Available online at www.sciencedirect.com
ScienceDirect
Journal of Sound and Vibration 320 (2009) 1164–1177
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

Morphological undecimated wavelet decomposition for fault diagnostics of rolling element bearings

Rujiang Hao[a,b], Fulei Chu[a,*]

[a]Department of Precision Instruments and Mechanology, Tsinghua University, Beijing 100084, China
[b]School of Computer and Information Engineering, Shijiazhuang Railway Institute, Hebei 050043, China

Received 11 October 2007; received in revised form 13 April 2008; accepted 7 September 2008
Handling Editor: L.G. Tham
Available online 21 October 2008

Abstract

This paper presents a novel morphological undecimated wavelet (MUDW) decomposition scheme for fault diagnostics of rolling element bearings. The MUDW scheme is developed based on the morphological wavelet (MW) theory for both the extraction of impulse features and noise smoothing in signal processing. The analysis operators and the synthesis operator of MUDW strictly satisfy the pyramid condition. The MUDW scheme is used to extract impulse features from rolling element bearing defect signals imposed with noise. The efficiency of the MUDW scheme used for noise smoothing and the extraction of impulse components is evaluated using the simulated data and measured signals from the bearing test rig. Compared with enveloping demodulation analysis, the MW transform and the traditional wavelet transform (WT), the MUDW decomposition scheme is more effective and suitable for the on-line diagnostics of bearings in rotating machines.

1. Introduction

Rolling element bearings are among the most important and frequently encountered components in the vast majority of rotating machines, their carrying capacity and reliability being prominent for the overall machine performance. Therefore, quite naturally, the fault identification of rolling element bearings has been the subject of extensive research.

When a fault in one surface of a bearing strikes another surface, a force impulse is generated which excites resonances in the bearing and the machine. The successive impacts produce a series of impulse responses which may be amplitude modulated as a result of the passage of the fault through the load zone or of the varying transmission path between the impact point and the vibration measurement point. This physical effect has been exploited by several vibration analysis methods [1], based either on detailed models of the vibration response, or on signal processing methods. A comprehensive model for the nature of vibrations induced by the fault in a rolling element bearing, taking the detailed account of imperfections, wear and lubrication, has been proposed in Refs. [2,3]. Several frequency domain signal processing methods have been developed to extract

*Corresponding author. Fax: +8610 6279 2842.
E-mail address: chufl@mail.tsinghua.edu.cn (F. Chu).

0022-460X/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2008.09.014

the useful information contained in the signals from the overall response, the envelope analysis being the most widely accepted one [4]. Improvement in this direction has been recently proposed, based on joint time–frequency domain methods, mainly on wavelet transforms (WTs), Hilbert–Huang transforms, as well as on cyclostationary analysis. This paper considers as a possible alternative, the application of a purely time domain analysis procedure, based on morphological wavelet (MW) decomposition concepts.

Morphological signal processing comprises a broad collection of theoretical concepts and mathematical tools for signal analysis, nonlinear signal operators, design methodologies and application systems that are related to mathematical morphology (MM) [5,6]. Morphological signal processing was firstly used to analyze binary image data and was then extended to gray-level images [7]. The traditional tools of linear systems and Fourier analysis are of limited use for solving geometry-based problems because they do not directly address the issues of how to quantify the shape and the size of the signals. Contrarily, morphological signal processing is perfectly able to quantify all aspects. However, applications of morphological filters in one-dimensional time series have been quite limited, restricted practically to biomedical EEG signals [8,9]. It has been recognized that multi-resolution signal decomposition schemes provide convenient and effective ways to process information. Most of the modern multi-resolution decomposition schemes are based on the theories of pyramid and wavelet, using the convolution and time–frequency domain transformations. However, the linear filtering approaches to multi-resolution signal decomposition have not been theoretically justified. In particular, the operators used for generating various levels of signal components in a pyramid must crucially depend on an application. Therefore, in recent years, a number of researchers have proposed nonlinear multi-resolution signal decomposition schemes based on morphological operators. However, until Goutsias and Heijmans presented a set of fundamental theories named morphological pyramid (MP) and MW, which were derived from traditional wavelet and pyramid theories, there are not a unified standpoint and framework for nonlinear pyramids, filter banks and wavelets, including MPs and wavelets construction [10,11]. MP and MW extend the original wavelet and pyramid from the linear domain to the nonlinear domain. Moreover, they do not require the time–frequency domain analysis.

Based on the MW theory, a multi-resolution signal decomposition scheme, the morphological undecimated wavelet (MUDW) decomposition scheme is presented in this paper. The analysis operators and synthesis operators of the MUDW scheme are constructed according to the morphological coupled wavelet theories. Such a scheme, composed of morphological operators, totally inherits the simple computation property of MM operators. One of the analysis operators in the MUDW scheme is constructed from two parts, one extracts the impulsive components and the other fulfills the noise reduction. Such a construction is efficient for extracting features from the defective bearing signals with noise disturbance. The characteristic frequency of the bearing defect is very obvious by the frequency spectrum analysis of the approximate signals. Compared with the enveloping method, the MUDW decomposition method is more valuable, as demonstrated by the results using the simulated data and measured signals in the rolling element bearing test rig.

This paper is organized as follows. In Section 2, we briefly introduce the concepts of the morphological operators, the MP condition and the morphological coupled wavelet. The construction of the proposed MUDW decomposition scheme is discussed and its properties are analyzed in Section 3. In Section 4, the effects of the MUDW decomposition scheme are examined using simulated impulsive signals with noises and harmonic components. The comparison is made with the current enveloping demodulation method. The proposed procedure is evaluated in Section 5, using three vibration signals from defective rolling element bearings measured in the bearing test rig, which presents an outer race fault, an inner race fault and a rolling element fault, respectively. Finally, Section 6 lays out the conclusive remarks.

2. Basic concepts of MUDW decomposition scheme

In contrast with Fourier transform and wavelet analysis, MM is developed from set theory and integral geometry, and is concerned with the shape of a signal waveform in the complete time domain rather than the frequency domain. MM is a nonlinear approach and has been widely used in the areas of image processing, machine vision and pattern recognition, due to its robustness in preserving the shape while suppressing noise. The mathematical calculation involved in MM includes only addition, subtraction, maximum and minimum operations without any multiplication and division. When acting upon complex shapes of signal,

MM operations are capable of decomposing a signal into meaningful parts and separating them from the background, as well as preserving the main shape characteristics.

2.1. Morphological preliminaries

Dilation and erosion are the two basic operations in MM. By combing these two operators in different orders, many other morphological operations are formed. They can be defined in terms of topology theory.

