Journal of Sound and Vibration 348 (2015) 381–396

Contents lists available at ScienceDirect

Journal of Sound and Vibration

journal homepage: www.elsevier.com/locate/jsvi

Rotating speed isolation and its application to rolling element bearing fault diagnosis under large speed variation conditions

Yi Wang [a], Guanghua Xu [a,b,*], Qing Zhang [a], Dan Liu [a], Kuosheng Jiang [a]

[a] School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, PR China
[b] State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, PR China

ARTICLE INFO

Article history:
Received 2 August 2014
Received in revised form
13 February 2015
Accepted 9 March 2015
Handling Editor: K. Shin
Available online 26 March 2015

ABSTRACT

During the past decades, the conventional envelope analysis has been one of the main approaches in vibration signal processing. However, the envelope analysis is based on stationary assumption, thus it is not applicable to the fault diagnosis of bearings under rotating speed variation conditions. This constraint limits the bearing diagnosis in industrial applications significantly. In order to extend the conventional diagnosis technique to speed variation cases, a rotating speed isolation method is proposed. This method consists of four main steps: (a) a low-pass filter is used to separate the rotating speed components and the resonance frequency band from the original signal; (b) the trend line of instantaneous rotating frequency (IRF) is extracted by ridge detection from the short-time spectrum of the low-pass filtered signal; (c) the envelope signal is obtained by fast kurtogram based resonance demodulation; (d) the trend line of instantaneous fault characteristic frequency (IFCF) is extracted by ridge detection from the short-time spectrum of the envelope signal; (e) the rotating speed is isolated and the instantaneous fault characteristic order (FCO), which is obtained by simply dividing the IFCF by IRF, can be used to identify the fault type. By rotating speed isolation, the bearing faults under speed variation conditions can be detected without additional tachometers. The effectiveness of the proposed method has been validated by both simulated and experimental bearing vibration signals. The results show that the proposed method outperforms the conventional envelope analysis method and is effective in bearing diagnosis under speed variation conditions.

1. Introduction

Rolling element bearings are critical mechanical components in rotating machinery. Faults occurring in bearings may lead to fatal breakdowns, and such failures can be catastrophic, resulting in costly downtime. Hence, the fault detection of rolling element bearing has attracted considerable attentions in recent years. Vibration signals collected from bearings contain rich information on machine health conditions [1,2]. Therefore, it is possible to obtain vital characteristic information from vibration signals through the use of advanced signal processing techniques due to their intrinsic advantage of revealing bearing failure [3].

* Corresponding author at: School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, PR China. Tel.:/fax: +86 29 8339 5054.
E-mail address: ghxu@mail.xjtu.edu.cn (G. Xu).

http://dx.doi.org/10.1016/j.jsv.2015.03.018
0022-460X/

If the rotating speed is constant, the transient impulses in the waveform of the vibration signal, which are excited by a localized fault, will occur periodically or quasi-periodically. The repetition frequency of the impulses is called fault characteristic frequency (FCF). At the same time, the impacts will excite the resonant frequencies of the bearing and adjacent components, and induce a modulating phenomenon. Therefore, the vibration signal of a defective bearing can be seen as an amplitude-modulated waveform with the resonant frequency as carrier frequency and FCF as modulation frequency. Demodulation of the original signal with envelope analysis can reveal fault-related signatures [4]. In this circumstance, the major challenge in the application of envelope analysis is how to choose the optimal frequency band for demodulation [5]. Several methods have been proposed to address this issue, among which spectral kurtosis was comprehensively investigated by Antoni [6,7] and the fast kurtogram method [8] was proposed. It is shown that spectral kurtosis can be used as a measure to discover the presence of transient components and to adaptively indicate in which frequency band the transient components occur. In recent years, improvements of both spectral kurtosis and the kurtogram have attracted a great deal of attention. Some other methods based on optimal filter frequency band selection have also been proposed, such as improved kurtogram method [2], enhanced kurtogram method [9], protrugram [10], sparsogram [11] and adaptive spectral kurtosis [12]. Therefore the optimal filter frequency band selection methods for demodulation provide a guideline for the optimal demodulation centre frequency and bandwidth selection in the envelope analysis. And bearing diagnosis can be conducted by simply identifying the fault characteristic frequencies (FCFs) in the envelope spectrum.

However, in practice, bearings often work under speed variation conditions. Some representative cases lie in wind turbines, mining equipment and rotating machinery during speed-up and ramp-down processes. In such operation conditions, the repetition frequencies of impulses also vary with time and hence the corresponding envelope signals are non-stationary in nature. Therefore, the direct application of traditional envelope analysis based methods and other techniques, with the assumption of constant rotating speed, will cause spectral smearing and false diagnosis.

Recently, fault diagnosis subject to non-stationary operating conditions has become a focus of many researchers [13–20]. Among those techniques, order tracking [21–23] is one of the most effective methods, which can be used to remove the effects of speed fluctuation and spectrum smearing for fault detection under speed variation conditions. The basic idea of this technique is resampling the original signal at a constant angular increment, which transforms the non-stationary signal (in time domain) to stationary one (in angular domain) [24]. In this way, the spectral smearing caused by speed variation will not be introduced in angular domain. Theoretically, the order tracking methods are straightforward. However these methods need an auxiliary device, such as tachometer or encoder, to provide phase reference information. It will not only increase the measurement cost, but also bring inconvenience in the installation and adjustment. As a result, the order tracking methods are not suitable for all rotating machinery.

Aiming at the shortcomings of the aforementioned traditional methods, a tacho-less method is proposed for bearing diagnosis under speed variation conditions. Generally, the vibration signals from rolling element bearings contain not only the resonance frequency (in high frequency range) excited by impacts, but also the fundamental rotating frequency and its harmonics (in low frequency range). Therefore, the instantaneous fault characteristic frequency (IFCF) can be obtained from the resonance frequency band by resonance demodulation technique (such as fast kurtogram method) incorporating with envelope analysis and the instantaneous rotating frequency (IRF) can be obtained from the low frequency range. The IFCF contains the instantaneous rotating frequency information along with the information of fault characteristic order (FCO-which is constant and determined by the geometry parameters of bearing) related to different fault locations. For bearings under speed variation conditions, it is impossible to use a detected IFCF for fault diagnosis, because it is difficult to relate IFCF to any bearing components, such as inner race, outer race or rolling ball, due to the lack of accurate instantaneous rotating speed information. However, if the instantaneous rotating speed information can be removed from IFCF, the FCO can be uncovered, and the bearing diagnosis under speed variation conditions can be accomplished. On this consideration, a rotating speed isolation method is proposed. The trend line of instantaneous rotating frequency (IRF) is firstly obtained by ridge detection from the short-time spectrum of the low-pass filtered signal. Then, the trend line of instantaneous fault characteristic frequency (IFCF) is extracted by ridge detection from the short-time spectrum of the envelope signal, which obtained by fast kurtogram based resonance demodulation. Finally, the rotating speed can be isolated by simply dividing the IFCF by IRF in the trend lines and the FCO can be uncovered. The fault type can be deduced by the detected FCO. Utilising rotating speed isolation, the proposed method is capable of detecting bearing faults under speed variation conditions without the use of a tachometer.

The outline of this paper is as follows. The proposed rotating speed isolation method is given in Section 2 in detail. And the effectiveness of the proposed method is verified by a simulated signal in Section 3. In Section 4, the proposed method is experimentally validated. Finally, conclusions are drawn in Section 5.

2. The proposed rotating speed isolation method

When we review the commonly used fault features in bearing diagnosis, the fault characteristic frequencies (FCFs) are probably the only kind of feature that is related to and thus varies proportionally with the rotating speed. The FCFs are usually calculated by Eqs. (1)–(4). The outer-race fault frequency f_O, the inner-race fault frequency f_I, the rolling element

fault frequency f_B and the fundamental cage frequency f_C were formulated as follows, respectively [21]:

$$f_O = \frac{n}{2}\left(1 - \frac{d}{D_p}\cos\theta\right) f_r \tag{1}$$

$$f_I = \frac{n}{2}\left(1 + \frac{d}{D_p}\cos\theta\right) f_r \tag{2}$$

$$f_B = \frac{D_p}{2d}\left[1 - \left(\frac{d}{D_p}\cos\theta\right)^2\right] f_r \tag{3}$$

$$f_C = \frac{1}{2}\left(1 - \frac{d}{D_p}\cos\theta\right) f_r \tag{4}$$

where n is the number of rolling balls, d is the diameter of the rolling element, D_p is the groove section size, θ is the contact angle, f_r is the rotating frequency.

The above equations demonstrate that the FCFs contain the instantaneous rotating frequency information along with the FCO information related to different fault locations. For bearings under speed variation conditions, FCFs also vary with time. In this circumstance, it is impossible to use a detected IFCF for fault identification, because it is difficult to relate IFCF to any bearing components, such as inner race, outer race or rolling ball, due to the lack of accurate instantaneous rotating speed information.

