



Subband averaging kurtogram with dual-tree complex wavelet packet transform for rotating machinery fault diagnosis

Lei Wang ^a, Zhiwen Liu ^{b,*}, Hongrui Cao ^a, Xin Zhang ^c



^aSchool of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, PR China

^bSchool of Automation Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, PR China

^cSchool of Aeronautics and Astronautics, Sichuan University, Chengdu, Sichuan 610065, PR China

ARTICLE INFO

Article history:

Received 30 August 2019

Received in revised form 3 February 2020

Accepted 19 February 2020

Keywords:

Subband averaging kurtogram
Dual-tree complex wavelet packet
transform
Rotating machinery
Fault diagnosis

ABSTRACT

This paper presents a method called subband averaging kurtogram (SAK), incorporating with dual-tree complex wavelet packet transform (DTCWPT), to improve performance of the fast kurtogram (FK) for rotating machinery fault diagnosis. The proposed method first segments a signal into M sub-signals by a sliding window, then computes the kurtosis of subbands obtained by DTCWPT of each sub-signal. Finally, average kurtosis of corresponding subbands are calculated to obtain the SAK, which indicates the optimal frequency band for the envelope analysis. The FK is easily misled by non-Gaussian noise (e.g., sporadic impulse interferences) whereas the SAK can overcome this problem. Moreover, the DTCWPT simultaneously subdivides bands at high and low frequencies, offers the desirable property of approximate shift-invariance and meanwhile remains less computationally expensive. When the original DTCWPT iterates filter banks on the high-pass channel, the obtained subbands of a signal are not arranged in monotone order of the center frequency. This problem can be resolved by exchanging the inverted filter banks based on their band-pass properties. The proposed method provides improved performance compared to FK, in particular, for extracting periodic transients from noisy signals containing a variety of interferences. A simulation case and two applications to fault diagnosis of a planetary gearbox and a rolling bearing validate the effectiveness and improvements of the proposed method.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

Rotating machinery fault diagnosis, typically for gears, bearings and rotors, remains a challenge in practice. It demands that the fault detection and diagnosis methods take into accounts of robustness and accuracy to balance missing alarm and false alarm. Because vibration signals from rotating machinery can be conveniently collected during operation and reflect their running conditions in real time, vibration-based fault diagnosis methods have become mainstream in rotating machinery fault diagnosis field [1–3]. Vibration signals from faulty rotating machinery often contain a variety of components, including the fault signal, Gaussian noise, additional harmonics, and non-Gaussian noise. A challenge of vibration-based fault detection and diagnosis is to accurately extract the fault signal, which is commonly modeled as periodic transient signal, from the collected complex signal [4,5].

* Corresponding author.

E-mail address: lzw1682007@126.com (Z. Liu).

To date, numerous advanced signal processing approaches have been developed and introduced for rotating machinery fault diagnosis. The short-time Fourier transform (STFT) bridges the gap between stationary and non-stationary signal analyses and remains the foundation of many modern time-frequency fault diagnosis methods [6,16]. But when the window function is selected, STFT can only achieve fixed time-frequency resolution in throughout frequency region. The wavelet transforms (WTs) are an inclusive toolbox for processing non-stationary signals and commonly used for rotating machinery fault diagnosis [7–9]. The continuous WT (CWT) provides relatively high time-frequency resolution, however, the method is computationally expensive and therefore, unsuitable for online fault monitoring. The discrete WT (DWT) is implemented through recursive convolution between a signal and wavelet filter banks known as the Mallat pyramidal algorithm. Performance of the WT relies heavily on the similarity between the target signal and wavelet basis. Thus, constructing wavelet basis that is better matched with the fault signal is consistently the focus of researches on wavelet-based fault diagnosis techniques. Sparse decomposition (SP) as an improvement of WT pursues a sparse representation of the signal from a redundant time-frequency dictionary [10,11]. More specifically, SP seeks the solution of the optimization model that represents the signal as a linear combination of a few atoms from a proper overcomplete dictionary based on sparse prior knowledge. It is always a challenge for one who uses WT or SP to diagnose rotating machinery fault to choose a proper wavelet basis or dictionary. Based on “sifting process”, empirical mode decomposition (EMD) and local mean decomposition (LMD) decompose a signal into a linear combination of waveforms embedded in the signal without prior knowledge [12–14]. However, sifting-based signal processing methods are restricted by problems of mode mixing and end effects. Although the ability of current fault diagnosis methods to extract fault features has been verified, they have their own shortcomings and advantages, and are constantly developing to overcome their existing shortcomings.

The spectral kurtosis (SK) or fast kurtogram (FK) can identify the frequency band that characterizes the transient vibration signature, thereby it enhances performance of the envelope analysis and makes itself a fully blind detection method. The SK originates from the concept of frequency domain kurtosis proposed by Dwyer [15] as a supplement to the power spectral density (PSD). Antoni [16] provided a specific mathematical definition of the SK based on the Wold-Gramér decomposition of a conditionally non-stationary process and therefore, further insighted into the SK. In addition, he developed the STFT-based SK [17] and multi-rate filter bank (MRFB)-based SK called the FK [18] for rotating machinery fault diagnosis. Further to this, Lei et al. [19] improved the FK by replacing the MRFB with the wavelet packet transform (WPT) to make use of its excellent local properties in the time-frequency domain. However, the regular WPT suffers from the time-variance problem due to the down-sampling operation, which influences the accuracy of the WPT-based FK. As other researchers have pointed out, the main drawback of the FK is the susceptibility to interference from non-Gaussian noise, which often exhibits a higher SK than the fault signal [20–22]. To overcome this, Barszcz et al. [20] developed the protrogram which fixes the bandwidth and finds the optimal center frequency that maximizes the kurtosis of the narrowed envelope spectrum amplitudes. However, to fix the bandwidth, the protrogram requires a priori knowledge about the sought fault, and is therefore, not a fully blind detection method. To eliminate the interference of non-Gaussian noise, Chen et al. [21] adopted the quasi-analytic wavelet tight frame as the filter and the spatial-spectral ensemble kurtosis as the indicator, however, the method is computationally expensive owing to the additional synthesis before calculating the indicator. Moshrefzadeh et al. [22] adopted the kurtosis of the unbiased autocorrelation of the squared envelope as an indicator and introduced the maximal overlap discrete wavelet packet transform (MODWPT) to the FK. In essence, the MODWPT is undecimated DWT and therefore, increases the computational costs considerably.

This paper proposes an alternative method called the subband averaging kurtogram (SAK) to enhance performance of the original FK. As mentioned above, non-Gaussian noise can mislead the FK into selecting an incorrect optimal frequency band resulting in failure of the fault diagnosis. The SAK segments a signal using a sliding window and averages the kurtosis of corresponding segment subbands to remove the influence of sporadic impulse interferences. An additional contribution of this paper is the use of the dual-tree complex wavelet packet transform (DTCWPT) to construct the SAK and extract the periodic transient signal, thereby taking advantages of the time-invariance and fast implementation of the DTCWPT.

The rest of the paper is organized as follows. In Section 2, the FK and DTCWPT theories are reviewed and some limitations of the FK are discussed. Then, two methods are proposed to rearrange subbands of the DTCWPT in order to address the subband disorder problem. In Section 3, the proposed SAK with DTCWPT is illustrated in detail. The subband averaging operation and DTCWPT play a collaborative role in enhancing performance of the FK. We perform a simulation validation to illustrate the use of the proposed method for extracting periodic transients from a noisy signal containing non-Gaussian noise and results are presented in Section 4. In Section 5, we describe the vibration signal analyses using the proposed method for a planetary gear in laboratory and a rolling bearing in a wind farm. Section 6 provides more detailed discussions of the results and finally, some conclusions are presented in Section 7.

2. Brief review of the fast kurtogram and the dual-tree complex wavelet transform

2.1. Fast kurtogram

The FK determines the frequency band with the maximum SK that best characterizes the transient signal for the envelope analysis. In this section, we briefly introduce basic principles of the FK and its limitations.

2.1.1. Principle and implementation

The kurtosis is defined as

$$\text{Kurtosis} = \frac{\frac{1}{N} \sum_1^N (x_i - E(x_i))^4}{\left[\frac{1}{N} \sum_1^N (x_i - E(x_i))^2 \right]^2} - 3 \quad (1)$$

where x_i are sampling points, $E(x_i)$ is the mean value of all sampling points and N is the length of the data. The kurtosis reflects the peakedness of the probability distribution of the data and is a sensitive indicator of machinery faults with impulsive features. As defined by Antoni [16], the SK is an extension of the kurtosis, given by

$$x(n) = \int_{-1/2}^{+1/2} e^{j2\pi fn} H(n, f) dZ_x(f) \quad (2)$$

$$K_x(f) = \frac{\langle |H(n, f)|^4 \rangle}{\langle |H(n, f)|^2 \rangle^2} - 2 \quad (3)$$

Eq. (2) is the Wold-Gramér decomposition of a zero-mean non-stationary signal $x(n)$. $e^{j2\pi fn} H(n, f) dZ_x(f)$ can be regarded as an infinite narrow-band filtered output of $x(n)$ centered on f . $H(n, f)$ is the complex envelope of $x(n)$ at f ; and $\langle \cdot \rangle$ denotes averaging operator. The essential idea is to calculate the kurtosis of a data in time series at each “centered frequency line”. For a signal $y(n) = x(n) + b(n)$ where $b(n)$ is a stationary Gaussian noise, the SK can be calculated as

$$K_y(f) = \frac{K_x(f)}{[1 + \rho(f)]^2} \quad (4)$$

where $\rho(f)$ is the noise-to-signal ratio at frequency f . Note that $K_y(f)$ increases as $\rho(f)$ decreases, which is a desirable property and requirement for identifying the location of non-stationary signals in the frequency domain. The kurtogram is implemented based on STFT and represents the SK as a function of center frequency, f , and frequency resolution, Δf . Then, the frequency resolution $\Delta f = 2f_s/N_w$ can be derived from the sampling frequency, f_s , and window length, N_w . Since mechanical fault signals are essentially narrowband signals, the frequency resolution (determined by window length) which maximizes the SK can be regarded as the optimal bandwidth of its location in frequency domain. The fast algorithm of kurtogram, so-called FK, presents the SK in the $(f, \Delta f)$ plane with filter banks using 1/2-binary tree and 1/3-binary tree structures. The frequency-scale paving of the FK is shown in Fig. 1. Although the frequency resolution must be reduced to decrease computational costs, performance of the FK is not visibly degraded compared to the full kurtogram.

