Research on a Signal Analysis Method based on Wavelet Theory and Approximate Entropy Algorithm

Xiaoyong Yu

School of Information Engineering, Suzhou University, Suzhou, 234000 China

Abstract

The vibration signal is one of the significant signals that reflects the fault. In allusion to the shortcomings of traditional signal analysis method in the high-frequency and non-stationary signal analysis, the wavelet theory and approximate entropy algorithm are introduced into the signal analysis in order to propose a new vibration signal analysis (WTAEAVSA) method in this paper. In the proposed WTAEAVSA method, the wavelet transform technology is used to reduce the noise and decompose the low and high frequency vibration signal in order to obtain the signal characteristics of different frequency bands. Then the approximate entropy algorithm is used to determine the complexity and irregular degree of vibration signal in the different scale and different frequency band, so as the non-stationary characteristics of vibration signal are extracted. At last, some simulated signals with time-domain and frequency-domain from the normal signal are used to test the characteristics of the proposed WTAEAVSA method. The simulation results show that the proposed WTAEAVSA method can extract the characteristic vector from vibration signal, visually and sharply reflect the changes of the mechanical states.

Keywords: signal analysis, wavelet theory, approximate entropy algorithm, wavelet packet decomposition, feature extraction

1. Introduction

Fault diagnosis of mechanical equipment is an effective method in mechanical fault diagnosis technology by using vibration signal [1]. Fault feature extraction is the key to fault diagnosis, which is related to the accuracy of fault diagnosis and the reliability of early prediction. So people analyze the signal form various angles including the time domain, frequency domain and time-frequency domain in order to propose a lot of feature parameters and achieve a certain effect [2-4].

