博士学位論文

Research on Intelligent Diagnosis Technology of Rotating Machinery by Fault Feature Extraction of Vibration Signal

（振動信号の故障特徴抽出による回転機械の知的診断技術に関する研究）

三重大学大学院 生物資源学研究科

共生環境学専攻 環境・生産科学講座

環境情報システム工学教育研究分野

段 棠少

2022 年 8 月

ABSTRACT

In rotating machinery, the rolling element bearing is one of the most widely used pieces of mechanical equipment, and meanwhile is also very easily destroyed. Over 40% faults of induction machines are caused by bearing component faults. Therefore, Fault Detection and Diagnosis (FDD) of potential bearing faults are very important and necessary when it comes to reliable operation of a given system. When fault occurs, such as fatigue, flake crack, and so on, since rolling element bearing is in a periodical rotating working model, in each rotation the excitation appears to be periodicity. So when a bearing is destroyed, the signal we measured, especially the vibration signal, generally causes strong periodicity in signal. However, due to the complex working environment of bearing, the low signal-to-noise ratio (SNR) of vibration signals causes difficulties in fault diagnosis.

To dynamically track bearing failure degradation，a generative fault diagnosis model based on state-space principal component tracking filtering (SPCTF) is proposed. The proposed method can dynamically track the bearing fault status and detect bearing faults in real time. Experiments on bearing data of three different speeds (500RPM, 1000RPM, and 1500RPM) and three fault states (inner race fault, outer race fault and roller fault) show the effectiveness of the proposed method. For the weak features of low-speed rolling bearing fault characteristics, an improved Teager energy operator (ITEO) is proposed. ITEO overcomes the shortcomings of sensitive to noise and vibration interference of the traditional Teager energy operator (TEO), through the method of amplifying the shock energy of the vibration signal, the fault features of the low-speed bearing are extracted. Then, an intelligent diagnosis model of low-speed bearings based on ITEO-AE is proposed, which can realize the fault diagnosis of low-speed bearings. The effectiveness of the proposed method is proved by the diagnostic experiments of bearing vibration signals for two different speeds (70RPM and 100RPM) and three fault states (inner race fault, outer race fault and roller fault). In order to prevent major accidents, early diagnosis of bearing faults is crucial. Therefore, the early

fault diagnosis model of rolling bearing based on CRBM-PCA is proposed， which can adaptively extract the early faults of rolling bearings, excellent for diagnosing faulty bearings with small dimensions. In order to more realistically simulate the early failure of bearings, three different sizes of faulty bearings (0.6W*0.3D, 1.2W*0.3D and 2.0W*0.3D) were used in the experiments.

The major contributions of this thesis are a variety of rolling bearing fault diagnosis methods for different speeds and different fault periods are proposed. All of proposed methods proposed in this thesis have been verified during the experiments, under different kinds of fault rolling bearings.

Content

1. Introduction

1.1 Background and significance

As one of the most important components in industrial equipment, rotating machinery is widely used in all walks of life. Large-scale rotating machinery such as fans, compressors, steam turbines and other equipment are key production tools in modern important production sectors such as petroleum, chemical, electric power, and aerospace. Real-time online monitoring and fault diagnosis of these equipment is ground foundation to ensure the healthy and stable operation of industrial systems. At the end of the 20th century, Japanese scholar Toyotaraifuru introduced the fault diagnosis technology of rotating machinery to Japan from the West, and then the intelligent fault diagnosis technology of rotating machinery was gradually promoted in Asia. Different from the traditional regular and planned maintenance, the intelligent rotating machinery fault diagnosis technology is to determine the fault location, cause and severity according to the operating status of the equipment under the condition of ensuring the normal operation of the equipment (without shutting down or disassembling the machine). Techniques for predicting the condition and reliability of equipment and proposing solutions. This technology can effectively avoid the waste of manpower and material resources caused by the blindness and uncertainty of traditional regular maintenance methods. And can reduce equipment downtime caused by sudden failure, improve production efficiency and safety.

The failure of rotating machinery is often accompanied by abnormal vibration and noise. The vibration signal can reflect the equipment fault information well [1-3]. The fault diagnosis technology of rotating machinery based on the analysis of vibration signal is a hot research topic at abroad in recent years. The main process of rotating machinery fault diagnosis based on vibration signal analysis is as follows: collect equipment vibration signals through sensors; preprocess the vibration signals to remove interference components in the signals, and obtain effective vibration signals containing fault information; Processing and extracting fault feature information; According to the extracted fault features, determine the fault type and analyze the cause of the fault, monitor and diagnose the rotating machinery online in real time, greatly reduce the loss caused by downtime maintenance and improve equipment safety and production efficiency [4-7]. Due to the harsh working environment of rolling bearings and the unavoidable interference of other components during operation, the measured vibration

signals often have the characteristics of non-stationarity and low signal-to-noise ratio. Therefore, the key to realize the fault diagnosis of rolling bearing is how to accurately extract more fault information from the strong background noise [8-10].

1.2 Purpose and content of this research

This research aiming at the non-stationary and nonlinear characteristics of the fault vibration signal of rolling bearings, intelligent analysis methods are used to remove the noise interference in the signal, such as: state tracking filtering, improved Teagear energy operator, principal element feature analysis, etc., to extract the bearing from the strong background noise [11-15]. At the same time, using neural network deep learning methods such as intelligent auto-encoding and class restricted Boltzmann machine, combined with rolling bearing fault mechanism and vibration signal characteristics, in-depth research on rolling bearing fault feature extraction and early fault diagnosis methods, in order to improve equipment work efficiency and service life. According to the different state characteristics of high-speed operation, low-speed operation and early fault bearings, the corresponding fault feature identification scheme is studied. All of proposed methods proposed in this thesis have been verified during the experiments, under different kinds of fault rolling bearings. Finally, we are committed to extending the intelligent diagnosis scheme proposed in this study to practical industrial applications.

1.3 Literature review

The fault in rolling bearing (one of the essential components in industrial equipment) results in equipment damage, industrial shutdowns, extensive property losses and even casualties. An urgent need exists for an effective monitoring system to ensure equipment's high-precision and stable operation [16], [17]. Vibration signal analysis has been widely used in the fault detection of rolling bearing. However, due to the complex working environment of bearing, the low signal-to-noise ratio (SNR) of vibration signals causes difficulties in fault diagnosis.

1.3.1 The review of bearings fault feature extraction

Effective fault feature extraction methods can overcome the problem of high SNR in bearing signal, and numerous fault feature extraction methods have been proposed in recent years [18], including signal decomposition, signal enhancement and signal filtering.

