Journal of Sound and Vibration (2000) **234**(1), 135–148
doi:10.1006/jsvi.2000.2864, available online at http://www.idealibrary.com on IDEAL®

FEATURE EXTRACTION BASED ON MORLET WAVELET AND ITS APPLICATION FOR MECHANICAL FAULT DIAGNOSIS

JING LIN

State Key Laboratory of Acoustics, Institute of Acoustics, Chinese Academy of Science, Beijing, 100080, *People's Republic of China*

AND

LIANGSHENG QU

Research Institute of Diagnostics & Cybernetics, School of Mechanical Engineering, Xi'an Jiaotong University, Shaanxi Province, People's Republic of China

(*Received* 14 *July* 1999, *and in final form* 14 *January* 2000)

The vibration signals of a machine always carry the dynamic information of the machine. These signals are very useful for the feature extraction and fault diagnosis. However, in many cases, because these signals have very low signal-to-noise ratio (SNR), to extract feature components becomes difficult and the applicability of information drops down. Wavelet analysis in an effective tool for signal processing and feature extraction. In this paper, a denoising method based on wavelet analysis is applied to feature extraction for mechanical vibration signals. This method is an advanced version of the famous "soft-thresholding denoising method" proposed by Donoho and Johnstone. Based on the Morlet wavelet, the time-frequency resolution can be adapted to different signals of interest. In this paper, this denoising method is introduced in detail. The results of the application in rolling bearing diagnosis and gear-box diagnosis are satisfactory.

1. INTRODUCTION

Vibration signals are always used for mechanical fault diagnosis, because they carry the dynamic information of the machines. However, these vibration signals sampled on the spot often contain a lot of noise. If the noise is too heavy, the useful information will be corrupted such that the working state cannot be recognized or even wrong conclusions will be drawn. Therefore, effective methods for feature extraction from these noisy signals should be used.

Wavelet analysis is one such powerful tool. It is especially suitable for non-stational signal processing. In 1990s, it has been successfully used in signal processing, such as image coding, compressing and edge detection. In the field of mechanical fault diagnosis, wavelet analysis has been used in rolling bearing diagnosis [1, 2], gear-box diagnosis [2, 3] and compressor diagnosis [4]. Wavelet has also been used for feature extraction and noise purging, such as *matching pursuits* developed by Mallat and his collaborators [5, 6], and *soft-thresholding denoising* developed by Donoho and Jonestone [7, 8]. The *soft-thresholding denoising* employs threshold in the wavelet domain and it can be shown to be asymptotically near optimal for many signals corrupted by additive white Gaussian noise. However, for many mechanical dynamic signals, the feature components are composed of

0022-460X/00/260135 + 14 $35.00/0

impulse components. But from some examples illustrated in section 3, this method is not effective for impulse component extraction.

To make up for the deficiency of Donoho's *soft-thresholding denoising*, a new denoising method based on Morlet wavelet is proposed in this paper, especially applicable to impulse components extraction. One application in mechanical fault diagnosis is introduced in this paper. The remaining parts are organized as follows. In section 2, the concept of wavelet is introduced and the review of wavelet transform is given. In section 3, the calculation for continuous wavelet transform of Morlet wavelet is given, which includes the discretization of scale parameter a and translation parameter b, and the optimization of parameter β that controls the time–frequency resolution of Morlet wavelet. A denoising method based on continous wavelet transform of Morlet wavelet is also established. In section 4, the denoising method is applied to rolling bearing diagnosis and gear-box diagnosis. The conclusion of this paper is given in section 5.

2. REVIEW OF WAVELET TRANSFORM

Wavelet transforms are inner products of the signal and a family of the wavelets. Let $\psi(t)$ be the mother wavelet or the wavelet "prototype". The corresponding family of wavelets consists of a series of son wavelets, which are generated by dilation and translation from the mother wavelet $\psi(t)$ shown as follows:

$$\psi_{a,b}(t) = |a|^{1/2}\psi\left(\frac{t-b}{a}\right), \tag{1}$$

where a is scale factor and b is time location: the factor $|a|^{-1/2}$ is used to ensure energy preservation.

The wavelet transform of signal $x(t)$ is defined as the inner product in the Hilbert space of the L^2 norm as follows:

$$W(a,b) = \langle \psi_{a,b}(t), x(t) \rangle = |a|^{-1/2} \int x(t)\psi^*_{a,b}\,\mathrm{d}t. \tag{2}$$

Here the asterisk stands for complex conjugate. Time parameter a and scale parameter b vary continuously. Mother wavelet $\psi(t)$ is assumed to lie in $L^2(C)$ and satisfies the admissibility condition

$$C_\psi = \int_{-\infty}^{\infty} |\hat{\psi}(\omega)|^2/|\omega|\,\mathrm{d}\omega < \infty, \tag{3}$$

where $L^2(C)$ is the space of square integrable complex functions, and

$$\hat{\psi}(\omega) = \int \psi(t)\exp(-\mathrm{j}\omega t)\,\mathrm{d}t. \tag{4}$$

The wavelet transform $W(a, b)$ can be considered as functions of translation b with each fixed scale a. $W(a,b)$ gives the information of $x(t)$ at different levels of resolution and also measures the similarity between the signal $x(t)$ and each son wavelet $\psi_{a,b}(t)$ ($W(a,b)$ is the convolution between $x(t)$ and the wavelet function). This implies that a wavelet can be used for feature discovery if the wavelet used is similar to the feature components hidden in the signal.

