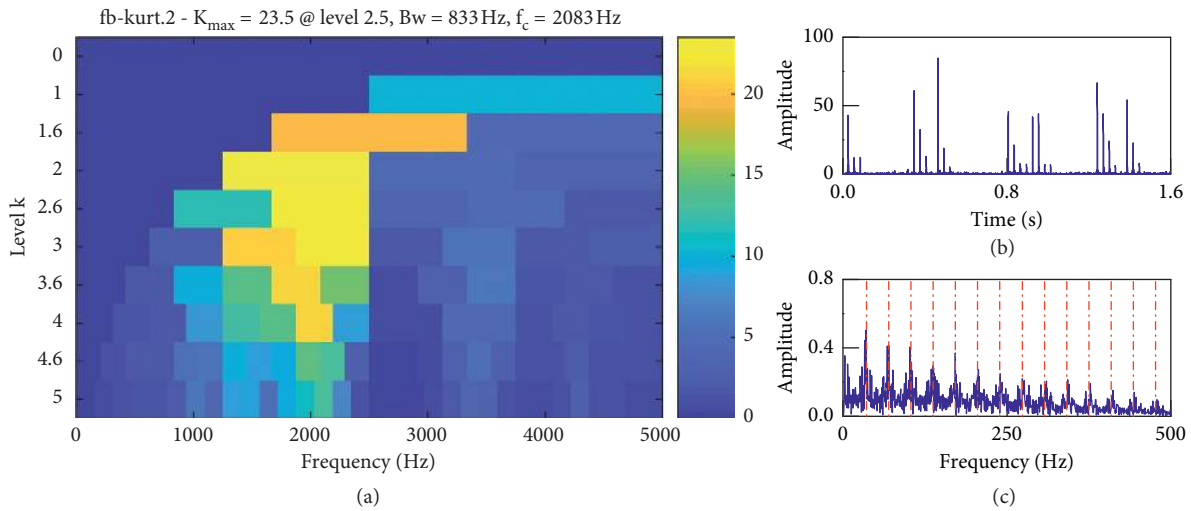

(a)

(b)

(c)

FIGURE 24: Results obtained by the fast kurtogram: (a) kurtogram; (b) squared envelope of the node with the largest kurtosis value; (c) squared envelope spectrum of (b) (red dash-dotted lines: f_{BSF} and its harmonics).

measured signal in Figure 25(a) for comparison. The SSP and threshold obtained by k-means are illustrated in Figure 26(a). According to the boundaries shown in Figure 26(e), the EWT decomposes the signal into 18 sub-band signals. Because the tenth, thirteenth, and sixteenth sub-band signals contain the most significant defective information about outer-race and roller faults; hence, only the corresponding filtered signal and squared envelope spectra of the tenth, thirteenth, and sixteenth sub-band signals are displayed in Figure 26. The rolling characteristic fault frequency f_{BSF} and its first three harmonics can be identified in Figure 26(f). Additionally, the outer-race fault characteristic frequency f_{BPFO} and its first six harmonics are clearly depicted in Figures 26(g) and 26(h). The EEMD is utilized to analyse the bearing multiple fault signal,

a total 14 IMFs is obtained. As the IMF5-14 are low-frequency narrow-band signals which would not be modulated by defect-related frequency, only the first four IMFs and their squared envelope spectra are depicted in Figure 27. Only the outer-race fault characteristic frequency f_{BPFO} and its first eleven harmonics are identified in Figure 27(c). The fast kurtogram is shown in Figure 28(a); the squared envelope and the squared envelope spectrum of the node with highest kurtosis have been displayed in Figures 28(b) and 28(c), respectively. From Figure 28(c), only the rolling characteristic fault frequency f_{BSF} and its first five harmonics are clearly identified.