In topology, a set Z with a partial ordering $\leqslant$ is defined as a complete lattice if every subset K of Z has a supremum (least upper bound) $\vee K$ and an infimum (great lower bound) $\wedge K$ [10]. Let Z and M be two complete lattices, and let $\varepsilon: Z \to M$ and $\delta: M \to Z$ be two operators, one can say that (ε, δ) constitutes an adjunction between Z and M if

$$\delta(y) \leqslant x \Leftrightarrow y \leqslant \varepsilon(x), \quad x \in Z, \ \ y \in M.$$

If (ε, δ) forms an adjunction between Z and M, then ε has the property

$$\varepsilon\left(\underset{i\in I}{\wedge} x_i\right) = \underset{i\in I}{\wedge} \varepsilon(x_i) \tag{1}$$

for any family $\{x_i | i \in I\} \subseteq Z$ of signals. Operator δ has the dual property

$$\delta\left(\underset{i\in I}{\vee} y_i\right) = \underset{i\in I}{\vee} \delta(y_i) \tag{2}$$

for any family $\{y_i | i \in I\} \subseteq M$ of signals. An operator ε that satisfies Eq. (1) is called an erosion, whereas an operator δ that satisfies Eq. (2) is called an dilation. We denote the identity operator on Z or M by id. With every erosion $\varepsilon: Z \to M$, there corresponds a unique dilation $\delta: Z \to M$ such that (ε, δ) constitutes an adjunction. If (ε, δ) is an adjunction between two complete lattices Z and M, then

$$\varepsilon\delta \geqslant \mathrm{id}, \ \ \delta\varepsilon \leqslant \mathrm{id} \quad \text{and} \quad \varepsilon\delta\varepsilon = \varepsilon, \ \ \delta\varepsilon\delta = \delta. \tag{3}$$

If ψ is an operator from a complete lattice Z into itself, then ψ is idempotent, if $\psi^2 = \psi\psi = \psi$. If ψ is increasing and idempotent, then ψ is called a morphological filter. A morphological filter ψ that satisfies $\psi \leqslant \mathrm{id}$ is an opening, whereas a morphological filter ψ that satisfies $\psi \geqslant \mathrm{id}$ is a closing. If (ε, δ) is an adjunction between two complete lattices Z and M, then $\varepsilon\delta$ is a closing on M and $\delta\varepsilon$ is an opening in Z. Given a complete lattice T and a nonempty set E, the set $\mathrm{Fun}(E, T) = T^E$ comprising all functions $x: E \to T$ is a complete lattice under the pointwise ordering

$$x \leqslant y, \quad \text{if } x(p) \leqslant y(p) \text{ and } \forall p \in E,$$

where y is a function in the set $\mathrm{Fun}(E, T)$. Based on the definition of $\mathrm{Fun}(E, T)$, the dilation δ_A and the erosion ε_A are given by

$$\delta_A(x)(n) = (x \oplus A)(n) = \underset{k\in A}{\vee} x(n-k), \tag{4}$$

$$\varepsilon_A(x)(n) = (x \Theta A)(n) = \underset{k\in A}{\wedge} x(n+k), \tag{5}$$

where $\oplus$ is a dilation operator and Θ is an erosion operator; n is an element of the function; A is a given set so-called structuring element (SE) and k is an element of A. Thus, we may conclude that the composition $\alpha_A(x) = \delta_A\varepsilon_A(x) = x \circ A$ is an opening operation, whereas the composition $\beta_A(x) = \varepsilon_A\delta_A(x) = x \bullet A$ is a closing operation. This proposition plays an important role in simplifying the analysis and synthesis operators of the MUDW decomposition scheme.

2.2. MP condition and coupled wavelet

Signal decomposition has been regarded as an important branch in information processing. The time–frequency domain analysis method is widely used in signal processing. It has also been investigated using MM in which the pyramid condition and coupled wavelet have been proposed [10,11]. In Goutsias and

Heijmans' work, they distinguished between two types of multi-resolution decomposition. One is the pyramid scheme and the other is the wavelet scheme. Both of them consist of a number of levels such that the amount of information decreases towards a higher level. Each step toward a higher level is implemented by an analysis operator, whereas each step toward a lower level is implemented by a synthesis operator.

The synthesis and analysis operators must satisfy the pyramid condition which plays an important role in constructing the operators of the decomposition scheme. Consider a family V_j of signal spaces where j is a finite or an infinite index set. Here the signals consist of two families of operators, a family $\psi_j^{\uparrow}$ of analysis operators mapping V_j into V_{j+1} and a family $\psi_j^{\downarrow}$ of synthesis operators mapping V_{j+1} back into V_j. The analysis and synthesis operators $\psi_j^{\uparrow}$, $\psi_j^{\downarrow}$ are said to satisfy the pyramid condition if $\psi_j^{\uparrow}\psi_j^{\downarrow} = \mathrm{id}$ on V_{j+1} where id is an identity operator. The pyramid condition guarantees that no information is lost in two consecutive steps: synthesis and analysis. It is the fundamental principle used to construct pyramid and wavelet operators.

The coupled wavelet is constructed according to the pyramid condition. A coupled wavelet comprises of two analysis operators and one synthesis operator. The analysis operators include a signal analysis operator and a detail analysis operator. Fig. 1 shows the coupled wavelet decomposition scheme. In Fig. 1, V_j and W_j are two sets: V_j is the signal space at level j and W_j is the detail space at level j; $\psi_j^{\uparrow} : V_j \to V_{j+1}$ is the signal analysis operator, $\omega_j^{\uparrow} : V_j \to W_{j+1}$ is the detail analysis operator and $\Psi_j^{\downarrow}$ is the synthesis operator mapping the information back to the lower level. In order to guarantee that no information is lost and the decomposition is non-redundant, the analysis operators $\psi_j^{\uparrow}, \omega_j^{\uparrow}$ and synthesis operator $\Psi_j^{\downarrow}$ of the coupled wavelet must satisfy the pyramid condition below

$$\psi_j^{\uparrow}(\Psi_j^{\downarrow}(x,y)) = x, \quad \text{if } x \in V_{j+1},\ \ y \in W_{j+1}, \tag{6}$$

$$\omega_j^{\uparrow}(\Psi_j^{\downarrow}(x,y)) = y, \quad \text{if } x \in V_{j+1},\ \ y \in W_{j+1}. \tag{7}$$

In order to yield a complete signal representation, the maps $(\psi_j^{\uparrow}, \omega_j^{\uparrow}) : V_j \to V_{j+1} \times W_{j+1}$ and $\Psi_j^{\downarrow} : V_{j+1} \times W_{j+1} \to V_j$ are the inverses of each other. This leads to the following perfect reconstruction condition:

$$\Psi_j^{\downarrow}(\psi_j^{\uparrow}(x), \omega_j^{\uparrow}(x)) = x, \quad \text{if } x \in V_j. \tag{8}$$

