Weak transient fault feature extraction based on an optimized Morlet wavelet and kurtosis

2016 Meas. Sci. Technol. 27 085003

(http://iopscience.iop.org/0957-0233/27/8/085003)

IOP Publishing Measurement Science and Technology
Meas. Sci. Technol. **27** (2016) 085003 (16pp) doi:10.1088/0957-0233/27/8/085003

Weak transient fault feature extraction based on an optimized Morlet wavelet and kurtosis

Yi Qin[1,2], **Jianfeng Xing**[1,2] **and Yongfang Mao**[3]

[1] School of Automotive Engineering, Chongqing University, Chongqing 400044, People's Republic of China
[2] State Key Laboratory of Mechanical Transmission, Chongqing University, Chongqing 400044, People's Republic of China
[3] College of Automation, Chongqing University, Chongqing 400044, People's Republic of China

E-mail: qy_808@aliyun.com (Yi Qin)

Received 29 January 2016, revised 22 May 2016
Accepted for publication 2 June 2016
Published 24 June 2016

Abstract

Aimed at solving the key problem in weak transient detection, the present study proposes a new transient feature extraction approach using the optimized Morlet wavelet transform, kurtosis index and soft-thresholding. Firstly, a fast optimization algorithm based on the Shannon entropy is developed to obtain the optimized Morlet wavelet parameter. Compared to the existing Morlet wavelet parameter optimization algorithm, this algorithm has lower computation complexity. After performing the optimized Morlet wavelet transform on the analyzed signal, the kurtosis index is used to select the characteristic scales and obtain the corresponding wavelet coefficients. From the time-frequency distribution of the periodic impulsive signal, it is found that the transient signal can be reconstructed by the wavelet coefficients at several characteristic scales, rather than the wavelet coefficients at just one characteristic scale, so as to improve the accuracy of transient detection. Due to the noise influence on the characteristic wavelet coefficients, the adaptive soft-thresholding method is applied to denoise these coefficients. With the denoised wavelet coefficients, the transient signal can be reconstructed. The proposed method was applied to the analysis of two simulated signals, and the diagnosis of a rolling bearing fault and a gearbox fault. The superiority of the method over the fast kurtogram method was verified by the results of simulation analysis and real experiments. It is concluded that the proposed method is extremely suitable for extracting the periodic impulsive feature from strong background noise.

Keywords: wavelet parameter optimization, Shannon entropy, kurtosis, periodic impulse, mechanical fault diagnosis

(Some figures may appear in colour only in the online journal)

1. Introduction

Rotating machinery is an important type of mechanical equipment. Due to the harsh working conditions, the typical rotating components, such as gears, bearings, rotors and shafts, are often the subject of various faults. In particular, when a bearing fault or gear fault occurs, frequency resonances will be excited at a specific rate. By detecting these vibration transients, we can diagnose the faults of bearings and gears. However, the measured vibration signals usually contain multiple vibration components and various noises. In such cases, the transient vibration signatures caused by the faulty parts are not easilyrecognized in the time-domain waveforms of the raw vibration signals.

An early developed and widely used technique to tackle this issue is envelope spectrum analysis, but the key point is

the proper selection of the center frequency and the bandwidth of the filter, which are mostly based on the type of mechanical structure. Thus this method requires professional knowledge and experience of the user. To extract the transient feature, we can also use non-stationary signal processing methods such as wavelet transform and empirical mode decomposition (EMD). The wavelet transform is well-known for its multi-resolution analysis capability and its good property of time-frequency localization, which are both extremely suitable for detecting transients. Various types of wavelet transform [1–5], such as the continuous wavelet transform, the discrete wavelet transform (DWT), the wavelet packet transform (WPT), the double density DWT, the higher density DWT, and the dense framelet transform, can be used to process the faulty vibration signals of rotating components. However, how to choose the proper wavelet basis and transform scale is a bigger challenge, and the adaptability is not satisfactory. As a data-driven method, EMD can decompose an arbitrary signal into the sums of intrinsic mode function [6], which has good local adaptability. Unfortunately, EMD has such disadvantages as the end effect and mode mixing, which can lead to serious distortion of the decomposition result [7], and the decomposition performance is easily influenced by strong noise. After signal decomposition or filtering, we need employ a proper measurement index for selecting the components that have transient features from the decomposition results or filtering results.

Kurtosis is one of the most important means of representing non-Gaussian transient features, which is useful for detecting fault-induced peaks (or transients) in vibration signals. The method of spectral kurtosis (SK) was first proposed by Dwyer [8]. Then Antoni gave the theoretical formulation of the SK for characterizing non-stationary signals [9] and proposed a method called fast kurtogram (FK) [10]. The extension and application of the SK method has been intensely researched in recent years. For example, Wang and Ming [11] proposed an adaptive SK technique based on window superposition and got a favorable result in the application for fault detection of rolling element bearings. Lei and Lin *et al* [12] adopted the WPT for the kurtogram. Wang and Tse *et al* [13] proposed an enhanced kurtogram, with kurtosis values calculated based on the power spectrum of the envelope of the signals extracted from wavelet packet nodes at different depths. Barszcz and Jabłoński [14] proposed a technique called the protrugram, which is based on the kurtosis of the envelope spectrum amplitudes of the demodulated signal instead of the kurtosis of the filtered time signal. Lee and Seo [15] applied spectral kurtosis to detect tip vortex cavitation in propellers. Luo *et al* [16] used tunable-Q wavelet transform and kurtosis to obtain the optimal filter for the Hilbert demodulation analysis. Wang and Lee [17] proposed an energy kurtosis demodulation technique for bearing fault detection. The basic idea of all the above methods is to use kurtosis to obtaini an optimum filter, and then the transient signal is detected by the optimum filter. Moreover, it is worth noting that the kurtosis of transients is related to the impulsive period, signal length and decay rate; thus its value is not fixed, but it is still much larger than the kurtosis of other typical signals, such as a sinusoidal signal and amplitude modulation–frequency modulation (AM–FM) signal. Thus, the kurtosis index can be used to choose the components that have transient features from signal transform results.

With the kurtosis, Lin proposed a method for optimizing Morlet wavelet parameters [18], but this method only determines the optimal bandwidth parameter and scale. As the kurtosis index lacks a meaningful benchmark, Bozchalooi and Liang proposed a new measurement index called the 'smoothed index' to quantify the impulsiveness, and used it to find the optimized wavelet shape parameter and scale [19]. The above methods did not consider the influence of wavelet central frequency on the result of wavelet transform, and the extracted impulsive signal is still disturbed by noise. In their simulation, there are no other signal components except for the impulses and noise. Under a strong noise background, both the kurtosis and smoothed indexes may be seriously disturbed by noise and other components, which may result in the optimal wavelet parameters not being correctly found. Moreover, the obtained optimal wavelet parameter is just suitable for a certain scale by using the kurtosis and smoothed indexes. In addition to these two indexes, the Shannon entropy is an effective approach to selecting the wavelet parameter [20]. Since the 'sparsity' is used as the rule for wavelet optimization, and all the wavelet coefficients will be used to compute the index, the wavelet corresponding to the fewest wavelet transformation coefficients of a signal is the best [18]. It follows that there is just one pair of optimal wavelet parameters for a given signal, which is very helpful for the inverse wavelet transform. Compared with the kurtosis and smoothed indexes, the Shannon entropy is less influenced by local noise and interference. On the whole, the Shannon entropy can better represent the similarity between the wavelet function and the analyzed signal. Therefore, the Shannon entropy is used to optimize the two Morlet wavelet parameters, i.e. bandwidth parameter and central frequency. To greatly increase the optimizing calculation speed, this paper firstly uses the Shannon entropy to calculate the optimum central frequency, and then find the bandwidth parameter with the central frequency obtained. With the Shannon entropy, we can obtain the wavelet that best represents the measured signal. However, the measured signal may have different vibration components, and the wavelet coefficients at different scales represent the corresponding components, and more importantly, we find that the extracted transient component with the maximum kurtosis does not match well with the original impulse through theoretical analysis and experiment. Therefore, kurtosis is applied to select the wavelet coefficients at some characteristic scales, which contain remarkable impulsive features, in order to accurately reconstruct the original periodical impulsive signal. As the chosen transient components are easily affected by noise, a soft-thresholding method is proposed to enhance the impulsive features and the threshold is adaptively determined by the characteristic wavelet coefficients. Finally, these denoised wavelet coefficients are used to reconstruct the transient signal. This proposed method combines the advantages of the optimized Morlet wavelet, kurtosis and adaptive

Table 1. The kurtosis values of four typical signals.

Signal	Impulsive signal	White noise	Sinusoidal signal	AM–FM signal
Kurtosis	99.14	3.00	1.50	1.75

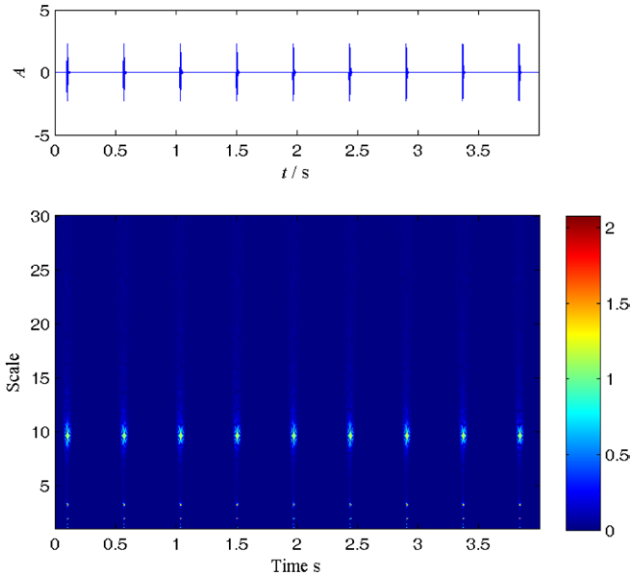

Figure 1. The waveform of the periodic impulsive signal and its wavelet scalogram.

soft-thresholding, and thus it can be applied to extract transient signatures (especially under strong background noise) and diagnose the weak faults of bearings and gears effectively. Compared with the current approaches, the transient signal obtained by the proposed method is closer to the original transient signal, and the weak transient feature can be more accurately extracted.