However, for a monitored bearing, the FCOs determined by the geometry parameters (n, d, D_p and θ) are all constant. Therefore, if the instantaneous rotating speed information can be removed from IFCF, the FCO can be uncovered, and bearing fault identification under speed variation conditions can be accomplished. Generally, the vibration signals obtained from defective bearings contain both of the resonance frequency in high frequency range excited by impacts and the fundamental rotating frequency and its harmonics in low frequency range. Then the IRF is obtained by ridge detection algorithm from the short-time spectrum of the low-pass filtered signal. As the fault signal is magnified in the resonance frequency band and the resonance frequency does not shift with rotating speed changes, the IFCF is therefore extracted by ridge detection algorithm from the short-time spectrum of the envelope signal demodulated by fast kurtogram, rather than the original signal. On this basis, by simply dividing the IFCF by IRF, the rotating speed can be isolated and the FCO can be uncovered, then the detected FCO is used for bearing fault diagnosis.

The main procedures of the proposed method are as follows. Envelope signal extraction based on fast kurtogram is described in Section 2.1. Section 2.2 introduces the algorithm of IRF and IFCF ridge detection. Rotating speed isolation and bearing fault diagnosis under speed variation conditions are illustrated in Section 2.3. The algorithm of the proposed method is summarized in Section 2.4.

2.1. Resonance demodulation via fast kurtogram

As the resonance frequency band does not shift with rotating speed changes, a constant narrow demodulation frequency band is sufficient to extract the train of impulses [22,26]. Therefore, for bearings operating under speed variation conditions, the fast kurtogram can be used for resonance demodulation, and obtaining the envelope signal with the highest impulsiveness, by optimal frequency band selection. The nature of the fast kurtogram based technique is to exploit the possibility of using the kurtosis as a measure to discover the presence of non-Gaussian components and to indicate in which frequency band the non-Gaussian components occur. Such non-Gaussian components can be considered as the structural response to an excitation, e.g. a local defect on the outer race in the case of rolling bearing faults.

It is well known that spectral kurtosis (SK) is a 4th order time–frequency statistics that can be used to reveal the distribution of the impulsive components of a signal in frequency domain. This makes spectral kurtosis an effective tool for determining the frequency band in which the transients take place even with strong additive noise. Recently, Antoni and Randall [7] proposed a re-defined SK which is more suitable for real world applications. The re-defined SK in [7] is:

$$SK(f_c) = \frac{\left\langle |X(t,f_c)|^4 \right\rangle}{\left\langle |X(t,f_c)|^2 \right\rangle^2} - 2 \tag{5}$$

where $\langle . \rangle$ represents the time–frequency averaging operator, $|X(t,f_c)|$ is the time-frequency envelope of the band-pass filtered signal at a centre frequency f_c, which can be calculated by some time–frequency transforms. Generally, the more impulsive signatures concentrate on a certain frequency band, the higher spectral kurtosis of the filtered signal will be. Therefore, $SK(f_c)$ can be used as a criterion to find the optimal filter frequency band to extract the most impulsive part of the signal in the frequency domain. $|X(t,f_c)|$ can be calculated by several algorithms, like short time Fourier (STFT) [7], wavelet transform (WT) [25] or the fast kurtogram algorithm [8].

In our research, the fast kurtogram algorithm is chosen to calculate $X(t,f_c)$ due to its high computational efficiency. A brief introduction about fast kurtogram is given here. The main idea is to design a tree-like multi-rate filter-bank structure with quasi-analytic band-pass filters, where the filter banks at different levels are composed of binary or 1/3-binary filter

structures. The filtering sequence from each band-pass filter denotes a complex envelope of the analysed signal, which is related to the filtering parameters, the centre frequency f_c and the bandwidth $B=2\Delta f$ of the band-pass filter. The envelope signals extracted from different filters are employed to calculate the spectral kurtosis values. As a result, the envelope signal filtered by the parameter combination $(f_c, \Delta f)$ of the band-pass filter with a maximum spectral kurtosis value is selected for further envelope analysis.

The details of this algorithm are summarized as follows:

(1) Decompose the original signal with binary tree structure and then extend it to 1/3-binary tree structure. This is implemented as follows. Firstly, construct the following quasi-analytic low-pass and high-pass filters based on $h(n)$ which is a low-pass prototype filter with a cut-off frequency $f_c=1/8+\varepsilon, \varepsilon\geq 0$:

$$h_0(n)=h(n)\times 2^{j\pi n/4} \tag{6}$$

$$h_1(n)=h(n)\times 2^{j3\pi n/4} \tag{7}$$

where $h_0(n)$ and $h_1(n)$ represent the low-pass and high-pass filters with normalized frequency band [0,1/4] and [1/4,1/2], respectively.

Then, the above process is repeated in a pyramidal manner to realize the binary-tree-like decomposition. In this way, the signal is decomposed into k levels where each level has 2^k bands. The iteration process between adjacent levels can be expressed by

$$C_{k+1}^{2i}=h_0(m)*C_k^i(2n) \tag{8}$$

$$C_{k+1}^{2i+1}=h_1(m)*C_k^i(2n) \tag{9}$$

where the asterisk $*$ denotes the convolution operation, $C_k^i(n)$ denotes the filtered sequence of original signal by using the ith filter, $i=0,\cdots,2^k-1$, at level k. If $k=0$, let $C_0^i(n)=x(n)$ represents the series of the original signal, $C_k^i(n)$ can be regarded as the complex envelope of signal $x(t)$ based on the centre frequency $f_{ci}=\left(i+2^{-1}\right)*2^{-k-1}$ and bandwidth $\Delta f_k=2^{-k-1}$. $|C_k^i(n)|$ represents the corresponding envelope signal of the band-pass sequences at the centre frequency f_{ci} and bandwidth Δf_k.

Finally, to obtain a finer resolution for narrow band, the 1/3-binary tree filter bank is added based on the existing levels. The procedure of the 1/3-binary decomposition is similar to that of plain binary decomposition, details about the algorithm can be found in [25].

(2) Calculate the $SK(f_{ci},(\Delta f)_k)$, which denotes the SK value of each envelope of the filtered signal $C_k^i(n)$ via Eq. (5).

(3) Determinate the envelope and demodulation frequency band. As aforementioned, SK is used to indicate the impulsiveness of the signal in related frequency band, and then the optimal frequency band can be determined based on the maximum SK. It can be expressed by the following equation:

$$[f_{oc},\Delta f_o,C_o(n)]=\arg\max\{SK(f_{ci},\Delta f_k)\}, \tag{10}$$

where the notion argmax is devoted to obtain the parameters corresponding to maximum SK value. The parameters in Eq. (10) are the optimal centre frequency f_{co}, optimal bandwidth Δf_o and optimal complex envelope $C_o(n)$ respectively.

(4) Select $C_o(n)$ as the final filtered signal highlighting the bearing fault signature.

2.2. IRF and IFCF ridge detection

In this subsection, the short-time Fourier transform (STFT) is applied to obtain the time- frequency representations (TFRs) of the low-pass filtered signal and the envelope signal, and the trend lines of IRF and IFCF are extracted from the TFRs by a ridge detection algorithm.

When bearings work under speed variation conditions, the peak amplitude at FCF may be no longer higher than those at the harmonics of the FCF in the envelope spectrum [26]. If we fix a time moment and examine its corresponding instantaneous envelope spectrum in the time – frequency representation (TFR), the amplitude of FCF and its harmonics may no longer decline monotonically. As a result, the FCF cannot be easily obtained by peak search algorithm from the envelope spectrum of a vibration signal sampled in a short period of time. Therefore, in this paper, STFT is employed to obtain the TFRs of the low-pass filtered signal and envelope signal, and thereby the trend lines of the IRF and the IFCF will be clearly revealed in the TFRs. The trend lines in the TFRs are extracted by ridge detection.

There are several algorithms available for ridge detection [27–30]. Aiming at the shortcomings of the direct maximum ridge detection algorithm which is easily contaminated by background noise, Liu et al. [30] proposed a noise robust ridge detection method, called the cost function ridge detection algorithm. For a given time-frequency matrix, such as the time-frequency distribution $TF(t,f)$ obtained by STFT, the cost function ridge detection algorithm is applied. The main idea of the cost function ridge detection algorithm is to find a path along the time axis to minimize the cost functions, and only the local maxima of each column are concerned to reduce the computational cost. For the TFD of a signal which is an $M\times N$ matrix,

there are $N-1$ cost functions as

$$CF_k = -\left[TF(t,f)\right]^2 + \left|f(k)-f(k-1)\right|^2, k=2,3,\ldots,N \tag{11}$$

On this basis, the n points on the final ridge are determined. By minimizing kth cost function CF_k, the kth ridge point $(k,f(k))$ is selected from the candidate points $\{(k,f(k))\}$ of the kth column. This is conducted after the $(k-1)$th ridge point $(k-1,f(k-1))$ is determined. In this way, the ridge points are determined step by step by scanning from $k=2$ to $k=N$. In the proposed method, the first point $k=1$ is the local maximum value in the approximate range of the beginning of the selected ridge. By applying the cost function ridge detection algorithm, the trend line of the IRF can be detected as IRF_i and the trend line of the IFCF can be extracted as $IFCF_i$ from the TFRs of the low-pass filtered signal and envelope signal respectively.