Furthermore, given an input signal $x_0 \in V_0$, considering the following recursive analysis scheme:

$$x_0 \to \{x_1, y_1\} \to \{x_2, y_2, y_1\} \to \{x_k, y_k, y_{k-1}, \ldots, y_1\} \to \cdots, \tag{9}$$

where

$$x_{j+1} = \psi_j^{\uparrow}(x_j) \in V_{j+1}, \tag{10}$$

$$y_{j+1} = \omega_j^{\uparrow}(x_j) \in W_{j+1}, \quad j \geqslant 0. \tag{11}$$

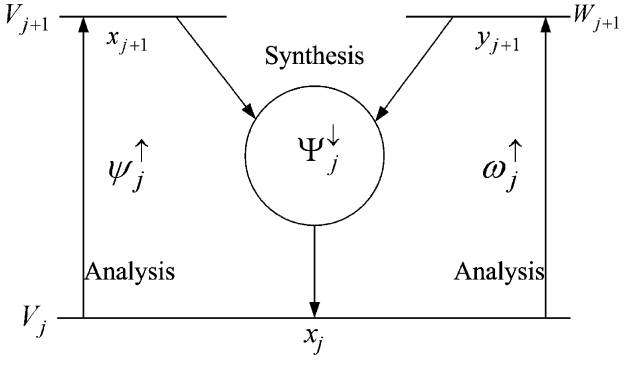

Fig. 1. One stage of coupled wavelet decomposition scheme.

The original signal x_0 can be exactly reconstructed from x_k and $y_1, y_2, \ldots, y_k$ by means of the following recursive synthesis scheme:

$$x_j = \Psi_j^{\downarrow}(x_{j+1}, y_{j+1}), \quad j = k-1, k-2, \ldots, 0, \tag{12}$$

which shows that the decomposition (9) and (12) are invertible. The signal representation scheme governed by Eqs. (6)–(12) is referred to as the coupled wavelet decomposition scheme. The coupled wavelet provides a standard structure for the multi-resolution signal decomposition schemes. Based on this structure, the MUDW scheme is presented by regenerating the analysis operators and synthesis operator.

3. MUDW decomposition scheme based on morphological gradient (MG)

The proposed MUDW decomposition scheme is based on the synthesis and analysis operators. It adopts not only multi-stage and varying-scale coupled wavelet, but also the undecimated algorithm. The undecimated algorithm is based on the idea of no decimation. It applies the WT and omits both down-sampling in the forward and up-sampling in the inverse transform. More precisely, it applies the transform at each point of the signal. In signal processing, this algorithm may give the best results, in terms of high quality filtering, by avoiding information loss and with less distortion that may be caused from noise removal.

The synthesis and analysis operators of the MUDW decomposition scheme are constructed using morphological dilation and morphological erosion, as well as morphological opening and closing, respectively. It is shown that

$$x_{j+1} = \psi_j^{\uparrow}(x_j) = \tfrac{1}{2}[(\gamma\phi + \phi\gamma)(\delta - \varepsilon)](x_j), \tag{13}$$

$$y_{j+1} = \omega_j^{\uparrow}(x_j) = [\mathrm{id} - \tfrac{1}{2}(\gamma\phi + \phi\gamma)(\delta - \varepsilon)](x_j), \tag{14}$$

$$\Psi_j^{\downarrow}(\psi_j^{\uparrow}(x_j), \omega_j^{\uparrow}(x_j)) = \psi_j^{\uparrow}(x_j) + \omega_j^{\uparrow}(x_j) = \mathrm{id}(x_j) = x_j. \tag{15}$$

It is obvious that the MUDW decomposition scheme satisfies the conditions (6)–(8). The signal analysis operator is composed of two parts: $(\gamma\phi + \phi\gamma)$ and $(\delta - \varepsilon)$. The first part is a mixture of morphological opening and closing filters and it can smooth and suppress the noises by selecting multi-scale SEs. The latter part is a MG operator that plays an important role in extracting the periodic impulse features. The way to determine the shape and length of the SEs can be referred in Ref. [12]. Then, the MUDW decomposition scheme not only smoothes the noises but also strengthens the impulsive components in the signals. Considering the idempotence property of morphological opening and closing, the multi-stage and varying-scale MW decomposition procedure was adopted. The approximate signal in the lowest stage of the decomposition was the expected signal that contained useful feature information. Moreover, the length of the approximate signal was the same with that of the original.

4. Simulation analysis

In order to verify the effectiveness of the MUDW decomposition scheme on noise suppression and impulsive feature extraction, a simulated signal is processed using the aforementioned method firstly. The simulated signal is formulated as follows (the sampling frequency is 1024 Hz and the sampling time is 1 s):

$$x(t) = 2x_1(t) + 9x_2(t) + x_3(t), \tag{16}$$

where $x_1(t)$ is the sum of two harmonic waves: $x_1(t) = \sin(2\pi \cdot 30t) + \cos(2\pi \cdot 50t)$; $x_2(t)$ is a typical series of exponentially decaying impulses with the impulse function of $f(t) = \mathrm{e}^{-20t}\sin(20\pi t)$ and the impulse frequency of 20 Hz; $x_3(t)$ is the Gauss white noises: $x_3(t) \propto N(0, 1)$. The purpose of the MUDW decomposition is to extract the impulse components and suppress the harmonic waves and white noises.

The time domain waveform of the simulated signal and its frequency spectrum by means of FFT are shown in Fig. 2(a) and (b). From Fig. 2(a) the impulse components cannot be clearly seen which were mixed with the harmonic waves and Gauss white noises. Fig. 2(b) is the partially enlarged frequency spectrum of the signal in Fig. 2(a), from which the harmonic waves of 30 and 50 Hz are more obvious than the impulse signal of 20 Hz