2.1.2. Limitations

Despite effectiveness of the FK, the following limitations should be mentioned: 1) FK calculates the kurtosis of the filtered signal and isolates the mechanical fault signal using STFT or MRFB which may not precisely capture the nonstationary characteristics of fault signals; 2) vibration signals from rotating machinery often contain multiple components, such as non-Gaussian noise which can lead to incorrect selection of the optimal frequency band for the envelope analysis, thereby

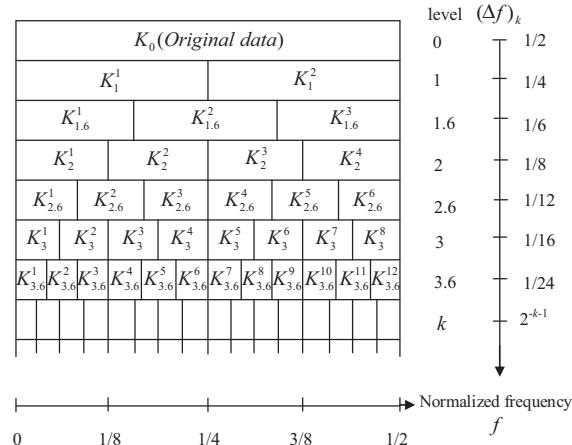


Fig. 1. The frequency-scale paving of the FK.

reducing robustness and accuracy of the FK; 3) the MRFB used in FK does not achieve an orthogonal decomposition, therefore it is inconvenient for interpreting the amplitude of the filtered signal and its squared envelope spectrum (SES).

Although the FK is a useful tool for detecting the location of fault signals in the frequency domain, the above shortcomings limit its robustness and accuracy. In [Section 3](#), we propose a new method to avoid the limitations and enhance performance of the FK for detecting periodic transients from complex signals.

2.2. Dual-tree complex wavelet transform

The dual-tree complex wavelet transform (DTCWT) consists of a pair wavelet filter trees where the wavelets form a Hilbert transform pair, and possesses several desired advantages, including approximate shift-invariance, aliasing reduction and low computational costs [[23,24](#)]. The DTCWPT is a packet form of the DTCWT, and can achieve a much finer frequency resolution in high frequency regions. Nevertheless, center frequencies of subbands obtained by original DTCWPT are not in order. In this section, we briefly review theory of the DTCWPT and propose two schemes to rearrange the subbands of DTCWPT in monotone order of the center frequency.

2.2.1. From the DTCWT to the DTCWPT

The DTCWT operates a pair of real filter bank trees in parallel on an input signal to form the separate real and imaginary parts of the complex WT. Wavelets corresponding to the real and imaginary trees, denoted as $\psi^{Re}(t)$ and $\psi^{Im}(t)$, satisfy the following two-scale relation

$$\psi^*(t) = \sqrt{2} \sum_n h_1^*(n) \phi^*(2t - n) \quad (5)$$

$$\phi^*(t) = \sqrt{2} \sum_n h_0^*(n) \phi^*(2t - n) \quad (6)$$

where $*$ denotes either the real or imaginary part. $\{h_0^{Re}, h_1^{Re}\}$ and $\{h_0^{Im}, h_1^{Im}\}$ are filter banks of the real and imaginary trees, respectively. The DTCWT ensures that the pair of wavelets satisfy the approximate Hilbert transform relationship (in the case of ideal Hilbert transform relationship, the wavelets are not compactly supported) given by

$$\psi^{Im}(t) \approx H\{\psi^{Re}(t)\} \quad (7)$$

where $H\{\cdot\}$ is the Hilbert transform operator, such that the complex wavelet $\psi(t) = \psi^{Re}(t) + j\psi^{Im}(t)$ is quasi-analytic. If the two real DWTs are orthogonal, then a sufficient and necessary condition of Eq. ([7](#)) is

$$h_0^{Im}(n) \approx h_0^{Re}(n - 0.5) \quad (8)$$

In the frequency domain, the condition is equivalent to

$$H_0^{Im}(e^{jw}) = e^{-j0.5w} H_0^{Re}(e^{jw}) \text{ for } |w| < \pi \quad (9)$$

$$H_1^{Im}(e^{jw}) = -j\text{sgn}(w) e^{j0.5w} H_1^{Re}(e^{jw}) \text{ for } |w| < \pi \quad (10)$$

$$\text{sgn}(w) = \begin{cases} 1 & \text{for } w > 0 \\ 0 & \text{for } w = 0 \\ 1 & \text{for } w < 0 \end{cases} \quad (11)$$

$H(\cdot)$ is the z-transform operator. In DTCWT theory, the condition presented in Eqs. ([9](#))–([11](#)) is called the half-sample delay condition. The key point of the DTCWT turns into designing a filter bank pair that satisfies the following conditions: ([1](#)) approximate half-sample delay property, ([2](#)) orthogonal or biorthogonal, ([3](#)) compactly supported, ([4](#)) vanishing moments and ([5](#)) linear-phase. Design of FIR dual-tree filters is comprehensively discussed in Ref. [[23](#)]. Correct implementation of DTCWT requires the first stage filters to be different from the following stages and any existing perfect reconstruction (orthogonal or biorthogonal) filters can be used (if one sample offset is taken, they are used as the imaginary tree filters). Moreover, the filters designed specifically for the DTCWT in Ref. [[25–27](#)] can be used for the following stages. Frequency responses of the DWT and DTCWT are illustrated in [Fig. 2\(a\)](#) and ([b](#)). Frequency responses of the DTCWT are only supported on the positive frequency region except the first stage wavelet and the scale function, indicating the DTCWT is quasi-analytic.

The straightforward idea behind constructing the DTCWPT is to extend the two real DWTs of the DTCWT into their packet forms by applying the same filters to the high frequency channels. However, notable energy leakage into the negative frequency region occurs as shown in [Fig. 2\(d\)](#), therefore, the implementation is not quasi-analytic. In this paper, we introduce the strategy proposed by Bayram [[28](#)] to implement the DTCWPT. The filter bank trees consist of three kinds of filters: ([1](#)) first stage filters $\{h_{10}^{Re}(n), h_{11}^{Re}(n)\}$ and $\{h_{10}^{Im}(n), h_{11}^{Im}(n)\}$, ([2](#)) dual tree filters $\{h_0^{Re}(n), h_1^{Re}(n)\}$ and $\{h_0^{Im}(n), h_1^{Im}(n)\}$, ([3](#)) general filters $\{f_0(n), f_1(n)\}$. The filter bank block of the DTCWPT is shown in [Fig. 3](#). Compared to the DTCWT, the DTCWPT only adds one

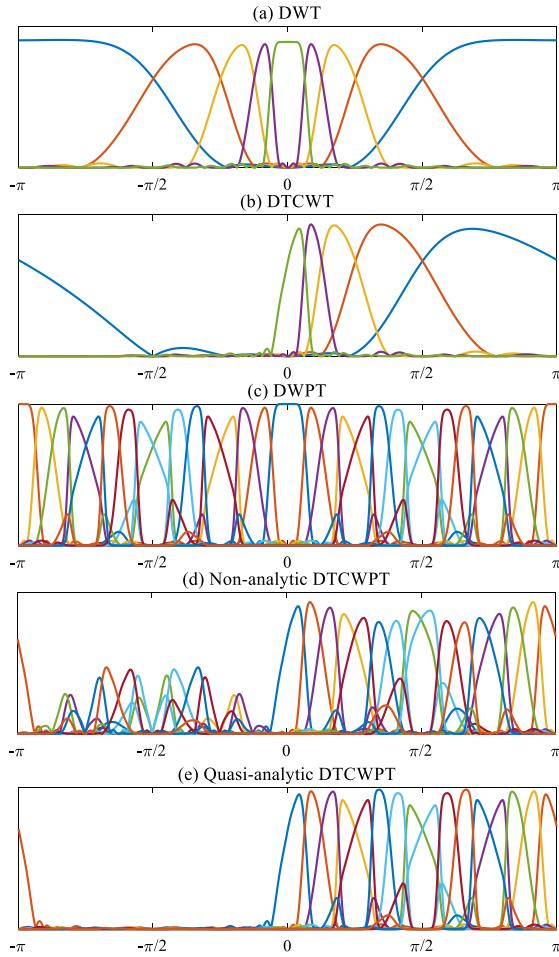


Fig. 2. The frequency responses of (a) DWT, (b) DTCWT, (c) DWPT, (d) Non-analytic DTCWPT and (e) Quasi-analytic DTCWPT.

additional rule: the same filter bank $\{f_0(n), f_1(n)\}$ should be applied to its own outputs as well as high-pass outputs of the dual tree filters in both real and imaginary trees to maintain the quasi-analytic transform. This does not constrain the general filters to be any particular filters, except the requirements for a regular DWT.

Considering the merits of Daubechies filters, 'db14' basis is adopted to construct the first stage filters and the general filters. In addition, Q-shift filters [27] of length 14 are used as the dual tree filters. Fig. 4 plots the wavelets and scale functions of the dual tree complex wavelet and 'db14'. The frequency responses of the DTCWPT is presented in Fig. 2(e). With the exception of the first and last subbands, frequency responses are approximately single sided on the positive frequency axis, confirming that the transform is quasi-analytic.

2.2.2. Subband rearranged DTCWPT

Although the DTCWPT decomposes a signal into octave subbands, their center frequencies are not in order. This phenomenon is referred to as subband disorder and is caused by the band-pass characteristics of the wavelet filter bank exchange in some frequency bands. According to the filter bank block shown in Fig. 3, the total filter corresponding to the node $d_{k,i}$ can be written as

$$W_{k,i}(e^{jw}) = H_{1n_0}(e^{jw}) \left[\prod_{u=1}^{N_1} H_{n_u}(e^{j2^u w}) \right] \left[\prod_{v=N_1+1}^{k-1} F_{n_v}(e^{j2^v w}) \right] \quad (12)$$

k denotes the decomposition level; i is the index of subbands; $[n_0 n_1 \dots n_{k-2} n_{k-1}]$ is the binary form of integer $i - 1$, namely, $i - 1 = \sum_{m=0}^{k-1} 2^{k-1-m} n_m$; N_1 is the index of the first nonzero bit in the sequence $[n_1 \dots n_{k-2} n_{k-1}]$, namely, $n_{N_1} = 1$ and $n_m = 0$ (for $0 < m < N_1$). $H_*(e^{jw})$, $H_s(e^{jw})$ and $F_*(e^{jw})$ are periodic function of period π , where $*$ denotes either the high-pass or low-pass filter. Fig. 5 illustrates the low-pass and high-pass filters of $F_*(e^{jw})$ and $F_*(e^{j2w})$. Note that the band-pass