The signal decomposition method, including wavelet analysis [19], [20], empirical

mode decomposition (EMD) [21]–[24], and variational mode decomposition (VMD) [25]–[27], decomposes and reconstructs the signal to reduce the influence of noise and then extracts the fault features according to the mathematical theory [28]. However, there are still some limitations within these methods: (I) Wavelet analysis is not precise when applied to high-frequency decomposition; (II)The edge phenomenon and mode mixing are inevitable in EMD; (III) It is difficult to determine the number of modes in VMD; and (IV) Those methods undermine the fault characteristics information to a certain extent when removing the noise. The stochastic resonance method can enhance gains of periodic fault feature signals with noise energy to extract the fault characteristic frequency. However, it is difficult for the stochastic resonance method to filter low-frequency interference [29–31] effectively. Furthermore, the structural parameters of the resonance system also have a significant influence on fault diagnosis results. Signal filtering methods include high pass filtering, adaptive filtering [32], [33], fractional filtering [34], and Kalman filtering [35], which extract the effective frequency band by analyzing the spectrum signals and performing a fault diagnosis. Liao [33] and Long [34] demonstrated the outperformance of the signal filtering method in bearing fault diagnosis through comparison with band-pass filtering and wavelet decomposition. However, it depends on the prior knowledge of bearing fault and is not conducive to practical industrial application.

1.3.2 The review of low-speed bearings fault feature extraction

In order to effectively extract the fault characteristic information of rolling bearings running at low speed and meet the requirements of real-time monitoring and fault diagnosis of working conditions to the greatest extent, many signal modulation methods based on spectrum analysis are widely used, including amplitude spectrum, cepstrum, power spectrum, refinement Pu, Hilbert demodulation and energy operators, etc. Due to the characteristics of non-stationarity, impact and modulation of bearing fault signal, the demodulation technology Hilbert Transform (HT) and Teager Energy Operator (TEO) technology are more favored by scholars. Compared with HT, TEO is used to analyze and track the energy of narrowband signals, and has the advantages of high demodulation accuracy and fast response speed. The demodulation of a combination of AM and FM signals is of great significance for the comparative analysis of different types of signals.。

According to the TEO theory, TEO has the ability to denoising and enhance the fault signal. TEO is a nonlinear difference operator, it estimates the total energy of the

signal through transient signal and signal differential. Liu Hongxing [36] summarized the demodulation principle of the TEO method and the comparison experiment verify it is better than Hilbert transform for mechanical signal demodulation. Tao jie [37] proposed bearing fault diagnosis method based on TEO and deep belief network, they perform the TEO to calculate the transient amplitudes and extract the transient energies of the vibration signal. TEO can enhance the transient features of the vibration signal, effectively detects shock components. Which has been widely applied in the fields of fault diagnosis to demodulate signal and extract the fault feature.

1.3.3 The review of early fault diagnosis techniques

Rotating machinery is the most commonly used type of machine in modern industry, civilian and military applications, such as compressors, steam turbines, automobiles, industrial fans, and aircraft engines [38]. Due to the high service load, harsh operating conditions or inevitable fatigue, faults may develop in rotating machinery [39-41]. If the fault cannot be diagnosed in a timely way, it may cause a shutdown of the whole system and even catastrophic failure. Therefore, it is significant to detect faults early and assess the fault level as early as possible to avoid catastrophic accidents and ensure the safe operation of the machinery [42-45]. Many researchers are committed to developing early fault diagnosis (EFD) techniques to monitor the health conditions of rotating machinery. At present, many EFD techniques have been successfully applied in the modern industries, such as the vibration-based EFD method, current-based EFD method, acoustic emission-based EFD method, sound-based EFD method, torque-based EFD method, and rotating encoder-based EFD method, etc. Among these sensor signals, vibration-based diagnostic method is the most commonly researched one because vibration signals can directly represent the dynamic behavior of rotating machinery. However, the early fault signal is often too low due to the fact that small localized damage may generate small periodical impulses. Meanwhile, the measured vibration signals at the early stage are interfered by strong noises [46-50].

To diagnose early faults as soon as possible, the feature extraction of vibration signals is very important in real engineering applications. Recently, the advanced signal processing-based weak feature extraction method has been becoming a hot research topic [51-53]. However, the measured vibration signals of rotating machinery often present nonlinear and non-stationary characteristics. Moreover, the vibration signals contain strong background noises during the early fault stage, making it difficult to extract fault information. In order to effectively analyze the non-stationary vibration

signals, massive research efforts have been made in last two decades to develop various signal processing technologies, including wavelet transformation (WT), empirical mode decomposition (EMD), ensemble empirical mode decomposition (EEMD), local mean decomposition (LMD), empirical wavelet transform (EWT), variational mode decomposition (VMD), stochastic resonance, sparse decomposition, etc. In addition, artificial intelligence-based methods are also attracting increasing attention.

1.4 Innovation and contribution of this thesis

In this thesis, a variety of bearing fault feature extraction and diagnosis methods have been proposed. These proposed algorithms and theories have been verified by various simulation experiments and case studies. The contributions of this article are described as follows.

1) The proposed method has better real-time and dynamics than the traditional filter methods when it innovatively adopts the switch Kalman filter to track optimal estimated state matrix consistent with more effective fault information in the process of model updating. Moreover, it effectively reduces the treated signal's dependence (with various filter methods) that has delay time compared with the proposed method.

2) The proposed method has a novel mechanism compared with the traditional filtering method, which directly removes unrelated fault information useful for the diagnosis. The reason for this contribution is that the SKF that established multiple filter models better describes the dynamic process of the system, preserving rich and valuable information in the optimal estimated state matrix.

3) Moreover, the proposed method has an outstanding anti-interference ability for time-series signals since the optimal estimated state matrix is calculated with the optimal estimates and observations during the model update process based on SKF. Therefore, a feature tracking matrix generated with the principal component analysis is extracted from the optimal estimated state matrix in real-time and dynamically, which has high SNR due to the excellent denoising ability of the proposed method.

4) It is essential to consider the influence of the computational burden in the optimal estimated state matrix, a high-dimensional matrix for the actual fault diagnosis. The reason is that the dynamic model based on SKF multiplies the amount of data, an advantage (multiple filter models) of the SKF. Therefore, the reduced-order method based on principal components analysis is utilized to extract the effective fault feature from the redundant information for generating a feature tracking matrix, which can

effectively reduce data dimensions and improve the computing efficiency of SPCTF.

5). Generalization is an important indicator of the diagnosis when changing the working condition. Therefore, the proposed method is applied to other different speeds at 500, 1000 and 1500 rpm. The result showed the effectiveness and strong generalization across various conditions, especially in 500rpm (low speed). Moreover, the comparative experiment with envelope and high pass filter is utilized to prove that SPCTF has an excellent ability to extract fault features from strong noise effectively. Furthermore, the comparative experiment with Kalman proves that the SKF has better real-time and dynamics to track the effective fault information in the signal. Additionally, it is effective to prove the advantage of PCA in order-reduction by comparing it with BTA, which can improve the efficiency of SPCTF when applied to fault diagnosis.

6). A new energy conditioning operator is proposed, the proposed ITEO is used to reduce the noise and enhance the transient features of the raw signal under the complex operating environments. Traditional forward difference is replaced by central-finite difference to smooth the raw signal, and the Hilbert transform is performed to redefine the analytic energy operator. ITEO enhances the robustness of the traditional TEO, and improves the feature extraction capability. Moreover, for low-speed bearing faults and early bearing faults, corresponding deep learning adaptive fault identification models are established.