Although the traditional EWT can extract f_{BSF} and f_{BPFO} and their harmonics, both the amplitudes of the

(a)

(b)

(c)

(d)

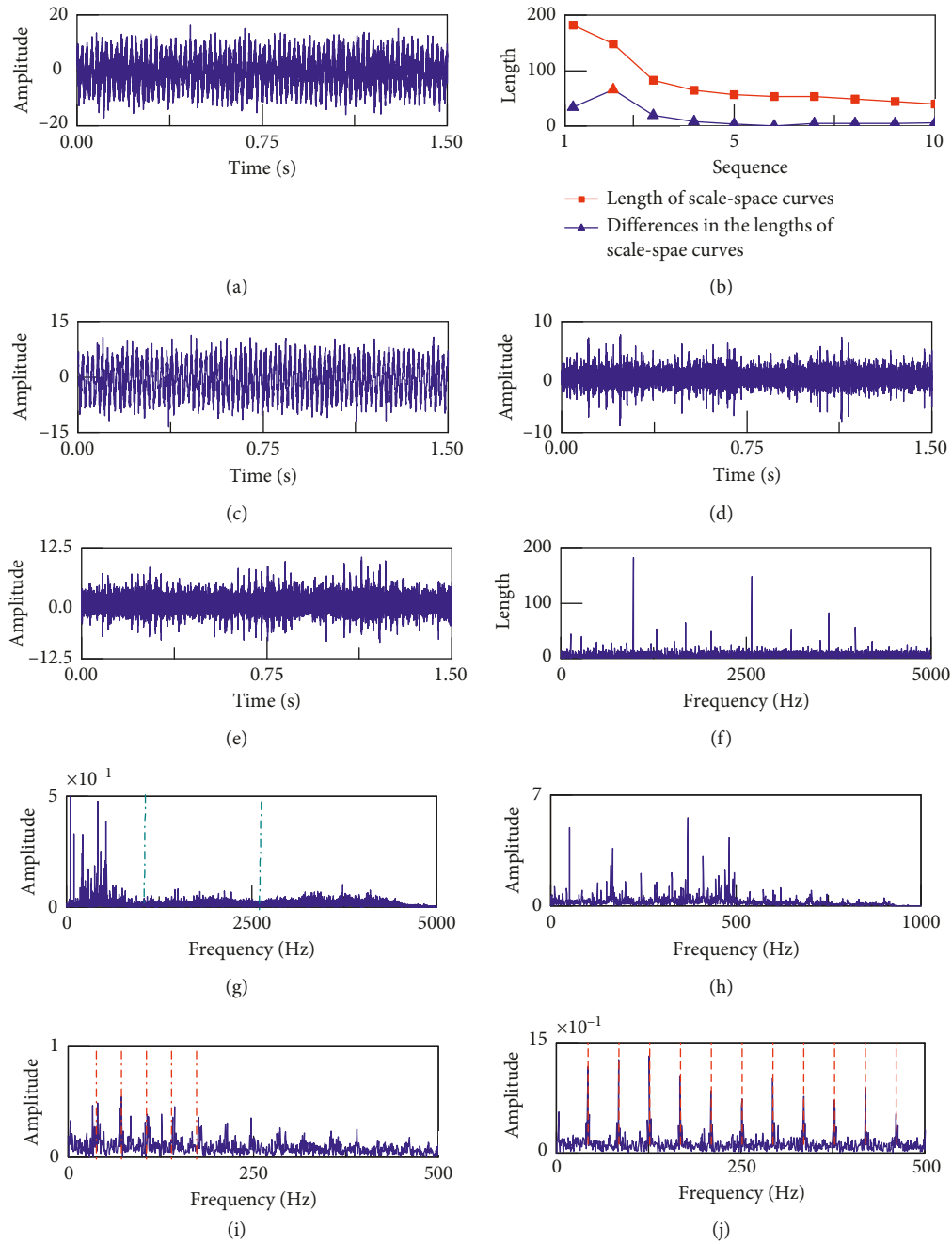(e)

(f)

(g)

(h)

(i)

(j)