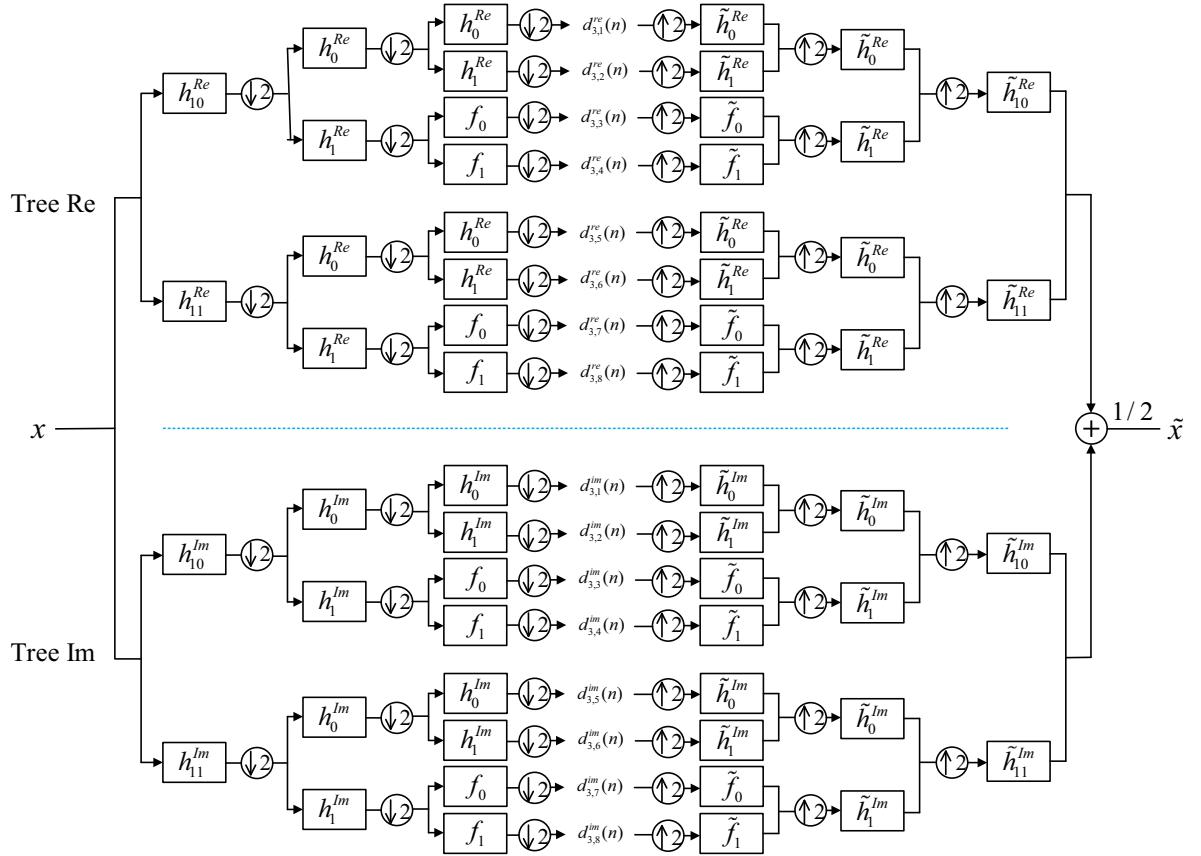


Fig. 3. The analysis and synthesis filter banks for the implement of DTCWPT.

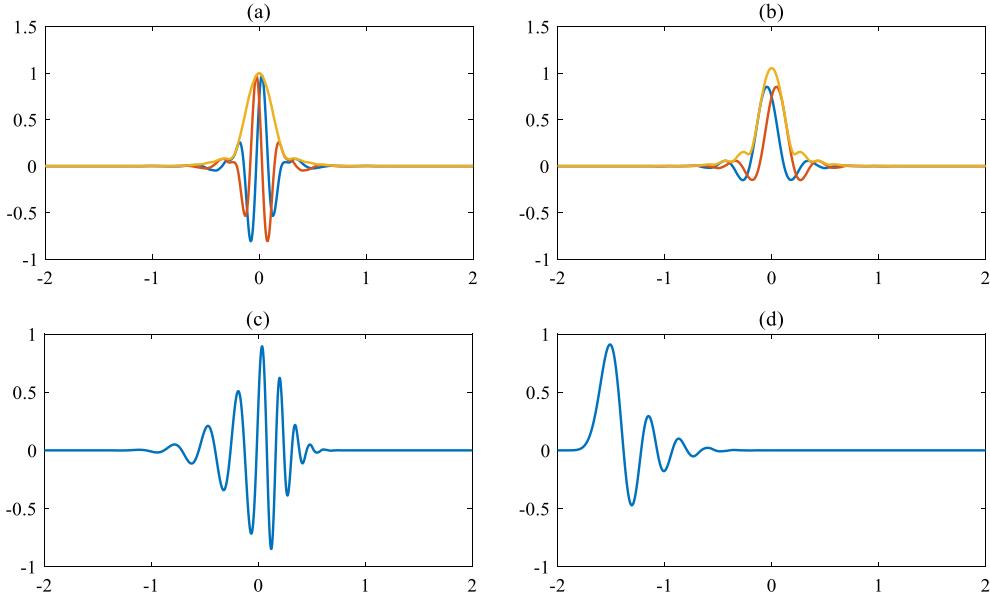


Fig. 4. (a) The wavelet of the dual tree complex wavelet, (b) the scale function of the dual tree complex wavelet, (c) the wavelet of 'db14' and (d) the scale function of 'db14'.

characteristics of $F_*(e^{j2w})$ exchange in $[-\pi, -0.5\pi] \cup [0.5\pi, \pi]$. In fact, it is still valid for either $F_*(e^{j2^k w})$ or $H_*(e^{j2^k \pi})$ in $[-\frac{i}{2^k}\pi, \frac{i-1}{2^k}\pi] \cup [\frac{i-1}{2^k}\pi, \frac{i}{2^k}\pi]$ if i is even. Based on this, we provide two schemes for solving the subband disorder problem.

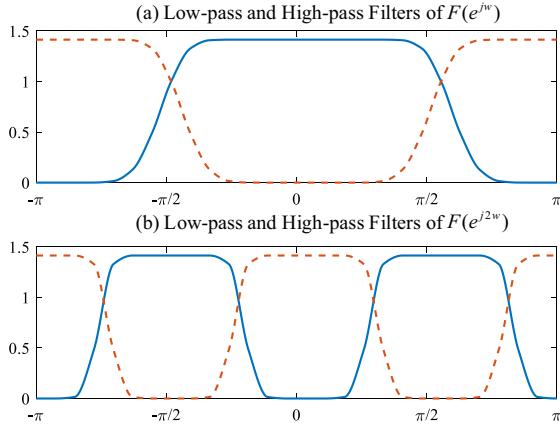


Fig. 5. Low-pass and high-pass filters of (a) $F_*(e^{jw})$ and (b) $F_*(e^{j2w})$.

Method 1: Subbands directly obtained by the DTCWPT are denoted as $\{d_{k,w} \mid w = 1, 2, \dots, 2^k\}$ and rearranged in decreasing order in respect to their center frequencies, $\{d_{k,r} \mid r = 1, 2, \dots, 2^k\}$. Table 1 lists examples of “wrong” and “right” order of the subbands obtained by the original DTCWPT. The binary form of integer $w - 1$ is $[n_{k-1} n_{k-2} \dots n_1 n_0]$, such that

$$i - 1 = \sum_{m=0}^{k-1} 2^m n_m \quad (13)$$

$[n'_{k-1} n'_{k-2} \dots n'_1 n'_0]$ is the binary form of integer $r - 1$, given by

$$n'_m = \begin{cases} n_m & m = k-1 \\ \text{mod}(n_m + n_{m+1}, 2) & 0 \leq m < k-1 \end{cases} \quad (14)$$

The decimal r can be written as

$$r = 1 + \sum_{m=0}^{k-1} 2^m n'_m \quad (15)$$

To reconstruct the signal, the set $\{d_{k,r} | r = 1, 2, \dots, 2^k\}$ should be rearranged as $\{d_{k,w} | w = 1, 2, \dots, 2^k\}$, illustrated as

$$n_m = \begin{cases} n'_m & m = k - 1 \\ mod(\sum_{s=m}^{k-1} n'_s, 2) & 0 \leq m < k - 1 \end{cases} \quad (16)$$

$$w = 1 + \sum_{m=0}^{k-1} 2^m n_m \quad (17)$$

Method 2: $\left\{ G_*^i(e^{j2^k w}) \mid i = 1, 2, \dots, 2^{k-1} \right\}$ denotes the k -stage i th filter bank. If i is even, exchange the low-pass and high-pass filters of $G_*^i(e^{j2^k w})$. If $i = 1$ or 2^{k-1} , $G_*^i(e^{j2^k w})$ should be $H_*(e^{j2^k w})$; If $1 < i < 2^{k-1}$, $G_*^i(e^{j2^k w})$ should be $F_*(e^{j2^k w})$ when $k > 1$. Similar to the decomposition process, exchange the low-pass and high-pass filters of $G_*^i(e^{j2^k w})$ in the reconstruction process if i is even.

The filter bank block of the subband rearranged DTCWPT is illustrated in Fig. 6. Since Method 2 rearranges subbands stage by stage, it can be conveniently incorporated into the SAK algorithm, which will be illustrated in Section 3.

Table 1

Examples of “wrong” and “right” order of the subbands obtained by the original DTCWPT.

Decomposition level	2																
"Wrong order"	1	2	3	4													
"Right order"	1	2	4	3													
Decomposition level	3																
"Wrong order"	1	2	3	4	5	6	7	8									
"Right order"	1	2	4	3	8	7	5	6									
Decomposition level	4																
"Wrong order"	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
"Right order"	1	2	4	3	8	7	5	6	16	15	13	14	9	10	12	11	

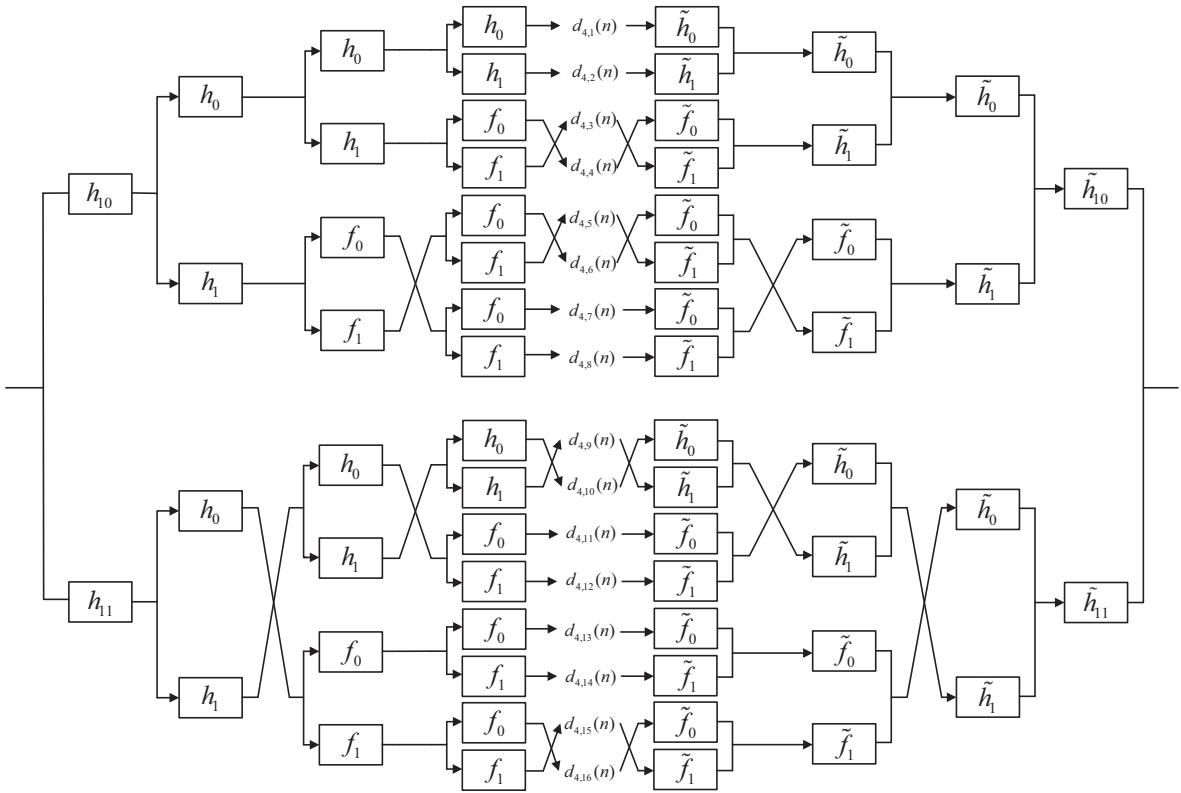


Fig. 6. The filter bank block of subband rearranged DTCWPT (the downsampling and upsampling operations are omitted).