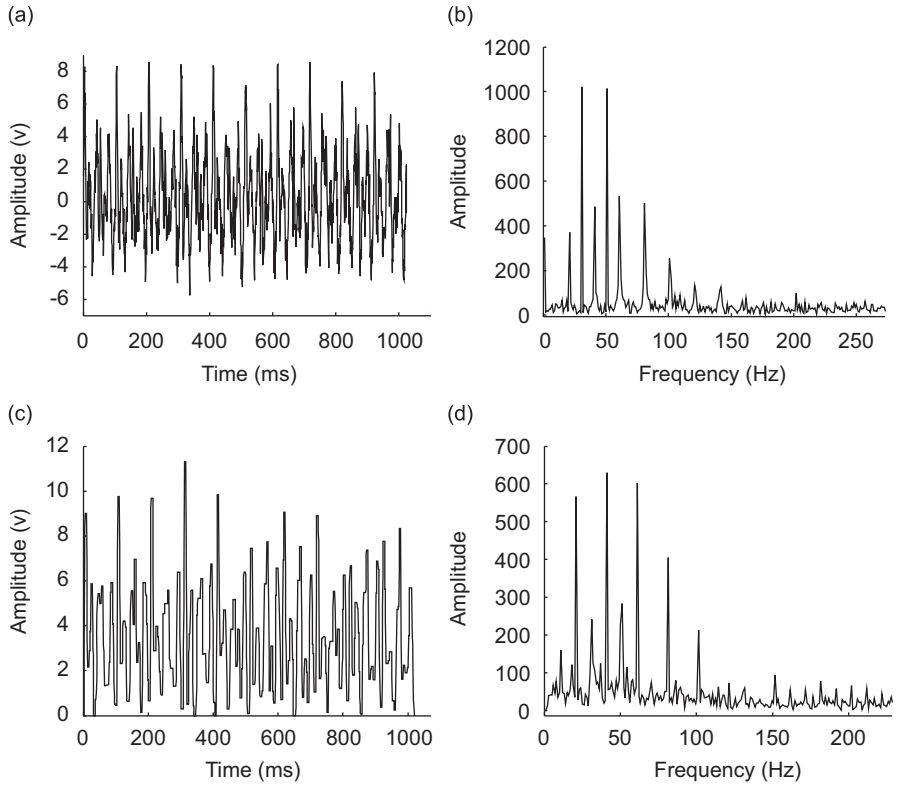

Fig. 2. (a) Time waveform of simulated signal, (b) frequency spectrum of (a), (c) time waveform of (a) with MUDW decomposition and (d) frequency spectrum of (c).

and its high-order harmonic components (40, 60, 80, 100 Hz). From the amplitude of frequency spectrum, it can be seen that the harmonic waves are main components and the periodic impulse signal is emerged in the harmonic waves and noises.

The MUDW decomposition procedure was applied to the simulated signal using three levels of analysis operation with varying scales. The flat SEs were selected, the value of each unit in the SE was zero and the element length was three, five and seven units, respectively. The determination of the SE length depended on distribution of the noise and the frequency of the harmonic waves. The positive impulse component was kept, moreover, the negative impulse and the noise, as well as the harmonic waves were suppressed by using the mixed opening–closing filters and MG operation. Fig. 2(c) is the approximate part of the decomposed signal. The length of the approximate signal is the same with that of the original signal. The frequency spectrum in Fig. 2(d) indicates clearly the effects of the MUDW decomposition. The components of harmonic waves and white noises are suppressed markedly compared with that in Fig. 2(b), the impulse characteristic frequency and its multiples are dominant. It can be seen that the results of the MUDW decomposition are useful to the impulsive feature extraction by designing proper morphological operators and wavelets.

The enveloping demodulation analysis is a widely used method for extracting the periodic impulse signal. In order to compare the effectiveness of the enveloping demodulation with that of the MUDW decomposition, the time waveform and frequency spectrum of the enveloped signal are shown in Fig. 3. The frequency of the impulse signal and its harmonic frequencies, as well as the fractional–multiple frequency (10 Hz) and the harmonic wave frequencies are all in evidence in Fig. 3(b), whereas, the impulse feature is not very clear. Compared with the results in Fig. 2(d), the effect of the enveloping demodulation method is not perfect enough for the simulated signal.

The WT is a traditional time–frequency analysis method. It can reveal the frequency distribution along the time with multi-resolution for a non-stationary signal. In order to extract the impulse characteristics, the Morlet wavelet has been employed for thresholding de-noising and time–frequency characteristic analysis that

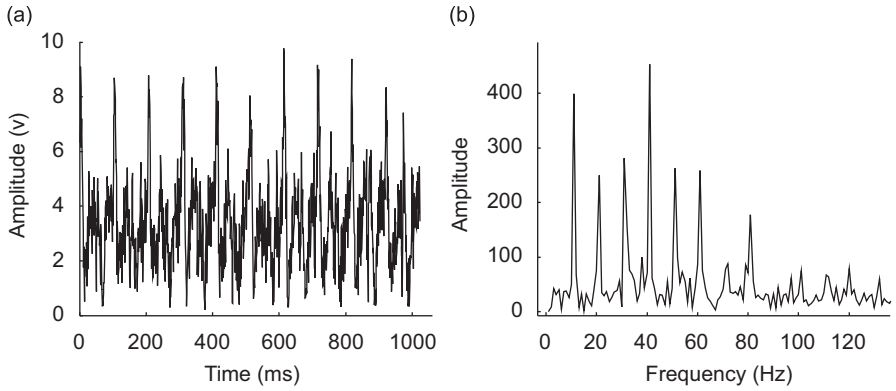

Fig. 3. (a) The enveloping waveform of the simulated signal and (b) the frequency spectrum of (a).

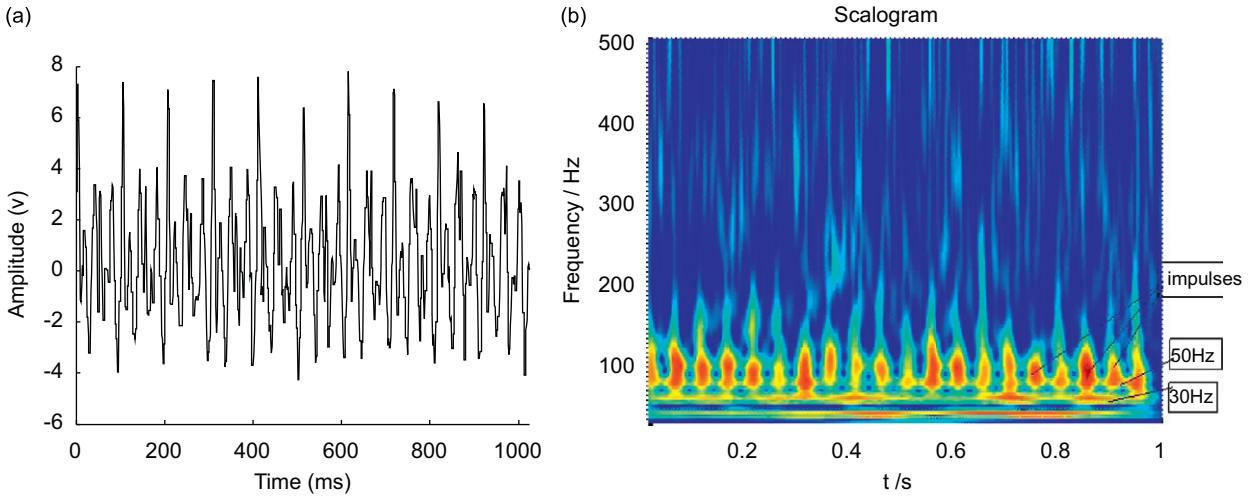

Fig. 4. (a) Time waveform of simulated signal with wavelet transform de-noising and (b) scalogram of (a).

utilizes the similarity between the Morlet wavelet and the impulse [13]. The so-called "sparse code shrinkage" de-noising method proposed by Hyvarinen was used to select the transformed wavelet coefficients, which was based on the maximum likelihood estimation (MLE) technique [14]. By reconstructing the filtered wavelet coefficients, the de-noised signal can then be obtained. Those methods that are based on orthogonal WTs are not suitable for vibration signal analysis from gearbox or rolling bearings because the impulses to be isolated are not smooth. The Morlet wavelet is non-orthogonal. Hyvarinen's thresholding rule with non-orthogonal WTs for impulse detection works better than soft-thresholding or hard-thresholding rule of orthogonal WTs [15]. The de-noised signal based on Morlet WT with MLE thresholding was shown in Fig. 4(a). The impulse components cannot be seen clearly which are mixed by the harmonic waves. Because the impulse frequency (20 Hz) is very close to the harmonic wave frequencies (30 and 50 Hz), the WT filtering just removes the noises and the harmonic waves are not suppressed. From the time domain waveform of the de-noised signal, it can be seen that the effect of the WT filtering is not ideal for the impulse extraction.