2.3. Rotating speed isolation and FCO detection for bearing fault diagnosis

In the traditional order tracking based methods, the FCO which is constant and just related to the structural parameters, is obtained by equal-angle resampling combined with envelope analysis and is used to diagnose the bearing fault. Given the FCFs in Eqs. (1)–(4), the corresponding FCOs are defined as the FCFs normalized by the rotating speed [31] and can be obtained by

$$FCO = \frac{60f_c}{N_r} = \frac{f_c}{f_r} \tag{12}$$

where f_c (Hz) represents the recurred frequency (such as the FCFs in Eqs. (1)–(4)) of the impulses excited by the defective component in bearing, N_r(rev/min) denotes the rotating speed of the reference shaft, f_r (Hz) is the rotating frequency.

Use R_r to define the mth harmonic of the extracted ridge of the instantaneous rotating frequency, and employ R_f to represent the nth harmonic of the detected ridge of the instantaneous fault characteristic frequency, and then the fundamental rotating frequency (i.e. IRF_i) and the fundamental fault characteristic frequency (i.e. $IFCF_i$) can be calculated as

$$IRF_i = R_r/m \tag{13}$$

$$IFCF_i = R_f/n \tag{14}$$

In the proposed method, the orders m, n are pre-settled, and the orders of the ridges in the TFR related to rotating frequency or fault characteristic frequency are automatically determined by harmonic relationship investigation. The determination of the orders of the detected ridges is described as follows. Firstly, in our ridge detection algorithm, as the rotating frequency and fault characteristic frequency are always with much higher amplitudes than the background noise or other modulation components, therefore, the beginnings of the ridges for detection are determined by the local maximums of the first slice of the corresponding TFR matrix. The schematic of local maximum detection based harmonic relationship determination is shown in Fig. 1. The local maximum values in the spectrum are indicated by the small circles. The orders of the fundamental frequency and its harmonics are all with integral values. However, for the background noise and other modulation components, the orders are all fractional and do not have the harmonic relationship with each other. Therefore, a preliminary judgment about the harmonic relationship can be made from Fig. 1. Secondly, the ridges which with the above determined beginning are automatically detected by the ridge detection algorithm. Thirdly, the harmonic relationships between the detected ridges are investigated and the orders of the ridges are automatically determined. Finally, the ridge with a pre-settled order is used for rotating speed isolation purpose.

As the $IFCF_i$ extracted from the TFR of the envelope signal contains not only the fluctuation trend of the instantaneous rotating frequency, i.e. IRF_i, but also the FCO information which is constant and determined by the structural parameters related to different fault locations. On this basis, the rotating speed can be isolated and the FCO can be uncovered by simply dividing $IFCF_i$ by IRF_i:

$$FCO = \frac{IFCF_i}{IRF_i} = \frac{mR_f}{nR_r}, i=1,2,\ldots,N \tag{15}$$

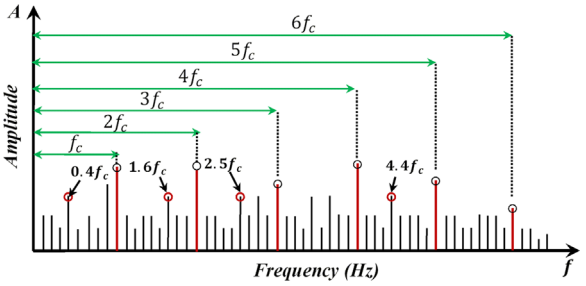

Fig. 1. The schematic of local maximum detection based harmonic relationship determination.

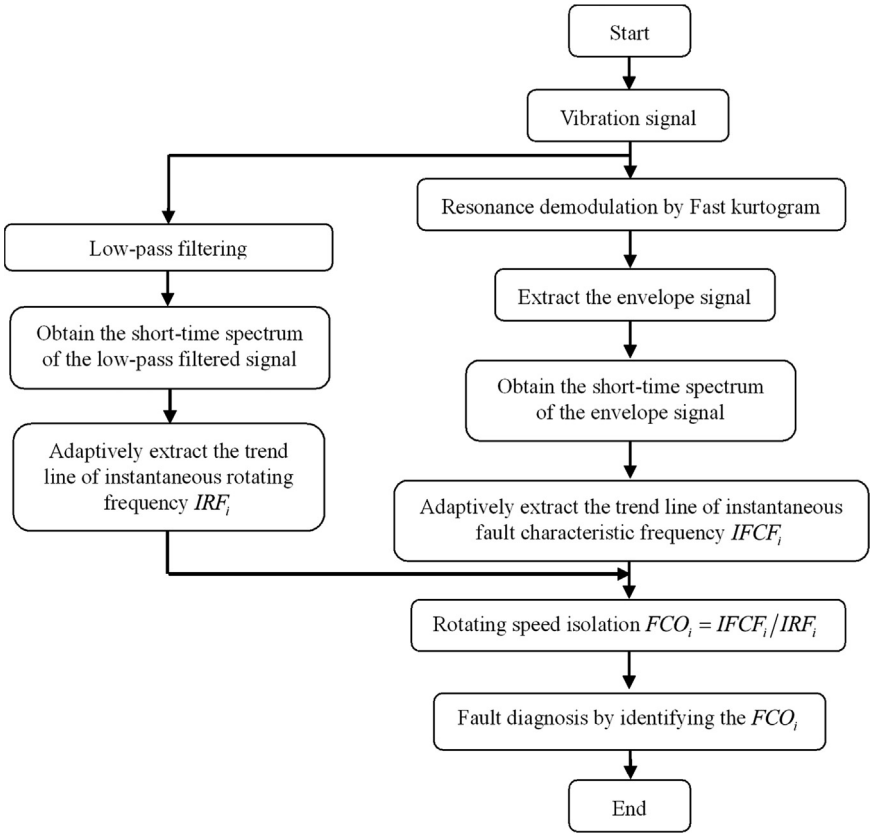

Fig. 2. Flow chart of the proposed rotating speed isolation method.

In this sense, by rotating speed isolation, the proposed method is capable of detecting bearing faults under speed variation conditions without the use of a tachometer. Besides, compared with the traditional order tracking methods, the proposed method does not need the acquisition of the instantaneous reference phase and the complicated equal-angle resampling, therefore, the algorithm is simplified significantly.

2.4. Summary of the proposed method

With the main steps described above, the schematic of the proposed method is shown in Fig. 2 and can be summarized as follows:

(1) The trend line of instantaneous rotating frequency is extracted as IRF_i, by the cost function ridge detection algorithm from the short-time spectrum of the low-pass filtered signal.
(2) The envelope signal $C_o(n)$ is obtained by fast kurtogram based resonance demodulation as described in Section 2.1.
(3) The trend line of instantaneous fault characteristic frequency is obtained as $IFCF_i$, by the cost function ridge detection algorithm from the short-time spectrum of the envelope signal $C_o(n)$.
(4) The rotating speed is isolated and the fault characteristic order (FCO) is obtained by simply dividing $IFCF_i$ by IRF_i, and the FCO is used to identify the fault type.

3. Simulation analysis

To validate the effectiveness of the proposed method, a simulated bearing fault signal under speed variation condition is generated according to the vibration model introduced in [22]

$$x(t) = \sum_i A_i S_i(t) * [1 + aM(t)] + R(t) + wgn(t) \tag{16}$$

The simulated signal $x(t)$ is composed of three terms. The first term denotes a series of impulses excited by fault, where A_i is the amplitude of the ith impulse, $S_i(t)$ represents the ith impulse, $1+aM(t)$ denotes the amplitude modulation term. As demonstrated in [25], the characteristics of the impulses depend on the location of the defect, that is, whether it is on the outer race, inner race or rolling element. For bearing outer race fault, the defect would be located in the load zone, and have

uniform conditions for the passage of each rolling element and the impulses are uniform. However, the bearings with rolling element and inner race faults experience a variation of the load while passing through the load zone. This has the modulation phenomenon on the impulse train. According to [25], for bearing outer race fault, $a=0$, that is, there is no modulation on the impulses train; for the bearing with rolling element fault, $a\neq 0$ and $M(t)$ is the cage speed; for the bearing with inner race fault, $a\neq 0$ and $M(t)$ is the shaft rotating speed. The second term $R(t)$ denotes the fundamental shaft rotating speed signal and its harmonics, which is caused by misalignment, eccentric or imbalance. The formula of the shaft rotating speed can be represented as,

$$R(t)=\sum_{n} B_n \ \sin\left\{2\pi n f(t)+\alpha_n\right\} \tag{17}$$

where B_n and α_n are the amplitude and initial phase of the nth harmonic, $f(t)$ is the instantaneous rotating frequency of the shaft. The third term $wgn(t)$ represents the background noise.