2.2.3. Approximate shift-invariance

Owing to the down-sampling operation, the DWT is shift-variant, which means small time translations in the input signal can result in visible variations of the wavelet coefficients. The continuous WT and undecimated WT can overcome the

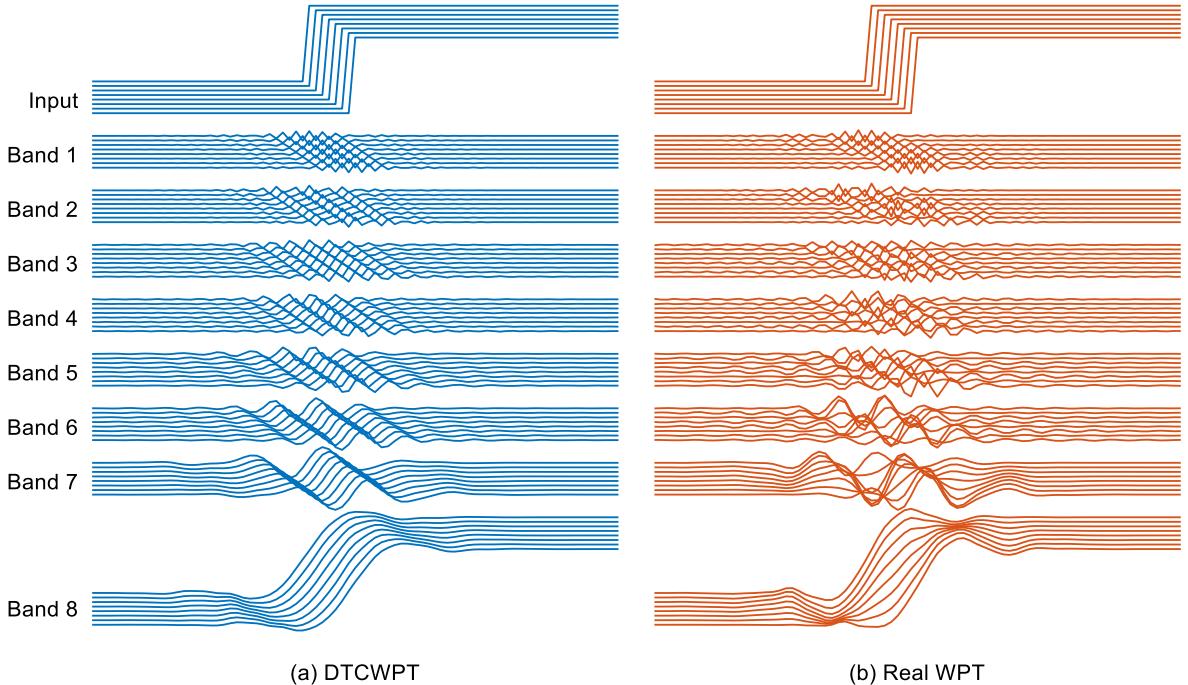


Fig. 7. Subband components (in decreasing order of center frequency) of 8 shifted step signals of (a) the DTCWPT and (b) the real WPT.

shift-variance problem but at the expense of increasing computational costs. The DTCWPT reduces aliasing caused by the down-sampling operation, thereby producing approximate shift-invariance, owing to the quasi-analytic filter bank. Moreover, computational costs of the DTCWPT are only twice the real WPT for a one-dimensional signal. The DTCWPT possesses the approximate shift-invariance property in contrast to the real WPT, as shown in Fig. 7.

3. Proposed method

This section presents a new method called the SAK to avoid the limitation that the FK is easy to be interfered by non-Gaussian noise. Furthermore, by incorporating the DTCWPT, the proposed method replaces the MRFB of the FK with quasi-analytic wavelet filter banks to obtain the kurtogram and extract periodic transients.

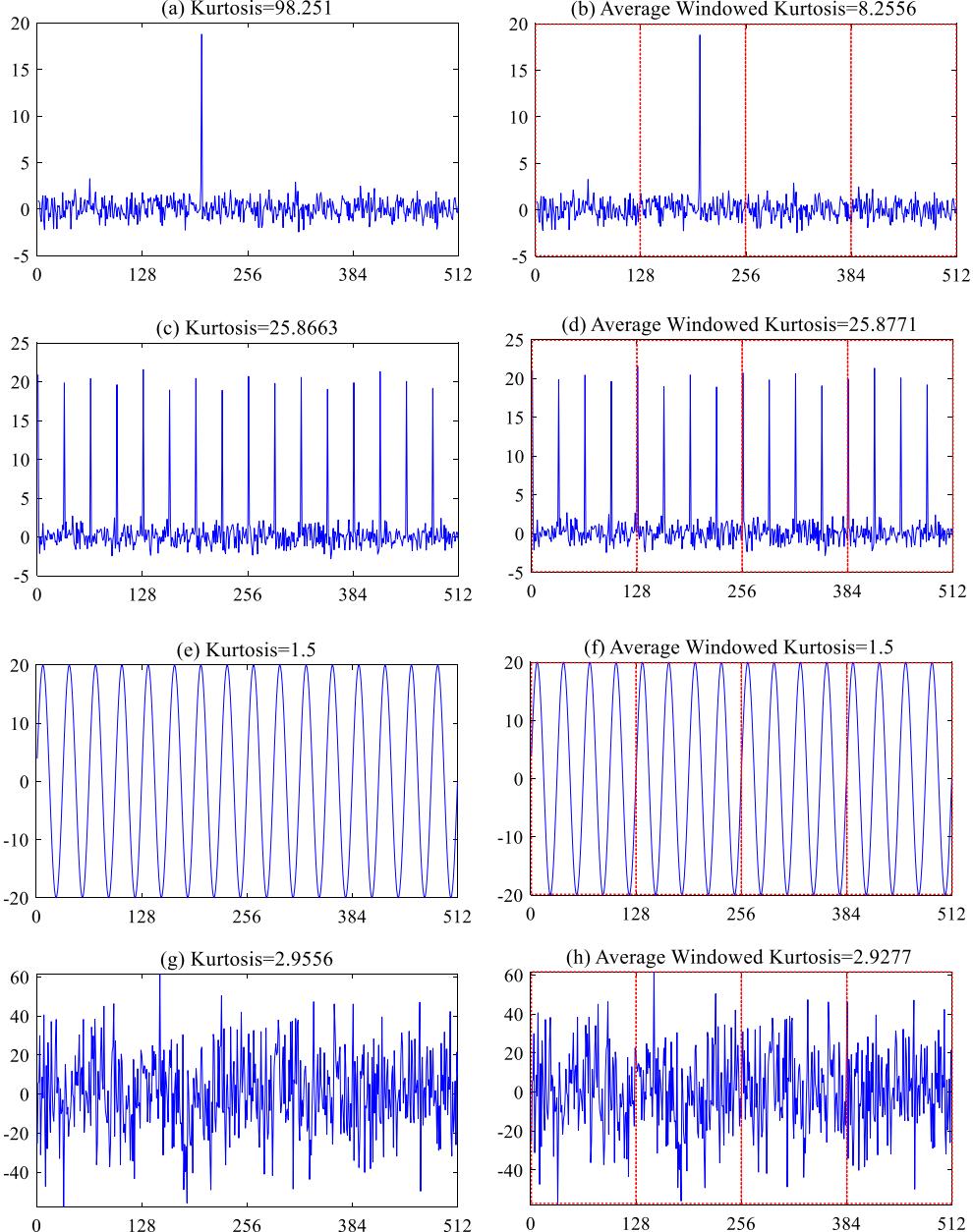


Fig. 8. (a) the kurtosis of noisy single impulse signal, (b) the average windowed kurtosis of noisy single impulse signal ($N_w = 4$), (c) the kurtosis of noisy periodic transient signal, (d) the average windowed kurtosis of noisy periodic transient signal ($N_w = 4$), (e) the kurtosis of harmonic signal, (f) the average windowed kurtosis of harmonic signal ($N_w = 4$), (g) the kurtosis of Gaussian white noise and (h) the average windowed kurtosis of Gaussian white noise ($N_w = 4$). (The kurtosis of the signal is calculated using Eq. (1). The average windowed kurtosis is calculated as $1/N_w \sum_{i=1}^{N_w} \text{Kurtosis}(x_i)$.