The time–frequency distribution can reveal more information about the signal. Fig. 4(b) is the scalogram of the de-noised signal with continuous Morlet transform. The reason for selecting the Morlet wavelet function as mother wavelet is the same as that of the de-noising process for the simulated signal aforementioned. The procedure of continuous Morlet WT is carried out by the function "Scalo_Morl(Sig,SampFreq, TimeSup,N,f0,iShow)" which was compiled by Matlab. From the scalogram, it can be seen that there are 20 impulses within 1 s. This is in accordance with the impulse frequency of 20 Hz. The harmonic

waves (50 and 30 Hz) are also visible in the scalogram which are mixed with the impulse frequencies. The time–frequency analysis based on Morlet WT for the de-noised signal can detect the impulse components in the simulated signal. Since the aim is to extract the impulse characteristics and suppress the other components in the simulated signal, the Morlet WT and the MUDW decomposition are all effective to impulse detection and noise removing. Moreover, the MUDW decomposition can suppress the harmonic waves in the meanwhile. Because the MUDW decomposition is nonlinear WT and its algorithm is simple, the MUDW method is more suitable to the non-stationary signal processing.

The wavelet packet transform is suitable to distinguish the close frequency components in the original signal by selecting the proper decomposition level. However, if the components of the original signal are not known beforehand, the number of the decomposition levels cannot be determined in advance. The WT analysis in this work is only used to remove the noise and show the time–frequency distribution characteristics. Its low-pass and band-pass filtering function is not utilized because the frequency components of the signals are unknown hypothetically.

5. Application on defective rolling element bearings

When a fault in one surface of a rolling element bearing strikes another surface, it produces an impact, which excites natural frequencies of the bearing and of the entire machine. Therefore, the typical response resulting from these periodic impacts, which are produced by bearing faults, usually comprises a sharp rise that corresponds to the impact between the rolling surfaces at the location of the defect and a gradual decay that corresponds to the vibration damping of the bearing outer ring. There are location-dependent characteristic frequencies of the fault which make it possible to identify the nature of a fault, like the outer race defect frequency (ORDF), the inner race defect frequency (IRDF), the rolling element defect frequency (REDF) or the cage train defect frequency (CTDF).

5.1. Data acquisition

The vibration data used in this study were collected from an experiment conducted on a test rig as shown in Fig. 5. The shaft of the test rig was supported by two journal bearings and driven by an AC motor through a V-belt. The test bearing was mounted at the end of the shaft. A two-step loading lever mechanism was employed to apply a load to the test bearing. A B&K 4096 accelerometer and a conditioning amplifier were used to pick up the vibration signals. Three bearings of the model GB6220 were tested. The specifications of these bearings were as follows: number of rolling element, 11; diameter of the rolling element, 22.8 mm; pitch diameter, 149 mm and the contact angle, 0°. The rolling element bearings were seeded with defects using electro-discharge machining (EDM). The defects with three different sizes of 1.0, 2.0 and 1.5 mm in diameter were introduced separately on the rolling element, inner raceway and outer raceway. Vibration signals were picked up from each test configuration at a fixed shaft rotation speed of 1680 rpm and a constant load of 5880 N. The signals were collected at the sampling rate of 10 kHz and the sampling time of 0.5 s. Under the

Fig. 5. Experimental setup for bearing test.

present test condition, the characteristic defect frequencies of the three defective bearings were computed using the formulas:

$$f_{\mathrm{ORDF}} = \frac{z}{2}\left(1 - \frac{d}{D}\cos\alpha\right) f_s, \tag{17}$$

$$f_{\mathrm{IRDF}} = \frac{z}{2}\left(1 + \frac{d}{D}\cos\alpha\right) f_s, \tag{18}$$

$$f_{\mathrm{CTDF}} = \frac{D}{2d}\left[1 - \left(\frac{d}{D}\right)^2 \cos^2\alpha\right] f_s, \tag{19}$$

where z is the number of rolling element, d the diameter of the rolling element, D the pitch diameter, α the contact angle and f_s the rotating frequency of the shaft. The corresponding characteristic defect frequencies were: IRDF = 158 Hz, REDF = 137 Hz, ORDF = 105 Hz and CTDF = 12 Hz.

5.2. Analysis of the MUDW decomposition for defective vibration signals

Fig. 6(a), (c) and (e) show the vibration signals with an inner race defect, a rolling element defect and an outer race defect, respectively, which were measured at the test rig. From the time domain waveforms, different defect characteristics cannot be distinguished easily. According to the sampling rate and the defect characteristic frequency, the MUDW decomposition aforementioned was applied to the signals with four stages of decomposition and the length of the flat SEs were 3, 6, 9 and 12 units. The frequency spectrums of the approximate signals of the MUDW decomposition were shown in Fig. 6(b), (d) and (f). It can be seen that the defect characteristic frequencies and their multiples are obvious. The values of the characteristic frequencies shown in Fig. 6 are close to the calculated IRDF, REDF and ORDF. Theoretically, the shaft rotating frequency f_s and the sidebands with the interval of f_s must be detected in Fig. 6(b) and the cage train rotating frequency f_c and the sidebands with interval of f_c must be detected in Fig. 6(d). However, the amplitude of the modulated signal caused by the damage and its sidebands cannot be detected properly in Fig. 6(b) and (d). The inconsistent phenomenon arises from the bearing test rig. Considering the assembly errors, the bearing house was not driven vertically by the drag bar of the lever mechanism. Therefore the tested bearing (6220 deep groove ball bearing) was acted by not only the radial load but also the axial load. For the inner race defect and rolling element defect, the impact region caused by the damage was not just the top of the ring but around the whole ring. So the f_s, f_c and sidebands were not clear in the frequency spectrum. If the bearing house was acted by only the radial load, this drawback would be overcome and the results would be better.