1.5 Structure of this thesis

This study takes rolling bearing as the research object, and studies the problems of fault feature extraction of rolling bearing and early fault diagnosis of low-speed bearing, forming the overall research framework of the paper. Fault feature extraction is the key to rolling bearing fault diagnosis. For the tracking and extraction of rolling bearing fault features, a novel method called state-space principal component tracking filtering (SPCTF) is proposed in chapter 2， SPCTF can recover the fault information through the sequence of ‘Prediction – Measurement –Correction - Optimal estimation - Principal component extraction’ from the polluted raw signals with noise. In chapter 3，in view of the problems of low signal-to-noise ratio of bearings running at low speeds and weak impact characteristics of fault signals. An improved Teager energy operator (ITEO) is proposed to reduce the noise of the raw signals and improve the signal to noise ratio (SNR), which provides a good foundation for effective feature extraction.

Chapter 4 proposed an early rolling bearing fault diagnosis based on Hilbert difference teager energy operator (HDTEO) and ClassRestricted Boltzmann machine (CRBM) model adaptive symptom parameters extraction. The last chapter 5 is the summary and prospect of proposed method in this study.

2. Fault diagnosis of high-speed bearings

2.1 Introduction

The fault in rolling bearing (one of the essential components in industrial equipment) results in equipment dam-age, industrial shutdowns, extensive property losses and even casualties. An urgent need exists for an effective monitoring system to ensure equipment's high-precision and stable operation [54], [55]. Vibration signal analysis has been widely used in the fault detection of rolling bearing. However, due to the complex working environment of bearing, the low signal-to-noise ratio (SNR) of vibration signals causes difficulties in fault diagnosis.

Fortunately, the Kalman filter, is an efficient recursive filtering algorithm that can estimate and track the system's state from a series of incomplete sequences and complete the filtering of signals and the tracking extraction of fault features. Singleton [56] established a data-based Kalman filter fault diagnosis model by using time-domain and time-frequency domain features to track the development of bearing fault features. These features are used to train the parameters of the established model. Anger proposed a bearing fault diagnosis method based on the Gaussian degradation model of an unscented Kalman filter, which improved the dynamic response speed and anti-interference ability of fault diagnosis. However, a single Kalman model is not enough to describe the dynamical fault degradation process, which reduces the model's universality.

The switch Kalman filter algorithm can establish multiple filter models to describe the dynamic processes of a system more effectively. In recent years, it has been widely used in biomedicine, power systems, life prediction, and especially in the fault diagnosis field. Lim proposed three filter models based on SKF and Bayesian estimation to perform failure prognostics. Although the diagnostic accuracy is improved, the amount of data increases several times, demanding a high calculative burden. Consequently, the rapid development of model order reduction approaches [57] is sought after. Moreover, the data recovery and real-time problems account for much in the fault diagnosis field [58]. This chapter establishes a generative fault diagnosis model based on state-space principal component tracking filtering (SPCTF) to address these challenges.

2.2 Vibration signal analysis based fault diagnosis

2.2.1 Rolling bearing basic structure

The working principle of rolling bearings is to replace sliding friction with rolling friction, and its structure is generally a mechanical base composed of two rings, a set of rolling elements and a cage. As shown in Fig.2-1, the most basic structure of rolling bearing is generally composed of inner ring, outer ring, rolling elements and cage. In practical applications, the outer ring is generally fixed on the bearing seat or on the body. The inner ring of the bearing is connected to the drive shaft and rotates with the shaft. The rolling elements are evenly distributed between the inner ring and the outer ring with the help of the cage [59-61]. The shape, size and number of the rolling elements directly affect the performance and life of the rolling bearing. The cage can make the rolling elements evenly distributed and prevent the rolling elements from falling off. , to guide the rolling elements to rotate for lubrication.

2.2.2 Characteristics of rolling bearing fault vibration signal

1) Non-stationary properties

When a rolling bearings fails, the measured vibration signal is usually characterized by non-stationarity. Because the mechanical system will generate a large number of excitation sources with time-varying characteristics during operation, the measured vibration signal is non-stationary [62-64]. For rolling bearings, when the bearing fails, the fault defect will cause a transient impact, which has non-stationary characteristics. For rolling bearing fault diagnosis, fault feature extraction is a key step. The traditional signal processing method is based on Fourier transform, which has certain limitations. Therefore, it is necessary to study a feature extraction method corresponding to the non-stationary signal characteristics.

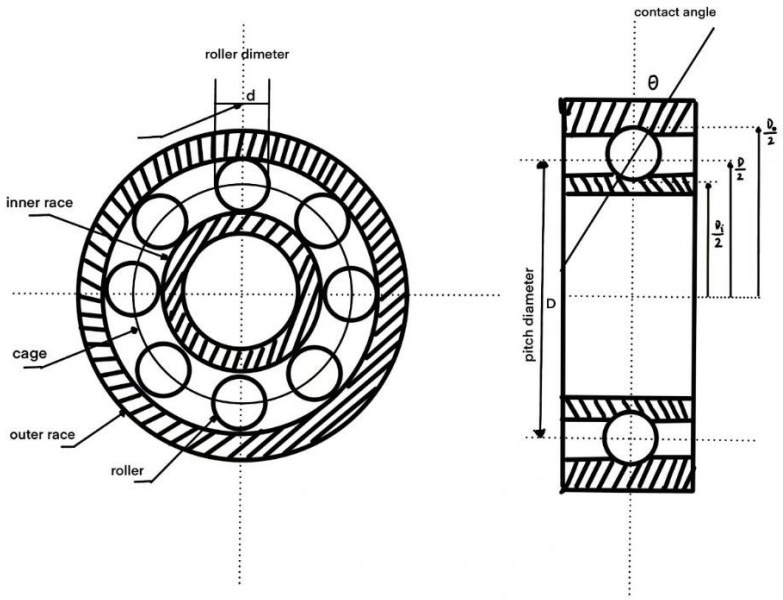


Fig.2-1 The diagrammatic sketch of the rolling bearing

2) Impact characteristics

When there is a partial failure of the rolling bearing, each contact during the operation will cause an instantaneous shock, resulting in a large number of shock waveforms in the measured vibration signal of the rolling bearing [65-68]. Therefore, the vibration signal of the rolling bearing exhibits a shock characteristic. However, when the local fault is small, the shock waveform caused by the fault is often easily submerged in the strong background noise, resulting in a low signal-to-noise ratio, and it is difficult to extract the shock features from the original signal, which brings great significance to the feature extraction of rolling bearings.

3) Modulation characteristics

The vibration signal of rolling bearing has typical modulation characteristics. The vibration signal caused by the local fault of the bearing can be regarded as a combination of low frequency part and high frequency part [69-72]. The vibration caused by the local fault of the bearing refers to the result of the modulation of the fault shock (the low frequency part) and the system resonant frequency (the high frequency part). Therefore, the fault signal of rolling bearing has obvious modulation characteristics, and it is necessary to study a suitable demodulation method to extract the fault impact of the vibration signal of rolling bearing.