FIGURE 25: Results obtained by the EWTOSSR: (a) measured signal; (b) differences in the lengths of scale-space curves; (c–e) sub-band signals; (f) OSSP; (g) detected boundaries (green dash-dotted lines); (h–j) squared envelope spectra (red dashed lines: f_{BPFO} and its harmonics; red dash-dotted lines: f_{BSF} and its harmonics).
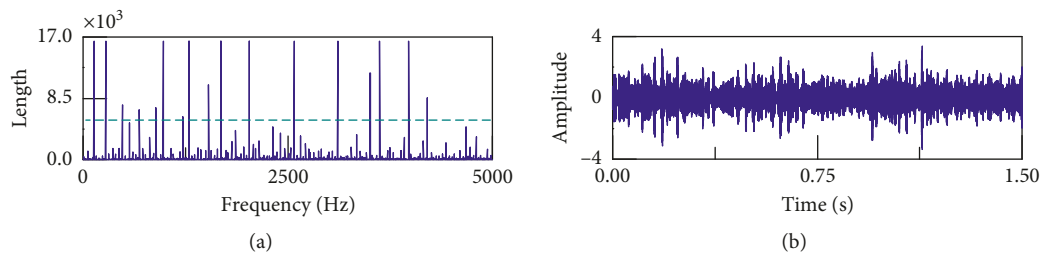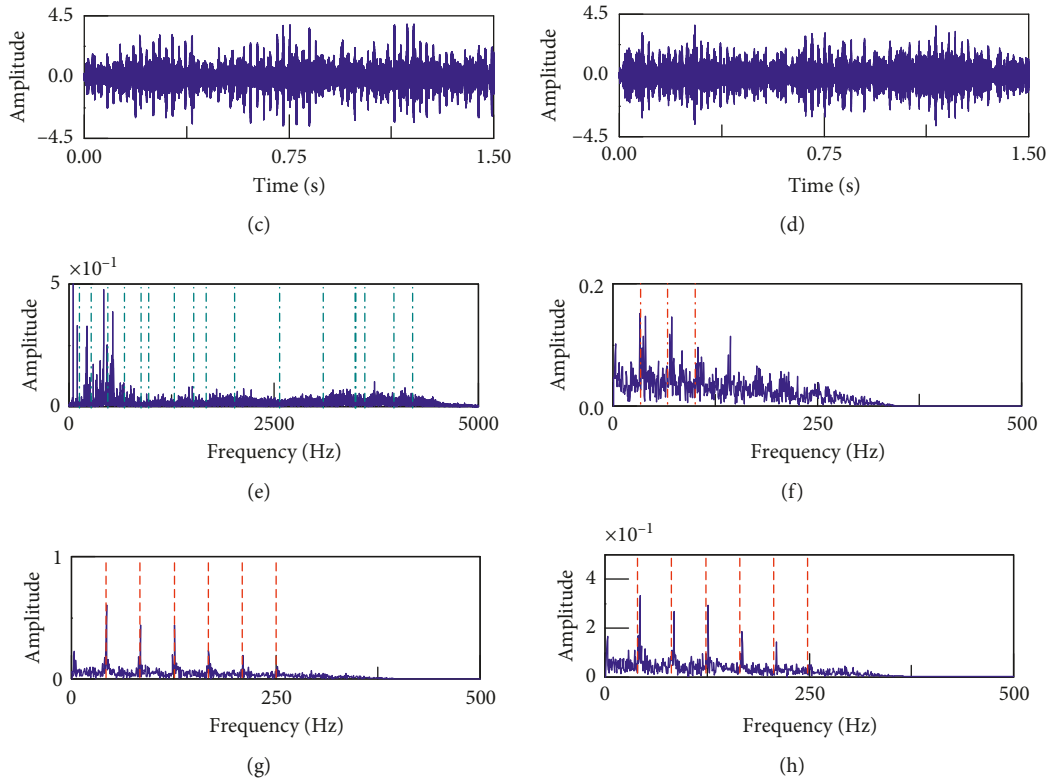
(a)

(b)

FIGURE 26: Continued.

