Research Article

Journal of Robotics and Automation Research

Application of Entropy Information on Bearing Fault Diagnosis

Zihan Wang and Yongjian Sun*

School of Electrical Engineering, University of Jinan, Jinan, Shandong, China.

***Corresponding author**
Yongjian Sun, School of Electrical Engineering, University of Jinan, Jinan, Shandong, China.

Submitted: 10 Aug 2021; **Accepted**: 30 Sep 2021; **Published**: 12 Oct 2021

Citation: *Yongjian Sun, Zihan Wang. (2021). Application of Entropy Information on Bearing Fault Diagnosis. J Robot Auto Res, 2(2), 31-37.*

Abstract
In order to solve the problem of nonlinear, nonstationary, complex components and redundant information of rolling bearing vibration signal in single scale, a rolling bearing fault feature extraction method based on wavelet packet decomposition and permutation entropy and sample entropy is proposed. Firstly, wavelet packet decomposition is used to decompose the original signal of rolling bearing into several sub bands with different frequencies, and the permutation entropy and sample entropy of signal data at different frequencies are calculated. Secondly, the sample entropy and permutation entropy of different frequency signals after decomposition and reconstruction are extracted to form a high-dimensional feature vector to complete the initial fault feature extraction. Finally, the extracted feature samples are randomly arranged for fault recognition. The experimental data of rolling bearing processed by this method are identified by extreme learning machine. The results show that the method can effectively identify the fault types of rolling bearing, and the classification effect is better than that of the original data set training, and the classification accuracy reaches 99.8

Keywords: Rolling Bearing, Wavelet Packet Decomposition, Entropy Characteristics, Elm, Feature Extraction

Introduction
The key work of fault diagnosis is mainly for rotating machinery, and rolling bearing fault diagnosis is the main part of rotating machinery. Rolling bearing is one of the parts of the machine which are easy to be damaged. Therefore, with the development of high technology and the increase of industrial production demand, rolling bearing fault diagnosis technology arises at the historic moment [1].

The original fault diagnosis method is to judge by hearing, which is similar to the method that doctors use stethoscope to diagnose the patient's condition. If there is a fault in the rolling bearing, teachers can listen to it with their ears, and then they can diagnose which kind of fault is combined with their long-term experience, and then they can hear which fault is according to the different sound [2]. This method was quite popular in China in the early years. Many people relied on this skill to eat, but the biggest problem of this method is that it requires years of experience. Another problem is that this method is easily affected by personal physical condition, and the quality of personal emotions often affects the judgment results, so this method was gradually eliminated. After that, with the continuous progress of science and technology, a variety of precision instruments were invented. Just like B ultrasound used in hospitals, workers no longer rely on their ears, but on professional mechanical equipment. Society gradually from artificial to mechanical, which greatly improves the efficiency and accuracy [3].

As one of the giants of rolling bearing manufacturing industry, SKF company of Sweden has never stopped the research on rolling bearing fault diagnosis. After long-term efforts, the company independently developed the impact pulse instrument in 1966 to detect the bearing damage. This improves the fault diagnosis technology to a great extent and drives the development of rolling bearing fault diagnosis. After that, there is an upsurge of using sensors to monitor bearings for a long time in enterprises, and even some equipment has been used in aerospace industry. After ten years of development, the rolling bearing fault diagnosis instrument which can detect the abnormal signal of rolling bearing in different frequency bands has been developed by Nippon Steel Co., Ltd. Along with the Japanese bearing fault detection instrument, the oil film detector, which uses ultrasonic or high frequency current to detect the lubrication state of the bearing, is also available. From 1976 to 1983, Japan Seiko Corporation (NSK) successively developed NB series bearing detector. Its principle is to measure RMS value and peak value of bearing vibration signal in the range of 1-15khz, so as to achieve the purpose of detecting bearing fault.

Bearing is the core component of mechanical structure. Condition monitoring and fault diagnosis are the necessary guarantee for the normal operation of mechanical equipment. The process of bearing fault diagnosis mainly includes fault feature extraction and fault classification. When the bearing fails, the frequency band energy of the vibration signal will change. If the frequency band signal features of bearings can be extracted accurately, the bearing faults can be classified. Therefore, accurate fault feature extraction is the key step of fault diagnosis technology.