To sum up, the vibration signal of rolling bearing has the characteristics of non-stationarity, nonlinearity, shock and modulation, which makes the vibration signal of

rolling bearing a complex signal that is difficult to deal with by traditional methods. Typically, bearing failure characteristics are closely related to rotational speed and failure severity. Under normal circumstances, when the bearing speed is higher or the fault is more serious, the periodic impact caused by the bearing fault is more obvious, and the harm to the operation of the equipment will also increase. If the abnormality can be detected in time when the bearing is running at a low speed, the maintenance cost of the equipment can be reduced, and the catastrophic accident caused by the development of the fault in the later stage can be effectively avoided. Therefore, low-speed running bearing fault detection is extremely important for the safe operation of mechanical equipment [73-76]. However, when the bearing is running at low speed, the vibration signal usually has the following characteristics: the fault feature is not obvious, the fault signal is weak and the signal-to-noise ratio is low. The reasons for the low signal-to-noise ratio of the bearing mainly include the following three aspects: First, during the low-speed operation of the bearing, due to the low rotational speed, the impact amplitude caused by the fault is small, and it is often easily overwhelmed by strong noise. Second, due to the bearing structure and installation method, the sensor cannot directly measure the bearing vibration signal [77-80]. The vibration shock caused by the fault often needs to be collected by the sensor through a complex transmission path, and the vibration will be mixed with a lot of noise during the transmission process. Third, the acquisition error will also be brought about in the process of vibration signal acquisition, such as the signal conversion and signal amplification process of the sensor. Therefore, the key problem of low-speed running bearing fault diagnosis is how to separate the shock waveform from the strong background noise and realize the effective extraction of early fault features of rolling bearings.

2.2.3 Rolling Bearing Failure Frequency

If the influence of assembly errors and external factors is not considered, bearing vibration is mainly divided into two categories: natural vibration of the bearing itself and vibration caused by bearing faults. The natural vibration of the bearing itself refers to the vibration caused by the inherent structure and physical characteristics of the bearing during the operation of the mechanical equipment, no matter whether the bearing is faulty or not, it will produce the attenuation movement of the natural frequency of the bearing element [81-85]. Vibration caused by bearing failure refers to the vibration generated when certain failures occur in the bearing (such as spalling,

pitting and wear of bearing components), such vibrations exhibit the characteristics of periodic transient shocks, and the location of bearing failures is different. Which will result in different shock characteristics. The vibration signal caused by bearing fault is an important data source and analysis basis for bearing fault diagnosis.

When the rolling bearing has a local fault during operation, each contact will cause an instantaneous impact, and stimulate the bearing and the system to perform high-frequency free damping vibration according to its natural frequency [86-92]. The instantaneous shock caused by the fault has obvious periodic characteristics, and the fault characteristic frequency of the shock is closely related to the parameters such as bearing speed, fault type and rotor number.

Assuming that the contact mode of each rolling element of the rolling bearing with the surface of the inner and outer rings is rolling friction contact, the bearing diameter corresponding to the center of the roller is called the bearing diameter. According to the geometric relationship given in Fig1-1, it can be known that the pitch diameter of the bearing is equal to:

$$\mathrm{D} = \frac{D_i + D_o}{2} \tag{2-1}$$

$D_i$is the diameter of the inner ring raceway of the bearing， $D_o$is the diameter of the outer ring raceway of the bearing, and the relationship between them and the diameter of the bearing can be expressed as:

$$D_i = D - d(\cos\theta) \tag{2-2}$$

$$D_o = D + d(\cos\theta) \tag{2-3}$$

Where d is the diameter of the roller, the rotational speed of the inner and outer raceways can be expressed as:：

$$V_i = \omega_i \frac{D_i}{2} \qquad V_o = \omega_o \frac{D_o}{2} \tag{2-4}$$

Where， ω_i is the rotational angular velocity of the inner raceway， ω_o is the rotational angular velocity of the outer raceway. If the rollers roll ideally on the raceway, the rotation speed of the cage is the average of the rotation speeds of the inner and outer races [93-98].

$$V_c = \frac{V_i + V_o}{2} = \frac{\omega_i(D - d(\cos\theta)) + \omega_o(D + d(\cos\theta))}{4} \tag{2-5}$$

It can also be expressed as:

$$f_c = \frac{f_i(1 - \frac{d}{D}(\cos\theta)) + f_o(1 + \frac{d}{D}(\cos\theta))}{2} \tag{2-6}$$

Normally, the outer ring of the bearing is stationary. Then the above formula can be simplified to:

$$f_c = \frac{f_i(1-\frac{d}{D}(\cos\theta))}{2} \quad (2\text{-}7)$$

The rolling frequency of the cage relative to the inner ring can be written as:

$$f_{ri} = f_c - f_i = \frac{(f_o-f_i)\,(1+\frac{d}{D}(\cos\theta))}{2} \quad (2\text{-}8)$$

Assuming that the number of rollers is Z, when the cage rolls one revolution relative to the inner ring, the rollers will pass through the fixed point of the inner ring Z times [99-105]. Then the passing frequency of the roller passing through the fixed point of the inner ring is:

$$f_{bpfi} = \frac{Z(f_o-f_i)(1+\frac{d}{D}(\cos\theta))}{2} \quad (2\text{-}9)$$

When the outer ring remains fixed, it simplifies to

$$f_{bpfi} = \frac{-Zf_i(1+\frac{d}{D}(\cos\theta))}{2} \quad (2\text{-}10)$$

The outer race fault characteristic frequency and the roller fault characteristic frequency can be calculated in turn, and the results are shown in Table 2-1.

Table 2-1 Bearing fault characteristic frequency calculation formula

Bearing parts	Calculation formula	Outer ring fixed
Inner race fault	$f_{bpfi} = \frac{Z(f_o-f_i)(1+\frac{d}{D}(\cos\theta))}{2}$	$f_{bpfi} = \frac{-Zf_i(1+\frac{d}{D}(\cos\theta))}{2}$
Outer race fault	$f_{bpfi} = \frac{Z(f_o-f_i)(1-\frac{d}{D}(\cos\theta))}{2}$	$f_{bpfi} = \frac{-Zf_i(1-\frac{d}{D}(\cos\theta))}{2}$
Roller fault	$f_{bsf} = \frac{f_o-f_i}{2}\frac{D}{d}\left\{1-\left[\frac{D}{d}\cos\theta\right]^2\right\}$	$f_{bsf} = -\frac{f_i}{2}\frac{D}{d}\left\{1-\left[\frac{D}{d}\cos\theta\right]^2\right\}$

It should be noted that due to the influence of external factors, in the actual bearing fault diagnosis, the measured fault frequency often deviates from the theoretical value [106-112]. Therefore, it is necessary to find the frequency value near the theoretical value as the fault frequency of the rolling bearing according to the actual situation.