Wavelet transform is a time-scale analysis method of signal, which has the characteristic of multi-resolution analysis. And it has the ability to represent the local features of the signal in the time-frequency domain. Wavelet packet transform is the further development based on wavelet transform, which can provide the higher resolution than wavelet transform. Wavelet packet decomposition is a more sophisticated method than the wavelet decomposition. It does not only decompose the low frequency of the signal, but also decompose the high frequency of the signal. When the fault occurs, the irregularity and complexity of the signal can reflect the occurrence and development of the fault. In recent years, a lot of new methods are proposed to analyze the signal. Masashi and Masato [5] proposed a method of extracting the desired signal from a noisy signal, addressing the problem of segregating two acoustic sources as a model of acoustic source segregation based on Auditory Scene Analysis. Angrisani et al. [6] proposed a digital signal-processing method for reliable analysis of ultrawide-band detected PD current signals. The method, exploiting the multiresolution approach peculiar to the Wavelet Transform, is capable of setting the detected signal free from most disturbances affecting its
true shape, thus ensuring the significance of the measurements carried out on the signal itself. Goumas et al. [7] proposed a new method for extracting features in the wavelet domain and uses them for classification of washing machines vibration transient signals. Peng et al. [8] proposed a reassignment method in order to improve the concentration of the scalogram and weaken the interference terms to a certain extent, and the readability of the scalogram. Dickenson and Ghassemlooy[9] proposed a novel receiver system based on the multiresolution time-frequency feature extraction capabilities of wavelet analysis, coupled with the well recognized pattern recognition performance of artificial neural networks for mitigating the effects of bandwidth limiting channel-induced distortion. Umaphy and Krishnan [10] proposed a novel approach is proposed to classify pathological speech signals using a local discriminant bases (LDB) algorithm and wavelet packet decompositions. Peng et al. [11] proposed a novel singularity based fault features. The wavelet transforms modulus maximal (WTMM) provides a simple but accurate method in calculating the Lipschitz exponents. Ubeyli[12] proposed the multiclass support vector machines(SVMs) with the error correcting output codes(ECOC) for the multiclass Doppler ultrasound signals(ophthalmic arterial Doppler signals and internal carotid arterial Doppler signals) classification problems. Chen et al. [13] proposed a new method for the fault degradation assessment of the water hydraulic motor based on Wavelet Packet Analysis (WPA) and KS test to analyze the impulsive energy of the vibration signal, which is used to detect the piston condition of water hydraulic motor. Andrade et al. [14] proposed a recent tool of the Hilbert spectrum (HS) for the analysis of nonlinear and no stationary time-series to study the electromyography (EMG) signals. Cheng et al. [15] proposed a novel generalized demodulation time-frequency analysis approach for processing non-stationary signals, especially for processing amplitude-modulated and frequency-modulated (AM-FM) signals. The novelty of this method is that the signal whose time-frequency distributions are curved paths can be transformed into that whose time-frequency distributions are linear paths paralleling time axis by using generalized demodulation. Ubeyli [16] proposed the probabilistic neural network for classification of electroencephalogram (EEG) signals. Zhu et al. [17] reviewed the state-of-the-art of wavelet analysis for tool condition monitoring (TCM). Wavelet analysis has been the most important non-stationary signal processing tool today, and popular in machining sensor signal analysis. Unser et al. [18] proposed to transpose the concept to the wavelet domain by considering a complexified version of the Riesz transform which has the remarkable property of mapping a real-valued (primary) wavelet basis of L2 into a complex one. Cherif et al. [19] proposed the extraction of features out of heart sounds in time-frequency (TF) domain for recognition of heart sounds through TF analysis. Bendjama et al.[20] proposed a new combined fault diagnosis method that uses Wavelet Transform (WT), Principal Component Analysis(PCA) and Neural Networks(NN) for rotating machinery vibration monitoring and analysis. The measurements from the vibration process are used to verify the WT-PCA-NN method for detecting and diagnosing faults under typical operating conditions. Shao et al. [21] proposed a single-channel BSS method based on wavelet transform and independent component analysis (ICA), and separated the source signals related to a milling cutter and spindle from a single-channel power signal. Liu [22] proposed a novel random and nonstationary vibratory signature analysis(R&N-VSA) technique to address challenge for extracting prominent features from random and nonstationary vibratory signals. Sun et al.[23] proposed a novel method of mixed programming which combines Matlab with Delphi based on COM(component Object Model) module technology to design a time-frequency analysis system for non-stationary vibration signals assembling fourteen methods. The advantages of two languages are shared with each other.
Glowacz [24] proposed a non-invasive method of early fault diagnostics of electric motors and a plan of study of acoustic signals of electric motors. Researches were carried out for faultless induction motor, induction motor with one faulty rotor bar, induction motor with two faulty rotor bars and flawless Direct Current, and Direct Current motor with shorted rotor coils. Bao et al. [25] proposed a method combined EMD with wavelet thresholding for fault signal analysis. Inspired by the wavelet translation invariant nonparametric signal denoising, the wavelet thresholding principle which has been the dominant technique in the area for many years is adopted in Intrinsic Mode Function (IMF) decomposition modes for fault feature extraction. Chen et al.[26] analyzed sensitivity analysis and experimental research on ball bearing early fault diagnosis based on testing signal from casing. The research results in this study provide an experimental basis for a ball bearing fault diagnosis based on a casing test signal from a practical aero-engine. Wang et al.[27] proposed an iterative extraction method for transient signal analysis based on transient modeling and parameter identification through Levenberg-Marquardt(LM) method, and eventually for fault feature extraction. Tang et al. [28] proposed a new feature extraction method that is based on multifractal detrended fluctuation analysis (MFDFA). UHF PD signals of four typical GIS discharge models that were collected in a laboratory were analyzed. Al-Qazza et al. [29] proposed a comparative study to select the efficient mother wavelet (MWT) basis functions that optimally represent the signal characteristics of the electrical activity of the human brain during a working memory (WM) task recorded through electroencephalography (EEG). Wang et al. [30] proposed a robust signal preprocessing framework for wrist pulse analysis. The cascade filter based on frequency-dependent analysis (FDA) is first introduced to remove the high frequency noises and to select the significant pulse intervals. Then the curve fitting method is developed to adjust the direction and the baseline drift with minimum signal distortion.

These proposed new methods for analyzing the signal can better analyze the original signal. But these methods exist some defects, such as low computation time and fault accuracy and so on. And the measuring complexity methods of the signal mainly include Lempel-Z complexity, approximate entropy and sample entropy and so on. Approximate entropy is used to quantitatively describe the repeatability of time series. If the time series are more complex, the approximate entropy is more. The signal is more complex, the system is more complex, and it contains the more abundant the frequency components. In this paper, the wavelet theory and approximate entropy algorithm are introduced into the signal analysis in order to propose a new vibration signal analysis (WTAEAVSA) method in this paper. At last, some simulated signals with time-domain and frequency-domain from the normal signal are used to test the characteristics of the proposed WTAEAVSA method.

The rest of this paper is organized as follows. Section 2 briefly introduces wavelet theory, including wavelet transform and wavelet packet transform. Section 3 briefly introduces the approximate entropy algorithm. Section 4 briefly feature extraction based on wavelet theory and approximate entropy algorithm. Section 5 gives Experiment result and analysis. Finally, the conclusions are discussed in Section 6.