In this simulation, we assume the outer race is kept fixed, and the inner race is rotating with shaft. Suppose there is a local defect on the inner race. In Eq. (16), a is set 5 and $M(t)$ is the rotating speed $R(t)$. The fault order frequency of the inner race fault is $f_{\text{order}}=2.7$ in angular domain, i.e. the average angular period of the impulse is $\theta_0=2\pi/f_{\text{order}}\approx 2.3$ radians. Suppose there are three harmonics in the rotating speed signal, and their amplitudes and initial phases are $\alpha_1=\pi/6$, $\alpha_2=-\pi/3$, $\alpha_3=\pi/2$, $B_1=0.18$, $B_2=0.2$, $B_3=0.18$ respectively. Suppose the bearing experiences a rotating speed fluctuation process with exponential oscillation decay, and the simulated rotating speed waveform is given as bellow:

$$f(t)=\left[2100+1300\cdot e^{-1.2t}\cdot \ \sin(2\pi\cdot 0.8\cdot t)\right]/60 \tag{18}$$

The impulse excited by inner race defect is simulated by an exponentially decaying sinusoid as follows:

$$S_i(t)=e^{-\beta(t-i\times T_i)} \ \sin\left(2\pi f_r t\right) \tag{19}$$

where $\beta=1000$ is the damping coefficient of the system, f_r is the resonance frequency, which is set 3500 Hz here, T_i is the occurrence time of the ith impulse. The occurrence time T_i is determined according to the rotating speed $f(t)$ and the order frequency f_{order}. The scheme of the determination of T_i is demonstrated in Fig. 3.

As the fault order frequency f_{order} indicates how many times the fault induced impulses occur during one revolution of the shaft. Therefore, $\theta_0=2\pi/f_{\text{order}}$ means the recurrence period of the impulse in angular domain. $\theta(t)$ is the instantaneous phase of the rotational shaft and is calculated by the following formula:

$$\theta(t)=2\pi\int_0^t f(t)dt \tag{20}$$

where $f(t)$ is the function of the rotating speed fluctuation process (the unit of $f(t)$ is Hz). As the fault order frequency f_{order} is constant and just related to the structural parameters, therefore, θ_0 is fixed and the impulse occurrence time T_i is calculated according to the time at which the instantaneous phase $\theta(t)$ is integer multiples of θ_0.

The sampling frequency is set 10 k and the time length of the simulated signal is 4 s. The impulse signal and the harmonics of rotating speed are shown in Fig. 4(a) and (b), respectively.

White Gaussian noise is added to obtain a noise-contaminated signal with signal-to-noise ratio of -7 dB and the noisy signal is presented in Fig. 5(a). The impulses generated by fault are seriously corrupted by the background noise. Besides, the impulses are not equally spaced in time domain due to speed variation, which will bring difficulties to the conventional diagnosis methods based on constant speed assumption. To demonstrate this problem, the conventional envelope analysis is applied to the simulated signal, and the corresponding envelope spectrum is shown in Fig. 5(b). As the range of the rotating speed is $f_r=26-51$ Hz i.e. $1560-3060$ rpm, therefore, the range of the fault characteristic frequency is $FCF=f_{\text{order}}*f_r=70-140$ Hz. The fluctuation range of the FCF is depicted by the two dashed red lines in Fig. 5(b). Due to the

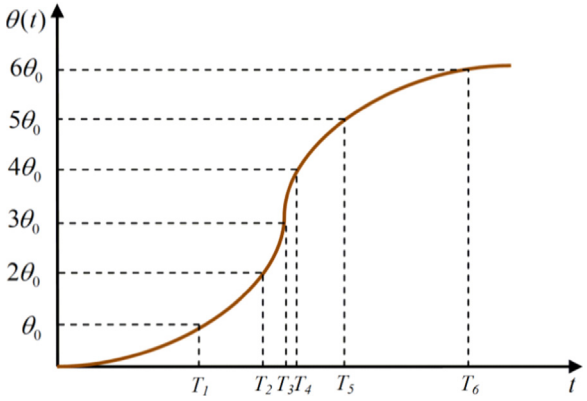

Fig. 3. The scheme of the determination of T_i.

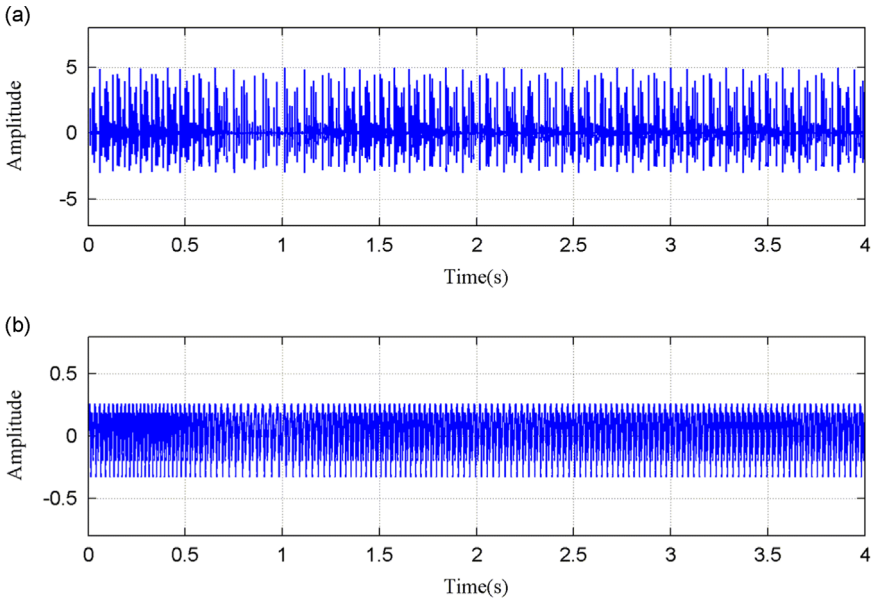

Fig. 4. The simulated signals: (a) the impulse signal under speed variation conditions and (b) the harmonics of rotating speed.

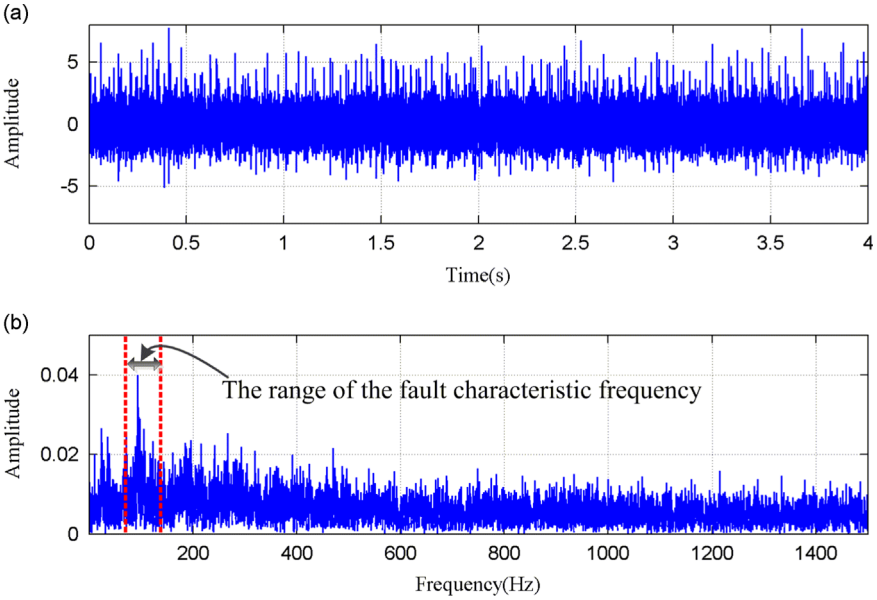

Fig. 5. The noise corrupted signal and its envelope spectrum: (a) The noise contaminated signal and (b) the envelope spectrum of the noisy signal. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

smearing phenomenon caused by the speed variation, there is no fault characteristic frequency in the envelope spectrum. In this circumstance, by direct envelope analysis, the accuracy and reliability of the diagnosis result decreases significantly.

For comparison, the proposed rotating speed isolation method is applied to analyze the simulated signal as presented in Fig. 5(a).

Firstly, fast kurtogram is applied to the simulated signal for resonance demodulation and the envelope signal with the highest impulsiveness can be obtained by optimal frequency band selection. The decomposition level in fast kurtogram is chose to 6. The corresponding kurtogram is demonstrated in Fig. 6, where an optimal filter frequency band with a centre frequency $f_{oc} = 3437$ Hz and a bandwidth $\Delta f_o = 625$ Hz, is indicated by black dashed rectangle. The filtered signal is plotted in Fig. 7(a), in which the transient impulses are successfully uncovered. The corresponding envelope signal is shown in Fig. 7(b).