FIGURE 26: Results obtained by the traditional EWT: (a) SSP (green dashed line: the detected threshold by k-means clustering); (b–d) sub-band signals; (e) detected boundaries (green dash-dotted lines); (f–h) squared envelope spectra (red dashed lines: f_{BPFO} and its harmonics; red dash-dotted lines: f_{BSF} and its harmonics).
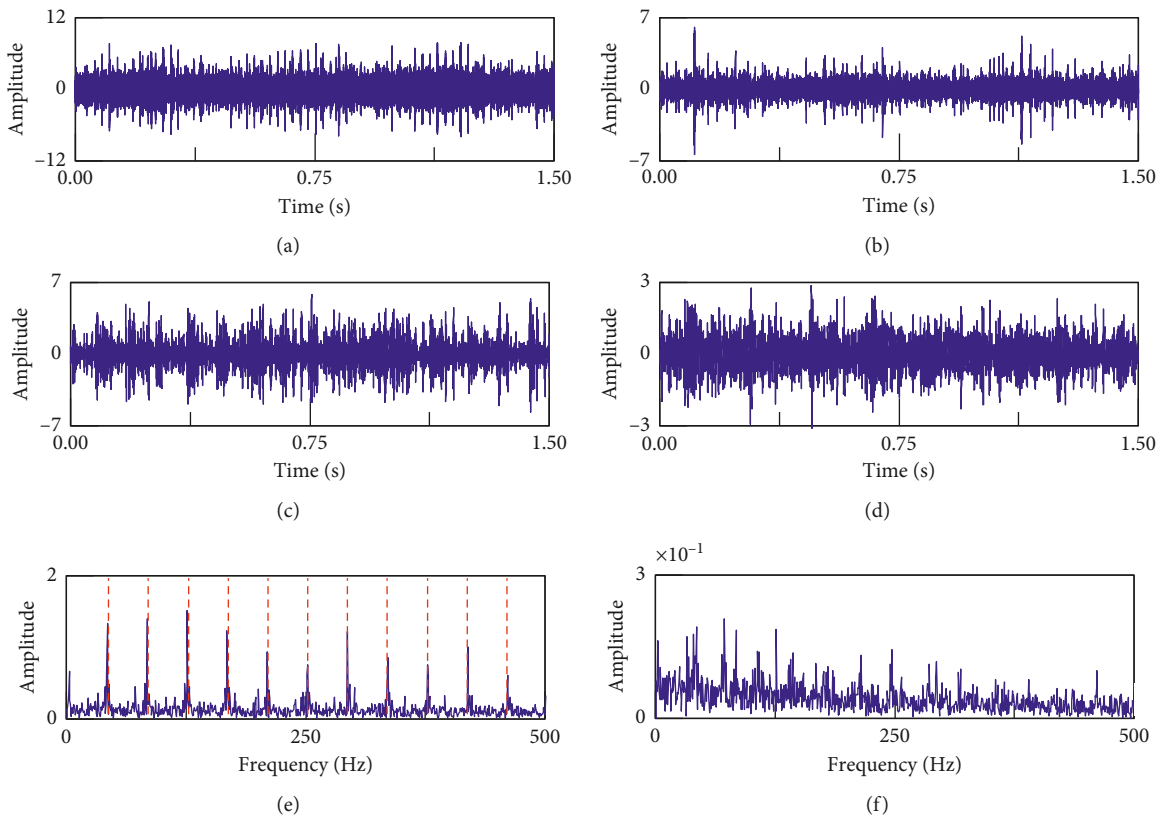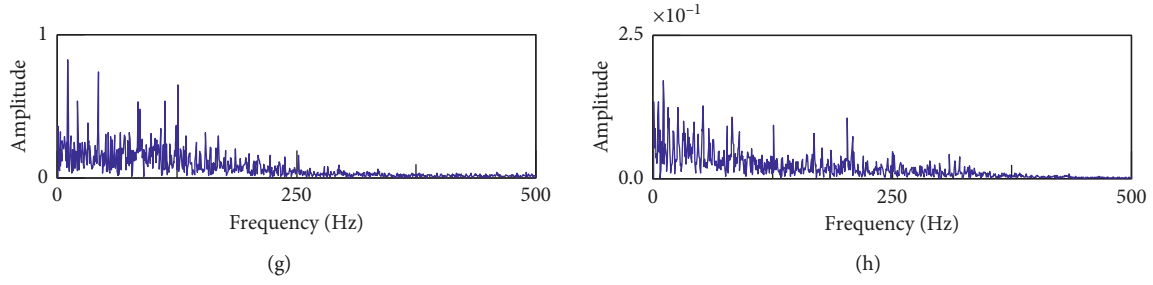