The wavelet transform introduced above is also called continuous wavelet transform (CWT). Before the calculation, the relevant parameters must be discretized for being computed by a computer. Dyadic discretization is the most popular method. In this method, $a = 2^j$, $b = k2^j$, $j, k \in Z$. It has fast algorithms, of which the famous Mallat algorithm is an example. Although this method saves a lot of computation time, it has three disadvantages making it unsuitable for feature extraction. First, it demands that the wavelet must be orthogonal. This restriction makes it rather difficult to find a proper wavelet for feature extraction. Second, the sampling grids in the time-scale plane are rather sparse. Usually, feature components cannot be separated from the irrelevant components by these sparse grids. Third, time invariant is very important for feature detection, while this algorithm does not meet requirement [9]. Thus, the dyadic discrete wavelet transform is not suitable for feature extraction. In the following section, CWT is introduced as a better method.

3. EXPLOITATION OF THE FEATURE EXTRACTION USING CWT

The scale parameter a and translation parameter b of CWT vary continuously. For convenience of calculation, they have to be discretized firstly. The details are given as follows.

3.1. PARAMETERS OPTIMIZATION

As stated in section 2, the basic wavelet in CWT may not be orthogonal, any one satisfying the admissible condition can be used as a basic wavelet. Thus, to find a proper wavelet function for feature extraction is simple. In this paper, Morlet wavelet is used, because in many mechanical dynamical signals, impulses are always the symptoms of faults and the Morlet wavelet is very similar to impulse component. A Morlet wavelet is defined as

$$\psi(t) = \exp(-\beta^2 t^2/2)\cos(\pi t). \tag{5}$$

By dilation with a and translation with b, a son wavelet can be acquired. As shown in equation (6)

$$\psi_{a,b}(t) = \exp\left[-\frac{\beta^2(t-b)^2}{a^2}\right]\cos\left[\frac{\pi(t-b)}{a}\right]. \tag{6}$$

Obviously, it is a cosine signal decaying exponentially on both sides. Figure 1 illustrates the shape of a Morlet wavelet, which appears as an impulse.

There are three parameters in a son Morlet wavelet altogether: a, b and β. We shall discuss the selection of a and b first. Different definition domains of a and b correspond to different segmentations of the time-scale plane of CWT. For a given digital signal, the sampling rate always follows the Nyquist sampling theory, generally the sampling rate can be considered high enough. Then it will have enough time resolution if the translation unit is equal to the sampling period. That is, the length of the time-scale grid along the time axis is equal to the sampling period. The length of the time-scale grid along the scale axis depends on the distribution and resolution of the feature components. Hence it may be different in various cases.

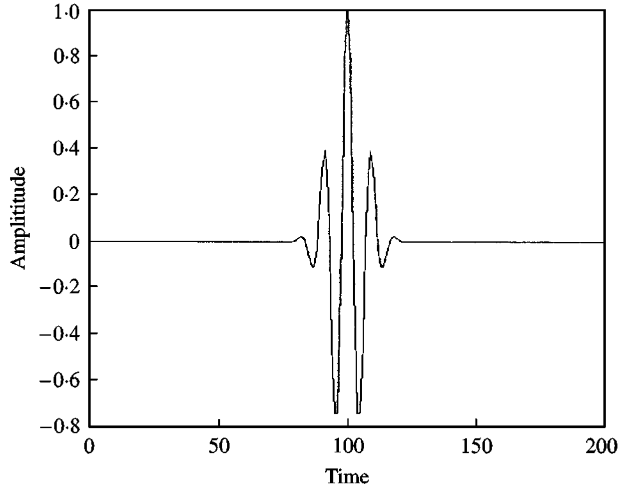

Figure 1. The shape of Morlet wavelet.

Then, there is only one parameter β left. From equation (5), it can be seen that parameter β controls the shape of the basic wavelet. Parameter β balances the time resolution and the frequency resolution of the Morlet wavelet. Decreasing β will increase the frequency resolution but it decreases the time resolution. When β tends to 0, the Morlet wavelet becomes a cosine function which has the finest frequency resolution, and when β tends to infinity, the Morlet wavelet becomes a Dirac function which has the finest time resolution. So, there always exists an optimal β that has the best time–frequency resolution for a certain signal localized in the time-frequency plane.

As for the wavelet base optimization, "sparsity" is usually used as the rule for evaluating the wavelet base. This means that the wavelet corresponding to the fewest wavelet transformation coefficients of a signal is the best. Therefore, the value of β can be determined according to whose wavelet coefficients are the sparsest. As we know, the diversity of a possibility series can be measured with Shannon Entropy. Thus, sparsity of wavelet coefficients may be measured with the entropy of those wavelet coefficients. The entropy here is termed *wavelet entropy*. Assume that $\{c_i\}_{i=1,\ldots,M}$ is the class of the coefficients, which must be normalized first. When it is divided by $\sum_{i=1}^{M} c_i$, $\{d_i\}_{i=1,\ldots,M}$ is obtained:

$$d_j = c_j \Big/ \sum_{i=1}^{M} c_i. \tag{7}$$

The wavelet entropy is calculated by

$$En = -\sum_{i=1}^{M} d_i \log \mathrm{d}_i. \tag{8}$$

A simulated signal is given to verify the validity of the wavelet entropy. The formula of the signal is defined as

$$f(t) = \exp[-(t-400)^2/200] \cos[\pi(t-400)/5]$$

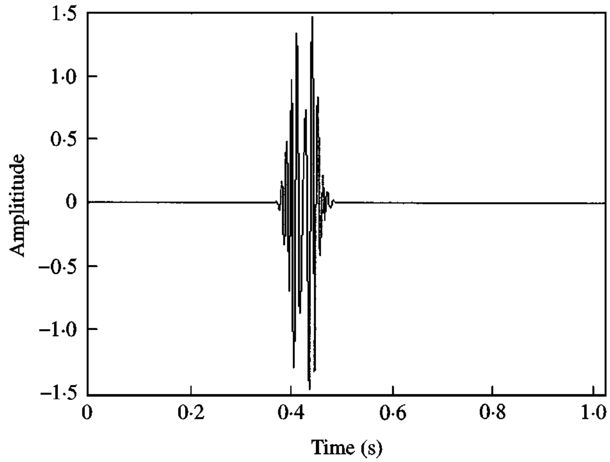

Figure 2. The waveform of the simulated signal.