Secondly, a low-pass filter is used to separate the rotating speed components and the resonance frequency band from the original signal. The short-time Fourier transform is applied to the low-pass filtered signal and the envelope signal which is

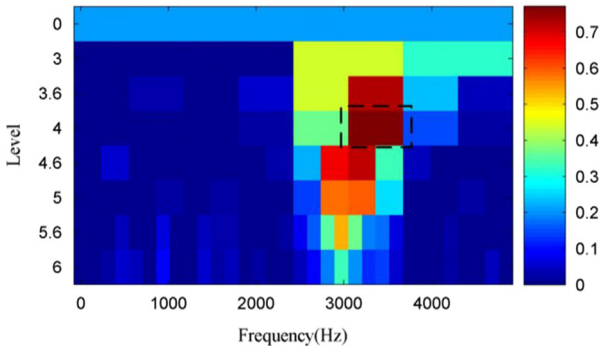

Fig. 6. The kurtogram of the original noise contaminated signal.

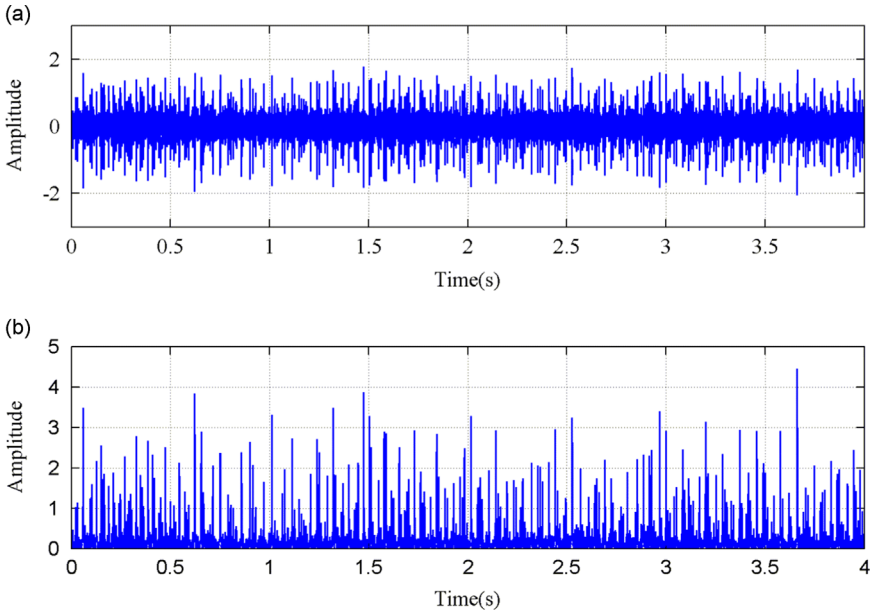

Fig. 7. The optimal frequency band filtered signal and its envelope: (a) the band-pass filtered signal and (b) the envelope signal of (a).

shown in Fig. 7(b). The TFR of the low-pass filtered signal is presented in Fig. 8(a), in which the fluctuation trend of IRF and its harmonics can be observed. The TFR of the envelope signal is depicted in Fig. 8(b) and the variation trend of IFCF and its harmonics are prominent, and there is also an amplitude modulation phenomenon around IFCF and its harmonics. The large fluctuation trend of the IFCF in Fig. 8(b) further illustrates that it is impossible to diagnose the bearing fault by directly using the envelope analysis method. Aiming at the shortcomings of the conventional envelope analysis method, the cost function ridge detection algorithm is applied to the TFRs. The fluctuation trend line of the second harmonic of instantaneous rotating frequency is detected as $R_r$and is shown in Fig. 8(c). Besides, the first harmonic of instantaneous fault characteristic frequency is extracted as R_f and is depicted in Fig. 8(d). It can be observed from Fig. 8(c) and (d) that the detected R_f is proportional to R_r. According to Eqs. (13) and (14), the fundamental instantaneous rotating frequency can be obtained as $IRF_i = R_r/2$ and the fundamental fault characteristic frequency can be represented as $IFCF_i = R_f$.

Finally, as the $IFCF_i$ extracted from the TFR of the envelope signal contains not only the fluctuation trend of the rotating frequency, i.e. IRF_i, but also the FCO information which is constant and fixed by the structural parameters related to different fault locations. On this basis, the rotating speed can be isolated and the FCO can be uncovered by simply dividing the $IFCF_i$ by IRF_i. By rotating speed isolation, the FCO_i can be uncovered as shown in Fig. 9 and the average value is around 2.7, which clearly indicates the existence of a fault on the inner race (an inner race fault with the order frequency equal to 2.7 is first introduced in the simulated signal). For comparison, the actual value of FCO_i is also plotted in the same figure. It is clear that the trend of FCO_i extracted by the proposed method in Fig. 9 agrees well with the actual value, with a slight fluctuation caused by the background noise. To quantify the comparison between the extracted FCO and the actual value, both of the maximum and average relative detection errors are calculated, they are 5.46% and 1.57% respectively. Compared with the conventional envelope analysis method, the proposed method is effective and reliable in bearing fault detection under speed variation conditions.

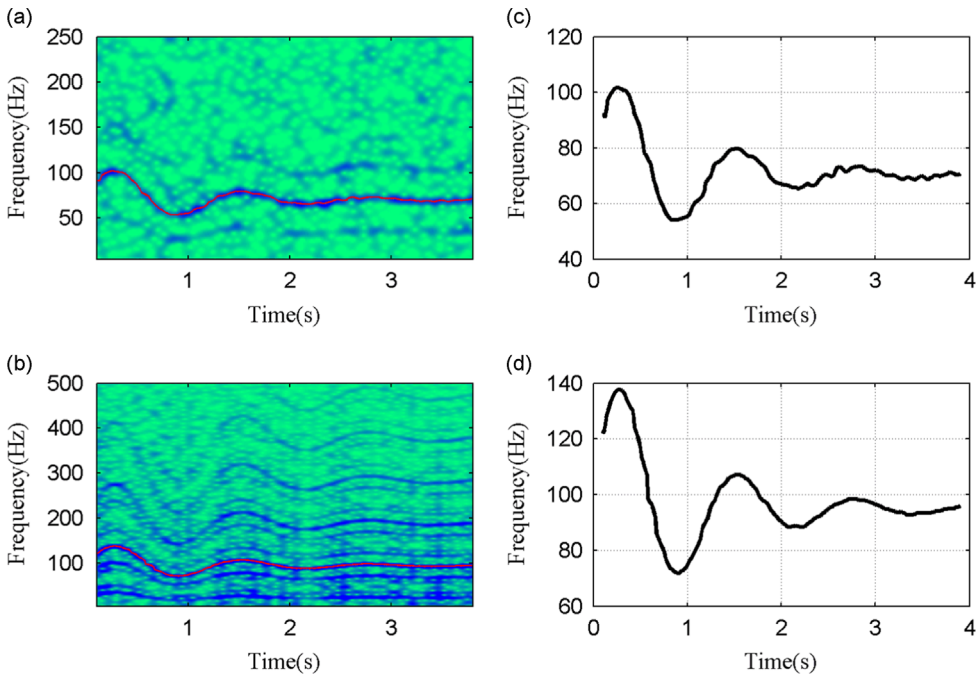

Fig. 8. The short-time spectrums and the ridge detection results: (a) the short-time spectrum of the low-pass filtered signal, (b) the short-time spectrum of the envelope signal, (c) the ridge of the second harmonic of the instantaneous rotating frequency, and (d) the ridge of the first harmonic of the instantaneous fault characteristic frequency.

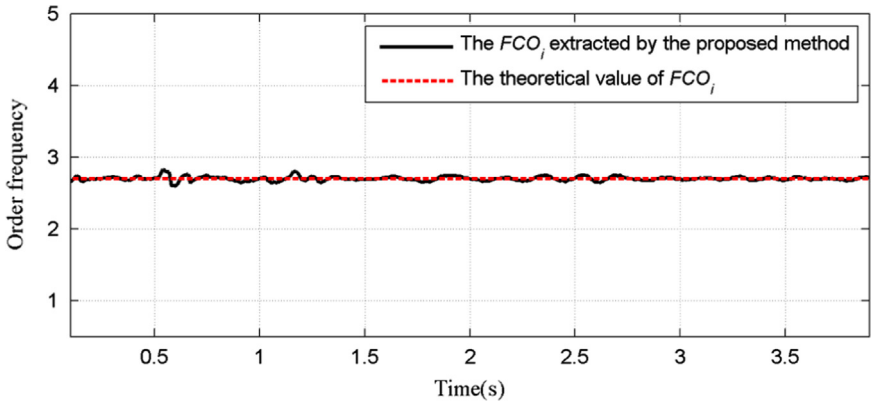

Fig. 9. The FCO_i obtained by rotating speed isolation.

In order to investigate the performance of the proposed method under different signal-to-noise ratios, the background noise with different intensity is added to the simulated signal. Fig. 10 shows the maximum relative error (the red line) and the average relative error (the dashed blue line) in relation to different signal-to-noise ratios. It can be observed from Fig. 10 that when the SNR decreases from -2 dB to -9 dB, both of the maximum and the average relative errors increase, with a slight fluctuation. When the SNR is higher than -8 dB (i.e. -7 dB, -6 dB, etc.), the maximum relative error is less than 6% while the average relative error is less than 1.6%. It indicates that the detection is quite accurate and the proposed algorithm can be regarded as a promising method for bearing fault detection under speed variation conditions.