FIGURE 27: Continued.

(g)

(h)

FIGURE 27: Results obtained by EEMD: (a–d) IMFs; (e–h) squared envelope spectra (red dashed lines: f_{BPFO} and its harmonics).
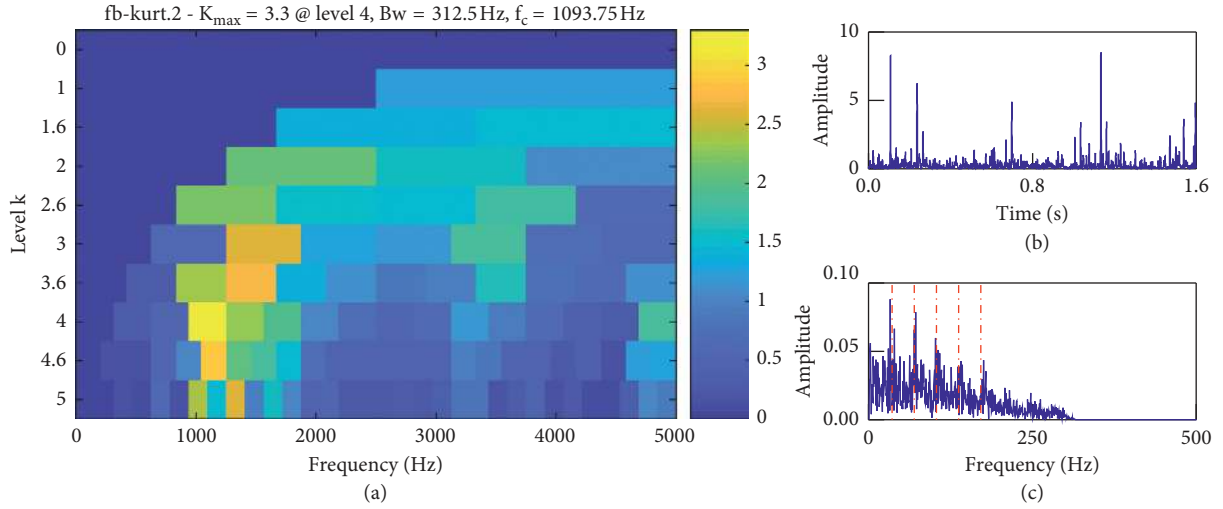

(a)

(b)

(c)

FIGURE 28: Results obtained by the fast kurtogram: (a) kurtogram; (b) squared envelope of the node with the largest kurtosis value; (c) squared envelope spectrum of (b) (red dash-dotted lines: f_{BSF} and its harmonics).

characteristic fault frequencies and the numbers of their harmonics are less than those obtained by the EWTOSSR method. In addition, the results obtained by EEMD and the fast kurtogram are shown in Figures 27 and 28, respectively. Neither the EEMD technique nor the fast kurtogram can extract the fault characteristic frequencies f_{BSF} and f_{BPFO} simultaneously. Therefore, the OSSR can determine a set of effective boundaries for extracting useful fault information, and it possesses higher capability in extracting weak faults from measured signals.