In signal analysis and recognition, wavelet packet decom- position is a relatively new time domain analysis method. By processing the measured signal with wavelet packet decom- position, we can decompose the signal into different frequency bands according to any time-frequency resolution without redundancy and overlap. In recent years, some scholars have done a lot of research on wavelet packet decom- position. Tang and others selected Hermitian wavelet to obtain the time wavelet energy spectrum containing bearing fault information through continuous wavelet packet transform, and then applied the method of sample on sample to obtain the time wavelet energy spectrum [4]. The sample entropy was used as the eigenvector to input the support vector machine to realize the classification of different bearing fault modes [5].

Proposed a feature separation method of bearing composite fault based on improved wavelet packet decomposition, which can decompose single channel composite fault signal of bearing into different channels and realize the feature separation of composite fault. Houssem proposed a bearing intelligent prediction method based on bidirectional long-term and short-term memory and wavelet packet decomposition, which successfully detected the bearing degradation process and predicted the remaining service life of the bearing.

Sample entropy is a measure of time series complexity, which eliminates the self-comparison problem of reconstructed state vector in approximate entropy, so it has stronger anti-interference ability, better statistical stability, and less dependence on the length of calculated data [6].

Yang proposed a feature extraction method based on variational mode decomposition and sample entropy, which has higher computational efficiency and strong ability to distinguish different fault states of rolling bearings [7]. Xie calculated the sample entropy feature based on ceemdan, and combined with support vector machine to classify rolling bearing fault diagnosis. Wei proposed a rolling bearing signal analysis and fault diagnosis method based on multiscale sample entropy of ensemble empirical mode decomposition. This method can accurately extract fault features and support simultaneous detection of multiple faults, which further im- proves the accuracy and efficiency of rolling bearing fault diagnosis.

Elm is a single hidden layer feedforward neural network algorithm proposed by Huang, which has the advantages of strong generalization learning ability and fast training speed, and has been successfully applied in the field of mechanical fault diagnosis. Wang and others used gray wolf optimization algorithm to improve extreme learning machine to realize fault classification of rolling bearing [8]. He proposed a prediction method of bearing residual life based on principal component analysis and multivariable extreme learning machine [9].

Scheme Design

Wavelet Packet Decomposition

Wavelet packet decomposition can also be called wavelet packet or sub-band tree and optimal sub-band tree structure. The concept is to represent wavelet packet with analysis tree, that is, to analyze the detail part of input signal by using wavelet transform of multiple generations. From the perspective of function theory, wavelet packet decomposition is to project the signal into the space formed by wavelet packet basis function. From the perspective of signal processing, it is to let the signal pass through a series of filters with different center frequencies but the same bandwidth. The specific formula is as follows:

$$L^2(R) = \oplus W_j \tag{1}$$

W_j is the wavelet subspace, wavelet packet decomposition further decomposes a, makes a better division of all frequency bands, and improves the frequency resolution. The wavelet packet decomposition can be expressed as:

$$W_j = U_{j-k}^{2^k} \oplus U_{j-k}^{2^k+1} \oplus \cdots \oplus U_{j-k}^{2^{k+1}-1}, j, k \in Z \tag{2}$$

When the number of decomposition layers increases, the dimension of the data to be processed will increase accordingly, so the signal cannot be decomposed indefinitely. In the process of application, a reasonable number of decomposition layers should be selected according to the actual needs. In view of the bearing fault signal data analysis and test to be processed in this paper, the three-level decomposition is the best.

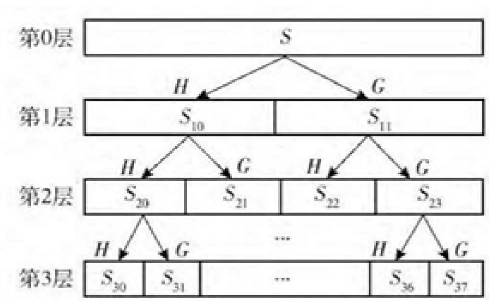

Figure 1: Three level wavelet packet decomposition tree

After wavelet packet decomposition, the original signal energy is divided into each sub-band

Permutation Entropy

Permutation entropy is the same as approximate entropy, sample entropy and fuzzy entropy, which are used to measure the complexity of time series. However, it introduces the idea of permutation when calculating the complexity between reconstructed subsequences.