2.3 Principle of the proposed algorithm

2.3.1 Kalman filter

A Kalman filter is an efficient recursive filtering algorithm, which estimates the

system state $\boldsymbol{X_n}$ through the state $\boldsymbol{Y_n}$ measured by the system [113] without relying on a large amount of system data. Define a stochastic system state $\boldsymbol{X_n}$, and then the stochastic difference equation and observation equation of the system can be expressed as:

$$X_n = F_{n,n-1}X_{n-1} + W_n \tag{2-11}$$

$$Y_n = C_nX_n + V_n \tag{2-12}$$

where $\boldsymbol{X_n}$ is the state vector of the system at a discrete time point $\boldsymbol{n}$; $\boldsymbol{Y_n}$ is the observation vector of discrete time point n; The matrix $\boldsymbol{F_{n,n-1}}$ and $\boldsymbol{C_n}$ are state transition matrix and observation matrix, respectively. The vectors $\boldsymbol{W_n}$ and $\boldsymbol{V_n}$ are uncorrelated zero mean Gaussian white noise, $\boldsymbol{W_n}\sim(\boldsymbol{0}, \boldsymbol{Q1_n})$ and $\boldsymbol{V_n}\sim(\boldsymbol{0}, \boldsymbol{Q2_n})$. The primary process of Kalman filtering is as follows:

State vector prediction equation:

$$\hat{X}_n = F_{n,n-1}X_{n-1} \tag{2-13}$$

where X_{n-1} is optimal estimation at discrete time point $n-1$.

State prediction correlation matrix:

$$\hat{P}_n = AF_{n,n-1}^T + Q1_n$$

$$A = F_{n,n-1}P_{n-1} \tag{2-14}$$

Filter gain:

$$G_n = \hat{P}_nC_n^T[\tau_n]^{-1}$$

$$\boldsymbol{\tau_n = C_n\hat{P}_nC_n^T + Q2_n} \tag{2-15}$$

Innovation process:

$$\partial_n = Y_n - C_n\hat{X}_n \tag{2-16}$$

One-step prediction and error correlation matrix:

$$X_n = \hat{X}_n + G_n\partial_n \tag{2-17}$$

$$P_n = \hat{P}_n - G_nC_n\hat{P}_n \tag{2-18}$$

Kalman filter estimation uses (3) and (4) to obtain prior estimations that are corrected through (5)-(6) to obtain the optimal estimation (7) and corresponding estimation error (8). The relevant prediction parameters are then updated according to the returned estimation error for the discrete time point n+1 system state prediction.

Kalman filters can use a small number of prior samples to predict the system state for removing the system noise. However, the accuracy of Kalman filters estimation will decrease with the increase of estimation step since the degradation of bearing is a dynamic process. Therefore, the switch Kalman filter algorithm is introduced to realize dynamic tracking of bearing fault features to solve the limitation of Kalman filters in

bearing fault diagnosis.

2.3.2 Switch kalman filter

The switch Kalman based on Baysian is shown in Fig.2-2. Firstly, the various Kalman filter models are established by analyzing various possible states of the system. The probability of each state at each moment is calculated, and then the most likely state of the system at that moment is analyzed. The detailed process is as follows:

There are m Kalman filter models to be established for a dynamic system based on Bayesian estimation theory. The model transformation probability:

$$S_n^{i|j} = \frac{Z_{ij} S_{n-1}^i}{\sum_{i=1}^m Z_{ij} S_{n-1}^i} \tag{2-19}$$

where the model switch variable $S_n^{i|j}$ describes the probability of the model switching from S_{n-1}^i to S_n^j. S_{n-1}^i is the probability of the model i in discrete time point n-1. Z_{ij} is the model transition probability. The weighted state and covariance estimates are:

$$\hat{X}_{n-1}^j = \sum_{i=1}^m S_n^{i|j} X_{n-1}^i \tag{2-20}$$

$$\hat{P}_{n-1}^j = \sum_{i=1}^m S_n^{i|j} \{P_{n-1}^i + [X_{n-1}^i - X_{n-1}^j][X_{n-1}^i - X_{n-1}^j]^T\} \tag{2-21}$$

where X_{n-1}^i is the state vector prediction value of ith model at discrete time point n-1. P_{n-1}^i is the error correlation matrix of the ith model at discrete time point n-1. Accordingly, the Likelihood estimation of each filter model:

$$L_n^i = N\left(\partial_n^i; 0, \tau_n^i\right),\ \partial_n = Y_n - C_n \hat{X}_n,\ \tau_n = C\hat{P}C^T + Q2_{(n)} \tag{2-22}$$

where ∂_n^i is the measurement residual, and τ_n^i the residual covariance. Probability of each model at discrete point n:

$$S_n^i = \frac{L_n^i \left(\sum_{i=1}^m Z_{ij} S_{n-1}^i\right)}{\sum_{i=1}^m \left(L_n^i \sum_{i=1}^m Z_{ij} S_{n-1}^i\right)} \tag{2-23}$$

Finally, the weighted state and covariance estimate updates are computed as follows:

$$X_n = \sum_{i=1}^m S_n^i X_n^i,\ P_n = \sum_{i=1}^m S_n^i \{P_n^i [X_n^i - X_n][X_n^i - X_n]^T\} \tag{2-24}$$

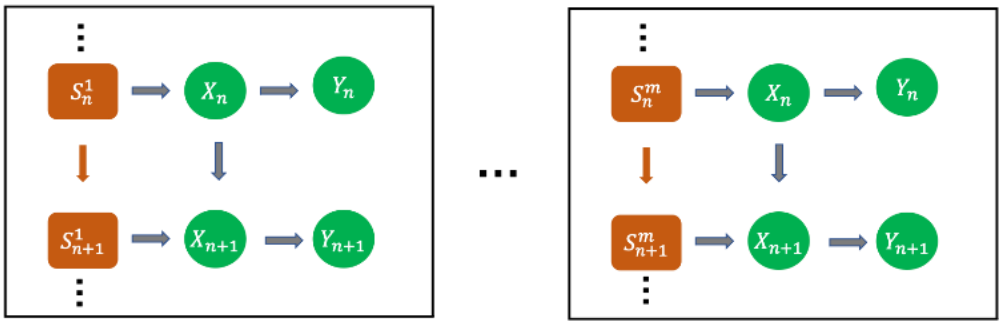


Fig.2-2 The switching Kalman is based on Baysian methods.

2.3.3 Feature extraction of bearing fault

The research shows that a pulse force will be excited when the rolling element rolls over the defect, such as pitting, spalling and scratching. The pulse force is a broadband signal, which can stimulate the high-frequency natural vibration of the bearing system, as shown in Fig. 2-3.