7. Conclusions

In this paper, a novel fault diagnosis method named the OSSR-based EWT, namely, the EWTOSSR method, was proposed. In the EWTOSSR technique, an optimized SSR replaces the traditional SSR to accelerate the calculation speed and increase the accuracy of the EWT with regard to boundary detection. The bandwidth determined by using these accurate boundaries can contain complete fault information with much less noise.

The EWTOSSR method was verified by simulated and experimental signals. The validation results show that the EWTOSSR approach is fairly suitable for extracting weak fault features from signals with strong noise.

Compared with three well-known fault diagnosis methods (i.e., the traditional EWT, EEMD, and the fast kurtogram), the EWTOSSR exhibits a much better fault detection performance.

Finally, although the EWTOSSR can effectively determine the resonant bands containing fruitful fault information, substantial in-band noise still exists and submerges the fault features to a certain extent. Therefore, further research needs to be conducted to reduce or eliminate the in-band noise following the implementation of the EWTOSSR method.

Nomenclature

$\widehat{\phi}_n(\omega)$:   Empirical scale function
$\widehat{\psi}_n(\omega)$:   Empirical wavelet function
ω_n: n-th detected boundary
T_n: n-th transition phase
τ_n: Width of T_n; $\widehat{\psi}_n(\omega)$
γ: $\tau_n = \gamma\omega_n$, where $0 < \gamma < 1$;
$\eta(x)$: $\eta(x) = x^4(35 - 84x + 70x^2 - 20x^3)$
$x(n)$:   Discrete signal
$X(f)$:   Discrete Fourier transform (DFT) of signal $x(n)$
$L(f, s)$:   Discrete scale-space representation
$g(f, s)$:   Kernel function

s: Scale parameter
M: Width of the kernel function
N_{max}: Maximum number of scale step of scale-space representation
WN: Number of local minimum of $L(f, s)$ in N-th scale step
C_i: i-th scale-space curve
Li: Length of the i-th scale-space curve
T: Threshold of the scale-space curve
A_j: Amplitude of the j-th fault impulse
T_p: Time period corresponding to the fault characteristic frequency
f_R: Resonance frequency
β: Structural damping coefficient
J: Number of fault impulses
$u(t)$: A unit step function
$s(t)$: Simulated signal
WT: The minimum threshold number of the optimized scale-space representation
N_e: Maximum number of scale steps of the optimized scale-space representation
L_i^d: Length of the i-th scale-space curve in descending order
ΔL_i^d: Difference in L_i^d
f_r: Rotation frequency
d: Roller diameter
D: Pitch diameter
Z: Number of rollers
α: Contact angle
f_{BPFO}: Outer-race fault characteristic frequency
f_{BSF}: Rolling fault characteristic frequency.

Data Availability

The experimental data are based on the CRRC project and are not suitable for publicity.

Conflicts of Interest

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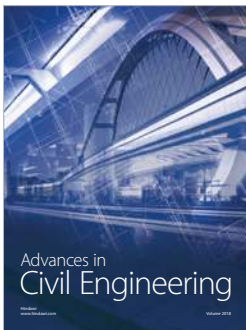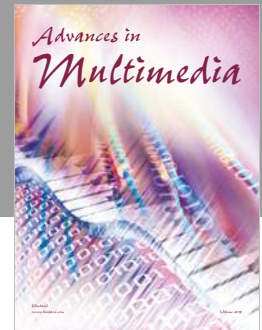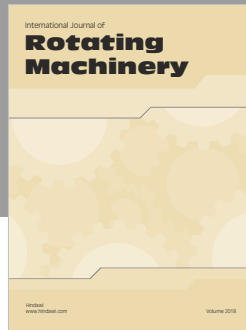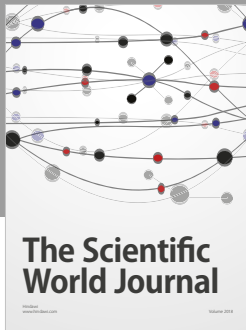
Acknowledgments

References

[1] Y. Zhang, C. Zhang, Y. Zhao, and S. Gao, "Wind speed prediction with RBF neural network based on PCA and ICA," *Journal of Electrical Engineering*, vol. 69, no. 2, pp. 148–155, 2018.