4. Experimental validation

4.1. Experiment setup of bearing defect test

In this section, a vibration signal collected from a bearing test bench will be used to demonstrate the effectiveness of the proposed method. A general view of the test bench is shown in Fig. 11. It consists of a shaft supported on two ball bearings where one bearing represents the test bearing and the other represents the support bearing. A variable-speed DC motor was used to drive the shaft. The load applied to the test bearing was set to be 250 N ($\pm$ 20 N). The inner ring of the test bearing

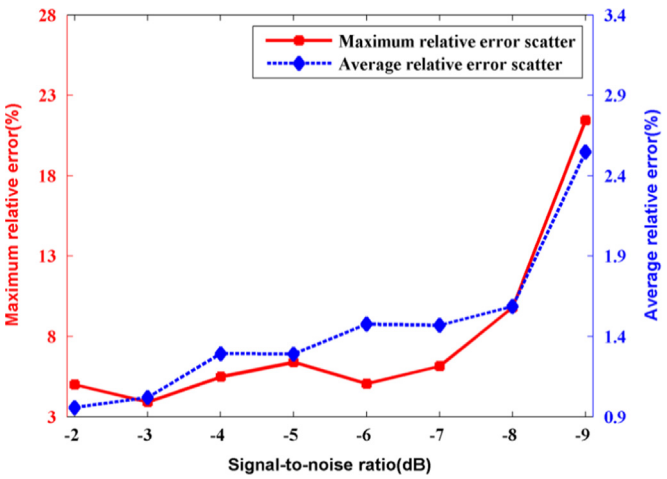

Fig. 10. The maximum and the average relative errors in relation to different signal-to-noise ratios. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

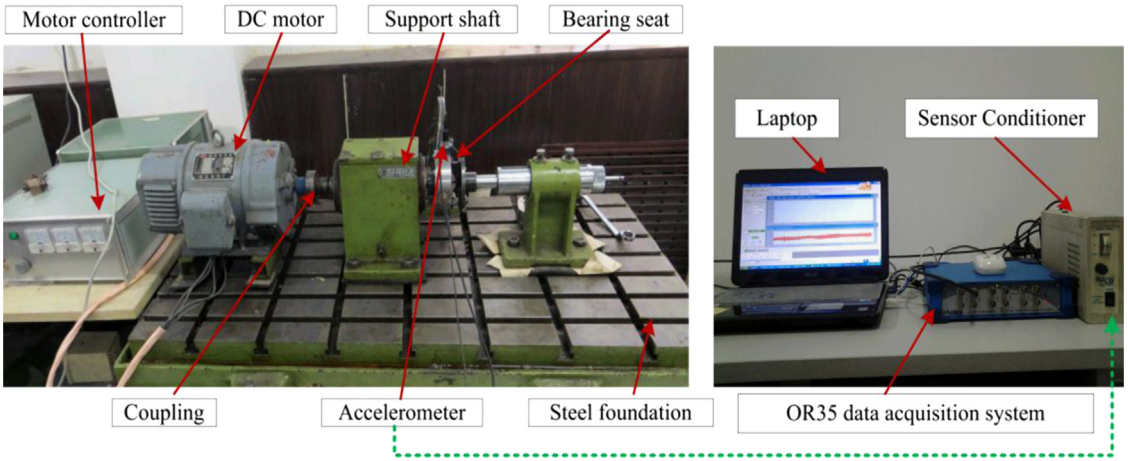

Fig. 11. The arrangement of the bearing test bench.

was free to rotate inside the outer ring. The geometric parameters of the bearing are listed in Table 1. According to those parameters, the corresponding FCOs obtained by normalizing the FCFs defined in Eqs. (1)–(4) by the rotating frequency are listed in Table 2.

The bearing with inner race defect, as depicted in Fig. 12, was installed in the bearing seat. An accelerometer was mounted on the bearing seat to collect the vibration signal generated by the defective bearing. In this experiment, the rotating speed was adjusted manually through the motor controller. The rotating speed experienced a speed-up and ramp-down process. The fluctuation range of the rotating speed was from 1260 RPM to 2520 RPM (the rotating frequency is range in $21-42$ Hz). The process of rotating speed fluctuation was made nonlinear intendedly and then the vibration signal is strongly non-stationary. The vibration signal was pre-processed by the sensor conditioner and subsequently collected by the OR35 data acquisition system. The sampling frequency is 12,800 Hz and the data length is 512,000.

4.2. Inner race defect identification under speed variation conditions

The vibration signal collected from the bearing seat is plotted in Fig. 13(a). In order to assess the health status of the bearing, conventional envelope analysis was first applied to the vibration signal, and the corresponding envelope spectrum is depicted in Fig. 13(b). As the original signal experiences a large speed variation process, as a result, the collected vibration signal from the bearing seat will be with strong non-stationary, which makes the envelope analysis method invalid based on the assumption that the analyzed signal should be stationary or quasi-stationary. The fluctuation range of the FCF is indicated by the two red dashed lines in Fig. 13(b). As discussed previously, the speed variation leads to smearing in the envelope spectrum, as a result, no fault characteristic frequency can be identified. As a consequence, the diagnostic result of

Table 1
The structural parameters of the test bearing.

Bearing designation	Ball numbers	Groove section size (mm)	Contact angle	Diameter of the rolling element (mm)
LYC6308	8	65	0°	15

Table 2
The FCOs obtained by normalizing the FCFs by rotating speed.

Fault location	Outer race	Inner race	Rolling ball	Bearing cage
The FCOs	3.07	4.92	2.05	0.38

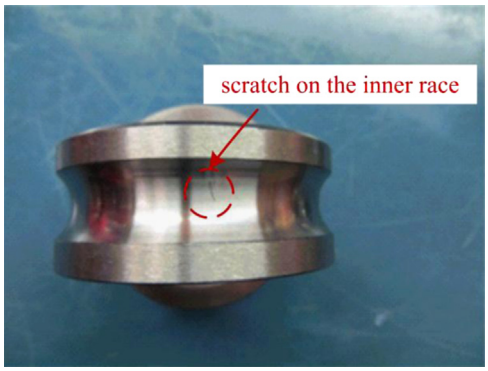

Fig. 12. The test bearing with inner race defect.

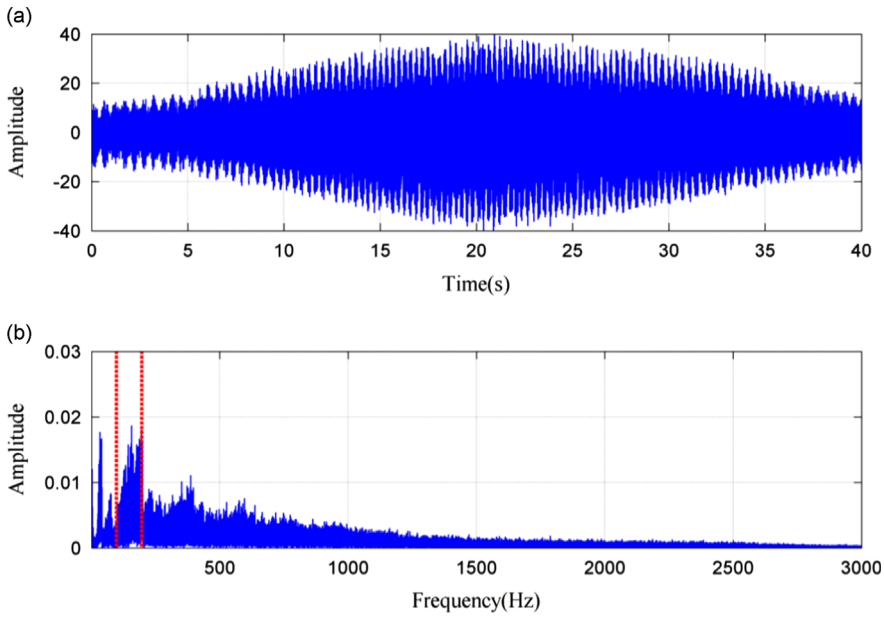

Fig. 13. The original signal and its envelope spectrum: (a) the original bearing vibration signal under speed variation conditions and (b) the envelope analysis result of the original signal.

conventional envelope analysis method is the bearing is in a healthy state, resulting in false diagnosis. Therefore, the proposed rotating speed isolation method is applied to evaluate the condition of the bearing.