The advantage of permutation entropy principle is that it does not need to consider the specific value of data, and it is an algorithm based on the comparison of adjacent data. The calculation method is described below.

For the time series of length N $\{x(i), i=1,2,...,N\}$, we reconstruct the space and get the following time series:

$$\begin{cases} X(1) = \{x(1), x(1+\lambda), \cdots, x(1+(m-1)\lambda)\} \\ \vdots \\ X(i) = \{x(i), x(i+\lambda), \cdots, x(i+(m-1)\lambda)\} \\ \vdots \\ X(N-(m-1)\lambda) = \{x(N-(m-1)\lambda), x(N-(m-2)\lambda), \cdots, x(N)\} \end{cases} \quad (3)$$

The m data of X(i) are rearranged from small to large.

Sample Entropy

In 2000, Richman et al. First proposed the concept of sample entropy, which is a more robust time series complexity measurement method similar to approximate entropy. Compared with approximate entropy, it has stronger anti- interference and anti-noise ability [7]. Sample entropy im- proves the algorithm of approximate entropy, which can re- duce the error of approximate entropy. It is similar to ap- proximate entropy, but its accuracy is better.

Sample entropy is a new measure of time series complexity proposed by Richman and moornan. The improvement of the sample entropy algorithm compared with the approximate entropy algorithm: compared with the approximate entropy, the sample entropy calculates the logarithm of sum. The purpose of sample entropy is to reduce the error of approximate entropy, which is more consistent with the known random part. sample entropy has better consistency. That is, if one-time series has a higher value than another time series, it also has a higher value for other m and R values.

Sample entropy is defined as

$$\mathrm{SampEn}(m,r) = \lim_{N\to\infty}\left\{-\ln\left[\frac{A^m(r)}{B^m(r)}\right]\right\} \quad (4)$$

It can be seen that the value of sample entropy is related to the value of m and r. Therefore, it is very important to determine the values of M and R for the calculation of sample entropy. Here, according to the research results of Pincus, m= 1 or 2, r = 1Std 0.25Std (Std is the standard deviation of the original data), the sample entropy calculated has more reasonable statistical characteristics, in this study, m=2, r = 0.2std.

Principle of ELM Extreme Learning Machine

ELM consists of input layer, hidden layer and output layer, as shown in the figure 2.

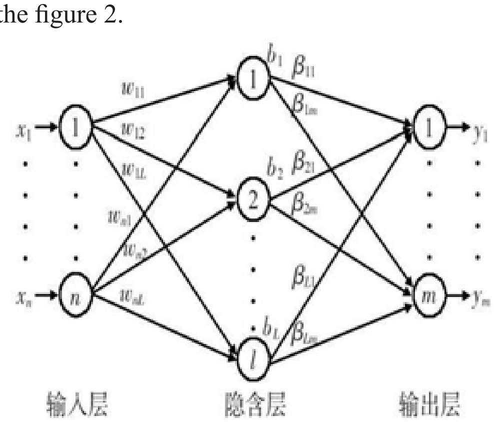

Figure 2: ELM structure

Diagnostic Process

The wavelet packet signal is decomposed and reconstructed for the rolling bearing signals under four different working conditions: normal state, inner ring fault, rolling body fault and outer ring fault. According to the minimum frequency criterion of effective signal, the vibration signal is divided into three layers by wavelet packet. DB4 wavelet basis function is selected to reconstruct the decomposition into the coefficient of three-layer nodes, and eight different sub band sequences are obtained, and make wavelet packet decomposition waveform

- Analyze the original signal.
- The original signal is decomposed by wavelet packet to reconstruct the signal.
- The sample entropy and permutation entropy were calculated.
- Sample entropy and permutation entropy are extracted to construct feature vector.
- Extreme learning machine training test.
- According to the absolute value of the difference with the entropy characteristic scale, which bearing working state is closest to the test data can be judged.

Simulation Experiment

Data Source

The data used in this paper are all from the bearing data center of Case Western Reserve University. The model of experimental bearing is 6205-2RS JEM SKF deep groove ball bearing.

In this paper, the data of normal state, inner ring fault, rolling element fault and outer ring fault with sampling frequency of 12khz and bearing fault diameter of 0.1778mm, 0.3556mm and 0.5334mm are selected for processing. There are 480000 normal state data and 120000 inner ring faults, rolling element fault and outer ring fault data.