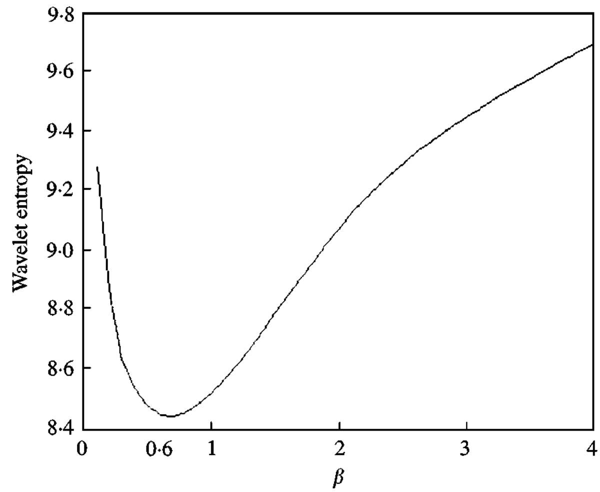

Figure 3. The relation curve between β and the wavelet entropy.

$$+ \exp[-(t-425)^2/600]\cos[\pi(t-425)/7{\cdot}2]$$

$$+ \exp[-(t-440)^2/400]\cos[\pi(t-440)/5{\cdot}3]. \qquad (9)$$

The sampling rate of t is 1 and the first 1000 points are illustrated in Figure 2. Increasing β from 0·1 to 20 and calculating the wavelet entropy of the coefficients, the relationship between β and wavelet entropy is obtained, as shown in Figure 3. There exists a minimal value of wavelet entropy when $\beta = 0{\cdot}6$, then 0·6 is the optimal value of β.

The CWT of the simulated signal is calculated, respectively, taking β as 0·2, 0·6 and 1·2, as shown in Figures 4(a)–4(c). From formula (9) it can be seen that the signal includes three close impulse components. For the sake of application in this paper, the CWT corresponding to the optimal value of β should separate the three parts effectively. From the Figures 4(a)–4(c), it can be seen that the CWT corresponding to $\beta = 0{\cdot}6$ divides the signal