The kurtosis value analysis is a very reliable time domain diagnostic technique for the impulse detection in the signals [13]. For the defective signals with MUDW transform, the kurtosis values were calculated by the kurtosis function in Matlab toolbox. They were 7.1472 for defective signals of inner race, 6.3596 for defective signals of rolling element and 5.1835 for defective signals of outer race, respectively. For the normal signal with no defects in the bearing, the kurtosis value was 2.8792 by using the same method. It can be seen that the kurtosis value is a good index to judge whether or not the defect is existent in the bearing. Moreover, by means of selecting many measured data as the learning samples, the kurtosis value can be used to bearing defect classification if it is combined with the artificial neural network (ANN) or support vector machine (SVM) algorithm. However, the accurate classification criterion based on kurtosis value for different types of defect cannot be established. It depends on the regional distribution of the kurtosis value. Whereas, the bearing defect characteristic frequency (IRDF, REDF and ORDF) is an exact value of reference which can be calculated by the shaft speed and the bearing structure specificity in advance. By using the FFT analysis of the MUDW transformed signal, the different types of defect can be obviously identified corresponding to the IRDF, REDF and ORDF. The FFT analysis based on defect characteristic frequency is an effective method for defect diagnosis, but its identification process is more complicated than that of kurtosis value analysis.

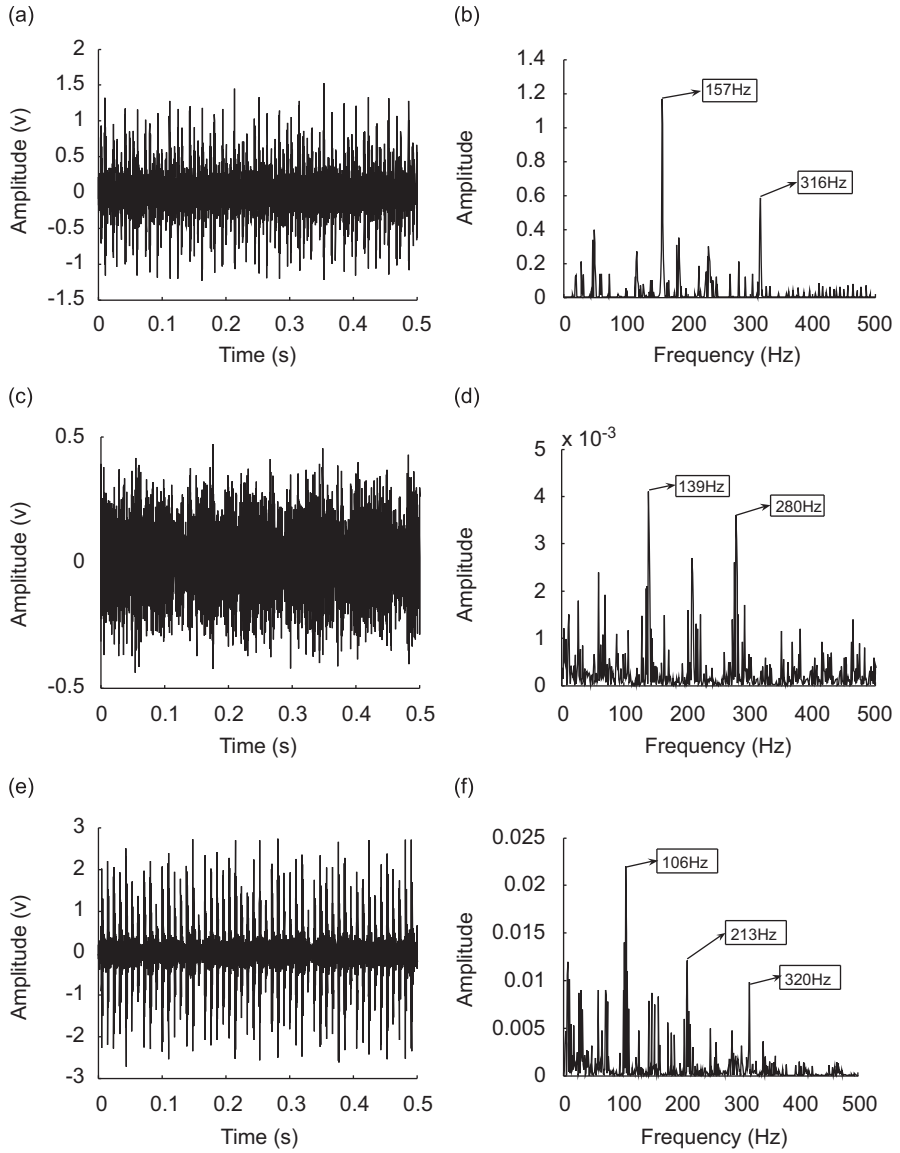

Fig. 6. The original vibration signals and local frequency spectral diagrams of defective rolling element bearings using the MUDW decomposition scheme: (a) the signal with an inner race defect, (b) the spectral diagram of (a) under the MUDW decomposition, (c) the signal with a rolling element defect, (d) the spectral diagram of (c) under the MUDW decomposition, (e) the signal with an outer race defect and (f) the spectral diagram of (e) under the MUDW decomposition.

5.3. Enveloping analysis for the defective vibration signals

In order to compare the enveloping demodulation analysis with the MUDW decomposition scheme, the vibration signals were preprocessed by a low-pass filter with 2.5 kHz and a band-pass filter with 800–2000 Hz according to the calculated natural frequency of the outer race. The enveloping spectrums were shown in Fig. 7(a), (b) and (c). Compared with the frequency spectrums in Fig. 6(b), (d) and (f), the performance of the enveloping demodulation method is not ideal enough, especially for the inner race and rolling element defects. The enveloping demodulation analysis procedure includes the low-pass filtering and band-pass filtering and the selection of the frequency band is determined by the natural frequency of the bearing parts, which is different for every part of the bearing, therefore the enveloping demodulation method is greatly affected by the man-made errors. Whereas, the MUDW decomposition method only includes addition, subtraction and