2. Wavelet Theory

Wavelet transform is a time-scale analysis method. It has the characteristic of multi-resolution analysis [31]. And it has the ability to represent the local features of the signal in the time-frequency domain. It is a kind of time frequency localization analysis method, which can change the time window and frequency window with fixed window area. So the wavelet transform is used to realize the fault diagnosis of dynamic system for
achieving good results.

2.1. Wavelet transform

Multi-sample-rate filter bank is a technology, which is developed in the field of digital signal processing in 1980s. It has a special similarity with wavelet transform. Wavelet transform of the signal is equivalent to the filtering of the signal by using a specific filter bank. A two-layer decomposition structure of wavelet transform is shown in Figure 1.

![Wavelet Transform Diagram](image)

**Figure 1. Two-layer Decomposition Structure of Wavelet Transform**

In the Figure 1, $S$ is the input signal, $H$ and $L$ are high pass filter and low pass filter. $\uparrow 2$ is sampler with the sampling interval 2. When $S$ passes the high and low pass filter, $cD_1$ and $cA_1$ are obtained by using the 2 sampler. The $cD_1$ and $cA_1$ are corresponding to the first layer of wavelet decomposition coefficient. The $cD_1$ is the detail signal of wavelet decomposition, it is corresponding to the high frequency of $S$. The $cA_1$ is corresponds to the trend signal of wavelet decomposition, it is corresponding to the low frequency of $S$. For the trend signal, the same filter is used to decompose, $cD_2$ and $cA_2$ are obtained by using the 2 sampler. The $cD_2$ and $cA_2$ are corresponding to the second layer of wavelet decomposition coefficient. The $cD_2$ is the detail signal and the $cA_2$ is corresponds to the trend signal. So the original signal $S$ are decomposed into the trend signal ($cA_2$) and a series of detail signals ($cD_1$ and $cD_2$). The $j^{th}$ layer decomposition is described as follow:

$$S = cA_j + \sum_{i=j} cD_i$$  \hspace{1cm} (1)

$$cA_{j-1} = cA_j + cD_j$$  \hspace{1cm} (2)

It can be seen that the wavelet transform is only the low frequency part of the signal to be decomposed, and the high frequency resolution is low.

2.2. Wavelet Packet Transform

Wavelet packet transform is a kind of wavelet analysis. It can increase with the frequency for further dividing the frequency window into the detail frequency. In the high frequency section, it has high time resolution and frequency resolution, so it has important application value in the frequency domain analysis of the mutation signal.

After the wavelet packet transform selects a wavelet function, the filter coefficients are selected:
\( h = \{ h_n \} \), suppose \( g_k = (-1)^k h_{-k} \), a recursive function is determined:

\[
W_{2n}(t) = \sqrt{2} \sum h_k W_n(2t - k) \\
W_{2n+1}(t) = \sqrt{2} \sum g_k W_n(2t - k)
\]

(3)

(4)

The wavelet packet \( \{ W_n(t) \} \) is determined by the above expressions. \( W_n(t) \) is the corresponding wavelet function.

The wavelet packet transform of the signal decomposes the high frequency signal and the low frequency signal. The wavelet packet decomposition structure is shown in Figure 2.

![Wavelet Packet Decomposition](image)

**Figure 2. Tree-layer Wavelet Packet Decomposition**

where the \( A \) expresses the low frequency, the \( D \) expresses the high frequency. The serial number is the layer number of wavelet decomposition. It can be seen that the number of decomposition level is larger, the wavelet packet size is larger, the spatial resolution of the wavelet packet coefficients. The spatial resolution of the corresponding space of wavelet packet coefficient is lower.

3. Approximate Entropy Algorithm

Approximate entropy is proposed to measure the complexity of signal sequence from the view by Pincus in 1991[32]. The approximate entropy is used to measure the probability of the generating new model in the signal. It uses a non-negative number to quantificationally describe the complexity and irregularity of one signal sequence. The entropy value of the approximate entropy is increased with the increasing of sequence complexity.