Specific Implementation Mode

The time domain and frequency domain features of the original signal data are extracted.

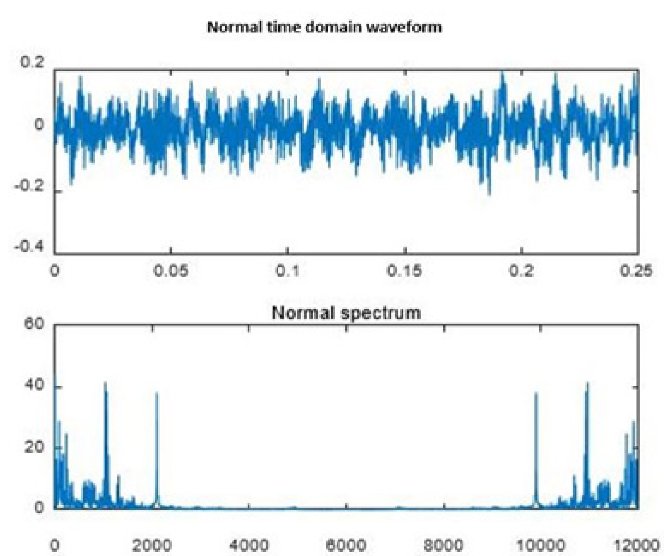

Figure 3: Normal

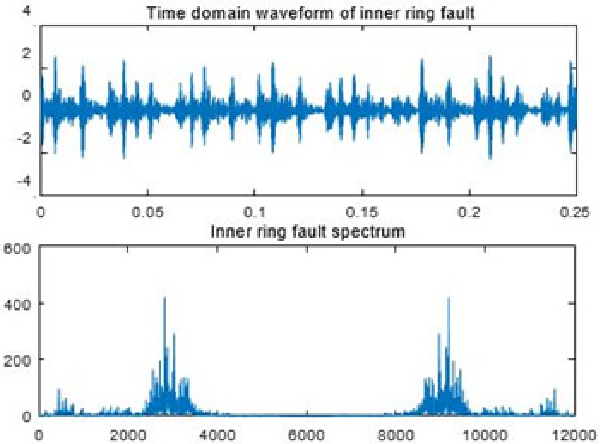

Figure 4: Inner ring fault

It can be seen from (Figure3-6) that the time domain and frequency range of the four working conditions of rolling bearing with fault diameter of 0.5334mm are different, the peak values of each frequency band are different, and there are too many repeated data. Fault feature extraction directly based on the original signal cannot accurately identify fault features, so it is necessary to decompose the original signal. According to the minimum frequency criterion of effective signal, the vibration signal is decomposed into three layers by wavelet packet. The DB4 wavelet basis function is selected to decompose and reconstruct the vibration signal into the coefficients of three-layer nodes, and eight different sub band sequences are obtained, and make wavelet packet decomposition waveform, as shown in (Figure 7-10).

The eight sub-band sequence of the reconstructed signal contains the information of each frequency band from the