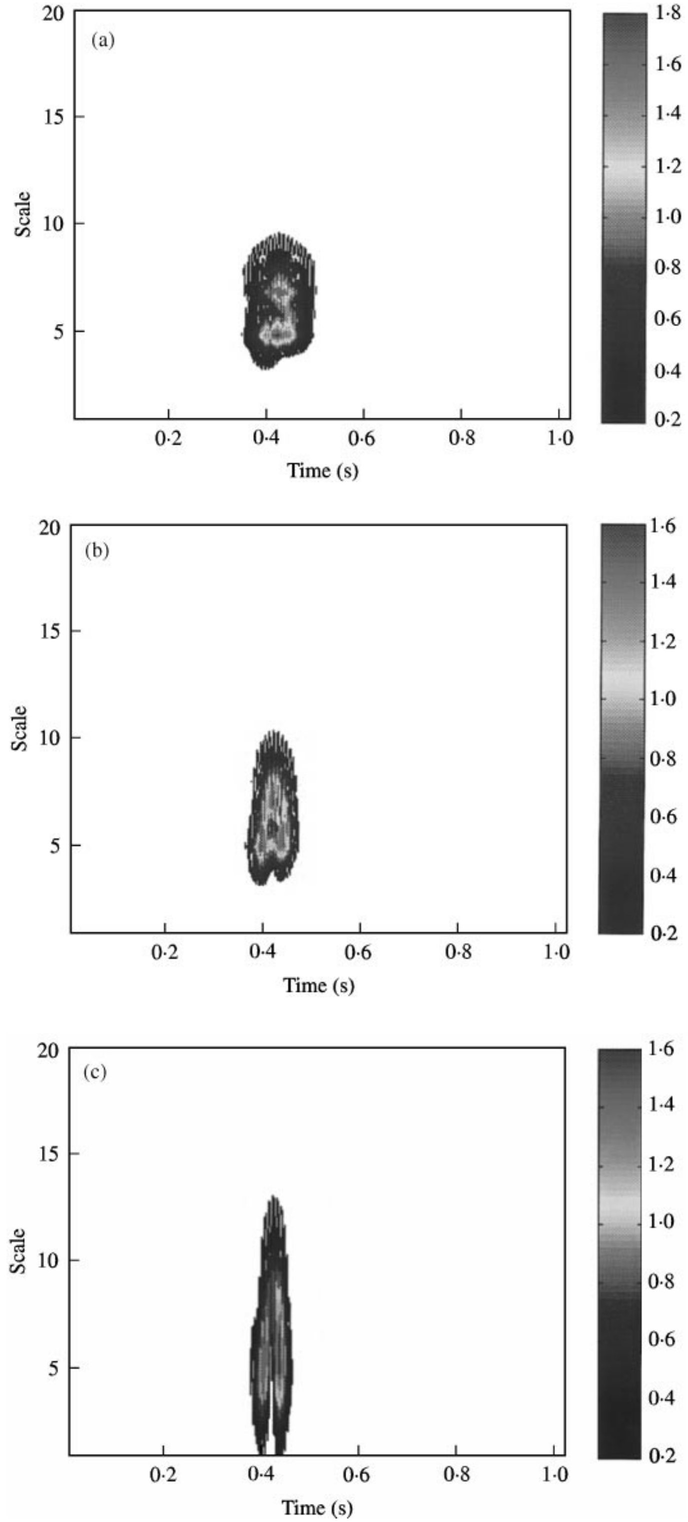

Figure 4(a). CWT of the signal when $\beta = 0{\cdot}2$; (b) CWT of the signal when $\beta = 0{\cdot}6$; (c) CWT of the signal when $\beta = 1{\cdot}2$.

into three parts, the CWT with $\beta = 0{\cdot}2$ divides the signal into two parts along the scale axis and the CWT with $\beta = 1{\cdot}2$ divides the signal into two parts along the time axis. So the method based on wavelet entropy is very effective for the selection of the wavelet bases.

3.2. FEATURE EXTRACTION USING MORLET CONTINUOUS WAVELET TRANSFORM

Although CWT brings about a lot of redundancy in the representation of the signal (a one-dimensional signal is mapped to a two-dimensional signal), it provides the possibility to reconstruct this signal. A classical inversion formula is

$$x(t) = C_{\psi}^{-1} \iint W(a,b)\psi_{a,b}(t) \frac{\mathrm{d}a}{a^2} \mathrm{d}b. \tag{10}$$

Another simple inverse way is to use Morlet's formula, which only requires a single integration. The formula is [10, 11]

$$x(t) = C_{1\psi}^{-1} \int W(a,b) \frac{\mathrm{d}a}{a^{3/2}}, \tag{11}$$

where

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

If all of the wavelet coefficients $W(a,b)$ corresponding to the feature components can be acquired, the purified signal can be obtained only by reconstructing these coefficients using the formula (11).

Here the principle of "soft-thresholding denoising" is introduced which was produced by Donoho and Johnstone [7, 8]. It provides a method on how to remove the irrelevant parts in the wavelet coefficients. The details of the method can be found in reference [7]. In this method, dyadic discrete orthogonal wavelet transform is used. Assuming that the background noise is additive Gaussian white noise, the wavelet coefficients caused by the noise are still the same independent distribution after the transform. The variance of the distribution depends on the variance of the Gaussian white noise and the number of the data according to the inference of Donoho [7]. Therefore, different thresholds may be used for the wavelet coefficients according to different demands for the risk. But this method has two disadvantages. First, the basic wavelet must be orthogonal. Usually, it is not similar to the feature components. Second, the background noise may not be Gaussian white noise. More generally, if all of their relevant components are considered as the noise needs to be removed, the purified signal obtained by this method will probably not be the true feature components.

These two deficiencies can be overcome by taking Morlet continuous wavelet transform instead of dyadic discrete wavelet transform. Wavelet coefficients measure the similarity between the signal and each of its son wavelets. The more the son wavelet is similar to feature component, the larger is the corresponding wavelet coefficient. So if the signal is transformed by the Morlet wavelet, those large wavelet coefficients are mainly caused by the impulse components contained in the signal. Reconstructing those large coefficients, the impulse components in the signal are obtained. Then, a threshold must be selected in order to obtain those large coefficients. The value of the threshold depends on the SNR of the signal. The greater the SNR, the lower the value. The theoretic instruction for the selection of the value of the threshold remains an open question.

The calculation of this denoising method can be summarized in the following steps:

(a) *Performing discrete wavelet transform*

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

where N is the data length.

(b) *Processing the coefficients using thresholding.* In this step, it is attempted to remove the components caused by noise. One method termed as "soft-thresholding" is used in Donoho's denoising method. It is an advanced version of the so-called "hard-thresholding" [7]. The "soft-thresholding" can be expressed in the following formula:

$$y_s=\begin{cases}\operatorname{sgn}(y)\,|y|-t, & |y|>t,\\ 0, & |y|<t.\end{cases} \tag{14}$$

where, sgn ( ) is the sign function, and t is the threshold. But when the SNR is so low that the threshold is beyond the half of the largest coefficient, all of the coefficients will be set to zero by this rule. To overcome this deficiency, a new way termed "generalized soft-thresholding" is established. It can be expressed by the following formula:

$$y_g=\begin{cases}\operatorname{sgn}(y)(|y|-\alpha t) & |y|>t,\\ 0, & |y|<t.\end{cases} \tag{15}$$

where α is constant and $0\leqslant\alpha\leqslant 1$. When $\alpha=0$, it becomes "hard-thresholding", and when $\alpha=1$, it becomes "soft-thresholding".

(c) *Reconstructing the revised coefficients.* Let $W'(a,b)$ be the revised coefficients. The purified signal can then be obtained using the following formula:

$$s(k)=\frac{1}{C_{1\psi}}\sum_{A}W'(a,k)a^{-3/2}, \tag{16}$$

where A is the definition domain of scale a.