Firstly, the resonance frequency band of the collected vibration signal under speed variation condition was investigated. The short-time spectrum of the original signal obtained by STFT is demonstrated in Fig. 14, in which the resonance frequency band with highest amplitude does not shift with the rotating speed changes, it is sufficient to extract the train of

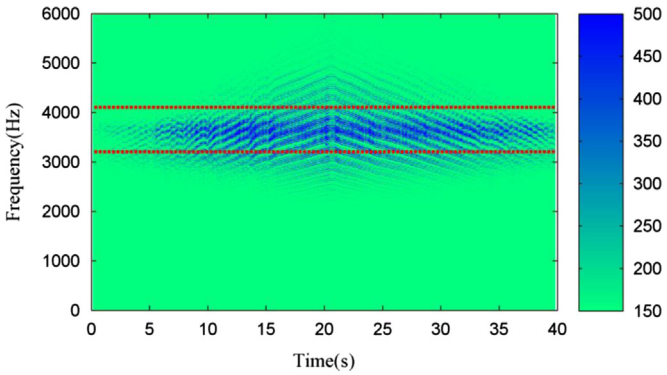

Fig. 14. The short-time spectrum of the original signal.

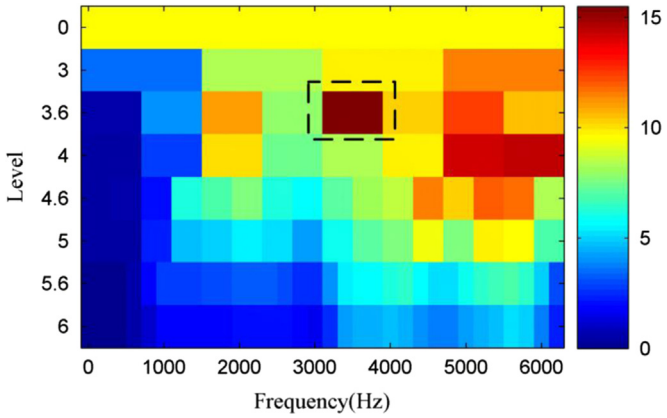

Fig. 15. The kurtogram of the original signal.

impulses with a constant narrow demodulation frequency band, as depicted by the two red dashed lines. It confirms that the band-pass based demodulation methods are applicable in the speed variation cases.

The fast kurtogram method was performed on the original raw signal for optimal frequency-band demodulation. Let the maximum decomposing level be 6, and the spectral kurtosis of the filtered signal in each frequency band at different levels were calculated. The kurtogram of the original signal can be obtained as shown in Fig. 15, where an optimal filter frequency band with a centre frequency $f_{oc} = 3600$ Hz and a bandwidth $\Delta f_o = 800$ Hz, indicated by black dashed rectangle, is automatically chosen. Then, the filtered signal and its corresponding envelope signal can be obtained as shown in Fig. 16(a) and (b) respectively.

Secondly, a low-pass filter was used to separate the rotating speed components and the resonance frequency band from the original signal. STFT was applied to the low-pass filtered signal and the envelope signal shown in Fig. 16(b). The TFR of the low-pass filtered signal is demonstrated in Fig. 17(a), in which the fluctuation trend of IRF and its harmonics are apparent. In Fig. 17(a), it can be observed that the rotating frequency first speeds up from 21 Hz (1260 rpm) to 42 Hz (2520 rpm) and then ramps down to 25 Hz (1500 rpm). The TFR of the envelope signal is depicted in Fig. 17(b) and the variation trend of IFCF and its harmonics are prominent, and there is also an amplitude modulation around IFCF and its harmonics, with the instantaneous rotating frequency as modulation frequency. The modulation phenomenon is similar to that in [25]. In Fig. 17(b), by visual inspection, the instantaneous frequency trend lines in the TFR of the envelope signal are proportional to the rotating frequency. According to Eqs. (1)–(4), a preliminary judgment can be made that there occurred a fault in the tested bearing. In order to locate the fault and confirm the judgement, the trend lines of the *IRF* and the second harmonic of *IFCF* are extracted for further rotating speed isolation analysis. Through the cost function based ridge detection algorithm, the fluctuation trend lines of the instantaneous rotating frequency IRF_i and the second order of instantaneous fault characteristic frequency $2IFCF_i$ are successfully detected and plotted in Fig. 17(c) and (d) respectively.

Finally, since the $IFCF_i$ contains both of the instantaneous rotating speed information and the bearing fault information, and in addition, as shown in Eqs. (1)–(4), the $IFCF_i$ is the product of the instantaneous rotating frequency IRF_i and the instantaneous fault characteristics order FCO_i, therefore, by simply dividing the $IFCF_i$ by IRF_i, the extracted FCO_i can be obtained as presented in Fig. 18. For comparison, the FCOs as demonstrated in Table 2 related to different fault types (such as outer race fault, inner race fault, rolling ball fault and bearing cage fault) of the tested bearing are also plotted in the same figure. The average value of FCO_i is calculated to be 4.89, which is very close to the inner race fault order frequency. Besides,

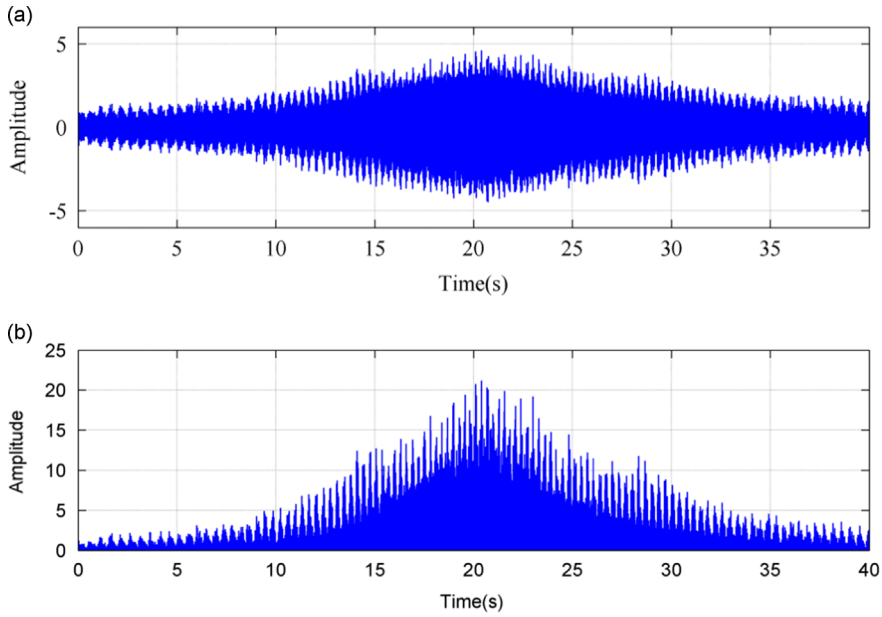

Fig. 16. The optimal frequency band filtered signal and its envelope: (a) the optimal band-pass filtered signal by the fast kurtogram method and (b) the corresponding envelope signal.

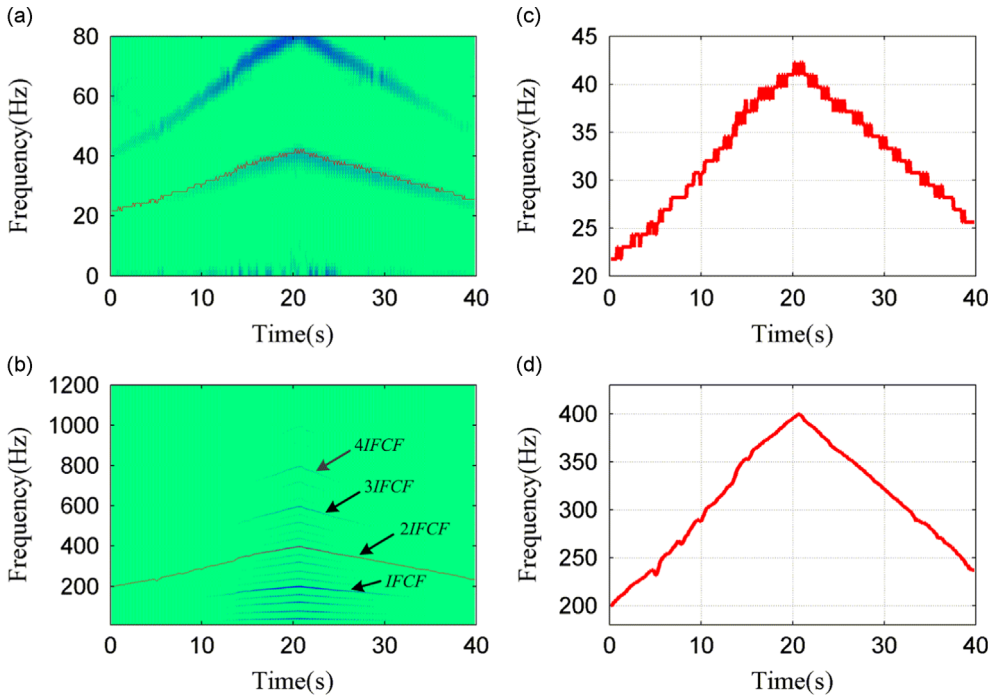

Fig. 17. The short-time spectrums and the ridge detection results: (a) the short-time spectrum of the low-pass filtered signal, (b) the short-time spectrum of the envelope signal, (c) the ridge of the instantaneous rotating frequency, and (d) the ridge of the instantaneous fault characteristic frequency.

by visual inspection, it is clear that the trend line of FCO_i extracted by the rotating speed isolation method provides a satisfactory result and agrees very well with the inner race fault order frequency. The experimental result demonstrates that the inner race defect is successfully detected by the proposed method. Therefore, it can be concluded that the proposed method outperforms the conventional envelope analysis method and can effectively detect the bearing fault under large speed variation conditions.