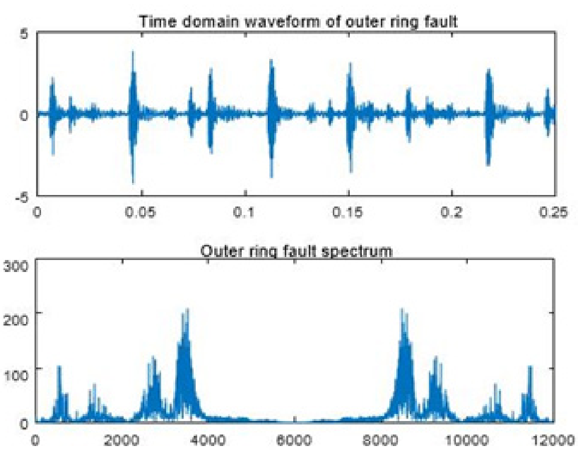

Figure 5: Outer ring fault

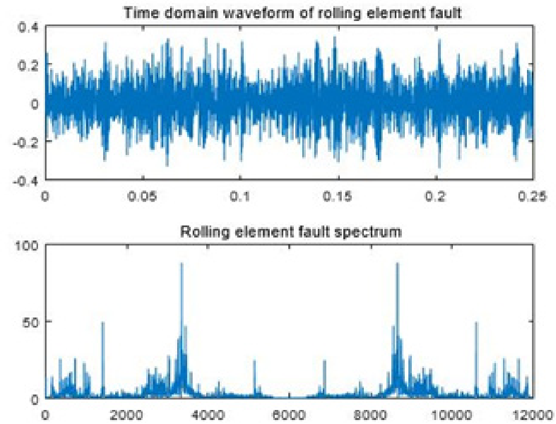

Figure 6: Rolling element failure

low frequency band to the high frequency band in the original signal, which ensures that the data contained in the reconstructed signal is neither repeated nor omitted, and further improves the accuracy of rolling bearing fault feature extraction. In this experiment, 2400 data points are taken as a sample, and 50 groups of data samples are selected for each working condition.

There are obvious differences between permutation entropy and sample entropy obtained by wavelet packet de- composition and reconstruction for vibration signals of rolling bearing under normal state, inner ring fault, rolling element fault and outer ring fault. Therefore, permutation entropy and sample entropy can be used as feature vectors to distinguish rolling bearing under different states.

The data signal after wavelet packet decomposition and reconstruction is extracted by sample entropy and permutation entropy to form 8-dimensional feature vector

$$S = \left[S_{30}, S_{31,}, S_{32}, S_{33}, \ S_{34}, S_{35}, S_{36}, S_{37}\right] \quad (5)$$

S_{3i} represents the sample entropy and permutation entropy of the rolling bearing signal extract- ed by the third wavelet packet coefficient of wavelet packet three-level decomposition. In this paper, the sample entropy and permutation entropy corresponding to the second and sixth wavelet packet coefficients are analyzed.

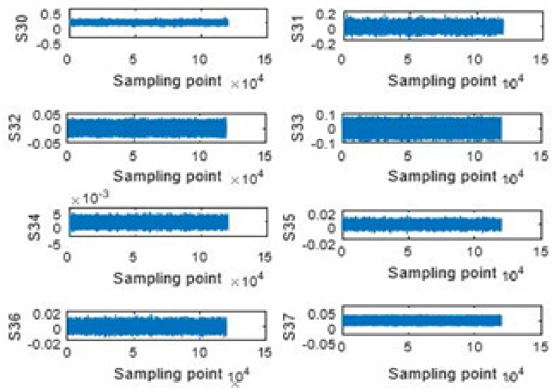

Figure 7: Normal

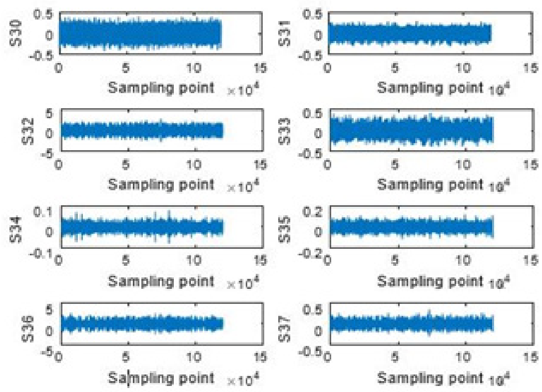

Figure 8: Inner ring fault

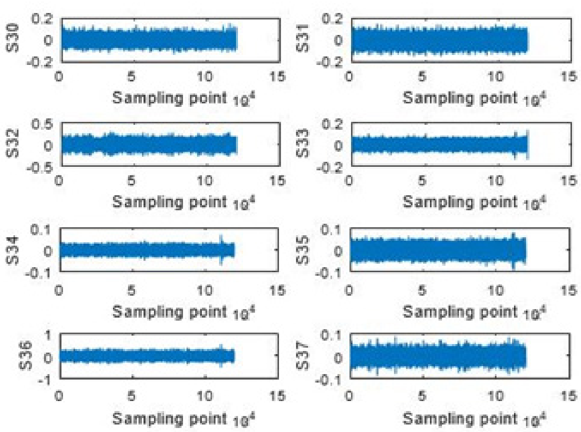

Figure 9: Rolling element failure

Using MATLAB software, the four different working conditions of rolling bearing normal state, inner ring fault, rolling element fault and outer ring fault are numbered as 1, 2, 3 and 4 respectively. In 50 groups of data samples, 25 groups were randomly selected as the elm training set, and the remaining 25 groups were randomly transformed again