As mentioned in Introduction, the challenge of rotating machinery fault diagnosis is to distinguish the periodic transient signal from complex environment noise. The FK is an efficient tool to detect the fault signal from a noisy vibration signal because the fault signal and noise exhibit different values of SK. Sporadic impulse interferences hinder the optimal frequency band selection of the FK because its SK is even higher than the fault signal. To avoid this problem, we develop the SAK based on the sliding window. As illustrated in Fig. 8, the average windowed kurtosis of the single impulse signal is far less than its kurtosis while the average windowed kurtosis and kurtosis of the periodic transient signal, harmonic signal and Gaussian white noise are approximately equal. It is believed that the fault signal is cyclostationary and the sporadic impulse interferences, at least those within same frequency band, do not spread over the entire time domain. Fig. 8 indicates that the sub-band averaging operation only significantly reduces the SK of sub-bands corresponding to the sporadic impulse interferences. Thus, the frequency band corresponding to the maximum SK is where the fault signal is located. Moreover,

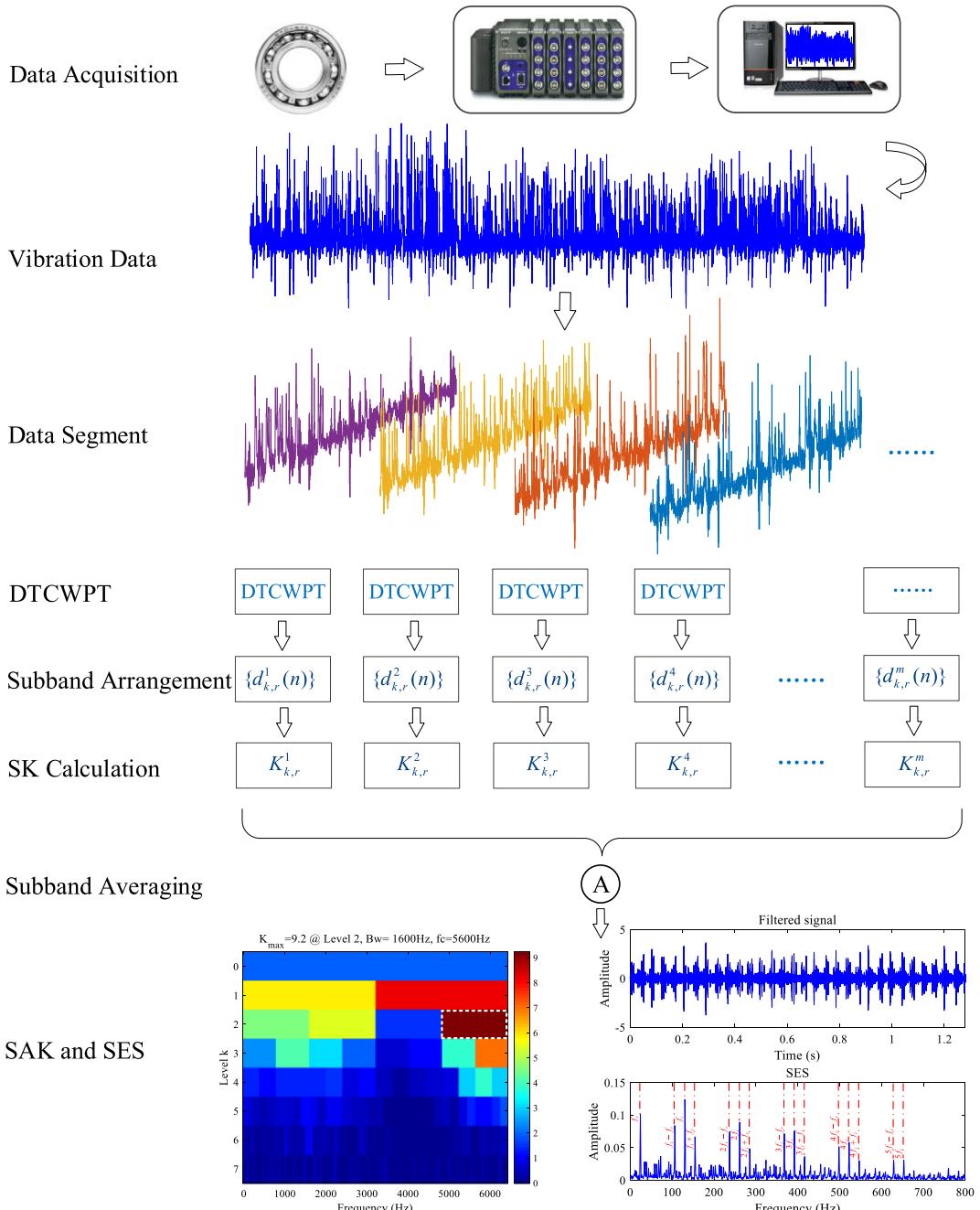


Fig. 9. The flowchart of the proposed method.

compared to the MRFB, the wavelet filter is more flexible for characterizing the local property of a nonstationary signal. However, one constraint of the kurtogram is the translation invariance [17] which the regular WPT does not satisfy. An advantage of the DTCWPT is that it provides an approximate shift-invariance property without significantly increasing computational costs. The DTCWPT also satisfies the other three constraints of the kurtogram discussed in Ref. [17]. Therefore, the proposed method applies the DTCWPT to construct the SAK and extract the periodic transient signal based on the optimal band identified by the SAK. A flowchart of the proposed method is presented in Fig. 9. The procedure can be described in detail as follows:

- (1) Segment the collected signal $x(t)$ into M sub-signals $\{x_m(t)|m = 1, 2, \dots, M\}$ by sliding window. The length of the sliding window is recommended according to the decomposition level and the length of the signal. In general, the recommended length is 2048 if the decomposition level is 7. The more segments, the less influence of non-Gaussian noise on the SAK. However, if the length of sub-signal is too short, the frequency resolution of the kurtogram will be unacceptably poor.
- (2) Decompose M sub-signals into M sets of subband wavelet coefficients $\{d_{k,w}^m(n)\}$. Rearrange $\{d_{k,w}^m(n)\}$ using Method 2 proposed in Section 2.2.2 to obtain $\{d_{k,r}^m(n)\}$.
- (3) Calculate the kurtosis of subbands, given by

$$K_{k,r}^m = \frac{\left\langle \left| d_{k,r}^m(n) \right|^4 \right\rangle}{\left\langle \left| d_{k,r}^m(n) \right|^2 \right\rangle^2} - 2 \quad (18)$$

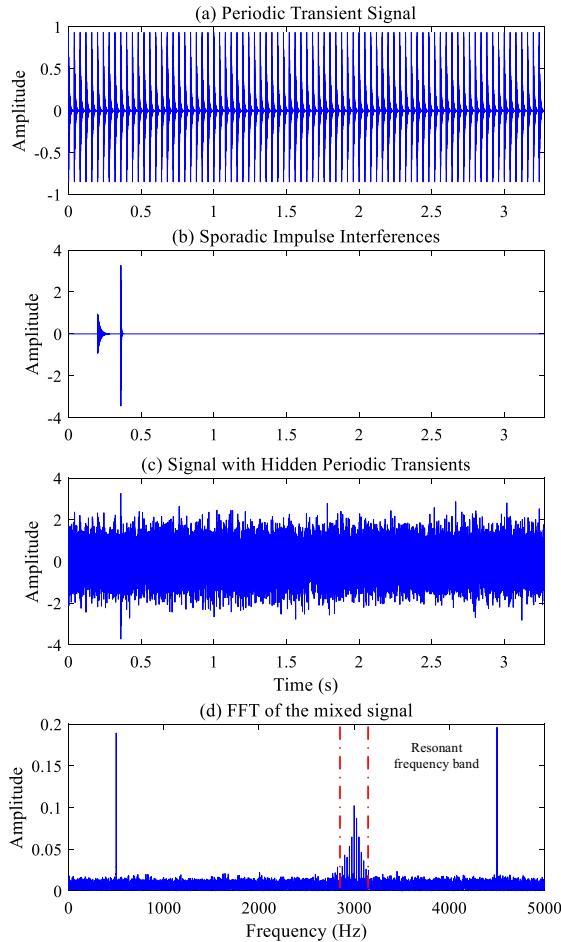


Fig. 10. Waveform in time domain of (a) the periodic transient signal, (b) sporadic impulse interferences, (c) the mixed signal and (d) FFT spectrum of the mixed signal.

(4) Average the kurtosis of corresponding subbands, as follows

$$\bar{K}_{k,r} = 1/M \sum_{m=1}^M K_{k,r}^m \quad (19)$$

(5) Obtain the SAK and select the optimal frequency band (f_c , Bw) for the envelope analysis.

(6) Obtain the purified signal and its SES. Finally, make the final diagnosis based on the obtained information.

4. Simulation validation

To validate the feasibility of the proposed method for extracting periodic transient signals, a multicomponent signal consisting of Gaussian white noise, harmonics, sporadic impulse interferences and a periodic transient signal is constructed. The simulated signal can be expressed as

$$\begin{aligned} s(t) &= x(t) + c(t) + i(t) + n(t) \\ \left\{ \begin{array}{l} x(t) = \sum_{k=0}^{81} e^{-60\pi(t-0.04k)} \sin(6000\pi\sqrt{1-0.01^2}(t-0.04k)) (0.04k \leq t < 0.04(k+1)) \\ c(t) = 0.25[\cos(1000\pi t + \pi/4) + \cos(9000\pi t + \pi/4)] \\ i(t) = e^{-20\pi(t-0.2)} \sin(4000\pi\sqrt{1-0.005^2}(t-0.2)) (0.2 \leq t < 0.28) \\ + 4e^{-160\pi(t-0.36)} \sin(8000\pi\sqrt{1-0.02^2}(t-0.36)) (0.36 \leq t < 0.4) \\ n(t) = 0.65rand(n) \end{array} \right. \end{aligned} \quad (20)$$

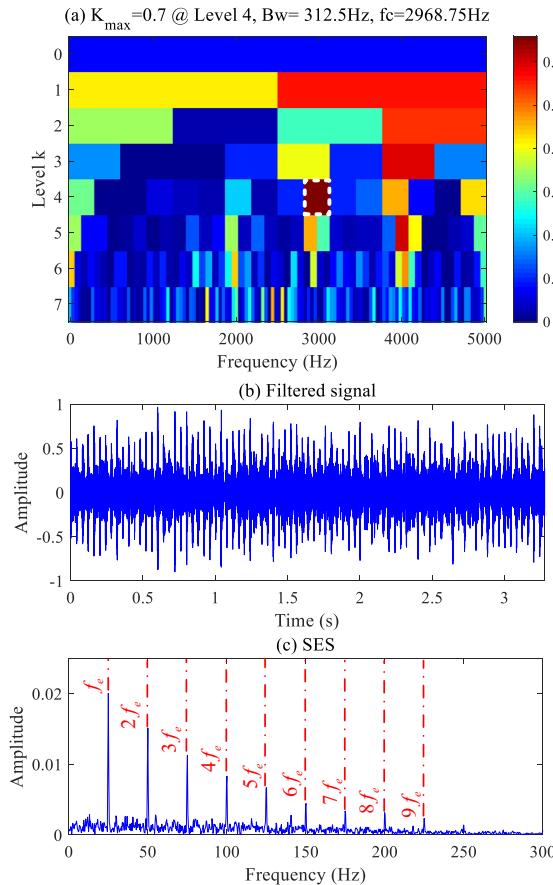


Fig. 11. The results of the proposed method: (a) the kurtogram, (b) the filtered signal and (c) the SES.