2. Optimized Morlet wavelet transform

Wavelet transforms are the inner products of the signal and a family of the wavelets. By dilation and translation from the mother wavelet $\psi(t)$, the wavelet transform of an arbitrary signal $x(t) \in L^2(\mathbf{R})$ is defined as [1]

$$W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\psi^*\left(\frac{t-b}{a}\right)\mathrm{d}t \tag{1}$$

where * denotes the complex conjugation, a is the scale parameter and b is the shifting parameter. $W(a,b)$ gives information on $x(t)$ at different levels of resolution and also measures the similarity between the signal $x(t)$ and the wavelet at each scale. Since the shape of the Morlet wavelet is similar to the impulsive signal, it has been widely applied to detect the faults of rotating components. The real Morlet wavelet can be defined as

$$\psi(t) = \frac{1}{\sqrt{\pi f_b}} \exp(-t^2/f_b)\cos(2\pi f_c t) \tag{2}$$

where f_b is the bandwidth parameter and f_c is the central wavelet frequency. These two parameters are closely

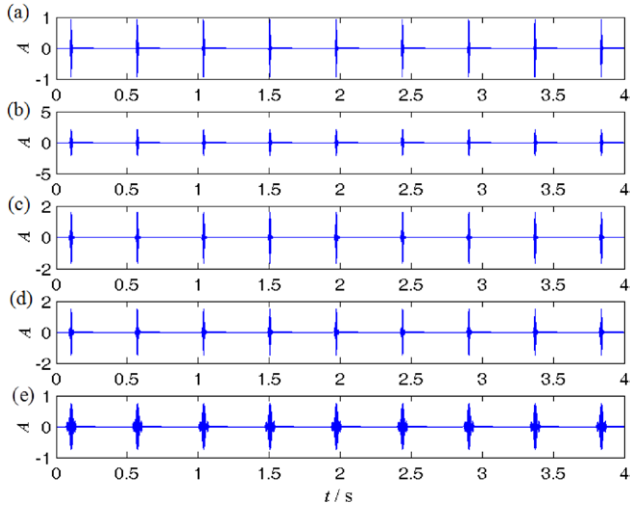

Figure 2. The waveforms of five characteristic components.

related to the shape and time-frequency resolution of the Morlet wavelet. The bandwidth parameter f_b controls the oscillation attenuation of the Morlet wavelet, while the central frequency f_c controls the oscillatory frequency of the Morlet wavelet. Increasing f_b will increase the oscillating property of the Morlet wavelet and improve its frequency resolution, while increasing f_c will decrease the frequency resolution of the Morlet wavelet. To detect the transient component in the measured signal, it is necessary to simultaneously optimize the parameters f_b and f_c of the Morlet wavelet.

As we know, 'sparsity' is usually used as the rule for wavelet optimization, and the diversity of a possibility series can be measured by wavelet energy entropy. The Shannon wavelet entropy will be used in this paper. To optimize both f_c and f_b of the Morlet wavelet, the Shannon entropy is calculated by [18]

$$E(f_c, f_b) = -\sum_{i=1}^{K} p_i \log p_i \sum_{i=1}^{K} p_i = 1 \tag{3}$$

where p_i is the distribution sequence obtained from wavelet coefficients, which is calculated by

$$p_i(f_c, f_b) = |W(a_i,t)| / \sum_{j=1}^{K} |W(a_j,t)|. \tag{4}$$

Obviously, for the given wavelet parameters and scales, the wavelet coefficients at all scales are used to calculate the entropy, while the wavelet coefficients at only one scale is used to calculate the kurtosis and smoothed indexes. The optimization algorithm proposed in [21] searches the optimal parameters in the setting range of (f_c, f_b), but if the searching numbers of f_c and f_b are, respectively, N and M, the number of entropies to be computed is up to $N \times M$; thus the calculation amount is quite large. To solve this problem, a fast optimization algorithm is proposed in this paper and its concrete steps are summarized as follows:

(1) Assume an initial range of central frequency $f_c \in [P, Q]$ and an initial step of central frequency τ_{f_c}. Then assume an initial value of f_b. As the impulsive fault vibration

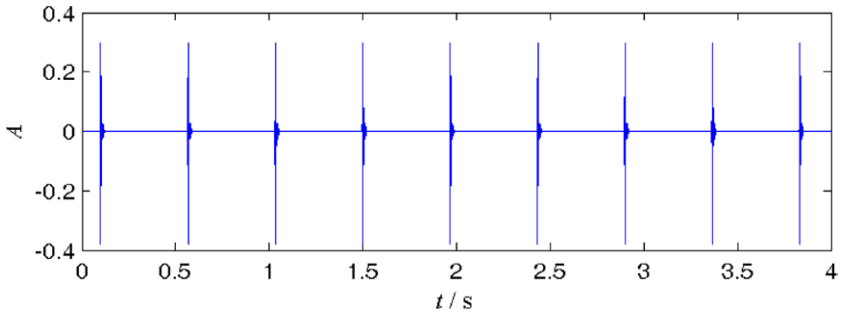

Figure 3. The waveform of the component with the largest kurtosis.

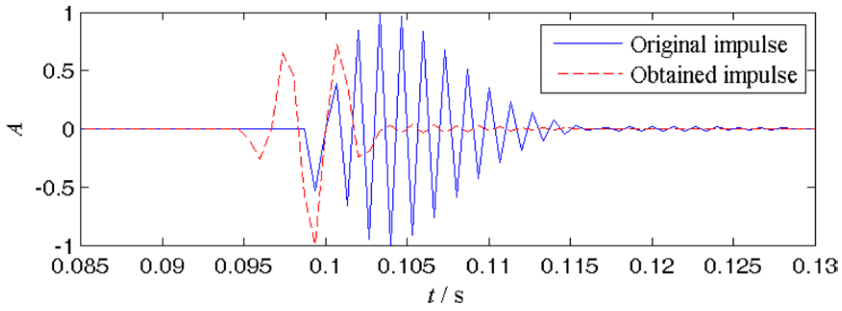

Figure 4. The comparison between one impulse of the component with the largest kurtosis and one impulse of the original signal.

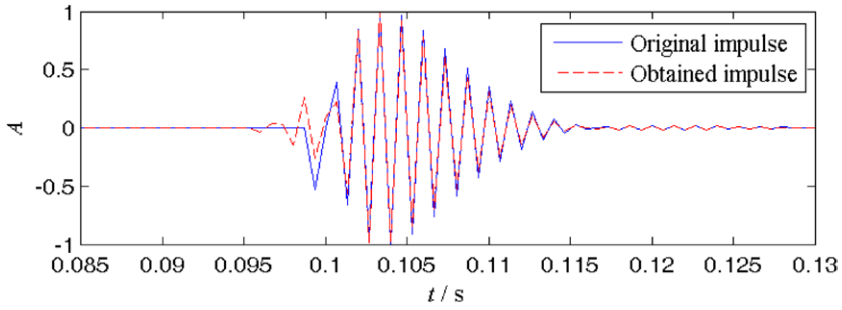

Figure 5. The comparison between the first impulse of the component at scale 1.1 and the first impulse of the original signal.

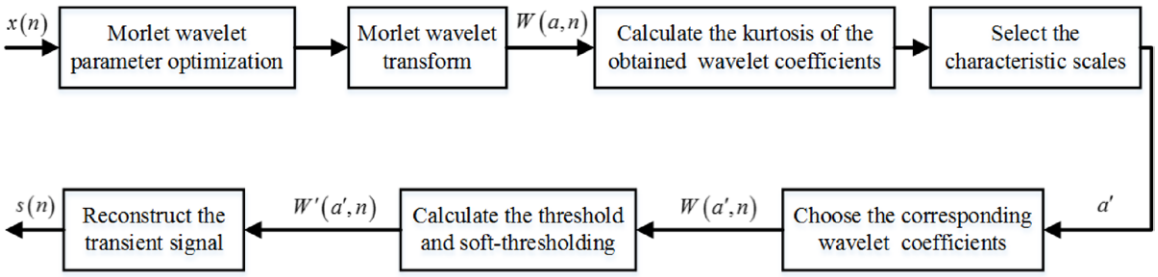

Figure 6. The flowchart of the proposed transient feature extraction algorithm.

signals usually have high oscillating property, the initial f_b can be set as a large value (we recommend 30).

(2) Increase f_c from P to Q with the step of τ_{f_c}, and calculate a series of wavelet entropies $E(f_c)$ using equations (1)–(4). Then the optimal central frequency f_{oc} can be found by the corresponding frequency of the minimum wavelet entropy $E_{min}(f_c)$.

(3) Set the central frequency as f_{oc}, and assume an initial range of bandwidth parameter $f_b \in [R, S]$. Increase f_b from R to S, and also calculate a series of wavelet entropies $E(f_b)$ with the given optimal central frequency. Similarly, the optimal bandwidth parameter f_{ob} can be found by the corresponding bandwidth parameter of the minimum wavelet entropy $E_{min}(f_b)$.

(4) The optimal Morlet wavelet parameter pair is set as (f_{oc}, f_{ob}).