[2] H. Cao, L. Niu, S. Xi, and X. Chen, "Mechanical model development of rolling bearing-rotor systems: a review," *Mechanical Systems and Signal Processing*, vol. 102, pp. 37–58, 2018.

[3] A. Khadersab and S. Shivakumar, "Vibration analysis techniques for rotating machinery and its effect on bearing faults," *Procedia Manufacturing*, vol. 20, pp. 247–252, 2018.

[4] G. Adam, W. Glowacz, and Z. Glowacz, "Early fault diagnosis of bearing and stator faults of the single-phase induction motor using acoustic signals," *Measurement*, vol. 113, pp. 1–9, 2018.

[5] G. Adam, "Fault diagnosis of single-phase induction motor based on acoustic signals," *Mechanical Systems and Signal Processing*, vol. 117, pp. 65–80, 2019.

[6] G. Adam, "Acoustic based fault diagnosis of three-phase induction motor," *Applied Acoustics*, vol. 137, pp. 82–89, 2018.

[7] L. Xu, W. Zhai, and Z. Chen, "On use of characteristic wavelengths of track irregularities to predict track portions with deteriorated wheel/rail forces," *Mechanical Systems and Signal Processing*, vol. 104, pp. 264–278, 2018.

[8] M. Liang and I. Soltani Bozchalooi, "An energy operator approach to joint application of amplitude and frequency-demodulations for bearing fault detection," *Mechanical Systems and Signal Processing*, vol. 24, no. 5, pp. 1473–1494, 2010.

[9] Y. Wang, J. Xiang, R. Markert, and M. Liang, "Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: a review with applications," *Mechanical Systems and Signal Processing*, vol. 66, no. 67, pp. 679–698, 2016.

[10] Y. Li, L. Wang, and J. Guan, "A spectrum detection approach for bearing fault signal based on spectral kurtosis," *Shock and Vibration*, vol. 2017, Article ID 6106103, 9 pages, 2017.

[11] J. Antoni, "Fast computation of the kurtogram for the detection of transient faults," *Mechanical Systems and Signal Processing*, vol. 21, no. 1, pp. 108–124, 2007.

[12] Y. Wang and M. Liang, "An adaptive SK technique and its application for fault detection of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 25, no. 5, pp. 1750–1764, 2011.

[13] D. Rudoy, P. Basu, and P. J. Wolfe, "Superposition frames for adaptive time-frequency analysis and fast reconstruction," *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2581–2596, 2010.

[14] J. Gilles, "Empirical wavelet transform," *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 3999–4010, 2013.

[15] K. C. D. Kompella, V. G. R. Mannam, and S. Rao, "DWT based bearing fault detection in induction motor using noise cancellation," *Journal of Electrical Systems and Information Technology*, vol. 3, no. 3, pp. 411–427, 2016.

[16] Y. Zhang, P. Wang, and P. Cheng, "Wind speed prediction with wavelet time series based on Lorenz disturbance," *Advances in Electrical and Computer Engineering*, vol. 17, no. 3, pp. 107–114, 2017.

[17] 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 A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[18] Y. Zhang, C. Zhang, and J. Sun, "Improved wind speed prediction using empirical mode decomposition," *Advances in Electrical and Computer Engineering*, vol. 18, no. 2, pp. 3–10, 2018.

[19] J. Hu and J. Wang, "Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression," *Energy*, vol. 93, pp. 1456–1466, 2015.

[20] C. Aneesh, S. Kumar, P. M. Hisham, and K. P. Soman, "Performance comparison of variational mode

decomposition over empirical wavelet transform for the classification of power quality disturbances using support vector machine," *Procedia Computer Science*, vol. 46, pp. 372–380, 2015.