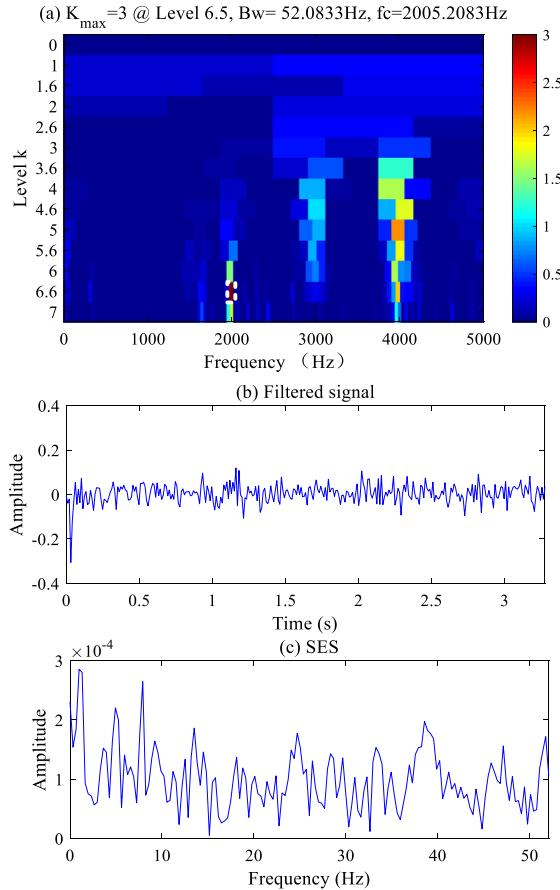


Fig. 12. The results of the original FK: (a) the kurtogram, (b) the filtered signal and (c) the SES.

Fig. 10 displays the waveform of the periodic transient signal, the sporadic impulse interferences, the mixed signal and the FFT spectrum of the mixed signal (sampling frequency of 10000 Hz and data length of 32768). The signal-to-noise (SNR) of $s(t)$ is -11.64 dB. The proposed method is employed to analyze $s(t)$ to extracting the periodic transient signal, $x(t)$. Fig. 11 describes the SAK, the filtered signal and its SES. The SAK indicates an optimal frequency band of (2968.75 Hz, 321.5 Hz). Periodic transients with period $t \approx 0.04$ s are visible in the time domain of the filtered signal. The envelope frequency of the transients ($f_e = 25$ Hz) and its harmonics can be recognized without ambiguity. To illustrate the enhancement of the proposed method, the original FK (see Fig. 12), the FK with DTCWPT (see Fig. 13) and the SAK with MRFB (see Fig. 14) are used to analyze the simulated signal. From the Figs, the three methods are misled by the sporadic impulse interferences and the periodic transients and its envelope frequency are hardly found in the time domain and SES, respectively. If we try the frequency band (3020.84 Hz, 208.33 Hz) which is the second possible frequency band indicated by the SAK with MRFB (marked by white square and the number 2), the filtered signal and its SES are presented in Fig. 15. Then, the periodic transients can be identified. However, note the differences in amplitude of the filtered signals and their SES between the proposed method and the SAK with MRFB. Table 2 reports the CPU time of the above four methods implemented in Matlab (version R2018a) on a laptop computer (i5-8300H CPU @ 2.30 GHz). The four methods are all computationally efficient (the FK with DTCWPT is the fastest) and thus suitable for online condition monitoring in industrial applications.

5. Applications to rotating machinery fault diagnosis

When faults (unbalance, pitting, crack, rubbing, etc.) occur in rotating machinery, they always arise periodic transient vibration signal due to the rotation of the machinery. Faults are usually assigned to corresponding known characteristic frequencies in a spectrum (Fourier spectrum or envelope spectrum), because fault responses are always related to the periods of the excitation caused by faults. Once the periodic transient signal corresponding to the fault is separated from the additive background noise, it is easy to evaluate the running state and diagnose the fault of the machinery. In this section, we present two examples of using the proposed method for rotating machinery fault diagnosis.

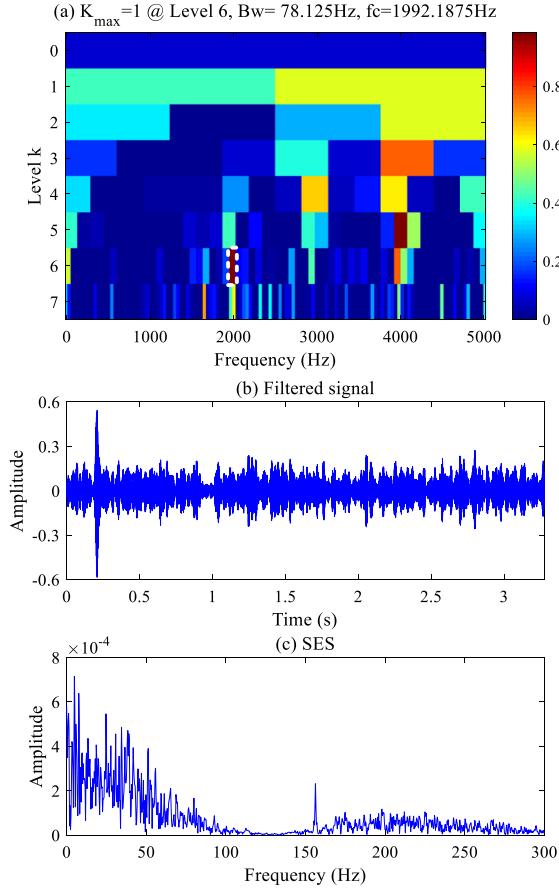


Fig. 13. The results of the FK with DTCWPT: (a) the kurtogram, (b) the filtered signal and (c) the SES.

5.1. Case 1: Planetary gearbox fault diagnosis

Owing to their advantages of high efficiency, compact structure and heavy load capacity, the planetary gearboxes are wildly applied to the transmission systems of helicopters, wind turbines and mining machinery, etc. Here we use the proposed method to analyze the vibration signal from a planetary gearbox fault simulation test rig.

5.1.1. Experimental setup and data acquisition

Fig. 16 (a) shows the planetary gearbox fault simulation test rig. Table 3 lists parameters of the planetary gearbox. To simulate gear damage, a crack was implanted into the root of the planet gear tooth, as shown in Fig. 16(b). The test rig was driven by an AC motor and loaded by a brake on the other end. The vibration signal was measured by an accelerometer mounted on top of the planetary gearbox housing. The rotational speed of the motor was 3000 r/min and the sampling frequency was 20480 Hz. According to Ref. [29], the fault characteristic frequencies of planetary gearbox with local damage can be calculated as

$$\begin{cases} f_{carrier} = f_{sunrot}/(1 + Z_{ring}/Z_{sun}) = 10 \text{ Hz} \\ f_{mesh} = f_{carrier} \times Z_{ring} = 840 \text{ Hz} \\ f_{sun} = f_{mesh} \times N_{planet}/Z_{sun} = 120 \text{ Hz} \\ f_{planet1} = f_{mesh}/Z_{planet} = 27.09 \text{ Hz} \\ f_{planet2} = 2f_{mesh}/Z_{planet} = 54.19 \text{ Hz} \\ f_{ring} = f_{mesh} \times N_{planet}/Z_{ring} = 30 \text{ Hz} \end{cases} \quad (21)$$

5.1.2. Results and analysis

Fig. 17 displays the waveform and the Fourier spectrum of the collected vibration signal. The proposed method is used to analyze the signal. The analysis results are illustrated in Fig. 18. According to the SAK, the signal in frequency band (2880 Hz, 640 Hz) is filtered out. The selected resonant frequency band is marked in red in Fig. 17(b). The filtered signal

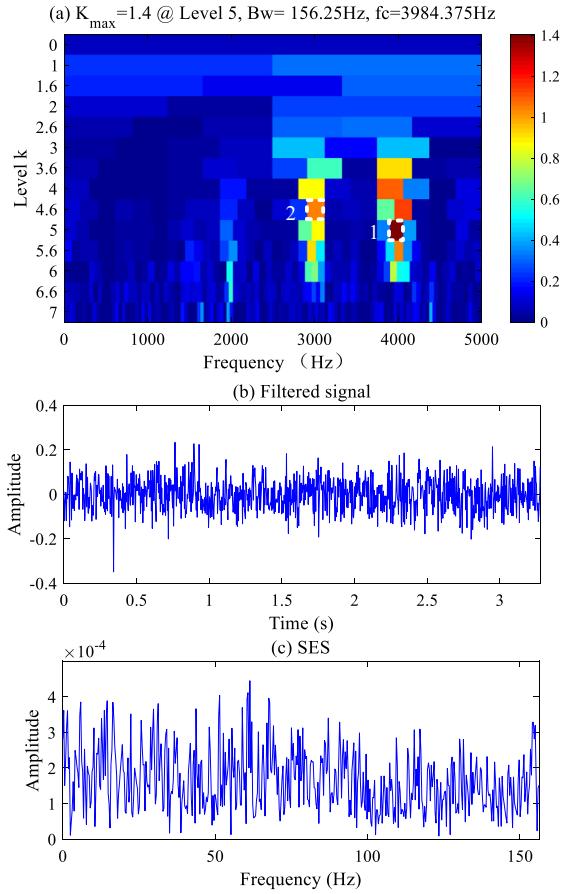


Fig. 14. The results of the SAK with MRFB: (a) the kurtogram, (b) the filtered signal and (c) the SES.

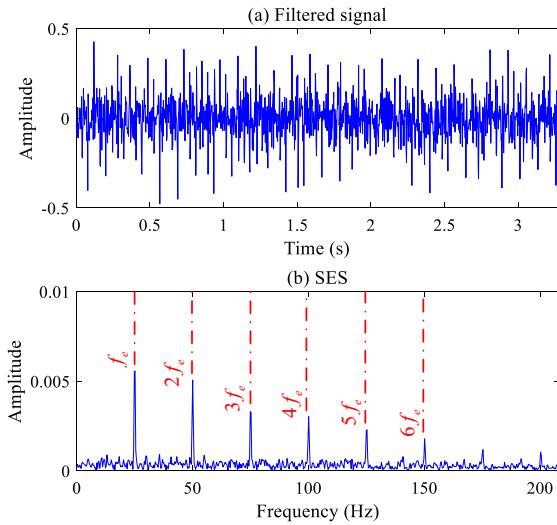


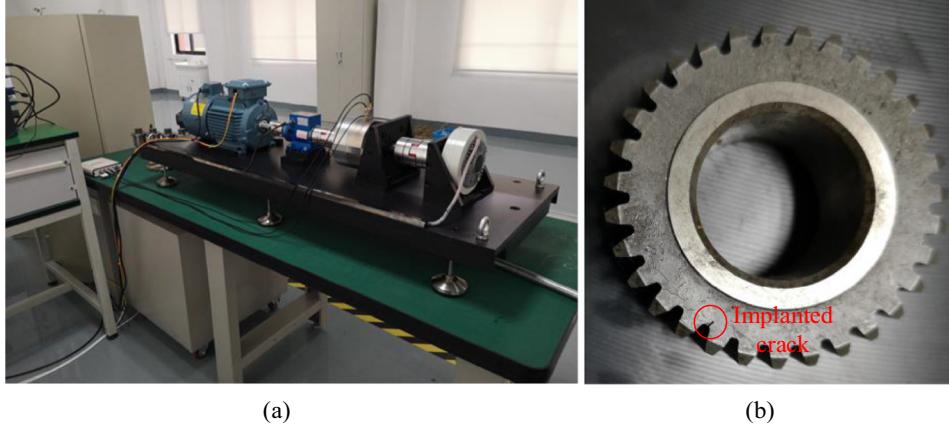
Fig. 15. The results obtained by the frequency band with fourth higher SK indicated by the SAK with MRFB (a) the filtered signal and (b) the SES.