It will be found that the method for noise canceling is more effective than Donoho's "soft-thresholding denoising" by two examples. Figures 5(a) and 5(b) are two simulated impulse signals with different decaying rates, and Figures 6(a) and 6(b) are the two signals with additive white noise. Then, using Donoho's method and the method introduced above, respectively, the purified signals were obtained, as shown in Figures 7(a), 7(b) and 8(a), 8(b). Obviously, the results from the denoising method based on Morlet wavelet is much better than those of Donoho's.

It should be pointed out that the parameter β of the Morlet wavelet may change in different cases. The optimal value of β can be obtained by the minimal entropy method introduced in section 2. Two examples illustrated in Figure 8 come from the denoising method based on a Morlet wavelet with different values of β. Figure 8(a) corresponds to $\beta=0{\cdot}3$, and Figure 8(b) corresponds to $\beta=3$.

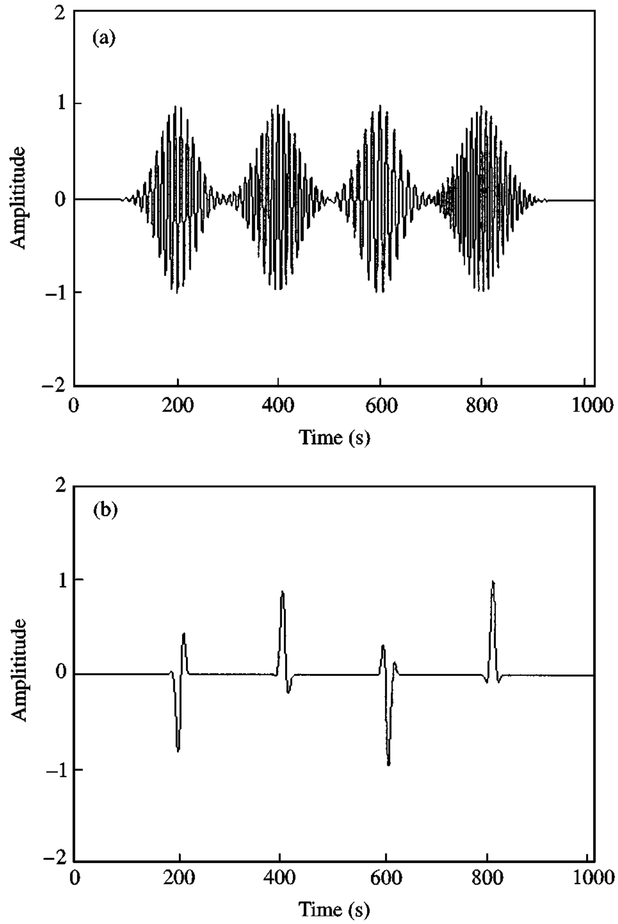

Figure 5. Two simulated impulse signals.

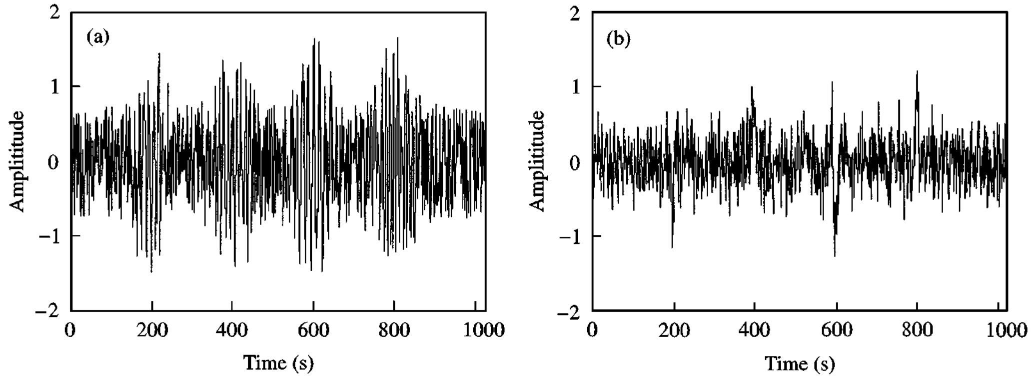

Figure 6. The two simulated signals with additive white noise.

4. MECHANICAL FAULT DIAGNOSIS USING THE DENOISING METHOD BASED ON MORLET WAVELET

For many mechanical dynamic signals, impulses usually indicate the occurrence of faults. In these cases, the impulses are covered by heavy background noise. Therefore, it is very

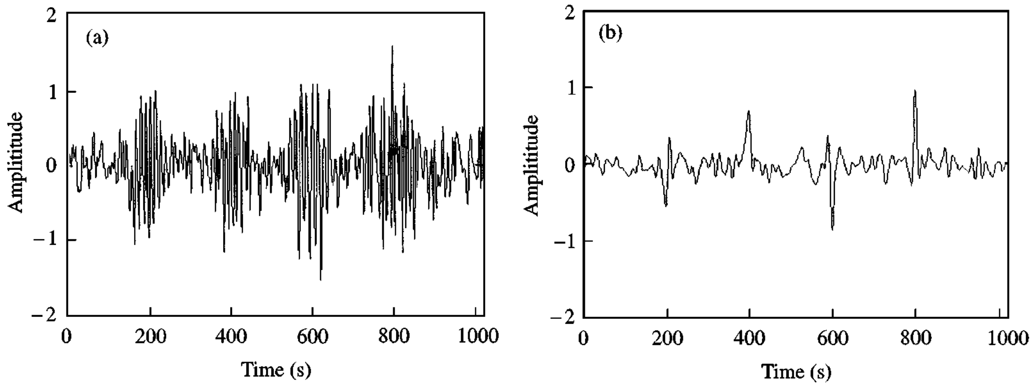

Figure 7. The purified signals obtained by Donoho's denoising method.