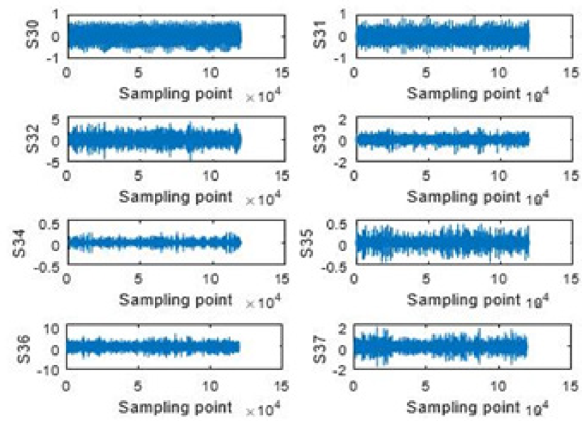

Figure 10: Outer ring fault

As the elm test set. The training and testing were repeated for 10 times to calculate the average classification accuracy.

The original data signals of rolling bearing with different fault diameters are decomposed and reconstructed by wavelet packet, and eight groups of signals with different frequencies are obtained. The sample entropy of each group of reconstructed data signals is calculated, and eight sample entropies are obtained for each group, forming an 8-dimensional eigenvector [S_{30},S_{31},S_{32},S_{33}, S_{34},S_{35},S_{36},S_{37}]. The sample entropy S_{33} and S_{37} extracted from the fourth and eighth wavelet packet coefficients are analyzed to determine whether they can be used as the ex- traction standard of bearing fault features.

As shown in Figure 11, the sample entropy values contained in the sample entropy S37 vector extracted from the eighth wavelet packet coefficients of 50 samples (200 sets of data) under four working conditions are arranged in ascending order, and the sample entropy ranges of four working conditions are obtained, which are: normal working condition 0.7935 0.8428, inner ring fault 0.2836 0.479, rolling element fault 0.5692 0.6604, outer ring fault 0.132 0.2435. The range of sample entropy of the four working conditions is different, which can be effectively used as the reference standard of sample entropy fault feature extraction.

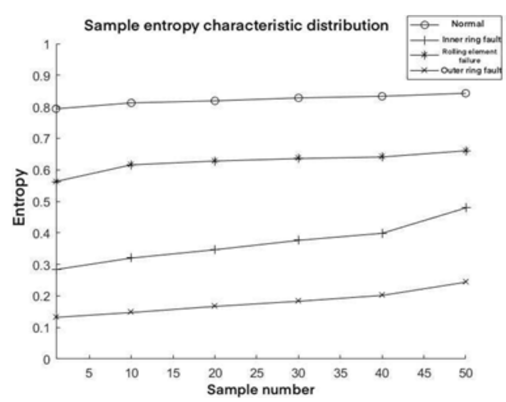

Figure 11: Sample entropy characteristic distribution

The original data signals of rolling bearing with different fault diameters are decomposed and reconstructed by wavelet packet, and eight groups of signals with different frequencies are obtained. The sample entropy of each group of reconstructed data signals is calculated, and are obtained for each group, forming an 8-dimensional eigenvector[S_{30},S_{31},S_{32},S_{33},S_{34},S_{35},S_{36}, S_{37}]. The sample entropy S33 and S37 extracted from the fourth and eighth wavelet packet coefficients are analyzed to determine whether they can be used as the extraction standard of bearing fault features.

As shown in Figure 12 the permutation entropy of S37 vector extracted from the eighth wavelet packet coefficient of 50 samples (200 sets of da-ta) under four working conditions is arranged in ascending order, and the range of four permutation entropy values is obtained, which are: normal working condition 1.7789 1.7820, inner ring fault 1.7293 1.7491, rolling element fault 1.7535 1.7614, outer ring fault 1.7623 1.7713. The range of permutation entropy of the four working conditions is different, which can be used as the reference standard of permutation entropy fault feature extraction.