Set the time series \( \{ u(i), i = 0, 1, 2, 3, ..., N \} \). The mode dimension \( m \) and similarity tolerance \( r \) are given in advance. The specific calculation method of approximate entropy is described as follows:

1. A set of \( m \) dimensional vector \( X(i) \) is composed according to the sequential order. That is,

\[
X(i) = [x(i), x(i + 1), ..., x(i + m - 1)]
\]

where \( i = 1 \sim N - m + 1 \)

2. The distance \( d[X(i), X(j)] \) between the vector \( X(i) \) and \( X(j) \) is obtained. This distance \( d[X(i), X(j)] \) is a largest absolute value of the corresponding element difference between the vector \( X(i) \) and other vector \( X(j) \). That is,

\[
d[X(i), X(j)] = \max_{k=0\sim m-1} |x(i+k) - x(j+k)|
\]

(5)

(6)
For each \( i \) value, the distance between the vector \( X(i) \) and other vector \( X(j) (j = 1 \sim N - m + 1, j \neq i) \) is computed.

(3) The similarity tolerance \( r \) is given in advance. For each \( i \) value, the number of the distance \( d[X(i), X(j)] \) less than the similarity tolerance \( r \) is calculated, and the ratio \( C_r^m (r) \) of \( N - m \) is described,

\[
C_r^m(r) = n/l(N - m)
\]

where \( i = 1 \sim N - m + 1, r = 0.1 \sim 0.25SD(X) \), \( n \) is the number of the distance \( d[X(i), X(j)] \) less than the similarity tolerance \( r \), \( SD(X) \) is the standard deviation of the original sequence. The value of \( m \) is 2, \( r = 0.025SD(X) \).

(4) The \( C_r^m (r) \) is solved the logarithm, then the mean value of all \( i \) is calculated in order to obtain the \( \phi^m (r) \).

\[
\phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln C_r^m(r)
\]

(5) The dimension is increased one, then the above steps are repeated in order to get the \( C_r^m (r) \) and \( \phi^m (r) \).

(6) The approximate entropy value (\( ApEn(m, r) \)) of this sequence is output

\[
ApEn(m, r) = \lim_{N \to \infty} [\phi^m(r) - \phi^{m+1}(r)]
\]

The ultimate value exists with the probability one. But in practice, the value of \( N \) could not be infinite. When the value of \( N \) is limited value, the sequence length is obtained according to the above-mentioned procedure, it is regarded as the estimated value of approximate entropy value (\( ApEn \)). It is denoted as follow.

\[
ApEn(m, r) = \phi^m(r) - \phi^{m+1}(r)
\]

The approximate entropy value (\( ApEn(m, r) \)) is clearly related to the value of \( m \) and \( r \).

The approximate entropy is a non-negative invariant, which is used to represent the complexity of the sequence. If the entropy value is larger, the complexity of the sequence is higher. Under different period, the approximate entropy reflects the self similarity state ability of all sequence phases. If the sequence is relatively stable, the approximate entropy is also relatively stable. If the sequence is suddenly changing, the approximate entropy is obviously changing. So the sliding window method is used to select a certain length subsequence from the original sequence, which to calculate the \( ApEn \) value of each subsequence. The changing of the \( ApEn \) value is used to analyze the mutability. The specific steps of the sliding window method are described as follows.

(1) Select sliding window length (\( L \)), sliding step length (\( h \)), the number of the dimension (\( m \)) and the allowable deviation (\( r \)).

(2) Select the subsequence of the window length (\( L \)) according to the \( i^{th} \) data of the sequence.

(3) Calculate the \( ApEn \) value of the subsequence.

(4) Keep the \( L \) is unchanged. Move the window with sliding step length (\( h \)), AND repeat step (1) to step (2) until the end of the original sequence.

(5) Obtain an \( ApEn \) sequence with the length of \( N - L + 1 \) by repeating step (1) to step (4).

(6) The obtained \( ApEn \) value is significantly tested in order to determine the mutability point or interval.
At present, the approximate entropy algorithm is widely applied in the feature extraction and fault diagnosis. Many characteristics of approximate entropy algorithm are suitable for analyzing the vibration signal. Because the approximate entropy algorithm not sensitive to the vibration signal, and better noise-resistant ability. The results of the approximate entropy algorithm can reflect the complexity of the time series signal, and the amplitude of the signal is independent. And the approximate entropy algorithm is suitable for analyzing the deterministic signals and random signals. So the approximate entropy algorithm can provide a new method for the fault diagnosis of the rotating equipment.

4. Feature Extraction Based On Wavelet Theory and Approximate Entropy Algorithm

The fault diagnosis method based on wavelet theory and sample entropy is to calculate the sample entropy value of each frequency band signal after wavelet packet transform, in order to realize the fault recognition. Feature extraction algorithm is described as follows:

(1) The obtained original vibration signal is decomposed into three layer wavelet packet in order to obtain eight frequency band coefficients from the low frequency to high frequency in the third layer.