In this algorithm, an initial range of central frequency can be set as (0.2 8), and the step of the central frequency can be set as 0.1; an initial range of bandwidth parameter can be set as (1 40), and the step of bandwidth parameter can be set as 0.2. The proposed algorithm first obtains the selected central

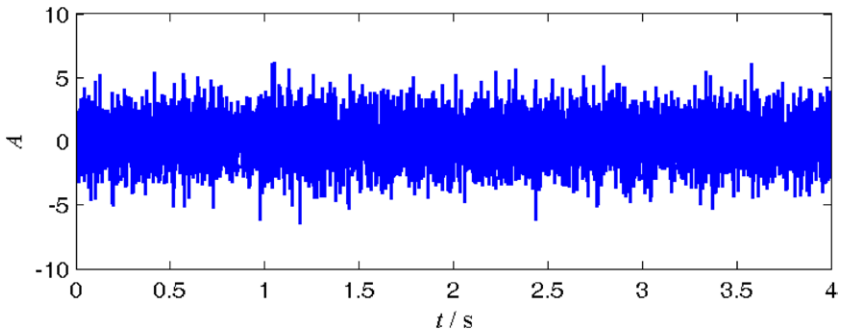

Figure 7. The time-domain waveform of the simulated signal.

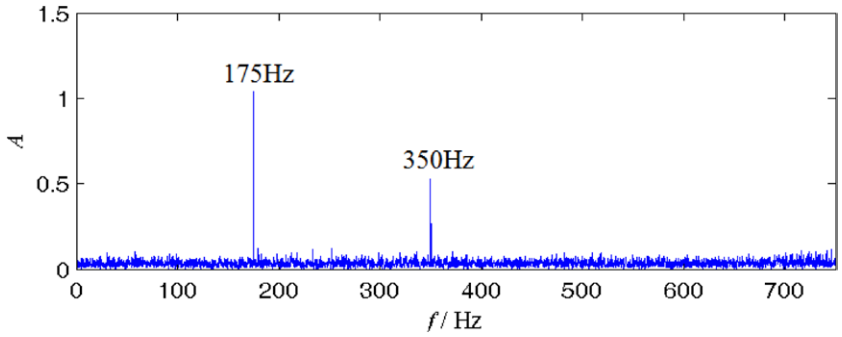

Figure 8. The Fourier frequency spectrum of the simulated signal.

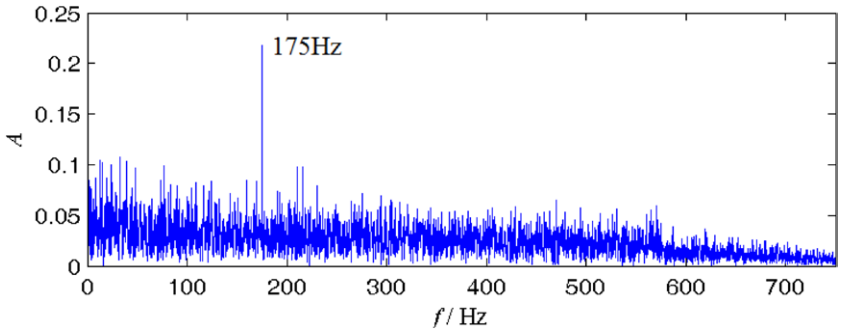

Figure 9. The envelope spectrum of the simulated signal without filtering.

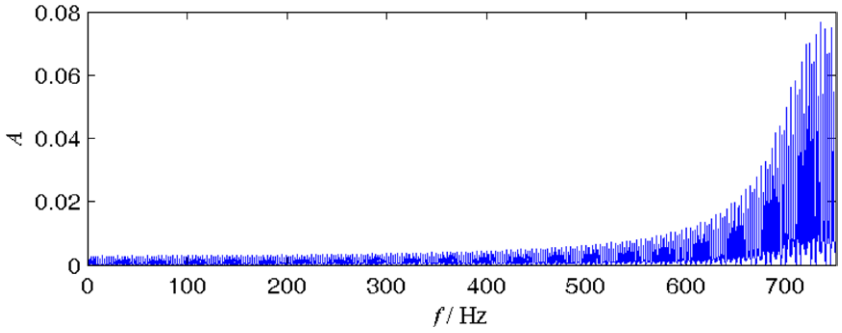

Figure 10. The Fourier frequency spectrum of the periodic impulsive signal.

frequency, and then finds the bandwidth parameter with the obtained f_{oc}. It is clear that the number of entropies to be computed by this proposed algorithm is just $N+M$, therefore the calculation efficiency is greatly enhanced. Moreover, by a great many of experiments, we find that most of the optimal wavelet parameters obtained by this proposed algorithm is the same as that obtained by the optimization algorithm in [21]. In fact, the obtained wavelet parameters are the result of global optimization.

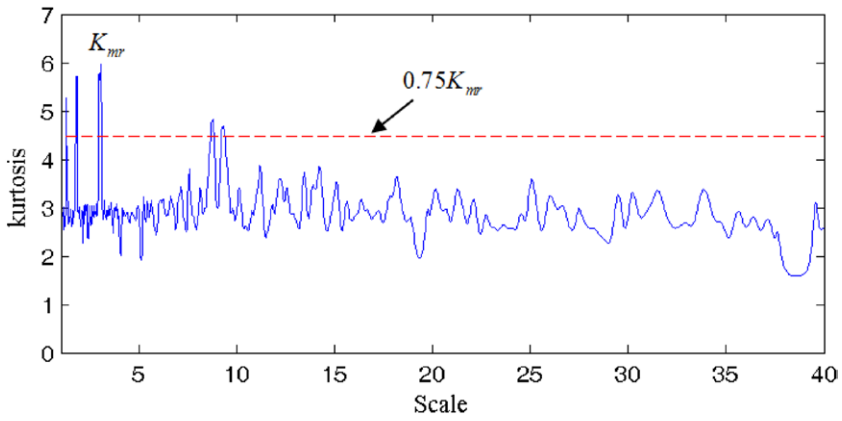

Figure 11. The kurtosis values of the obtained wavelet coefficients at all scales.

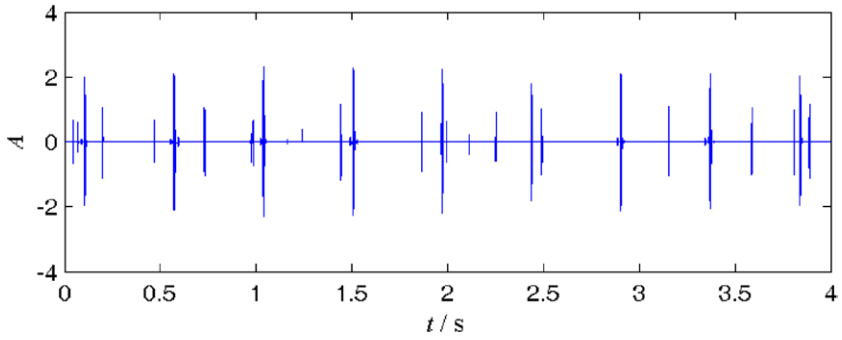

Figure 12. The periodic impulsive signal obtained by the proposed method.

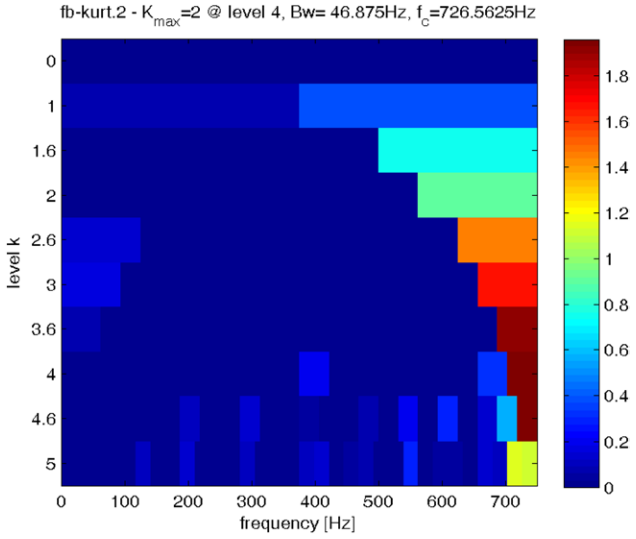

Figure 13. The kurtogram of the FK method.

3. The principle of transient detection with kurtosis

How to identify the impulsive components from all the signal components is an important issue for transient feature extraction. As a non-dimensional index, the kurtosis index is usually applied to quantitative measurement of impulsive components in vibration signals because it is sensitive to sharp variant structures, such as impulses. The kurtosis index can be calculated by [22]:

$$K_r = \frac{E(x-\bar{x})^4}{[E(x-\bar{x})^2]^2} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^4}{\left[\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2\right]^2} \tag{5}$$

where x is the signal sequence, N is the length of signal, and $\bar{x}$ is the mean value.

Let us first consider the simulated transient signal

$$h(t) = \sum_k A_k z(t-kT) \tag{6}$$

where A_k represents the amplitude of the kth impulse, T represents the impulsive period and

$$z(t) = \exp(-\sigma\omega_n t)\sin(\omega_n\sqrt{1-\sigma^2}\,t) \tag{7}$$

where ω_n is the natural frequency of the system, and σ is the damping parameter. The sampling frequency is 2000 Hz, and the length of signal is 5120. If A_k, T, ω_n and σ are respectively set as 4.10, 0.35, 4529.65 and 0.05, the kurtosis of the simulated transient signal can be calculated to be 99.14 using equation (5). Via a large number of experiments, we find that the kurtosis of the periodic impulsive signal will be increased with an increase in T. Similarly, the kurtosis of Gaussian white noise can be calculated as 3. Then we compute the kurtosis of other typical mechanical vibration signals, such as sinusoidal signal and amplitude modulation–frequency modulation (AM–FM) signal, which are respectively given by

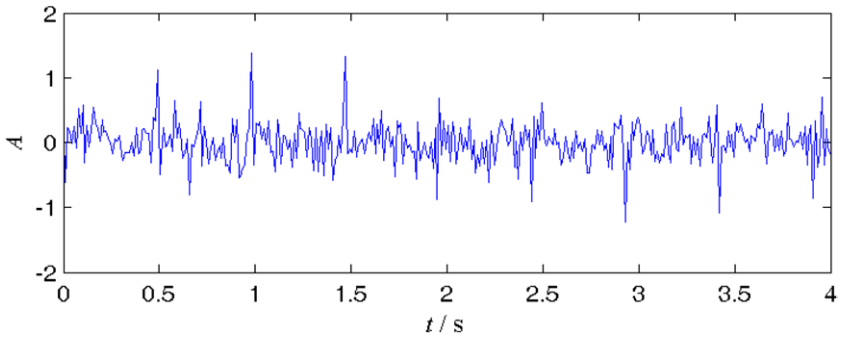

Figure 14. The periodic impulsive signal obtained by the FK method.