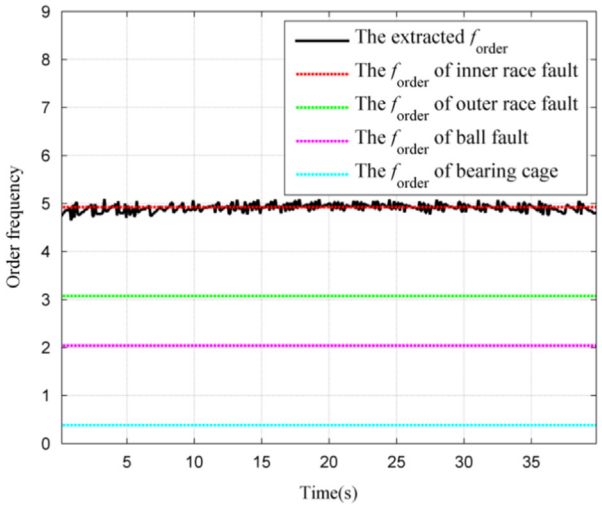

Fig. 18. The f_{order} obtained by the proposed rotating speed isolation method.

5. Conclusions

Aiming at the shortcomings of the conventional envelope analysis method for bearing fault diagnosis under speed variation conditions, a rotating speed isolation method is proposed.

The major innovative features of the proposed method are reflected by (a) the proposed technique is capable of detecting bearing faults under speed variation conditions without the use of a tachometer, thus the measurement cost is reduced and the inconvenience in the auxiliary device installation is avoided; (b) compared with the traditional order tracking methods, the proposed method does not need the acquisition of the instantaneous reference phase and the complicated equal-angle resampling, therefore, the algorithm is simplified significantly; (c) as the cost function ridge detection method is robust to noise, which makes the proposed method reliable and applicable to industry applications. The comparative studies on the simulated and real bearing fault signals indicate that the proposed method outperforms the conventional envelope analysis method. It is shown that the proposed method can effectively detect the bearing faults and provide a promising approach for bearing fault diagnosis under speed variation conditions.

Acknowledgements

This research is supported by the National Natural Science Foundation of China (No. 51105296), which is highly appreciated by the authors. Special thanks should be expressed to the anonymous reviewers for their valuable suggestions.

References

[1] Q.B. He, Vibration signal classification by wavelet packet energy flow manifold learning, *Journal of Sound and Vibration* 332 (2013) 1881–1894.
[2] Y.G. Lei, J. Lin, Z.J He, et al., Application of an improved kurtogram method for fault diagnosis of rolling element bearings, *Mechanical Systems and Signal Processing* 25 (2011) 1738–1749.
[3] H. Qiu, J. Lee, J. Lin, G. Yu, Wavelet filter-based weak signature detection method and its application on roller bearing prognostics, *Journal of Sound and Vibration* 289 (2006) 1066–1090.
[4] P.W. Tse, Y.H. Peng, R. Yan, Wavelet analysis and envelop detection for rolling element bearing fault diagnosis—their effectiveness and Flexibilities, *Journal of Vibration and Acoustics* 123 (2001) 303–310.
[5] Y. Wang, M. Liang, An adaptive SK technique and its application for fault detection of rolling element bearings, *Mechanical Systems and Signal Processing* 25 (2011) 1750–1764.
[6] J. Antoni, The spectral kurtosis: a useful tool for characterising non-stationary signals, *Mechanical Systems and Signal Processing* 20 (2006) 282–307.
[7] J. Antoni, R.B. Randall, The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines, *Mechanical Systems and Signal Processing* 20 (2006) 308–331.
[8] J. Antoni, Fast computation of the kurtogram for the detection of transient faults, *Mechanical Systems and Signal Processing* 21 (2007) 108–124.
[9] D Wang, P.W. Tse, K.L. Tsui, An enhanced kurtogram method for fault diagnosis of rolling element bearings, *Mechanical Systems and Signal Processing* 35 (2013) 176–199.
[10] T. Barszcz, A. JabLonski, A novel method for the optimal band selection for vibration signal demodulation and comparison with the kurtogram, *Mechanical Systems and Signal Processing* 25 (2011) 431–451.
[11] P.W. Tse, D. Wang, The design of a new sparsogram for fast bearing fault diagnosis: Part 1 of the two related manuscripts that have a joint title as "Two automatic vibration-based fault diagnostic methods using the novel sparsity measurement – Parts 1 and 2", *Mechanical Systems and Signal Processing* 40 (2013) 520–544.
[12] Y. Wang, M. Liang, Identification of multiple transient faults based on the adaptive spectral kurtosis method, *Journal of Sound and Vibration* 331 (2012) 470–486.
[13] Z. Li, Z. Wu, Y. He, C. Fulei, Hidden Markov model-based fault diagnostics method in speed-up and speed-down process for rotating machinery, *Mechanical Systems and Signal Processing* 19 (2005) 329–339.

[14] Z. Li, Y. He, F. Chu, J. Han, W. Hao, Fault recognition method for speed-up and speed-down process of rotating machinery based on independent component analysis and factorial hidden markov model, *Journal of Sound and Vibration* 291 (2006) 60–71.
[15] P. Borghesani, P. Pennacchi, S. Chatterton, R. Ricci, The velocity synchronous discrete Fourier transform for order tracking in the field of rotating machinery, *Mechanical Systems and Signal Processing* 44 (2013) 118–133.
[16] W. Bartelmus, R. Zimroz, Vibration condition monitoring of planetary gearbox under varying external load, *Mechanical Systems and Signal Processing* 23 (2009) 246–257.
[17] L. Renaudin, F. Bonnardot, O. Musy, J.B. Doray, et al., Natural roller bearing fault detection by angular measurement of true instantaneous angular speed, *Mechanical Systems and Signal Processing* 24 (2010) 1998–2011.
[18] J. Urbanek, T. Barszcz, J. Antoni, A two-step procedure for estimation of instantaneous rotational speed with large fluctuations, *Mechanical Systems and Signal Processing* 38 (2013) 96–102.
[19] F. Combet, R. Zimroz, A new method for the estimation of the instantaneous speed relative fluctuation in a vibration signal based on the short time scale transform, *Mechanical Systems and Signal Processing* 23 (2009) 1382–1397.
[20] J. Luo, D. Yu, M. Liang, Application of multi-scale chirplet path pursuit and fractional Fourier transform for gear fault detection in speed up and speed down processes, *Journal of Sound and Vibration* 331 (2012) 4971–4986.
[21] R.B. Randall, J. Antoni, Rolling element bearing diagnostics – a tutorial, *Mechanical Systems and Signal Processing* 25 (2011) 485–520.
[22] P. Borghesani, R. Ricci, S. Chatterton, et al., A new procedure for using envelope analysis for rolling element bearing diagnostics in variable operating conditions, *Mechanical Systems and Signal Processing* 38 (2013) 23–35.
[23] P.D. McFadden, M.M. Toozhy, Application of synchronous averaging to vibration monitoring of rolling element bearings, *Mechanical Systems and Signal Processing* 14 (2000) 891–906.
[24] K.R. Fyfe, E.D.S. Munck, Analysis of computed order tracking, *Mechanical Systems and Signal Processing* 11 (1997) 187–205.
[25] N. Sawalhi, R.B. Randall, H. Endo, The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis, *Mechanical Systems and Signal Processing* 21 (2007) 2616–2633.
[26] T.Y. Wang, M. Liang, J.Y. Li, et al., Rolling element bearing fault diagnosis via fault characteristic order (FCO) analysis, *Mechanical Systems and Signal Processing* 45 (2014) 139–153.
[27] R.A. Carmona, W.L. Hwang, B. Torrésani, Characterization of signals by the ridges of their wavelet transforms, *IEEE Transactions on Signal Processing* 45 (1997) 2586–2590.
[28] N. Delprat, B. Escudié, P. Guillemain, et al., Asymptotic wavelet and Gabor analysis: extraction of instantaneous frequencies, *IEEE Transactions on Information Theory* 38 (1992) 644–664.
[29] R.A. Carmona, W.L. Hwang, B. Torrésani, Multiridge detection and time–frequency reconstruction, *IEEE Transactions on Signal Processing* 47 (1999) 480–492.
[30] H. Liu, A.N. Cartwright, C. Basaran, Moiré interferogram phase extraction: a ridge detection algorithm for continuous wavelet transforms, *Applied Optics* 43 (2004) 850–857.
[31] M. Bai, J Huang, M. Hong, F. Su, Fault diagnosis of rotating machinery using an intelligent order tracking system, *Journal of Sound and Vibration* 280 (2005) 699–718.