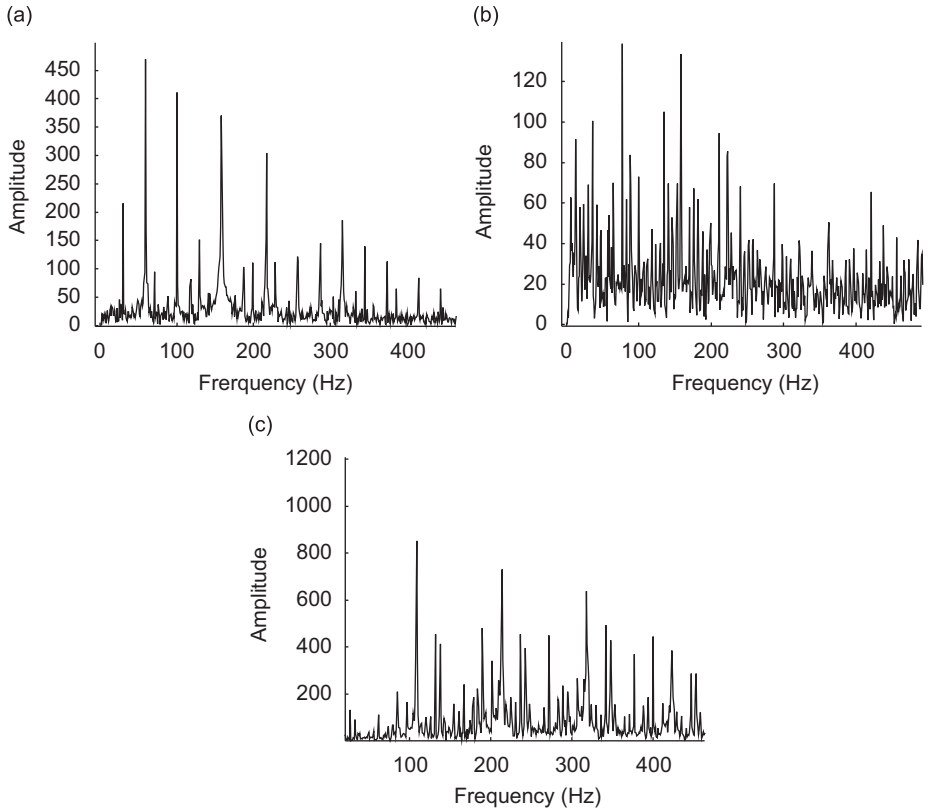

Fig. 7. The enveloping spectral diagram of the original vibration signals: (a) the inner race defect, (b) the rolling element defect and (c) the outer race defect.

extreme value operations, the algorithm is not affected by the properties of the bearing parts and it not only smoothes the noises but also extracts the impulse features in the signals. The simplicity and validity of the MUDW decomposition is preponderant for the diagnostics of rolling element bearing defects, especially for real-time processing in the on-line diagnostics.

5.4. Morphological decimated wavelet (MDW) transform for the defective vibration signals

The MDW transform was widely used in the image processing. In one-dimensional signal analysis, the MDW transform is a two-based decomposition method, and the decomposed signal length is decreased by 2^n. The major difference with the classical linear wavelet is that the linear signal analysis filter of the latter is replaced by an erosion (or dilation), i.e., by taking the minimum (or maximum) over two samples. The typical morphological Haar wavelet decomposition is defined as follows:

$$\psi^{\uparrow}(x)(n) = x(2n) \wedge x(2n+1), \tag{20}$$

$$\omega^{\uparrow}(x)(n) = x(2n) - x(2n+1), \tag{21}$$

$$\psi^{\downarrow}(x)(2n) = \psi^{\downarrow}(x)(2n+1) = x(n), \tag{22}$$

$$\omega^{\downarrow}(y)(2n) = y(n) \vee 0, \quad \omega^{\downarrow}(y)(2n+1) = -(y(n) \wedge 0). \tag{23}$$

Here "$\wedge$" denotes minimum and "$\vee$" denotes maximum, the analysis and synthesis operators satisfy the pyramid conditions. The detail parts $\omega^{\uparrow}(x)(n)$ contain the signal information with different scales. By reconstructing the detail parts using the formulas (22) and (23), the noises are removed and the edge

information is preserved. The MDW transform was applied to the vibration signals of the defective bearings. The frequency spectrums of the reconstructed signals are shown in Fig. 8(a), (b) and (c).

Compared with the analysis results in Fig. 6(b), (d) and (f), the defect characteristic frequency and its multiples of the MDW transform are not as obvious as those of the MUDW decomposition for the defective bearing signals. According to the property of analysis operators of the morphological Haar WT, the noises were removed in the reconstruction procedure, whereas, the impulse components reflecting the defect characteristics were not emphasized. It can be seen that the MDW transform does not as perfectly fit for the impulse feature extraction as the MUDW decomposition does.

5.5. Traditional WT for the defective vibration signals

The same methods as those for the simulated signal aforementioned are used to analyze the measured data. After the Morlet wavelet de-nosing processing, the scalograms are obtained and shown in Fig. 9. Among the three plots, the periodic impulse characteristics caused by the outer race defect are the most obvious. The frequency of the impulses is equal to the ORDF. The impulses caused by inner race defect are not very clear and the periodicity characteristics are not as ideal as those of the outer race defect signal. The impulses caused by the rolling element defect are blurry and the impulse frequency corresponding to REDF cannot be found at all. Because the components of the measured signal are very complicated, which comprise the noises, periodic impulses, the natural vibration of different parts of the bearing with high frequencies and other vibration components from the machine, the impulses caused by the defects are mixed with the other non-stationary signals from the test rig. Even though the noises are removed greatly, the results of the impulse extraction are not very good, especially for the rolling element defect and the inner race defect. Compared with MUDW decomposition and the enveloping demodulation analysis, the effect of Morlet WT is not so good.

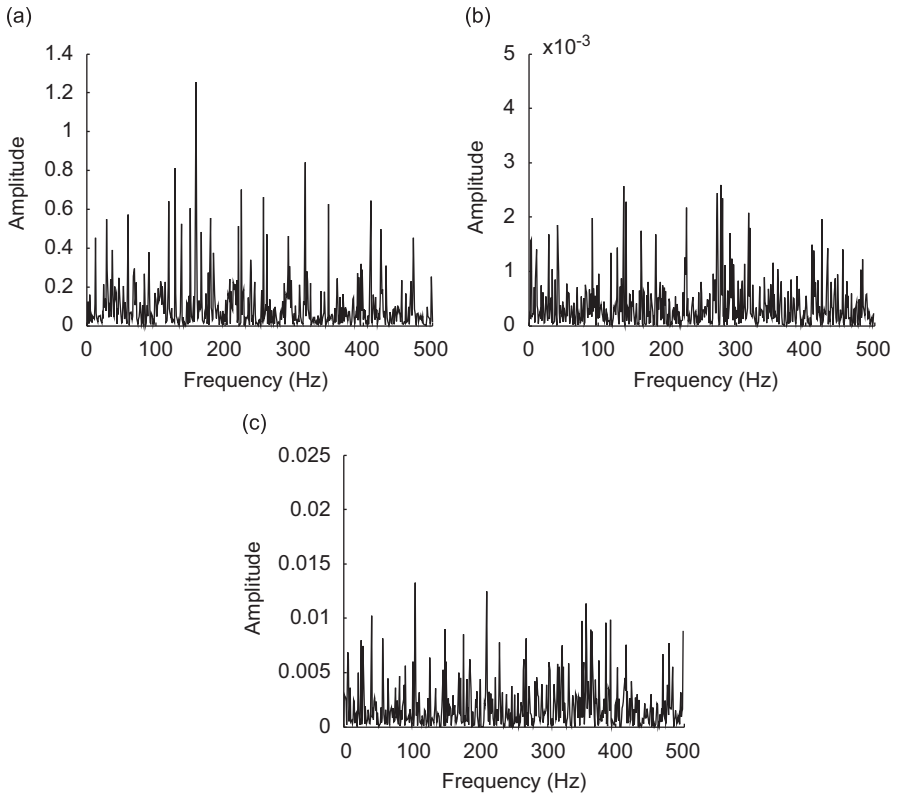

Fig. 8. Frequency spectral diagram of the vibration signals with the MDW transform: (a) the inner race defect, (b) the rolling element defect and (c) the outer race defect.