Table 2. The comparative results between the proposed method and the FK method.

The standard deviation of noise	The proposed method		The FK method	
	SNR	Kurtosis	SNR	Kurtosis
1.1	5.14	52.59	−3.04	6.94
1.3	4.89	50.61	−2.71	5.17
1.5	4.63	55.91	−2.67	5.68
1.7	4.23	54.26	−1.86	3.73
1.9	3.48	44.35	−2.80	4.34

$$s(t) = \cos(2\pi f t) \tag{8}$$

$$x(t) = [1 + 0.3\sin(2\pi \times 5t)] \cos[2\pi \times 400t + \sin(2\pi \times 10t)]. \tag{9}$$

For the sinusoidal signal, no matter what the frequency f is, its kurtosis equals 1.5. For the AM–FM signal, the kurtosis equals 1.75. If we change the carrier frequency and modulation frequency, the kurtosis remains almost the same. Table 1 lists the kurtosis values of four typical signals. It is easily noted from this table that the kurtosis value of the periodic impulsive signal is much larger than that of other typical signals [22, 23], therefore kurtosis can be quite effectively applied to recognizing the transients obtained by the wavelet filtering.

In fact, kurtosis represents the similarity between the signal and the Dirac function. The more quickly the impulse is attenuated, the larger the kurtosis is. For example, if we increase the damping parameter σ from 0.05 to 0.1, the kurtosis value will be increased from 99.14 to 196.76.

4. Transient detection using the optimized Morlet wavelet and kurtosis

4.1. The wavelet scalogram of the periodic impulsive signal

For the periodic impulsive signal, which is given by equation (6), we use the proposed fast algorithm to find the optimized Morlet wavelet parameters. The searching numbers of f_c and f_b are respectively set as 78 and 200, then the calculation time of the proposed optimization algorithm is 78.9s on the Matlab platform with Intel CPU i5-4570 and 8G RAM. However, if we use the traditional optimization algorithm, i.e. simultaneously optimize the combination of central frequency and bandwidth parameter, the calculation time is up to 4448.1s on the same platform. It can be seen that the computation speed of the traditional optimization algorithm is extremely slow, which does harm to engineering applications. After optimization, we get $f_{ob} = 10.6$ and $f_{oc} = 4.6$. With these parameters, the Morlet wavelet transform is performed. The waveform of this signal and its wavelet scalogram are illustrated in figure 1. From this figure, we can see that the energy of the periodic impulsive signal mainly distributes at some low scales. The wavelet coefficients at scales 1.1, 1.9, 3.1, 3.3 and 9.4, are respectively illustrated in figures 2(a)–(e). In these figures, we can see that all the five components have an obvious impulsive feature, but the decay of wavelet coefficients at scale 9.4 is the slowest. From the last section, we know that the component at one scale with the largest kurtosis has the shortest duration, and this duration is probably shorter than that of the original impulse. The wavelet coefficients at scale 9.4 represent the component with the largest kurtosis, whose waveform is illustrated in figure 3. It can be seen from this figure that this component has low energy. Via normalization, one impulse of the component with the largest kurtosis and one impulse of the original signal are simultaneously shown in figure 4. From this figure, it is easy to note that there is an obvious phase shift between the obtained impulse and the original impulse, and the obtained impulse has a faster attenuation. For comparison, figure 5 illustrates the first impulse of the component at scale 1.1 and the first impulse of the original signal. From figure 5, we can see that the obtained impulse matches the original impulse better. Therefore, it is not appropriate to only use the wavelet coefficients at one scale with maximum kurtosis for representing the original periodic impulsive signal. To reconstruct the real periodic impulsive component more accurately we should use other wavelet coefficients with large kurtosis (less than the maximum kurtosis) at a number of characteristic scales.

4.2. Weak transient detection algorithm

Since kurtosis is the sensitive index of transients, it can be effectively used to select the characteristic scales. If the wavelet coefficients at one scale have a large kurtosis value, this scale will be chosen as characteristic scale. The wavelet coefficients at the characteristic scale may be polluted by strong noise, thus it is necessary to denoise the wavelet coefficients by soft thresholding so as to enhance the transient feature. Compared with hard thresholding, soft thresholding can better retain data

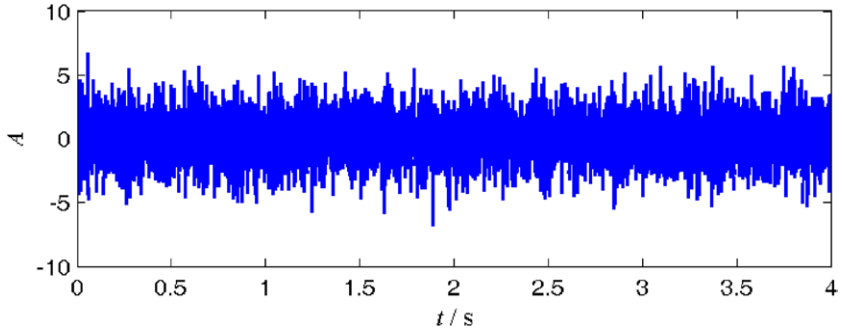

Figure 15. The time-domain waveform of the simulated signal with AM–FM components.

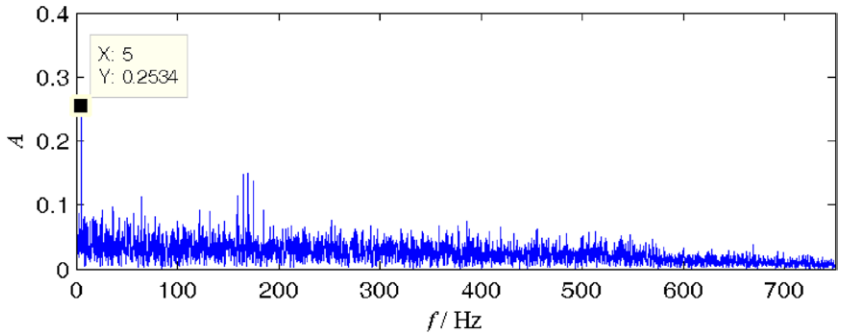

Figure 16. The envelope spectrum of the simulated signal with AM–FM components without filtering.

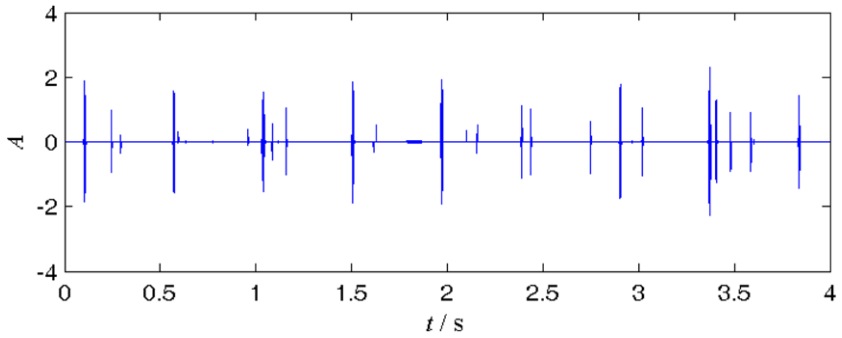

Figure 17. The periodic impulsive component extracted from a non-stationary AM–FM signal by the proposed method.

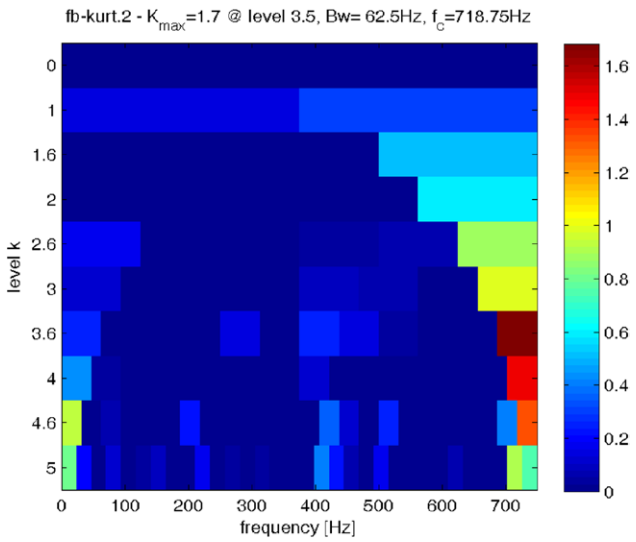

Figure 18. The kurtogram of the FK method for the non-stationary AM–FM signal.

during the impulse attenuation process. The soft-thresholding method can be expressed by the following formula [1]:

$$Y = \begin{cases} \text{sign}(X)(|X| - T) & |X| > T \\ 0 & |X| \leqslant T \end{cases} \tag{10}$$

where X denotes the wavelet coefficients and T denotes the threshold. Then the denoised wavelet coefficients will be used to reconstruct the transient signal $s(n)$ by

$$s(n) = \frac{1}{C_{1\psi}} \sum_{a} W'(a,n) a^{-3/2} \tag{11}$$

where $W'(a,n)$ denotes the selected and denoised wavelet coefficients, and $C_{1\psi}$ is given by

$$C_{1\psi} = \int_{-\infty}^{\infty} \widehat{\psi}^{*}(\omega)/|\omega| \mathrm{d}\omega. \tag{12}$$

Based on the above principle, we propose an algorithm for extracting the transient feature, and its concrete steps are summarized as follows:

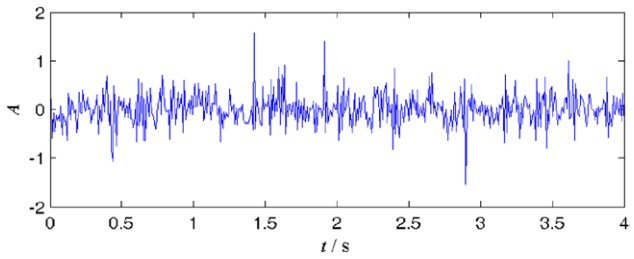

Figure 19. The periodic impulsive signal obtained by the FK method for the non-stationary AM–FM signal.