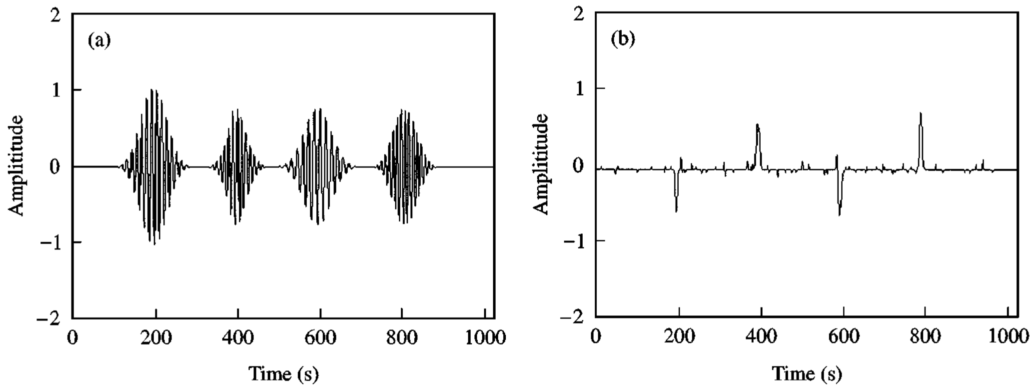

Figure 8. The purified signals obtained by the denoising method based on Morlet wavelet.

important to remove the noise or extract the feature components for following diagnosis. In the field of mechanical engineering, diagnosis for rolling bearings, gear boxes and reciprocating machines are especially difficult, because their vibration and acoustic signals often have low SNR. It will be very helpful to diagnose these machines when an effective method for feature extraction is used. As stated above, a Morlet wavelet can be used for extracting impulse components. The reason is that a Morlet wavelet is more similar to an impulse. Besides, parameter β of a Morlet wavelet can be adjusted to adapt to those impulses with any decaying rate. For Fourier transformation, the base is $e^{-j\omega t}$. The real and the imaginary parts are both triangular functions that are not impulses. Those feature components cannot be revealed using Fourier transformation. In the following, two examples are given to show how the denoising method based on Morlet wavelet is applied to mechanical fault diagnosis.

4.1. DIAGNOSIS FOR A ROLLING BEARING

Rolling bearings are installed in many kinds of machinery. A lot of problems of those machines may be caused by rolling bearings. Generally, local defects occur on outer-race, inner-race or rollers of bearings. When the rollers pass through the defect, an impulse may appear. According to the period of the impulse, we can judge the location of the defect using

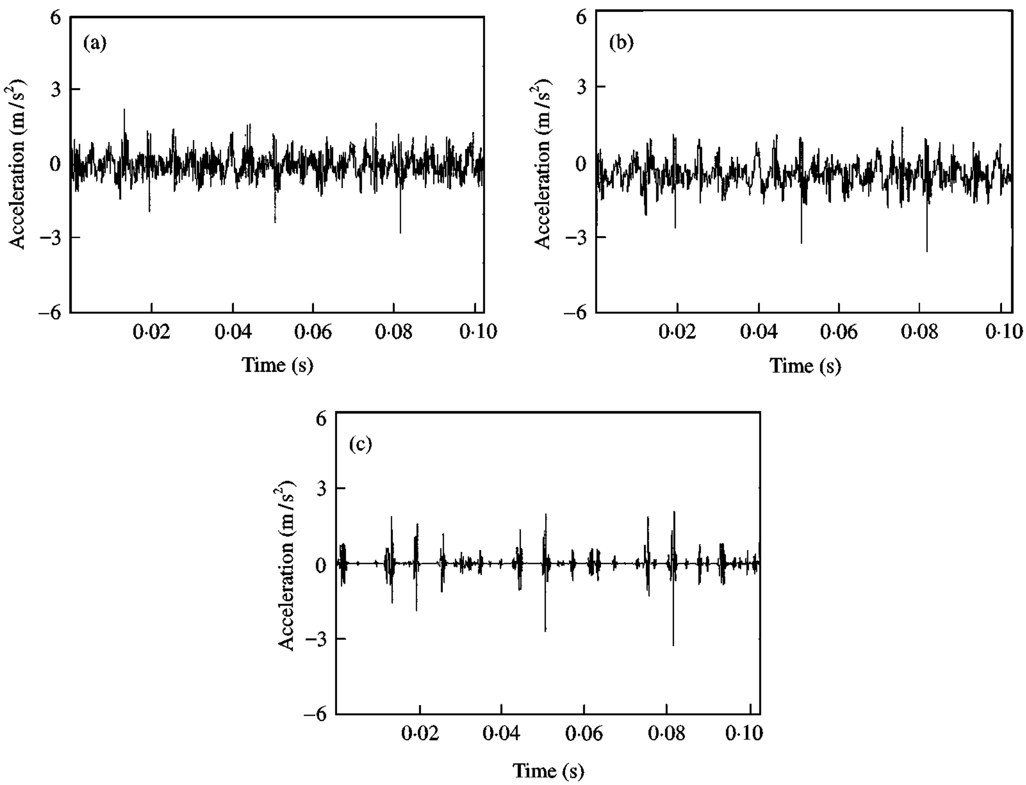

Figure 9. The vibration signal of an inner-race damaged rolling bearing. (a) The original vibration signal; (b) the purified signal obtained by Donoho's denoising method; (c) the purified signal obtained by the denoising method based on Morlet wavelet.