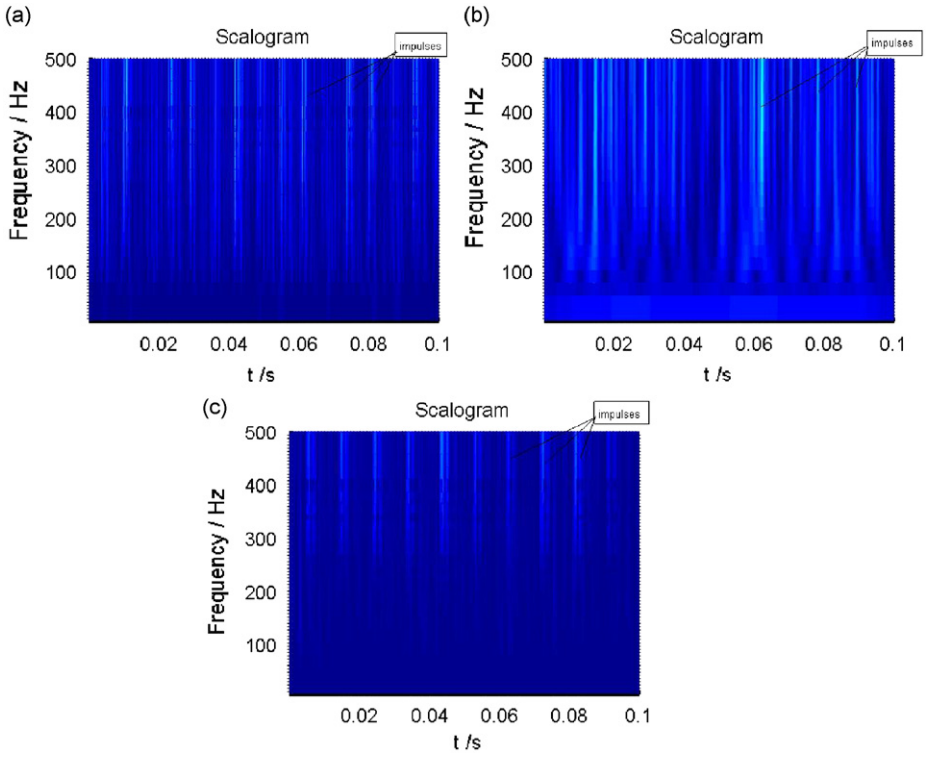

Fig. 9. Scalogram of the vibration signals with WT: (a) the inner race defect, (b) the rolling element defect and (c) the outer race defect.

6. Conclusions

In this paper, a MM-based MUDW decomposition scheme has been proposed to effectively smooth noise and extract the impulse components in the vibration signals of defective rolling element bearings. The development of its analysis operator has been discussed in detail and the multi-stage and varying-scale MUDW decomposition procedure has been analyzed. The efficiency of the MUDW decomposition scheme has been evaluated in simulation studies and the experimental signals measured in the bearing test rig. Compared with the enveloping demodulation analysis method, the MDW transform and the traditional WT, the results show that the MUDW decomposition scheme performs satisfactorily and is able to extract the impulse feature of defective rolling element bearing signals. Because of the simplicity of the MUDW decomposition algorithm, this method is suitable for the on-line diagnostics of rolling element bearings in rotating machines.

Acknowledgments

This research is supported by Natural Science Foundation of China (Grant No. 50425516 and 10732060) and the "863" High-Tech Scheme (2006AA04Z438) by the Ministry of Science and Technology. Comments and suggestions from the referees and editors are very much appreciated.

References

[1] N. Tandon, A. Choudhury, A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings, *Tribology International* 32 (1999) 469–480.

[2] Y. Su, M.H. Lin, M.S. Lee, The effects of surface irregularities on roller bearing vibrations, *Journal of Sound and Vibration* 165 (1993) 455–466.

[3] A. Choudhury, N. Tandon, A theoretical model to predict vibration response of rolling bearings to distributed defects under radial load, *ASME Transactions, Journal of Vibration and Acoustics* 120 (1998) 214–220.
[4] D. Ho, R.B. Randall, Optimization of bearing diagnostic techniques using simulated and actual bearing fault signals, *Mechanical Systems and Signal Processing* 14 (2000) 763–788.
[5] P. Maragos, *Morphological Signal and Image Processing. The Digital Signal Processing Handbook*, CRC Press, Boca Raton, FL, 1998, pp. 74.1–74.26.
[6] P. Maragos, R. Schafer, Morphological filters—Part I: their set-theoretic analysis and relations to linear shift invariant filters, *IEEE Transactions on Acoustics, Speech and Signal Processing* 35 (1987) 1153–1169.
[7] F. Meyer, P. Maragos, Nonlinear scale space representation with morphological levelings, *Journal of Visual Communication and Image Representation* 11 (2000) 245–265.
[8] S. Nishida, M. Nakamura, M. Miyazaki, S. Suwazono, M. Honda, Construction of a morphological filter for detecting an event related potential P300 in signal sweep EEG record in children, *Medical Engineering Physics* 17 (1995) 425–430.
[9] S. Nishida, M. Nakamura, A. Ikeda, H. Shibasaki, Signal separation of background EEG and spike by using morphological filter, *Medical Engineering Physics* 21 (1999) 601–608.
[10] J. Goutsias, H.J.A.M. Heijmans, Multiresolution signal decomposition schemes—part 1: linear and morphological pyramids, Technical Report PNA-R9810, CWI, Amsterdam, The Netherlands, October 1998.
[11] H.J.A.M. Heijmans, J. Goutsias, Multiresolution signal decomposition schemes—part 2: morphological wavelets, Technical Report PNA-R9905, CWI, Amsterdam, The Netherlands, July 1999.
[12] N.G. Nikolaou, I.A. Antoniadis, Application of morphological operators as envelope extractors for impulsive-type periodic signals, *Mechanical Systems and Signal Processing* 17 (2003) 1147–1162.
[13] W.Q. Wang, F. Ismail, M.F. Golnaraghi, Assessment of gear damage monitoring techniques using vibration measurements, *Mechanical Systems and Signal Processing* 15 (5) (2001) 905–922.
[14] A. Hyvarinen, Sparse code shrinkage: de-noising of nonGaussian data by maximum likelihood estimation, *Neural Computation* 11 (7) (1999) 1739–1768.
[15] J. Lin, M.J. Zuo, K.R. Fyfe, Mechanical fault detection based on the wavelet de-noising technique, *ASME Transactions, Journal of Vibration and Acoustics* 126 (2004) 9–16.