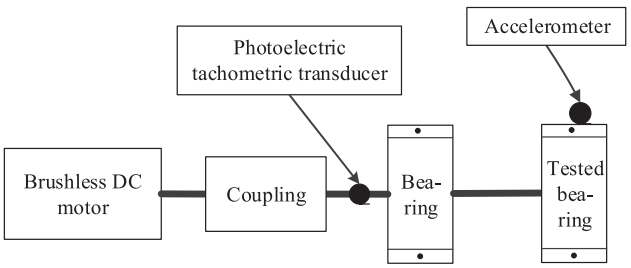

Figure 20. Schematic diagram of the test rig of a faulty rolling bearing.

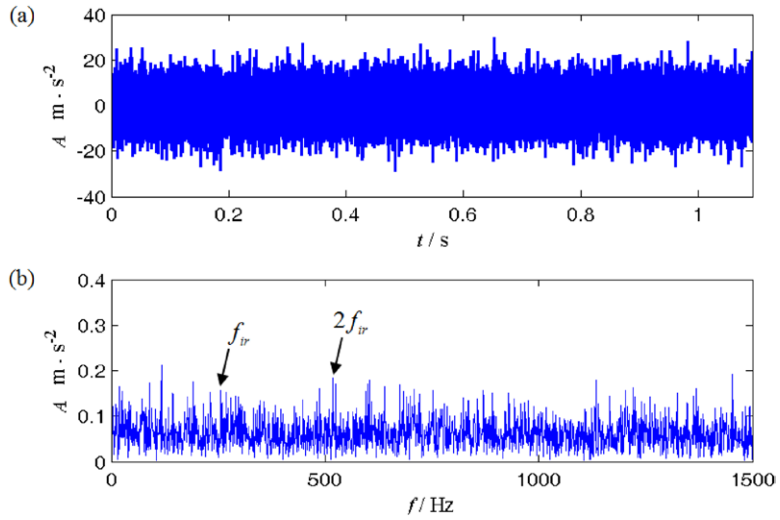

Figure 21. The waveform and its envelope spectrum of a faulty rolling bearing signal: (a) waveform and (b) envelope spectrum.

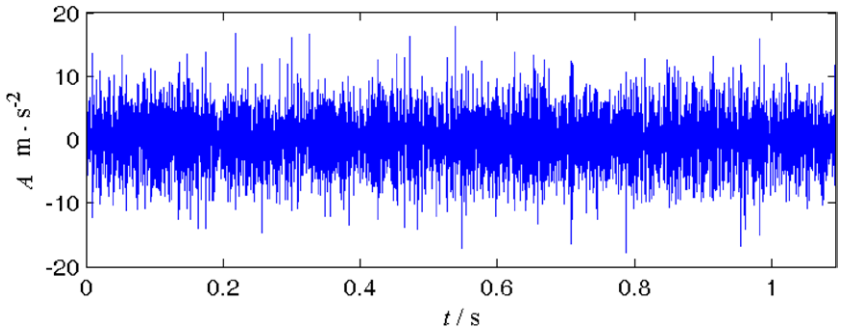

Figure 22. The transient component of a faulty rolling bearing vibration signal obtained by the proposed method.

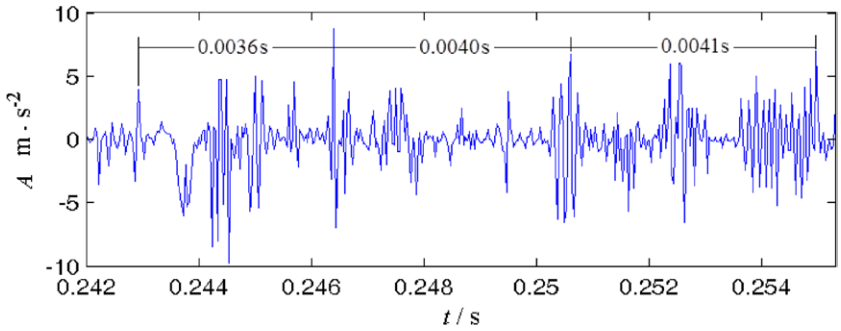

Figure 23. The zoom plot of the extracted transient component of a faulty rolling bearing vibration signal.

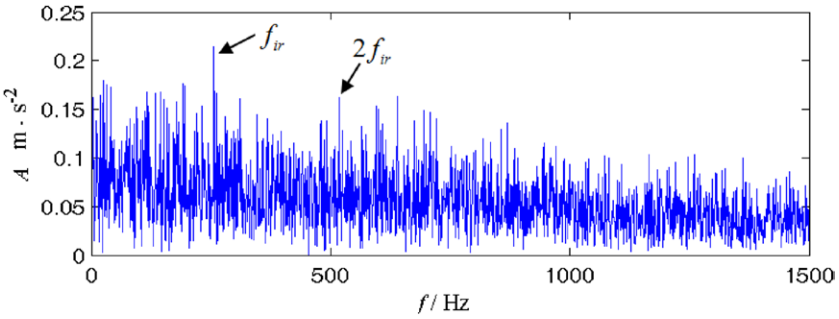

Figure 24. The envelope spectrum of the transient component of a faulty rolling bearing obtained by the proposed method.

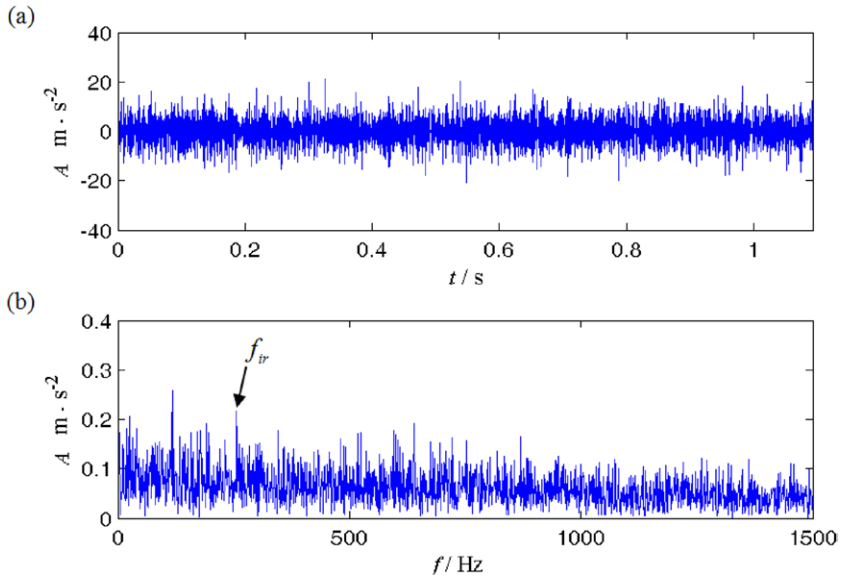

Figure 25. The result of the proposed method without wavelet optimization for a faulty rolling bearing vibration signal: (a) the transient and (b) its envelope spectrum.

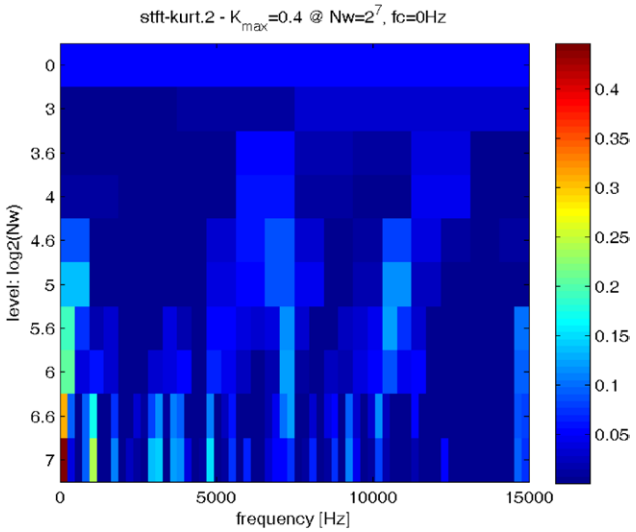

Figure 26. The kurtogram of a faulty rolling bearing vibration signal.

(1) For the analyzed signal $x(n)$, set the transform scales and perform the proposed optimization algorithm, then get the optimized Morlet wavelet parameters f_{oc} and f_{ob}.
(2) Perform the optimized Morlet wavelet transform using

$$W(a,n) = \frac{1}{\sqrt{a}} \sum_k x(k)\psi^*\left(\frac{k-n}{a}\right) \tag{13}$$

where $W(a,n)$ denotes the wavelet coefficient at the scale a and the shift n.
(3) Calculate the kurtosis values of the obtained wavelet coefficients at all scales using equation (5).
(4) Calculate the maximum kurtosis K_{mr}, and find the characteristic scales a' which satisfy

$$K_r(a') > 0.75K_{mr}. \tag{14}$$

(5) According to the characteristic scales, choose the corresponding wavelet coefficients $W(a',n)$.
(6) To retain more useful attenuation information, the threshold T is defined as

$$T = 2 \times \text{std}\,[W(a',n)] \tag{15}$$

where std denotes the standard deviation operator. With this threshold, the soft-thresholding method is performed on $W(a',n)$ and the denoised wavelet coefficient $W'(a',n)$ is obtained.
(7) Use denoised wavelet coefficient $W'(a',n)$ to reconstruct the transient signal $s(n)$ using equation (11).