characteristic frequency formulae [1]. Because inner-race damage have more transfer segments when transmitting the impulse to the outer surface of the case, usually the impulse components are rather weak in the vibration signal. Thus, diagnosis for inner-race damage is very difficult. The rolling bearing tested has a pit on inner-race. The speed of the spindle is 2000 r.p.m., that is, the rotating frequency $f = 33{\cdot}3$ Hz. There are eight rollers in a bearing and the contact angle $\alpha = 0$, roller diameter $d = 15$ mm, bearing pitch diameter $E = 65$ mm, and the number of the roller $z = 8$. The characteristic frequency of the inner-race damage can be calculated by the formula.

$$f_i = 0{\cdot}5z\left(1 + \frac{d}{E}\cos\alpha\right)f. \tag{17}$$

Using the formula, the characteristic frequency for inner-race damage is calculated to be at 164 Hz. That is, the characteristic impulse period is 0·0061 s for inner-race damage.

Figures 9(a)–9(c) illustrate the three pictures of an inner-race damaged rolling bearing. The signal is acquired by an accelerometer mounted on the case of the bearing. It is sampled at 40 kHz with a 15 kHz filter in advance. We can hardly find any periodic impulses in the original signal. Even after Donoho's "soft-thresholding denoising" has been used to process the signal, the periodic impulses do not appear. Here the denoising method based on Morlet wavelet is used, the periodic impulses appear clearly, as shown in Figure 9(c). In this case,

$\beta = 0{\cdot}9$. The period is just about 0·006 s, which is in accordance with the characteristic frequency for the inner-race damaged rolling bearing.

4.2. DIAGNOSIS FOR A GEAR BOX

Gear boxes are very popular in industrial applications. A broken gear tooth failure may cause many fatal accidents, so the recognition of gear tooth cracks is very important for the safety of a gear box. They can be avoided by catching the early symptoms. Signals of a gear box are always noisy, and so it remains difficult to detect gear tooth crack effectively.

The following example concerns the life test of an automobile transmission box shown in Figure 10. Its transmission path is shown as follows:

$$input \rightarrow \left(\frac{Z28}{Z48}\right) \rightarrow \left(\frac{Z20}{Z44}\right) \rightarrow \left(\frac{Z30}{Z36}\right) \rightarrow \left(\frac{Z15}{Z42}\right) \rightarrow output.$$

In order to detect the gear tooth crack, many signal processing methods, such as zoom FFT cepstrum analysis, Hilbert transformation, etc., have been recommended. But the results are not very effective.

In this paper, the signal was acquired by an accelerometer mounted on the outer case of a gear box. It was sampled at 5 kHz with a 2 kHz low-pass filter in advance. Because of its long transmission path and multi-sources of excitation, the picked-up vibration signal was very complex. The vibration signal was picked up before the tooth was broken near the end of the life test.

The rotating speed of the input shaft was 1600 r/min, i.e., 26·67 Hz, and the rotating frequency of the output shaft was 2·1 Hz. The fault gear was the 42-tooth-gear on the output shaft. The cracks happened in two symmetrical teeth. Therefore, the signal in this case should include the impulse components whose period equals 0·24 s.

Figure 11(a) is the waveform of the transmission box. We cannot find any periodic impulses in it. Figure 11(b) is the purified signal obtained by the denoising method based on Morlet wavelet. Figure 11(c) is the purified signal obtained by Donoho's method. Periodic impulses appear in Figure 11(b). It can be easily found that the period is around 0·24 s. Thus,

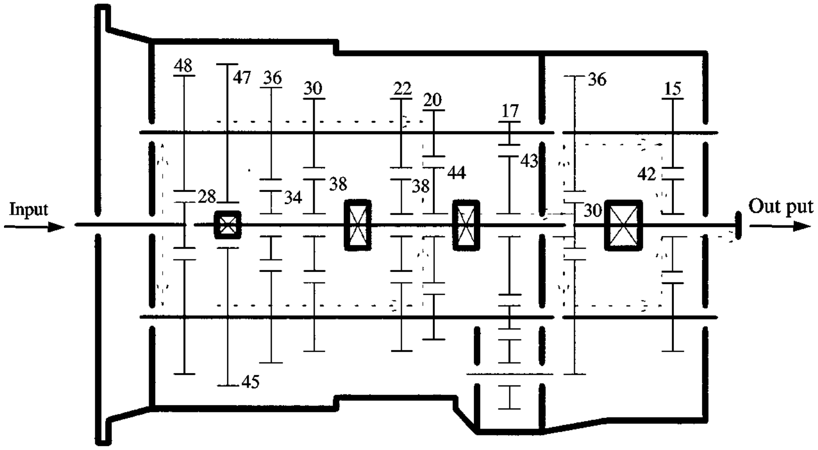

Figure 10. The transmitting routine of the automobile transmission box.