and its SES are plotted in Fig. 18(b) and (c), respectively. The SES clearly exhibits the rotating frequency of the carrier, $f_{carrier}$, and the fault characteristic frequency of planet gear, $f_{planet2}$, its second-harmonic, $2f_{planet2}$, as well as their combinations $n f_{planet2} \pm f_{carrier}$. All frequencies are associated with the planet gear fault, and consistent with the known fault of the

Table 2

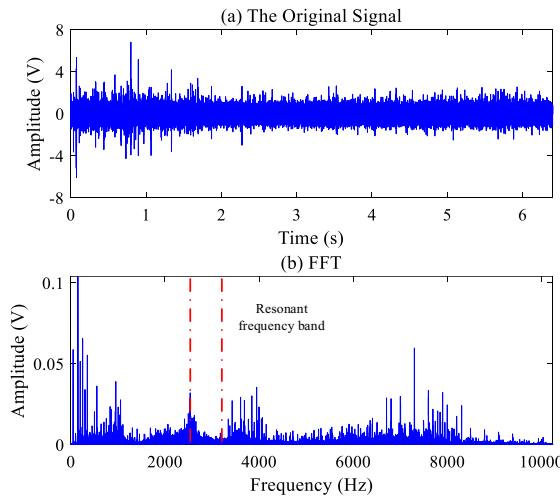
CPU time required for the four methods in simulation validation.

Method	The proposed method	the original FK	The FK with DTCWPT	the SAK with MRFB
CPU time (s)	0.3278	0.3087	0.2834	0.3923

**Fig. 16.** Case1: (a) the planetary gearbox fault simulation test rig and (b) the planet gear with implanted crack.**Table 3**

The parameters of the planetary gearbox in case 1.

Gear	Number of teeth
Sun	21
Ring	84
Planet (Number)	31(3)

**Fig. 17.** Case1: (a) the time domain waveform of the original signal and (b) its FFT spectrum.

planetary gearbox. As a comparison, the original FK is also adopted to analyze the signal (see Fig. 19). Frequencies related to the planet gear fault cannot be identified in the SES obtained by the FK, suggesting that the FK fails to detect the planet gear fault. The FK is a powerful tool for identifying the resonance frequency band of mechanical faults in most of cases. Here, we provide a case where the FK fails while the proposed method successfully diagnoses the fault in order to illustrate the proposed method can promote the robustness and accuracy of the original FK.

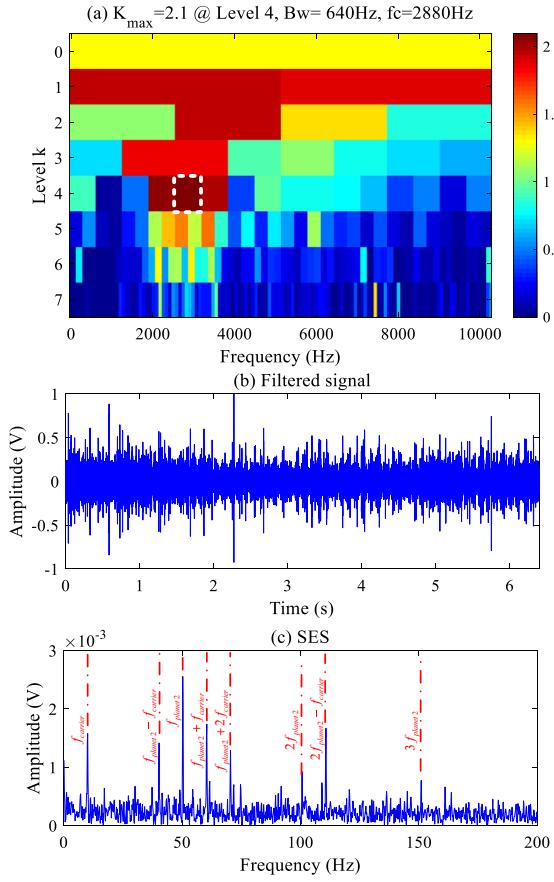


Fig. 18. The results of the proposed method in case 1: (a) the kurtogram, (b) the filtered signal and (c) the SES.

5.2. Case 2: Bearing fault diagnosis

Rolling bearings are common components of most transmission systems and also one of the most common origins of faults in mechanical systems. This section provides an example of using the proposed method for fault diagnosis of a rolling bearing in a wind farm.

5.2.1. Experimental setup and data acquisition

Vibration signals were measured from a wind farm in Nan'ao Island on Shantou, China. Vibration-based technique was applied to the fault detection and diagnosis for the wind turbine transmission system. Fig. 20 illustrates the sensor layer of the wind turbine transmission system and a field measuring point. During the test, it was found that the amplitude of the vibration signal measured from the generator front bearing (measuring point 10 in Fig. 20(a)) of No. 9 wind turbine was larger than at other measuring points, so it was suspected that the generator front bearing had failed and the vibration signal required to be further analyzed. According to the pulse key phase signal, the average rotating speed of the generator at that time was 1453 r/min. Parameters of the generator front bearing are listed in Table 4. Fault characteristic frequencies are calculated as $f_o = 86.8$ Hz, $f_i = 131.1$ Hz, $f_b = 57.1$ Hz and $f_c = 9.6$ Hz. The sampling frequency was 12800 Hz and the sampling points was 16384.

5.2.2. Results and analysis

The time domain waveform and the Fourier spectrum of the raw signal are displayed in Fig. 21. Although apparent transients can be observed in the time domain, the period of the transients can hardly be identified. Messy Fourier spectrum indicates the raw signal contain various components. The proposed method (see Fig. 22) and the FK (see Fig. 23) are adopted to diagnose the fault of the bearing. There are dominated peaks at the rotating frequency, the inner race fault characteristic frequency and its harmonics as well as the sidebands in the SES obtained by both the proposed method and the FK, which indicates that fault is located on the inner race of the bearing. However, the resonant frequency band is more precisely identified by the proposed method and the fault characteristic frequencies are more dominant with larger amplitudes compared

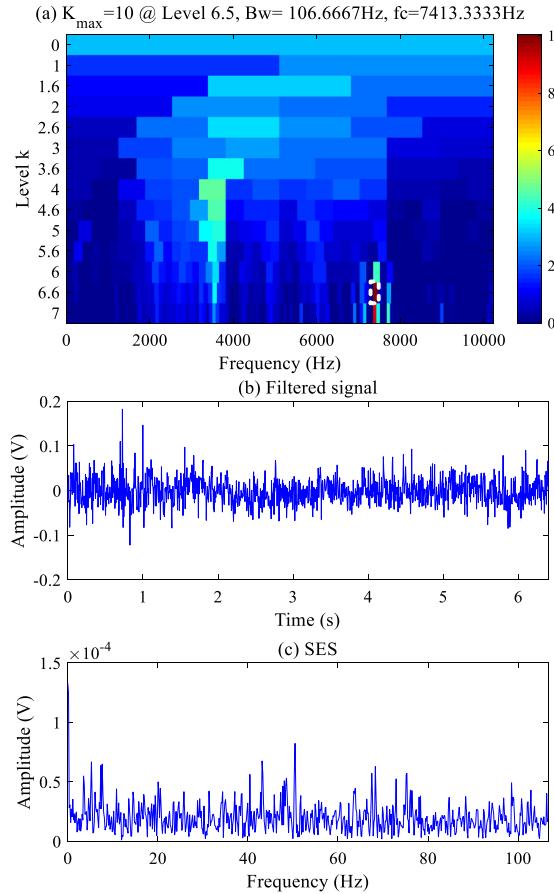


Fig. 19. The results of the original SK in case 1: (a) the kurtogram, (b) the filtered signal and (c) the SES.



Fig. 20. Case2: (a) the measuring points of the wind turbine transmission and (b) the field measuring point.

Table 4
Bearing parameters in case 2.

Bearing Parameters	Value
Type	6332M_FAG
Roller diameter (d/mm)	50.8
Pitch diameter (D/mm)	250
Roller number (N)	9
Contact angle ($\alpha/^\circ$)	0

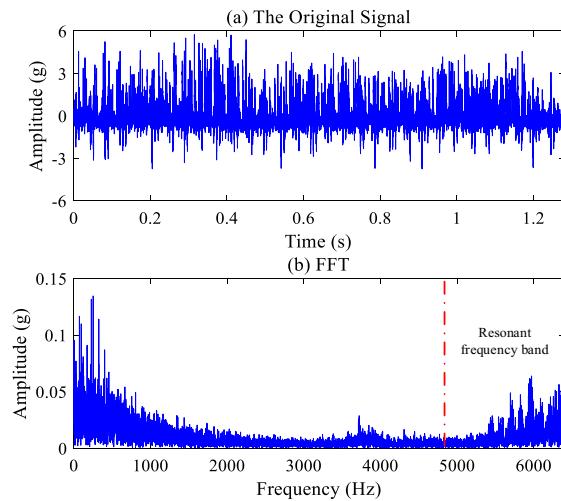


Fig. 21. Case2: (a) the time domain waveform of the original signal and (b) its FFT spectrum.

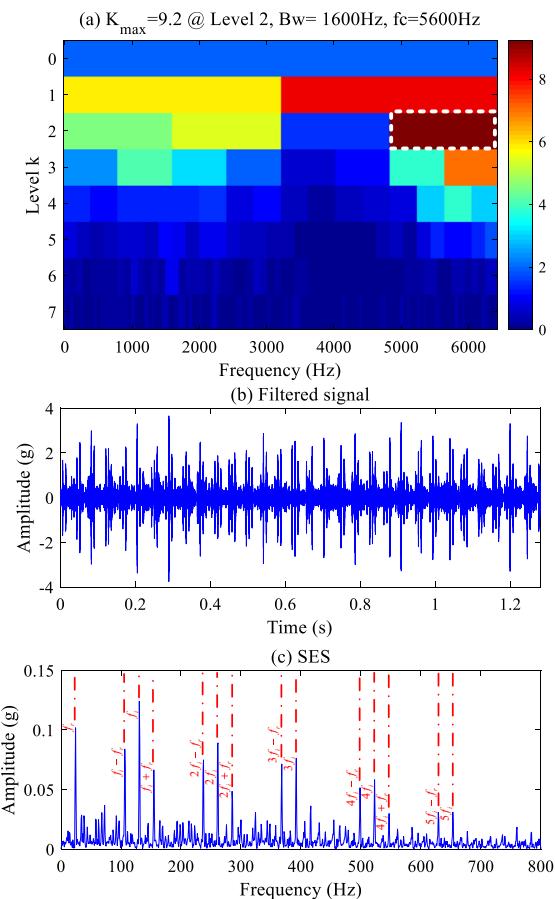


Fig. 22. The results of the proposed method in case 2: (a) the kurtogram, (b) the filtered signal and (c) the SES.