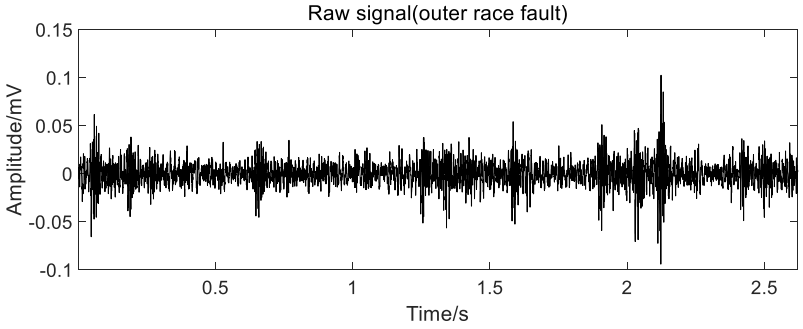


Fig.2-3. The vibration signals of fault bearings.

Recently, the time series modeling autoregressive (AR) model, known as the high-resolution parametric spectrum analysis method and capable of revealing the occurrence of rotating machinery fault, has been applied to vibration signal analysis. Therefore, the AR is utilized to establish an autoregressive state equation for the switched Kalman to obtain the optimal estimated state matrix.

Considering $\boldsymbol{A_n} = [\boldsymbol{a_{1(n)}}, \boldsymbol{a_{2(n)}}, \ldots \boldsymbol{a_{p(n)}}]$ as the weighted parameter of stochastic difference equation, the estimate $\boldsymbol{x_n}$ can be described:

$$x_n = \sum_{k=1}^{m} a_p x_{n-p} + e_n \tag{2-25}$$

where $\boldsymbol{e_n}$ is the optimal estimation error.

Conduct $\boldsymbol{X_n} = [\boldsymbol{x_{n-1}}, \boldsymbol{x_{n-2}}, \ldots \boldsymbol{x_{n-p}}]$ as the input vector (bearing vibration signal) of SKF at discrete time point n, taking into account the characteristics that the optimal

parameters will change with iteration, the process noise $\boldsymbol{V_n}$ is introduced. The expected output of each SKF model $\boldsymbol{x_n = X_n A_n + V_n}$. Furthermore, $\boldsymbol{A_n}$ is the key parameter to be estimated in the process of applying SKF into fault diagnosis. The state equation and prediction equation is:

$$\boldsymbol{A_n = F_{n,n-1} A_{n-1} + W_n, x_n = X_n A_n + V_n} \quad (2\text{-}26)$$

According to (1, 2), $\boldsymbol{W_n}$ and $\boldsymbol{V_n}$ are uncorrelated zero mean Gaussian white noise, $\boldsymbol{W_n \sim (0, Q1_n)}$ and $\boldsymbol{V_n \sim (0, Q2_n)}$.

Initialize $\boldsymbol{F_n}$ is the identity matrix, $\boldsymbol{q = 10^{-4}}$ is the variance of $\boldsymbol{W_1}$, and thus $\boldsymbol{Q1_1 = qI}$, taking the variance of $\boldsymbol{X_1 = [x_1, x_2, \ldots x_p]}$ as $\boldsymbol{V_1}$.The final state vector $\boldsymbol{A_n}$ will converge to its best estimate value through performing multiple iterations and operations on the filter model with (2-25).

SKF can not only effectively remove the interference of noise, but the state vector-matrix $\boldsymbol{A_n}$ retains the key fault features of the signal during the updating of the SKF model. However, the $\boldsymbol{A_n}$ will increase the data dimension by p times. Thus PCA is used to extract $\boldsymbol{\mu}$ principal component parameter features from $\boldsymbol{A_n}$. It not only reduces the data dimensions but also retains the useful fault features. The detailed process is illustrated schematically in Fig. 2-4.

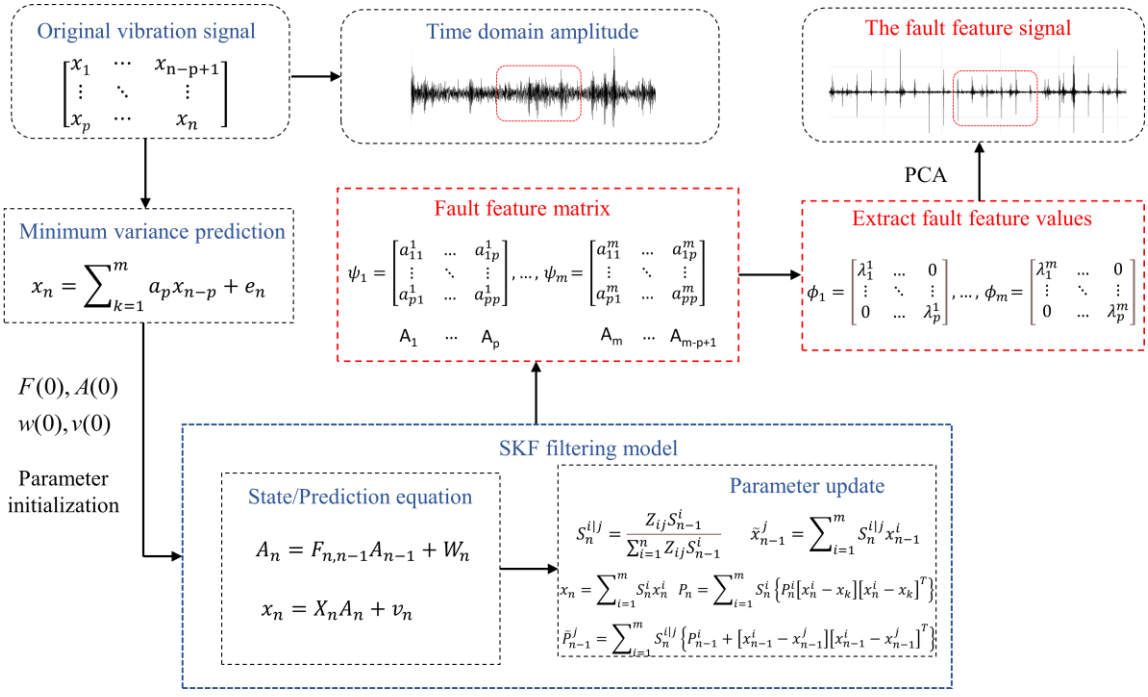


Fig.2-4 SKF based fault feature extraction of bearing fault signals.

2.4 Bearing fault diagnosis based on SPCTF

In this section, a novel method called state-space principal component tracking filtering (SPCTF) is described to use in fault diagnosis. Firstly, the switch Kalman filter (SKF) is utilized to establish a dynamic filter model for time-series signals. Then a feature tracking matrix is generated from the principal component features extracted from the optimal estimated state matrix of each time in the model. The optimal estimated state matrix is calculated with the optimal estimation at the previous time and the observation at the present time of the system. Thirdly, the principal component analysis extracts the effective fault feature from the redundant information generated in the feature tracking matrix. Therefore, the proposed method can effectively eliminate the random interference to recover fault features from polluted signals through the sequence of "Prediction – Measurement – Correction - Optimal estimation - Principal component extraction". Finally, the envelope spectrum analysis is used to extract fault characteristic frequencies to achieve high-precision fault diagnosis. The scheme of the proposed method based on SPCTF is briefly described in Fig.2-5.