The flow chart of the proposed algorithm is depicted in figure 6. Its computation complexity mainly depends on the

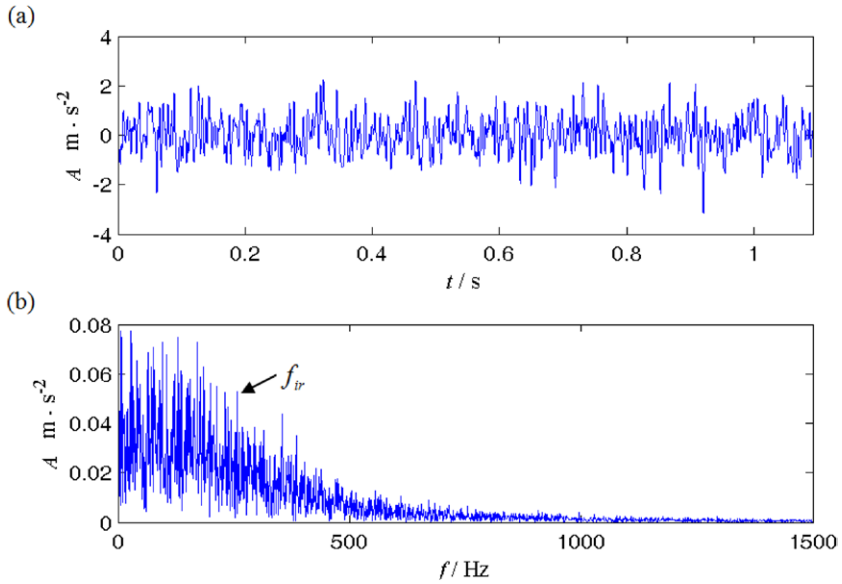

Figure 27. The result of the FK method for a faulty rolling bearing vibration signal: (a) the transient and (b) its envelope spectrum.

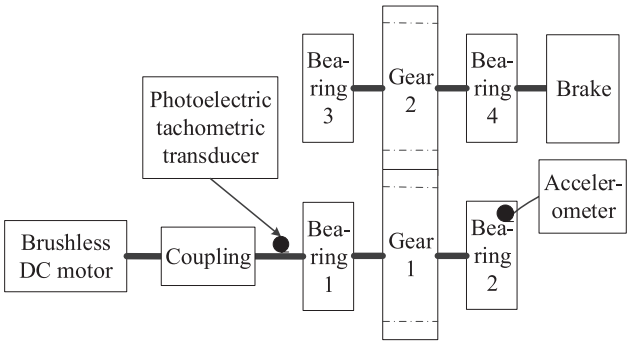

Figure 28. Schematic diagram of the test rig of a faulty single stage gearbox.

calculated amount of the Morlet wavelet parameter optimization process. Since the continuous wavelet transform is implemented by a convolution operation, its computation complexity is related to the signal length (P) and the number of scales (Q). It then follows that the computation complexity of the proposed method is $O((N+M)PQ)$. For comparison, the computation complexity of the Morlet wavelet parameter algorithm in [21] reaches up to $O(NMPQ)$.

5. Simulation and comparison

In this section, the proposed method for extracting transient features will be applied to analyze a noisy signal and compared with the FK method to illustrate its superiority. The program used of the FK method was downloaded from [24].

Let us consider such a noisy signal that

$$x(t) = \cos(2\pi \times 175t) + 0.5\cos(2\pi \times 350t) + h(t) + n(t) \tag{16}$$

where $n(t)$ is an additive white Gaussian noise (AWGN) with a standard deviation of 1.5, and $h(t)$ is a periodic impulsive signal given by equation (6) (A_k, T, ω_n and σ are respectively set as 8.36, 0.47, 4529.65 and 0.05). The sampling frequency is 1500 Hz, and the length of signal is 6000. By calculation, its signal-to-noise ratio (SNR) is −5.05 dB. The waveform of the simulated signal is illustrated in figure 7. Due to the strong noise, it is difficult to find the transient feature from the original time-domain waveform. We then calculate the Fourier frequency spectrum and Hilbert envelope spectrum of the simulated signal without filtering, which are respectively illustrated in figures 8 and 9. From figure 8, we can just see the spectral lines of the two harmonics, and it is very difficult to determine whether there is a periodic impulsive signal. The Fourier frequency spectrum of the periodic impulsive signal $h(t)$ is illustrated in figure 10. Obviously, the Fourier frequency spectrum of $h(t)$ has small amplitude, thus it is easily submerged in the Fourier frequency spectrum of a strong noise. Furthermore, the impulsive frequency (2.13 Hz) cannot be directly obtained from figure 10. From figure 9, we also cannot see a clear spectral line at the impulsive frequency due to the strong noise.

Consequently, the proposed method is applied to process this signal. The kurtosis values of the obtained wavelet coefficients at all scales are shown in figure 11. It can be seen from this figure that there are large kurtosis peaks at several scales. Using equation (14), we can obtain the characteristic scales. The corresponding wavelet coefficients are denoised by soft-thresholding, and then the denoised wavelet coefficients are used to reconstruct the transient signal, which is illustrated in figure 12. We can see from figure 12 that the periodic impulsive signal is effectively extracted.

The FK method based on the decimated filterbank tree is used for comparison. The kurtogram of the FK method is shown in figure 13. With the optimal carrier frequency and level, the transient signal is filtered out, which is shown in figure 14. Comparing figure 12 with figure 14, we can see that the periodic impulsive feature extracted by the proposed method is more obvious than that extracted by the FK method. From figure 14, it can be also seen that the periodic impulsive signal is disturbed by some noise. To further demonstrate the superiority of the proposed method, we use different levels of

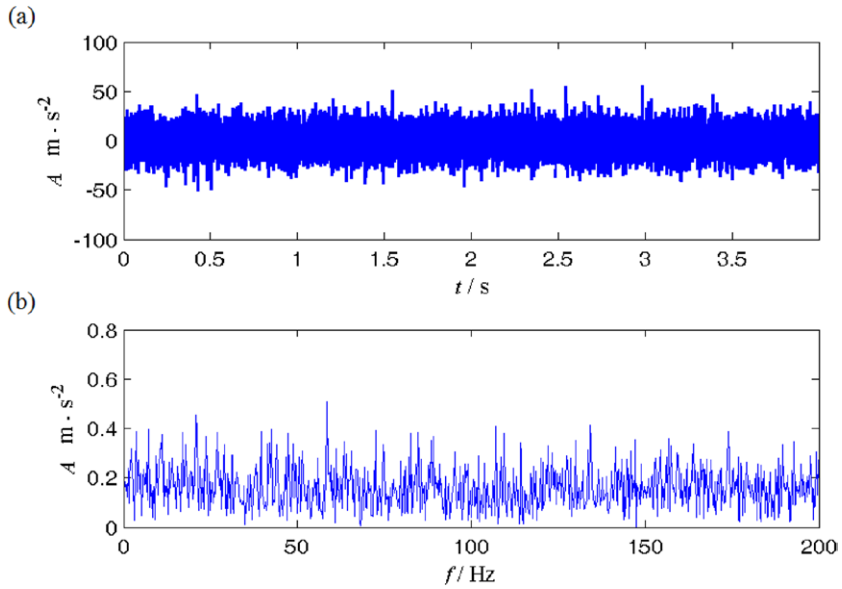

Figure 29. The waveform and its envelope spectrum of a faulty gearbox signal: (a) the waveform and (b) the envelope spectrum.

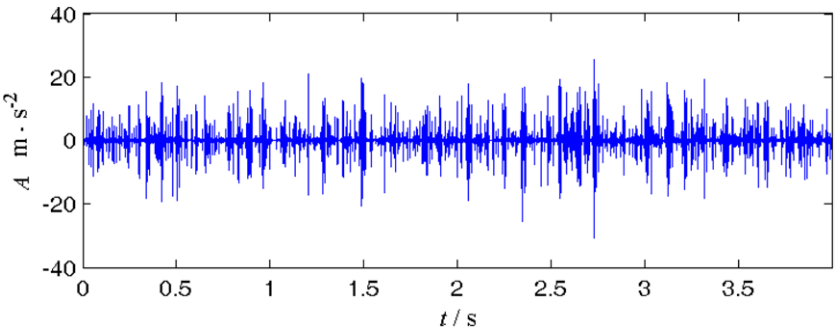

Figure 30. The transient component of a faulty gearbox vibration signal obtained using the proposed method.

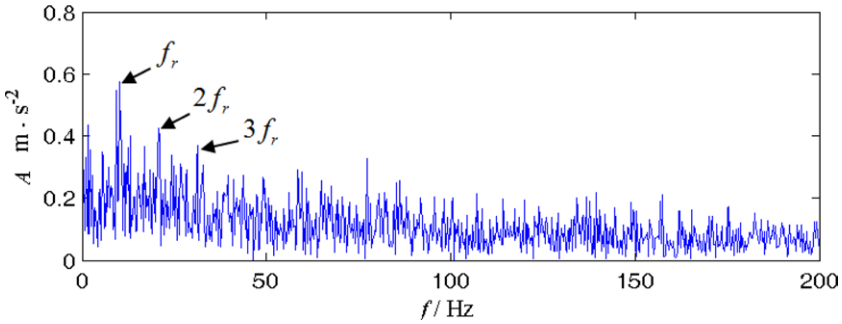

Figure 31. The envelope spectrum of the transient component of a faulty gearbox obtained using the proposed method.

AWGN for comparison. In equation (16), the standard deviations of different AWGNs are respectively set as 1.1, 1.3, 1.5, 1.7 and 1.9. The SNR and kurtosis of the extracted transient signals are used as the quantitative indexes, and then the comparative results between the proposed method and the FK method are obtained, which are listed in table 2. It can be seen from this table that the proposed method has better performance on transient feature extraction.