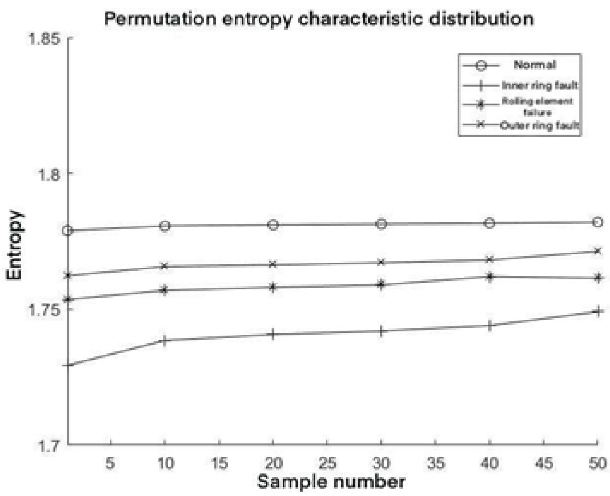

Figure 12: Permutation entropy characteristic distribution

50 groups of samples of normal state, inner ring fault, rolling element fault and outer ring fault of rolling bearing with fault diameter of 0.5334mm are input into the extreme learning machine model. The first 25 groups of samples of each working condition are input as training samples for learning, and the new samples obtained from the random arrangement of the last 25 groups are used as test samples, the fault classification accuracy of rolling bearings with three different fault diameters is calculated when the sample entropy is used as the fault feature and the permutation entropy is used as the fault feature. The training test is repeated for 10 times, and the average accuracy is obtained.

ELM training results and test results of rolling bearing fault feature extraction based on sample entropy.1

Conclusion

In this paper, a fault feature extraction method of rolling bearing based on sample entropy and permutation entropy of wavelet packet extreme learning machine is proposed to detect the fault types of rolling bearing in complex environment.

Table 1: Entropy characteristic matrix

Fault diameter	Number of tests	Average accuracy
0.1178	10	100
0.3556	10	97.1
0.5334	10	99.8

Single analysis of the bearing original signal data in time domain and frequency domain image features, the signal contains noise, large repeatability, cannot accurately extract the fault features of the bearing signal.

Using wavelet packet, the one-dimensional vibration signal is decomposed into eight different frequency vibration signals, and the more accurate entropy feature is used to construct the eight-dimensional feature vector, which is used as the rolling bearing fault feature extraction standard, and further improves the accuracy of rolling bearing fault feature classification.

The introduction of extreme learning machine (ELM) improves the intelligence of rolling bearing fault diagnosis.

The experimental results show that the wavelet packet decomposition can extract the time-domain features of non-stationary signals without redundancy, and the time-frequency resolution is more precise, which effectively solves the shortage of high frequency and low resolution of wavelet transform. Compared with one-dimensional signal, the time-frequency image can display more effective features. Taking the sample entropy and permutation entropy as the feature vectors can effectively be used as the fault feature classification criteria, and permutation entropy can more accurately identify the fault type of rolling bearing.

However, there are still some room for improvement, such as improving the signal decomposition method, increasing the number of extracted fault features, extracting multiple fault features synchronously, and using multi-scale permutation entropy and sample entropy to process data.

References

1. Yan, R., Liu, Y., & Gao, R. X. (2012). Permutation entropy: A nonlinear statistical measure for status characterization of rotary machines. Mechanical Systems and Signal Processing, 29, 474-484.
2. Ding, X., & He, Q. (2016). Time–frequency manifold sparse reconstruction: A novel method for bearing fault feature extraction. Mechanical Systems and Signal Processing, 80, 392-413.
3. Poor, H. V. (2013). An introduction to signal detection and estimation. Springer Science & Business Media.
4. Smith, B. (1995). An approach to graphs of linear forms (Unpublished work style).
5. Cheng, D. (2005, June). On logic-based intelligent systems. In 2005 International Conference on Control and Automation (Vol. 1, pp. 71-76). IEEE.
6. Cheng, D., Ortega, R., Panteley, E., Sun, Y., Shen, T., & Ohmori, H. (2005). On port controlled hamiltonian systems. In Advanced Robust and Adaptive Control-Theory and Appli-

cations (pp. 3-16). Beijing: Tsinghua University Press.
7. Huang, W., Kong, F., & Zhao, X. (2018). Spur bevel gearbox fault diagnosis using wavelet packet transform and rough set theory. Journal of Intelligent Manufacturing, 29(6), 1257-1271.
8. Bandt, C., & Pompe, B. (2002). Permutation entropy: a natural complexity measure for time series. Physical review letters, 88(17), 174102.
9. Cheng, D. (2005). Controllability of switched bilinear systems. IEEE Transactions on Automatic Control, 50(4), 511-515.