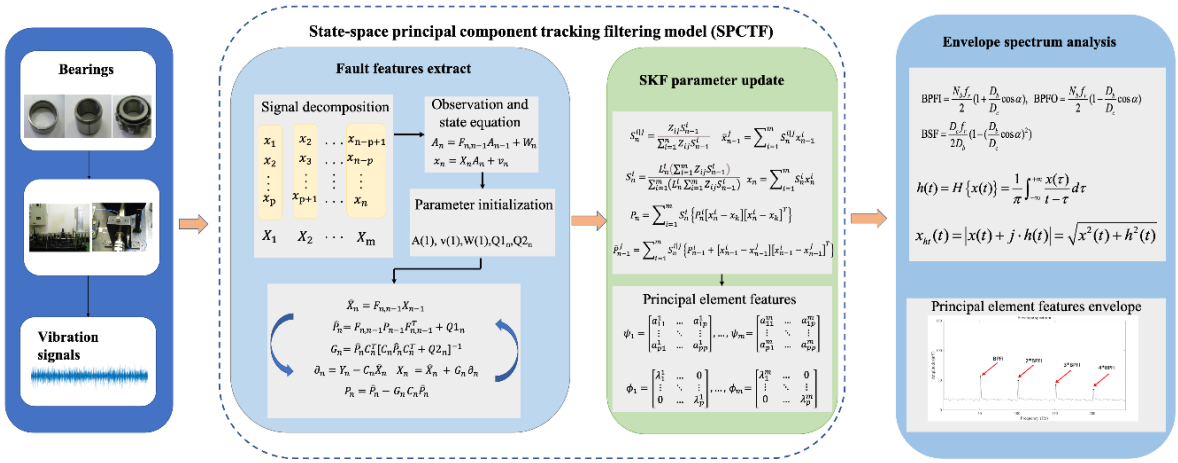


Fig.2-5. Scheme of the proposed SPCTF for bearing fault diagnosis

2.5 Simulated signal verification

2.5.1 The simulation signals

In order to verify the accuracy and reliability of the proposed method, there are two impulse response simulation signals as shown in (17) and (18). They are the raw outer race fault and outer race fault with noise, respectively.

$$x_o(t) = \sum_{i=1}^{N_o} \begin{matrix}[e^{-\zeta_o 2\pi f_{o1}\left(t-\frac{i}{f_o}\right)} sin\left(2\pi f_{o1}\left(t-\frac{i}{f_o}\right)\right)\sqrt{1-\zeta_o^2} + \\ e^{-\zeta_o 2\pi f_{o2}\left(t-\frac{i}{f_o}\right)} sin(2\pi f_{o2}\left(t-\frac{i}{f_o}\right))\sqrt{1-\zeta_o^2}]\end{matrix} \quad (2\text{-}27)$$

$$x_o(t) = \sum_{i=1}^{N_o} \begin{matrix}[e^{-\zeta_o 2\pi f_{o1}\left(t-\frac{i}{f_o}\right)} sin\left(2\pi f_{o1}\left(t-\frac{i}{f_o}\right)\right)\sqrt{1-\zeta_o^2} + \\ e^{-\zeta_o 2\pi f_{o2}\left(t-\frac{i}{f_o}\right)} sin(2\pi f_{o2}\left(t-\frac{i}{f_o}\right))\sqrt{1-\zeta_o^2}]\end{matrix} + w \quad (2\text{-}28)$$

where the sampling frequency is $\boldsymbol{f_o}$= 20000 Hz, and the number of sampling points is $\boldsymbol{N_o}$= 16384. $\boldsymbol{x_o(t)}$ is the simulation signal of a rolling bearing with an outer race fault, and the frequencies of its carrier centers are $\boldsymbol{f_{o1}}$= 2000 Hz and $\boldsymbol{f_{o2}}$=5200Hz. The damping ratio is $\boldsymbol{\zeta}$=0.05. $\mathbf{w} = \mathbf{0.8} \times \boldsymbol{random}(\mathbf{1}, \boldsymbol{n})$ is the random white noise. The simulated signal according to (17) and (18) is shown in Fig.2-6.

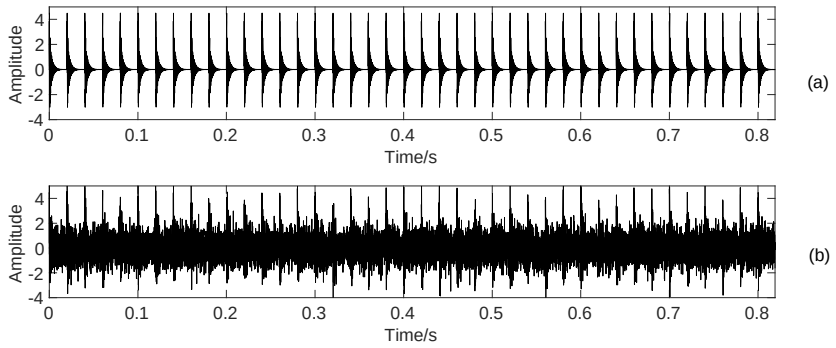


Fig.2-6. The simulation fault signals. (a) Raw outer race fault; (b) Outer race fault with noise.

2.5.2 Results and discussion

This section explores comparative experiments, including envelope analysis and the reduced-order aggregate method based on the balanced truncation approach (BTA), which is implemented to prove the effectiveness of the proposed method. The process of the proposed method is as follows:

(i) 16384 data points are divided into 16380 groups. Each group contains $\boldsymbol{p = 5}$ data points (The number $\boldsymbol{p}$ is obtained from experience). (ii) the system state and prediction equations of the principal component tracking filter model are established (16). $\boldsymbol{A = [a_1, a_2, a_3, a_4, a_5]^T}$ is retained through 16380 generations with (9)-(14). (iii) the PCA is used to complete the final fault feature extraction and generate a fault feature matrix. (iv) the extracted fault features are analyzed by an envelope spectrum to achieve high-precision fault diagnoses.

1) Evaluation metric.

There are two dimension parameters and three dimensionless parameters,

including maximum value, peak-to-peak value, kurtosis, skewness and crest factor that are selected to be treated as an evaluation metric. The parameters are calculated according to Table 2-2. The $\boldsymbol{x_{min}}$ is the minimum value of the data.

Table 2-2 Dimensionless parameters.

Symbol	Parameter	Formula
K	kurtosis	$\frac{\mathrm{E}[(X-\mu)^4]}{(\mathrm{E}[(X-\mu)^2])^2}$
S	skewness	$\frac{\mathrm{E}[(X-\mu)^3]}{(\mathrm{E}[(X-\mu)^2])^{3/2}}$
C	crest	$\frac{x_{peak}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}}$
Max	maximum	x_{peak}
Ptp	peak to peak	$x_{peak} - x_{min}$

2) Experimental results

Table 2-3 shows the comparison of dimension and dimensionless parameter characteristics of fault simulation signals processed by SPCTF. As a result, the evaluation metric representing impact components has changed significantly after pre-processing with SPCTF. Experiments show that the interference components in all signals are well suppressed, and the proposed method is effective for the bearing fault diagnosis.