In order to validate the effectiveness of the proposed method on different signals, we consider the following non-stationary signal

$$\begin{aligned} x(t) = {} & [1 + 0.5\,\mathrm{in}(2\pi \times 5t)] \times \cos[2\pi \times 175t + \sin(2\pi \times 5t)] \\ & + [0.5 + 0.3\,\mathrm{in}(2\pi \times 5t)] \\ & \times \cos[2\pi \times 350t + 1.2\sin(2\pi \times 10t)] + h(t) + n(t) \end{aligned} \tag{17}$$

where $n(t)$ and $h(t)$ are the same as those in equation (16). The sampling frequency and the length of signal are also the same. By calculation, its SNR is −4.56 dB. The waveform of the simulated signal with amplitude modulation and frequency modulation (AM–FM) components and its envelope spectrum without filtering are respectively illustrated

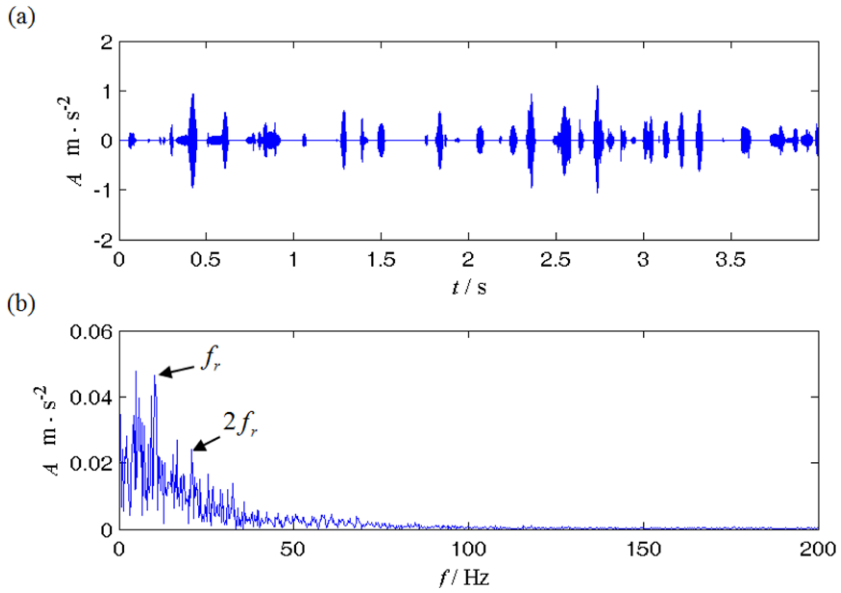

Figure 32. The result of the proposed method without wavelet optimization for a faulty gearbox vibration signal: (a) the transient and (b) its envelope spectrum.

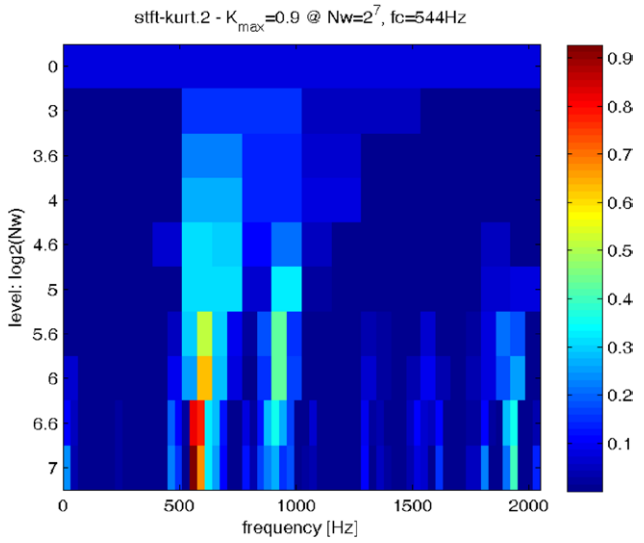

Figure 33. The kurtogram of a faulty gearbox vibration signal.

in figures 15 and 16. From these two figures, we cannot see the periodic impulsive signal and its impulsive frequency. Therefore, the proposed method is applied to extract the periodic impulsive signal $h(t)$. The obtained result is illustrated in figure 17. Obviously $h(t)$ is well extracted. This implies that the proposed method can also effectively extract the periodic impulsive signal from the non-stationary AM–FM signal.

Similarly, the FK method based on the decimated filterbank tree is used for comparison. The kurtogram of the FK method is shown in figure 18, and the transient signal extracted from the kurtogram is shown in figure 19. Comparing figure 19 with figure 17, we can see that the transients extracted by the proposed method are more obvious and have better periodicity. It can be also seen from figure 19 that the periodic impulsive signal is disturbed by some noise, and the impulses at some time instants are not clear.

6. Application to mechanical fault diagnosis

6.1. Application to rolling bearing fault diagnosis

Rolling bearings usually consist of an inner race, an outer race, several rollers and a cage. When the surface of one of these components develops a localized fault, a force impulse with the structure resonant frequency is produced in short time duration [25]. According to the frequency of the periodic impulsive vibration, we can identify the fault type. However, in some engineering applications, the impulsive vibration may be submerged in much background noise and vibration signals. Therefore, the proposed method is applied to extract the transient feature.

A simple rolling bearing test rig was set up, and a schematic diagram of the setup is shown in figure 20. The test rig mainly consists of a motor, a coupling, a bearing and a test faulty bearing. The motor is an AC servo motor (EMJ-04APB22), and its rated power and speed are respectively 0.4 kW and 3000 rpm. The two bearings are of the same type, 22NU15EC. The photoelectric tachometric transducer is used to measure the motor speed. The piezoelectric accelerometer, which is placed on the case of the faulty bearing, is used to acquire the vibration signal. In the experiment, the tested rolling bearing with an inner race defect is used, and the signal conditioning device is removed to increase the noise intensity. The sampling frequency is 30 000 Hz and the number of samples is 32 768. The rotating frequency of the motor is set as 25 Hz. According to the structure of the rolling bearing, the fault characteristic frequency f_{ir} was calculated as 257.6 Hz. The sampled vibration signal and its envelope spectrum without filtering are respectively illustrated in figures 21(a) and (b). From figure 21(a), the fault characteristic component with impulsive feature is embedded in strong background noise. The spectral peaks at the fault characteristic frequency and its multiple frequencies are vague in figure 21(b). In particular,

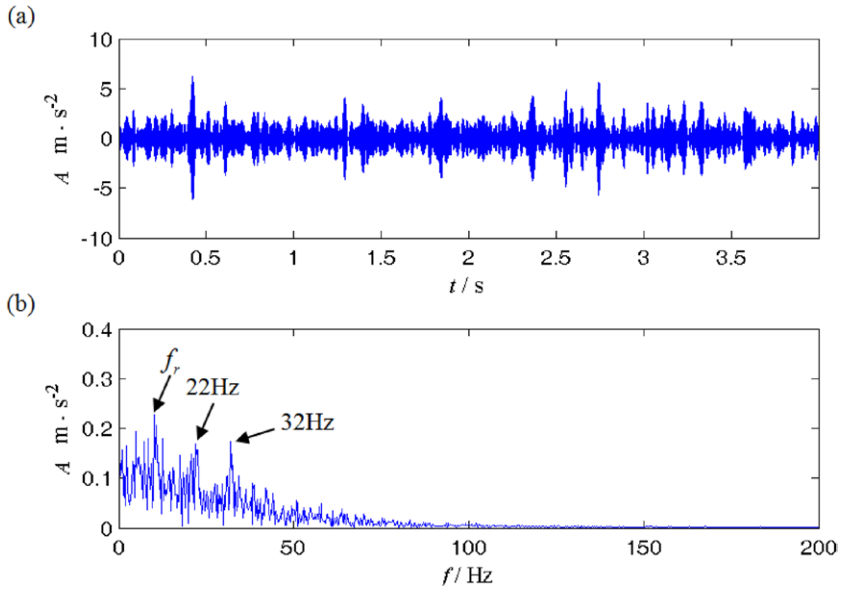

Figure 34. The result of the FK method for a faulty gearbox vibration signal: (a) the transient and (b) its envelope spectrum.

the spectral line at f_{ir} is not distinct. The proposed method is therefore applied to analyze this faulty vibration signal. Via the optimization algorithm of the wavelet parameter, the optimal bandwidth parameter f_{ob} and central frequency f_{oc} should be respectively set as 29.2 and 3.8. With the optimal wavelet parameter pair, the extracted transient signal is illustrated in figure 22. Since the mean impulsive frequency is high and the impulses caused by the fault are not obvious, the quasi-periodic transients [26] are very intensive and are difficult to directly observe from figure 22. Then we give the zoom plot of figure 22 in (0.2420 0.2553) s, which is shown in figure 23. It can be seen from figure 23 that there actually exists quasi-periodic impulses, and all the three time intervals between two adjacent impulses are close to the reciprocal of f_{ir}, i.e. 0.0039 s. To accurately obtain the mean impulsive frequency, we calculate the envelope spectrum of the extracted transient signal as shown in figure 24. From this figure, we can see that the spectral peak at f_{ir} is obvious, and there is also a spectral peak at $2f_{ir}$. Thus we can judge that the inner race of the tested rolling bearing has a local fault.

To demonstrate the important role of the wavelet parameter optimization algorithm, if we set the wavelet parameter as an arbitrary value, e.g. $f_b = 10$ and $f_c = 3$, the obtained transient signal and its envelope spectrum are respectively illustrated in figures 25(a) and (b). Comparing figure 24 with figure 25(b), we can see that two clear spectral peaks exist at f_{ir} and $2f_{ir}$ in figure 24, but there is only one spectral peak at f_{ir} in figure 25(b) and it is not the spectral line with the largest amplitude. As the sub-band signals at different levels obtained by the FK method based on the decimated filterbank tree have different sampling frequency, the FK method based on the short-time Fourier transform is applied to process the same faulty vibration signal, and its Matlab code also comes from [24]. This type of FK method can make the obtained transients have the same sampling frequency as the original signal, so it is beneficial to increase the analyzed frequency range of the envelope spectrum. After using the FK method based on the short-time Fourier transform, the obtained kurtogram is illustrated in figure 26, and then the extracted transient from the kurtogram and its envelope spectrum are respectively illustrated in figures 27(a) and (b). It is easy to note from figure 27(a) that the impulsive feature is obscure. And we can also see from figure 27(b) that the spectral peak at f_{ir} is not very clear. Therefore, the proposed method is more effective for extracting the weak transient fault feature than the FK method.