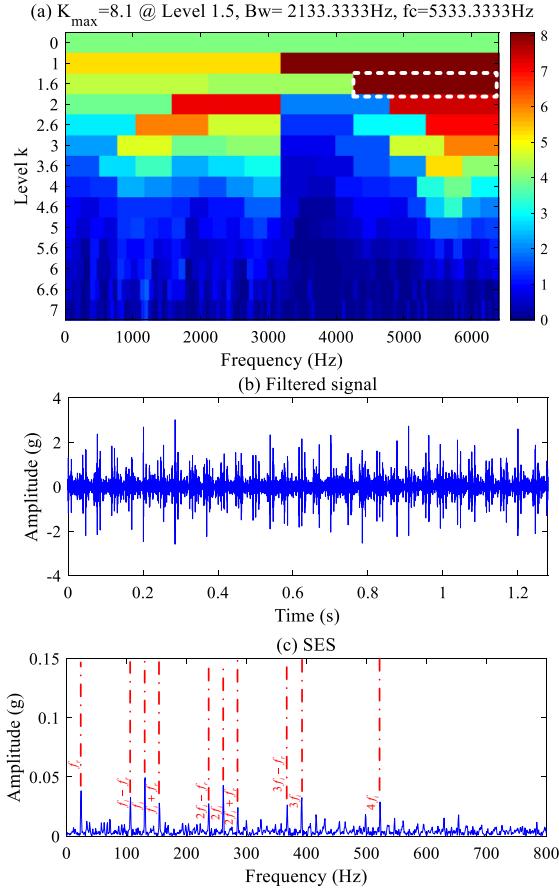


Fig. 23. The results of the original FK in case 2: (a) the kurtogram, (b) the filtered signal and (c) the SES.

to the FK. Thus, the proposed method is again validated and shown to be superior to the FK for rotating machinery fault diagnosis.

6. Discussions

- (1) The idea of subband averaging operation can be extended to most of the SK-based methods to suppress the influence of non-Gaussian noise on the kurtogram. Although simple, the validity of SAK has been verified by a simulation case and two fault diagnosis applications: a planetary gearbox and a bearing.
- (2) Wavelet filters better match local features of fault signals owing to their excellent local properties in the time-frequency space. Furthermore, compared to real wavelet filters, dual-tree wavelet filters yield approximately shift-invariant property without significantly sacrificing computational efficiency and can extract the mechanical fault signatures more precisely.
- (3) A disadvantage of the proposed method compared to the FK is that it does not have the same delicate band divisions since the FK simultaneously involves 1/2-binary tree and 1/3-binary tree band divisions whereas the DTCWPT is only equivalent to the 1/2-binary tree filter banks. Fineness of the band division can be improved by designing M-band (in case of $M = 3$, the frequency resolution is equivalent to the FK) dual-tree wavelet filters, but is beyond the scope of this paper.
- (4) On the other hand, the proposed method has high computational efficiency as illustrated in Section 4, which is important for online monitoring. In the online applications, the vibration signal data stream measured from running rotating machinery can be intercepted through sliding window to update the sub-window kurtogram as an early bird warning. Meanwhile, the SAK is updated as a more reliable and accurate detection and diagnosis result.

7. Conclusions

The SAK algorithm provides a simple solution for suppressing the influence of non-Gaussian noise on the kurtogram and enhancing performance of the FK. The DTCWPT can be used to decompose a signal into different frequency subbands. Two

methods are proposed to rearrange the subbands of the DTCWPT to put them arranged in monotone order of center frequency in each decomposition level. The proposed method robustly identifies the optimal frequency band where the fault signal is located and can further accurately extract fault features. Furthermore, the method is a fully blind detection method and therefore, can be combined with the SES to automatically generate a single spectrum containing clear fault information without any prior knowledge. Results of the simulation show that the proposed method can accurately extract periodic transients in the presence of non-Gaussian noise, whereas the FK cannot. Finally, the effectiveness of the proposed method is further validated by fault diagnosis applications where it is applied to analyze two vibration signals from faulty rotating machineries.

Acknowledgments

This work was supported by the joint fund of Civil Aviation research of China Administration of China project (No. U1733107), and the Aeronautical Science Foundation of China (No. 20173319003).

References

- [1] J. Li, M. Li, J. Zhang, et al, Frequency-shift multiscale noise tuning stochastic resonance method for fault diagnosis of generator bearing in wind turbine, *Measurement* 133 (2019) 421–432.
- [2] J. Yuan, Y. Wang, Y. Peng, et al, Weak fault detection and health degradation monitoring using customized standard multiwavelets, *Mech. Syst. Sig. Process.* 94 (2017) 384–399.
- [3] Y. Xu, K. Zhang, C. Ma, et al, Adaptive Kurtogram and its applications in rolling bearing fault diagnosis, *Mech. Syst. Sig. Process.* 130 (2019) 87–107.
- [4] J. Yao, J. Zhao, Y. Deng, et al, Weak fault feature extraction of rotating machinery based on double-window spectrum fusion enhancement, *IEEE Trans. Instrum. Meas.* (2019).
- [5] Y. Xu, Y. Deng, J. Zhao, et al, A novel rolling bearing fault diagnosis method based on empirical wavelet transform and spectral trend, *IEEE Trans. Instrum. Meas.* (2019), <https://doi.org/10.1109/TIM.2019.2928534>.
- [6] S. Wang, X. Chen, G. Cai, et al, Matching demodulation transform and synchrosqueezing in time-frequency analysis, *IEEE Trans. Signal Process.* 62 (1) (2013) 69–84.
- [7] B. Chen, B. Shen, F. Chen, et al, Fault diagnosis method based on integration of RSSD and wavelet transform to rolling bearing, *Measurement* 131 (2019) 400–411.
- [8] S. Hemamalini, Rational-dilation wavelet transform based torque estimation from acoustic signals for fault diagnosis in a three phase induction motor, *IEEE Trans. Ind. Inf.* (2018).
- [9] W.T. Peter, Y.H. Peng, R. Yam, Wavelet analysis and envelope detection for rolling element bearing fault diagnosis—their effectiveness and flexibilities, *J. Vib. Acoust.* 123 (3) (2001) 303–310.
- [10] R.B. Sun, Z.B. Yang, Z. Zhai, et al, Sparse representation based on parametric impulsive dictionary design for bearing fault diagnosis, *Mech. Syst. Sig. Process.* 122 (2019) 737–753.
- [11] G. He, K. Ding, H. Lin, Fault feature extraction of rolling element bearings using sparse representation, *J. Sound Vib.* 366 (2016) 514–527.
- [12] J. Yuan, F. Ji, Y. Gao, et al, Integrated ensemble noise-reconstructed empirical mode decomposition for mechanical fault detection, *Mech. Syst. Sig. Process.* 104 (2018) 323–346.
- [13] Z. Feng, M. Liang, Y. Zhang, et al, Fault diagnosis for wind turbine planetary gearboxes via demodulation analysis based on ensemble empirical mode decomposition and energy separation, *Renewable Energy* 47 (2012) 112–126.
- [14] L. Wang, Z. Liu, Q. Miao, et al, Time-frequency analysis based on ensemble local mean decomposition and fast kurtogram for rotating machinery fault diagnosis, *Mech. Syst. Sig. Process.* 103 (2018) 60–75.
- [15] Dwyer R. Detection of non-Gaussian signals by frequency domain kurtosis estimation[C]//ICASSP'83. IEEE International Conference on Acoustics, Speech, and Signal Processing. IEEE, 1983, 8: 607-610.
- [16] J. Antoni, The spectral kurtosis: a useful tool for characterising non-stationary signals, *Mech. Syst. Sig. Process.* 20 (2) (2006) 282–307.
- [17] J. Antoni, R.B. Randall, The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines, *Mech. Syst. Sig. Process.* 20 (2) (2006) 308–331.
- [18] J. Antoni, Fast computation of the kurtogram for the detection of transient faults, *Mech. Syst. Sig. Process.* 21 (1) (2007) 108–124.
- [19] T. Barszcz, A. Jabłoński, A novel method for the optimal band selection for vibration signal demodulation and comparison with the Kurtogram, *Mech. Syst. Sig. Process.* 25 (1) (2011) 431–451.
- [20] Y. Lei, J. Lin, Z. He, et al, Application of an improved kurtogram method for fault diagnosis of rolling element bearings, *Mech. Syst. Sig. Process.* 25 (5) (2011) 1738–1749.
- [21] B.Q. Chen, Z.S. Zhang, Y.Y. Zi, et al, Detecting of transient vibration signatures using an improved fast spatial-spectral ensemble kurtosis kurtogram and its applications to mechanical signature analysis of short duration data from rotating machinery, *Mech. Syst. Sig. Process.* 40 (1) (2013) 1–37.
- [22] A. Moshrefzadeh, A. Fasana, The Autogram: An effective approach for selecting the optimal demodulation band in rolling element bearings diagnosis, *Mech. Syst. Sig. Process.* 105 (2018) 294–318.
- [23] I.W. Selesnick, R.G. Baraniuk, N.G. Kingsbury, The dual-tree complex wavelet transform, *IEEE Signal Process Mag.* 22 (6) (2005) 123–151.
- [24] Y. Wang, Z. He, Y. Zi, Enhancement of signal denoising and multiple fault signatures detecting in rotating machinery using dual-tree complex wavelet transform, *Mech. Syst. Sig. Process.* 24 (1) (2010) 119–137.
- [25] I.W. Selesnick, The design of approximate Hilbert transform pairs of wavelet bases, *IEEE Trans. Signal Process.* 50 (5) (2002) 1144–1152.
- [26] H. Shi, B. Hu, J.Q. Zhang, A novel scheme for the design of approximate Hilbert transform pairs of orthonormal wavelet bases, *IEEE Trans. Signal Process.* 56 (6) (2008) 2289–2297.
- [27] Kingsbury N. Design of Q-shift complex wavelets for image processing using frequency domain energy minimization[C]//Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429). IEEE, 2003, 1: I-1013.
- [28] I. Bayram, I.W. Selesnick, On the dual-tree complex wavelet packet and M-band transforms, *IEEE Trans. Signal Process.* 56 (6) (2008) 2298–2310.
- [29] Z. Feng, M.J. Zuo, J. Qu, et al, Joint amplitude and frequency demodulation analysis based on local mean decomposition for fault diagnosis of planetary gearboxes, *Mech. Syst. Sig. Process.* 40 (1) (2013) 56–75.