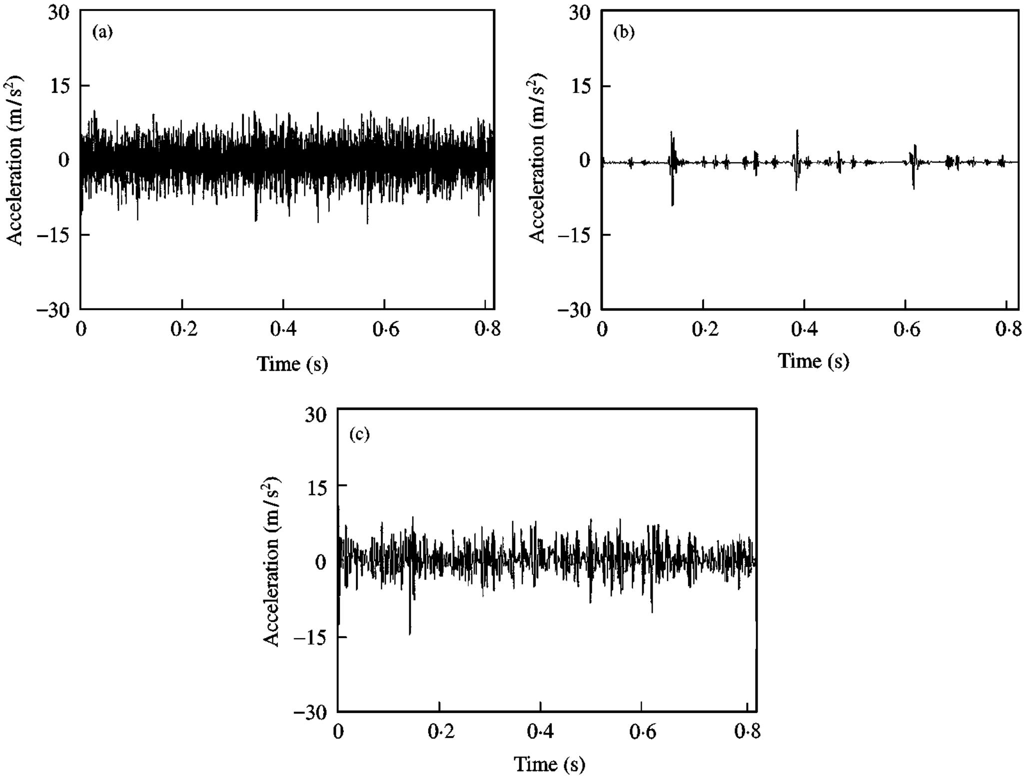

Figure 11. The vibrational signal of a gear box. (a) The original vibration signal of a gear box; (b) the purified signal obtained by the denoising based on Morlet wavelet; (c) the purified signal obtained by Donoho's method.

it can be judged by this purified signal that the tooth crack failure has happened. In this case, $\beta = 1{\cdot}8$.

From these two examples it can be seen that the denoising method based on Morlet wavelet is very effective for feature extraction. Although the impulses in the signal of roll bearing and those in the signal of gear box are different, they still can be extracted by this method. And though Donoho's "soft-thresholding denoising" has a lot of excellent mathematic attributes, it does not behave well in feature extraction for mechanical dynamic signals.

5. CONCLUSION

In the process of feature extraction, it is required that the time-scale structure of a wavelet should be consistent with the feature components. For a Morlet wavelet, its time and frequency resolution can be altered by adjusting the value of β. The optimal value of β can be obtained using minimal wavelet entropy method, In this paper, one denoising method based on Morlet method is proposed. It uses Morlet wavelet as the basic wavelet and compromises its time resolution and frequency resolution by adjusting the value of β to adapt to different signals. Having been tested using two simulated signals, it can be seen that this denoising method is more effective than Donoho's "soft-thresholding denoising" method. In this paper, this method is also applied to mechanical fault diagnosis—feature extraction from the signals of rolling bearings and a gear box. Among the final extracted features, it can be seen that the feature components, periodic impulses are very clear. In other words, the information immersed in the noise has been extracted completely. However, Donoho's "soft-thresholding denoising" method does not behave well in these two applications. Therefore, this denoising method based on Morlet wavelet has more advantages than Donoho's "soft-thresholding denoising" method in feature extraction from these impulse signals. It is a stronger tool for feature extraction and mechanical fault diagnosis.

REFERENCES

1. C. J. Li and J. Ma 1997 *NDT&E International* **30**, 143–149. Wavelet decomposition of vibrations for detection of bearing-localized defect.
2. Z. Geng and L. Qu 1994 *The British Journal of NDT* **36**, 11–15. Vibrational diagnosis of machine parts using the wavelet packet technique.
3. W. J. Staszewski 1998 *Journal of Sound and Vibration* **211**, 736–760. Wavelet based compression and feature selection for vibration analysis.
4. J. Zhao 1997 *Doctoral dissertation, School of Mechanical Engineering, Xi'an Jiaotong University*. Study for practical diagnosis techniques based on wavelets and neural networks.
5. S. Mallat and Z. Zhang 1993 *IEEE Transactions on Signal Processing* **45**, 3397–3415. Matching pursuits with time–frequency dictionaries.
6. G. Davis, S. Mallat and Z. Zhang 1994 *Optical Engineering* **33**, 2183–2191. Adaptive time–frequency decompositions.
7. D. L. Donoho 1995 *IEEE Transactions on Information Theory* **41**, 613–627. Denoising by soft-thresholding.
8. D. L. Donoho and I. M. Jonestone 1994 *Biometrika* **81**, 425–455. Ideal spatial adaptation via wavelet shrinkage.
9. H. Guo, H. Odegard, J. E. Burrus and C. S. Wells 1996 *IEEE Transactions on Signal Processing Letters* **3**, 10–12. Noise reduction using an undecimated discrete wavelet transform.
10. O. Rioul and P. Duhamel 1992 *IEEE Transactions on Information Theory* **38**, 569–586. Fast algorithms for discrete for discrete and continuous wavelet transform.
11. M. J. Shensa 1996 *IEEE Transactions on Signal Processing* **44**, 798–807. Discrete inverses for nonorthogonal wavelet transforms.