6.2. Application to gear fault diagnosis

When the gear suffers from an obvious fault, the vibration of the damaged gear is usually represented as impulses. Theoretically, periodic impulses, including the intensity and the time interval, are the robust symptom of tooth damage. However, the useful fault feature may be drowned by noise and normal gear vibration. In order to accurately extract the fault feature, the proposed method has to be used.

To perform faulty gear diagnosis experiments, a gearbox rig was established, and a schematic diagram of the setup is illustrated in figure 28. The test rig mainly includes a single stage gearbox, a brushless direct current (DC) motor for driving the gearboxes, a coupling and a magnetic powder brake for loading. The rated power and speed of the DC motor (HAE4895) were respectively 0.6 kW and 3400 rpm. In the experiment, a magnetic powder brake (TZ6K-3) was used to provide a load of 3 N·m. The accelerometer was placed on the gearbox shell, whose position corresponds to that of bearing 2. The gears 1 and 2 respectively have 34 and 65 teeth. Gear 1 of the input shaft had a broken tooth, and the signal conditioning device was also removed. Using the photoelectric tachometric transducer, the rotating speed of the motor was measured as 623 rpm, i.e. the rotating frequency of the input shaft (f_r) was calculated as 10.38 Hz. The vibration acceleration signal was sampled at a frequency of 4096 Hz, which is illustrated in figure 29(a). The number of samples was 16 384.

By calculation, the envelope spectrum without filtering is illustrated in figure 29(b). Since the impulsive component is submerged in strong noise, the periodic impulsive feature can hardly be seen in figure 29(a) and there are no clear spectral peaks at the fault characteristic frequency f_r and its multiple frequencies in figure 29(b). Consequently, we use the proposed method to extract the transient component. In this case, the optimal bandwidth parameter f_{ob} and central frequency f_{oc} are respectively 33.4 and 6.3, and the obtained impulsive component is illustrated in figure 30. In this figure, we can see the clear periodic impulsive feature. Then, by using the Hilbert transform and Fourier transform, its envelope spectrum is calculated, which is shown in figure 31. It can be seen that there are three clear spectral peaks at f_r, $2f_r$ and $3f_r$. This immediately indicates that the gear of the input shaft has a local fault.

To again validate the role of wavelet optimization, we set the bandwidth parameter and central frequency as two arbitrary values, i.e. $f_b = 6$ and $f_c = 4$, and then the obtained transient signal and its envelope spectrum are respectively illustrated in figures 32(a) and (b). It can be seen from figure 32(a) that a certain number of the impacts caused by the gear fault are not correctly filtered out. In figure 32(b), there are two spectral peaks at f_r and $2f_r$, but the spectral peak at f_r does not have the largest amplitude. Then, the FK method based on the short-time Fourier transform is also applied to process the same faulty vibration signal. The obtained kurtogram is illustrated in figure 33, and then the transient component is filtered out by the optimal parameter. Its waveform and envelope spectrum are respectively illustrated in figures 34(a) and (b). From figure 34(a) we can see that the periodicity of the obtained impulsive component is not very obvious. In figure 34(b) there is only a clear spectral peak at f_r. Moreover, the spectral amplitude at f_r in figure 34(b) is smaller than that in figure 31. The contrasting result of this experiment again proves that the proposed method is superior to the FK method for extracting the weak periodic impulsive feature.

7. Conclusions

This paper explores a weak transient feature extraction method with the optimized Morlet wavelet transform and the kurtosis index. To improve the accuracy of transient feature extraction, the Shannon entropy is used to find the optimal Morlet wavelet parameter. A fast optimization algorithm has been proposed to obtain the optimal central frequency and bandwidth parameter, which has much lower computation complexity than the existing Morlet wavelet parameter optimization algorithm. Via theoretical analysis and experiments, we have revealed that the transient signal should be reconstructed by the wavelet coefficients at several characteristic scales. Since the kurtosis index is sensitive to the impulsive feature, it can be used to select characteristic scales and obtain the corresponding wavelet coefficients. Moreover, as the wavelet coefficients at the characteristic scales may be seriously influenced by noise, the soft-thresholding method is applied to denoise the characteristic wavelet coefficients, by which the corresponding threshold is adaptively calculated. Finally, the transient signal can be reconstructed by the denoised wavelet coefficients. Compared with the FK method, it has been proved by simulation and application results that the proposed method is more suitable for extracting the transients from strong background noise. The necessity of wavelet optimization is also elaborated by the experiments. Therefore, the proposed method has great potential in the weak fault diagnosis of rotating components. Moreover, for some important rotatory components it is helpful to research the dynamic response caused by their faults and the relevant fault mechanism

Acknowledgments

The work described in this paper was supported by the Fundamental Research Funds for the Central Universities (No. CDJZR14285501), China Postdoctoral Science Foundation funded project (No. 2012M521690), National Natural Science Foundation of China (Nos. 50905191 and 51435001), and State Key Laboratory of Mechanical Transmission funded project (no. SKLMT-ZZKT-2015Z14). The authors also thank the anonymous reviewers for their valuable comments that helped improve this paper.

References

[1] Mallat S 2008 *A Wavelet Tour on Signal Processing* 3rd edn (New York: Academic)
[2] Selesnick I W 2001 The double density DWT *Wavelets in Signal and Image Analysis: From Theory to Practice* ed A Petrosian and F G Meyer (Boston: Kluwer)
[3] Yan R and Gao R X 2010 Harmonic wavelet-based data filtering for enhanced machine defect identification *J. Sound Vib.* **329** 3203–17
[4] Qin Y, Wang J X and Mao Y F 2013 Dense framelets with two generators and their application in mechanical fault diagnosis *Mech. Syst. Signal Process.* **40** 483–98
[5] Qin Y, Tang B P and Wang J X 2010 Higher-density dyadic wavelet transform and its application *Mech. Syst. Signal Process.* **24** 823–34
[6] Huang N E, Shen Z, Long S R, Wu M C, Shih H H, Zheng Q, Yen N C, Tung C C and Liu H H 1998 The empirical mode decomposition and Hilbert spectrum for nonlinear and nonstationary time series analysis *Proc. R. Soc.* A **454** 903–95
[7] Rato R T, Ortigueira M D and Batista A G 2008 On the HHT, its problems, and some solutions *Mech. Syst. Signal Process.* **22** 1374–94
[8] Dwyer R F 1983 Detection of non-Gaussian signals by frequency domain kurtosis estimation *Int. Conf. on Acoustic, Speech, and Signal Processing (Boston)* pp 607–10
[9] Antoni J 2006 The spectral kurtosis: a useful tool for characterising non-stationary signals *Mech. Syst. Signal Process.* **20** 282–307
[10] Antoni J 2007 Fast computation of the kurtogram for the detection of transient faults *Mech. Syst. Signal Process.* **21** 108–24
[11] Wang Y X and Ming L 2011 Anadaptive SK technique and its application for fault detection of rolling element bearings *Mech. Syst. Signal Process.* **25** 1750–64
[12] Lei Y G, Lin J, He Z J and Zi Y Y 2011 Application of an improved kurtogram method for fault diagnosis of rolling element bearings *Mech. Syst. Signal Process.* **25** 1738–49

[13] Wang D, Tse P W and Tsui K L 2013 An enhanced Kurtogram method for fault diagnosis of rolling element bearings *Mech. Syst. Signal Process.* **35** 176–99
[14] Barszcz T and Jabłoński A 2011 A novel method for the optimal band selection for vibration signal demodulation and comparison with the Kurtogram *Mech. Syst. Signal Process.* **25** 431–51
[15] Lee J H and Seo J S 2013 Application of spectral kurtosis to the detection of tip vortex cavitation noise in marine propeller *Mech. Syst. Signal Process.* **40** 222–36
[16] Luo J S, Yu D J and Liang M 2013 A kurtosis-guided adaptive demodulation technique for bearing fault detection based on tunable-Q wavelet transform *Meas. Sci. Technol.* **24** 055009
[17] Wang W and Lee H 2013 An energy kurtosis demodulation technique for signal denoising and bearing fault detection *Meas. Sci. Technol.* **24** 025601
[18] Lin J and Qu L S 2000 Feature extraction based on Morlet wavelet and its application for mechanical fault diagnosis *J. Sound Vib.* **234** 135–48
[19] Ozchalooi I S and Liang M 2007 A smoothness index-guided approach to wavelet parameter selection in signal de-noising and fault detection *J. Sound Vib.* **308** 246–67
[20] Nikolaou N G and Antoniadis I A 2002 Demodulation of vibration signals generated by defects in rolling element bearings using complex shifted Morlet wavelets *Mech. Syst. Signal Process.* **16** 677–94
[21] Jiang Y H, Tang B P, Qin Y and Liu W Y 2011 Feature extraction method of wind turbine based on adaptive Morlet wavelet and SVD *Renew. Energy* **36** 2146–53
[22] Qu L S and He Z J 1986 *Mechanical Diagnostics* (Shanghai: Science and Technology Press)
[23] Samanta B and Al-Balushi K R 2003 Artificial neural network based fault diagnostics of rolling element bearings using time-domain features *Mech. Syst. Signal Process.* **17** 317–28
[24] Antoni J webpage; www.utc.fr/_antoni/S
[25] McFadden P D and Smith J D 1983 Model for the vibration produced by a single point defect in a rolling element bearing *J. Sound Vib.* **96** 69–82
[26] Antoni J 2004 Cyclostationary modelling of rotating machine vibration signals *Mech. Syst. Signal Process.* **18** 1285–314