(2) The decomposition coefficients of wavelet packet are reconstructed, the sample entropy value of each frequency band signal is calculated.

(3) Construct feature vectors. The sample entropy value of each frequency band signal is regarded as the elements for constructing feature vectors.

5. Experiment Result and Analysis

In order to prove the performance of the proposed WTAEAVSA method, the bearing of rotating machinery is selected in this paper. The corresponding experiment conditions for obtaining the data are shown in Table 1. In the experiment, the vibration signal includes the normal signal and fault signal. The wavelet packet decomposition method is used to precisely decompose the signal for obtaining the characteristic information of the signal. Then the approximate entropy algorithm is used to judge the fault signal.

| Table 1. The Corresponding Experiment Conditions |
|-----------------|----------|----------|----------|
| Source          | Time/div | Volts/Pos | Volts/div |
| CH              | 110.0us  | 0.001V   | 3.00V    |
| Trig/Type       |          |          |          |
| Trig/Level      |          |          |          |
| Trig/Slope      |          |          |          |
| Trig/Source     |          |          |          |
| Edge            | -60.0mV  | Rising   | CH       |

From the Table 1, the normal signal and the fault signal are included in here. If the signal contains the noise signal, the waveform, amplitude frequency and spectral characteristics are shown in Figure 1 and Figure 2.
As can be seen from the Figure 1 and Figure 2, if the signal mixes the noise, its original waveform, amplitude frequency and spectrum characteristics take on the obvious changes in the experiment.

The fault signal and normal signal are respectively decomposed in order to obtain the waveform. The results for the second layer are shown in Figure 3 and Figure 4.
The fault waveform figure is obtained by using the proposed method. In general, it is very difficult to discover the obvious difference between the normal signal and fault signal. Therefore, the approximate entropy algorithm is used to analyze the experiment results.

Table 2. The Entropy Value between the Normal Signal and Fault Signal

<table>
<thead>
<tr>
<th></th>
<th>Ap(\text{En})</th>
<th>Ap(\text{En})</th>
<th>Ap(\text{En})</th>
<th>Ap(\text{En})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal signal</td>
<td>0.434 523</td>
<td>0.690 527</td>
<td>0.376 732</td>
<td>0.728 630</td>
</tr>
<tr>
<td>Fault signal</td>
<td>0.464 362</td>
<td>0.657 921</td>
<td>0.395 274</td>
<td>0.684 753</td>
</tr>
</tbody>
</table>

The computation results show that the approximate entropy value and characteristic value are different in the different condition. The proposed WTAEAVSA method takes on a strong ability for describing the complexity of the signal and a very good effect for judging the running states of mechanical equipments. It can extract the characteristic vector from vibration signal, visually and sharply reflect the changes of the mechanical states.

6. Conclusion

Wavelet transform is a time-scale analysis method with the characteristic of multi-resolution analysis. It is a kind of time frequency localization analysis method, and change the time window and frequency window with fixed window area. Approximate entropy algorithm is used to measure the probability of the generating new model in the signal. It uses a non-negative number to quantificationally describe the complexity and irregularity of one signal sequence. So the wavelet theory and approximate entropy algorithm are introduced into the signal analysis in order to propose a new vibration signal analysis (WTAEAVSA) method for overcoming the shortcomings of traditional signal analysis method in the high-frequency and non-stationary signal analysis. The wavelet transform technology is used to reduce the noise and decompose the low and high frequency vibration signal.
in order to obtain the signal characteristics of different frequency bands. Then the approximate entropy algorithm is used to determine the complexity and irregular degree of vibration signal in the different scale and different frequency band, so as the non-stationary characteristics of vibration signal are extracted. In order to prove the performance of the proposed WTAEAVSA method, the bearing of rotating machinery is selected. The computation results show that proposed WTAEAVSA method takes on a strong ability for describing the complexity of the signal and a very good effect for judging the running states of mechanical equipments. It can extract the characteristic vector from vibration signal, visually and sharply reflect the changes of the mechanical states.

Acknowledgments

This research was supported by Colleges and university Natural Science Research Project, Anhui Province, China, Project Number: KJ2012Z396. The authors would like to thank all the reviewers for their constructive comments. The program for the initialization, training, and simulation of the proposed algorithm in this article was written with the tool-box of MATLAB 2012 produced by the Math-Works, Inc.

References


**Authors**

Xiaoyong Yu, he is a lecturer, received the Master degree in Computer Application Technology from Hefei University of Technology in 2008, Hefei, China. The main research directions: Computer Architecture, Artificial Intelligence and Information process.