[21] J. Chen, J. Pan, Z. Li, Y. Zi, and X. Chen, "Generator bearing fault diagnosis for wind turbine via empirical wavelet transform using measured vibration signals," *Renewable Energy*, vol. 89, pp. 80–92, 2016.

[22] H. Cao, F. Fan, K. Zhou, and Z. He, "Wheel-bearing fault diagnosis of trains using empirical wavelet transform," *Measurement*, vol. 82, pp. 439–449, 2016.

[23] M. Kedadouche, M. Thomas, and A. Tahan, "A comparative study between empirical wavelet transforms and empirical mode decomposition Methods : application to bearing defect diagnosis," *Mechanical Systems and Signal Processing*, vol. 81, pp. 88–107, 2016.

[24] M. Kedadouche, Z. Liu, and V. Vu, "A new approach based on OMA-empirical wavelet transforms for bearing fault diagnosis," *Measurement*, vol. 90, pp. 292–308, 2016.

[25] Y. Hu, X. Tu, F. Li, H. Li, and G. Meng, "An adaptive and tacholess order analysis method based on enhanced empirical wavelet transform for fault detection of bearings with varying speeds," *Journal of Sound and Vibration*, vol. 409, pp. 241–255, 2017.

[26] D. Wang, Y. Zhao, C. Yi, K. Tsui, and J. Lin, "Sparsity guided empirical wavelet transform for fault diagnosis of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 101, pp. 292–308, 2018.

[27] J. Pan, J. Chen, Y. Zi, Y. Li, and Z. He, "Mono-component feature extraction for mechanical fault diagnosis using modi fi ed empirical wavelet transform via data-driven adaptive Fourier spectrum segment," *Mechanical Systems and Signal Processing*, vol. 72-73, pp. 160–183, 2016.

[28] J. Gilles and K. Heal, "A parameterless scale-space approach to find meaningful modes in histograms—application to image and spectrum segmentation," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 12, no. 6, pp. 1–17, 2014.

[29] Y. Song, S. Zeng, J. Ma, and J. Guo, "a fault diagnosis method for roller bearing based on empirical wavelet transform decomposition with adaptive empirical mode segmentation," *Measurement*, vol. 117, pp. 266–276, 2017.

[30] J. Zheng, H. Pan, S. Yang, and J. Cheng, "Adaptive parameterless empirical wavelet transform based time- frequency analysis method and its application to rotor rubbing fault diagnosis," *Signal Processing*, vol. 130, pp. 305–314, 2017.

[31] J. Babaud, A. P. Witkin, and M. Baudin, "Uniqueness of the Gaussian kernel for scale-space filterng," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 1, pp. 26–33, 1986.

[32] T. Lindeberg, "Scale-space for discrete signals," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 3, 1990.

[33] J. Weickert, "Linear scale-space has first been proposed in Japan," *Journal of Mathematical Imaging and Vision*, vol. 252, pp. 237–252, 1999.

[34] W. L. Zhao, C. H. Deng, and C. W. Ngo, "*k*-means: a revisit," *Neurocomputing*, vol. 291, pp. 195–206, 2018.

[35] P. Smith, D. B. Reid, C. Environment, L. Palo, P. Alto, and P. L. Smith, "A Threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[36] C. Sha, J. Hou, and H. Cui, "A robust 2D Otsu's thresholding method in image segmentation," *Journal of Visual Communication and Image Representation*, vol. 41, pp. 339–351, 2016.

[37] C. Mishra, A. K. Samantaray, and G. Chakraborty, "Ball bearing defect models : a study of simulated and experimental fault signatures," *Journal of Sound and Vibration*, vol. 400, pp. 86–112, 2017.

[38] 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, 2009.

[39] J. Ding, "Fault detection of a wheelset bearing in a high-speed train using the shock-response convolutional sparse-coding technique," *Measurement*, vol. 117, pp. 108–124, 2017.