Table 2-3 SPCTF dimensionless parameters comparison.

Signal	K	S	C	Max	Ptp
Raw (17)	9.379	1.397	8.453	4.46	8.29
SPCTF	19.81↑	1.500↑	12.36↑	5.7↑	10.7↑
Noisy (18)	2.716	0.2415	5.623	4.75	9.35
SPCTF	5.448↑	0.3378↑	11.03↑	5.21↑	13.17↑

To further verify the superiority of the proposed method, the signal of state space is compared with the raw signal (out race fault simulation signals with noise). As shown in Fig.2-7, the model tracking signals established by SPCTF can reduce the noise and retain the valuable impact characteristics of the fault signal, which lays a good foundation for the next step to extract the effective fault features.

The first principal component of PCA is retained to extract fault features. Fig.2-8

shows the cumulative contribution rate of different principal components in the PCA model in bearing fault simulation signal with noise. It can be concluded that the contribution rate of the first principal component has reached 89.3%, which contains most of the valid information in the model. Therefore, the first principal component is used as a fault signal.

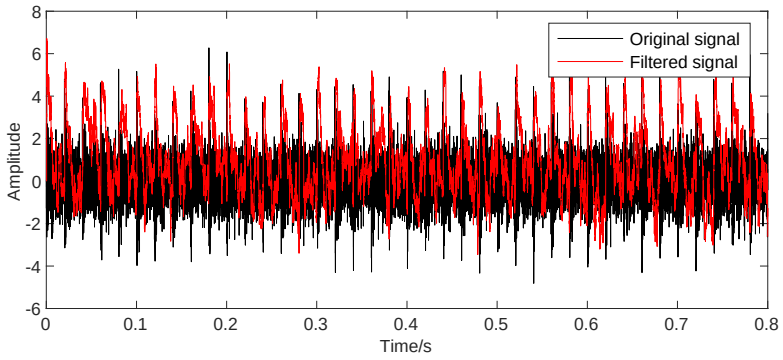


Fig.2-7 The SPCTF tracking simulation signals comparison.

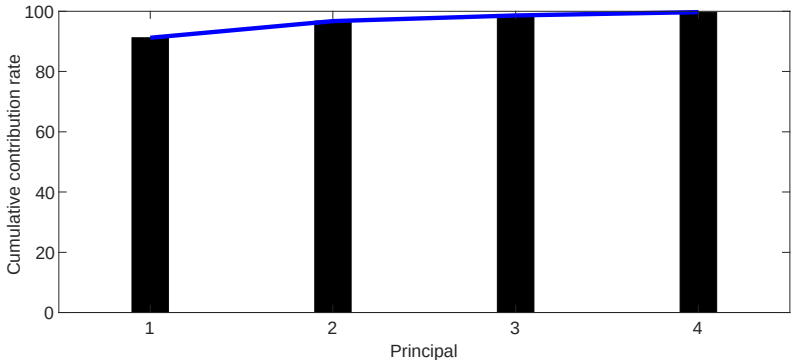


Fig.2-8 Cumulative contribution rate.

Furthermore, the original simulation signal, SPCTF tracking signal and the first principal component characteristic signal are shown in Fig. 2-9. The results show that the first principal component accurately reflects the fault characteristics of the signal extracted by the principal component tracking filter model. The above experimental results confirm that SPCTF can effectively track the system's state without pre-information and effectively extracts fault information embedded within extensive background noise, based on anti-interference ability and real-time performance.

3) Comparative experiment

The comparative experiment with envelope is performed to validate the effectiveness of the proposed method, with the envelope enabling analysis of the raw signal. The experimental results are shown in Fig.2-10, where (a) the eigenvalues of the

principal component of the fault characteristic matrix have evident periodic burst; the fault characteristic frequency (50 Hz) is vividly shown in Fig.2-10(b). The outer race fault simulation signal with noise is used to carry out the fault diagnosis for the proposed method's effectiveness with high background noise levels. As shown in Fig. 2-10(c), the eigenvalues of principal components of the signal fault characteristic matrix are affected by noise.

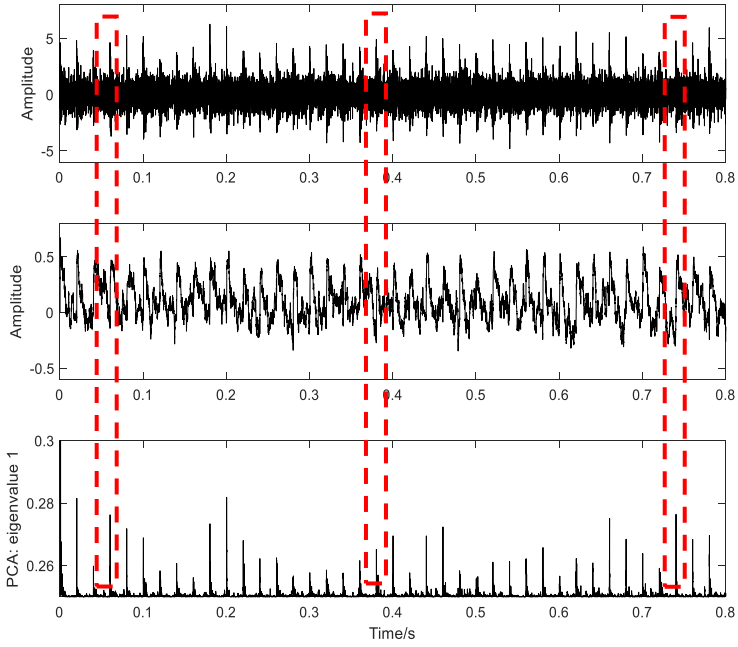


Fig.2-9.The fault bearing signals. (a) Original simulation signal; (b) SPCTF tracking signal; (c) First principal component characteristic signal.

The envelope spectrum analysis results of principal components are shown in Fig. 2-10(d), with obvious peaks at the fault frequency and its multiple frequencies (50 Hz, 100 Hz, 150 Hz and 200 Hz) from the out-race fault bearings, despite the noise causing independent amplitude vibrations of the signal. Fig.2-9(e) shows the envelope analysis of the unprocessed noise signals, where identifying the fault characteristic frequency is challenging. The results show that the proposed method is effective for bearing fault diagnoses.

Moreover, the reduced-order method based on principal components analysis is utilized to extract the first principal component of the optimal estimated state matrix, treated as a feature tracking matrix for generating characteristic parameters. For fairness, the proposed method is compared with the reduced-order aggregate model based on the balanced truncation approach (BTA). The result is shown in Fig.2-11, where the noise

signal is more evident following BTA, and the bearing fault features can be extracted. These experimental results are due to the advantage of PCA where: (i) PCA only needs to measure the amount of information by variance, which is not affected by factors other than the data set, and (ii) The interactive factors between the components of the original data can be eliminated since the principal components are orthogonal.

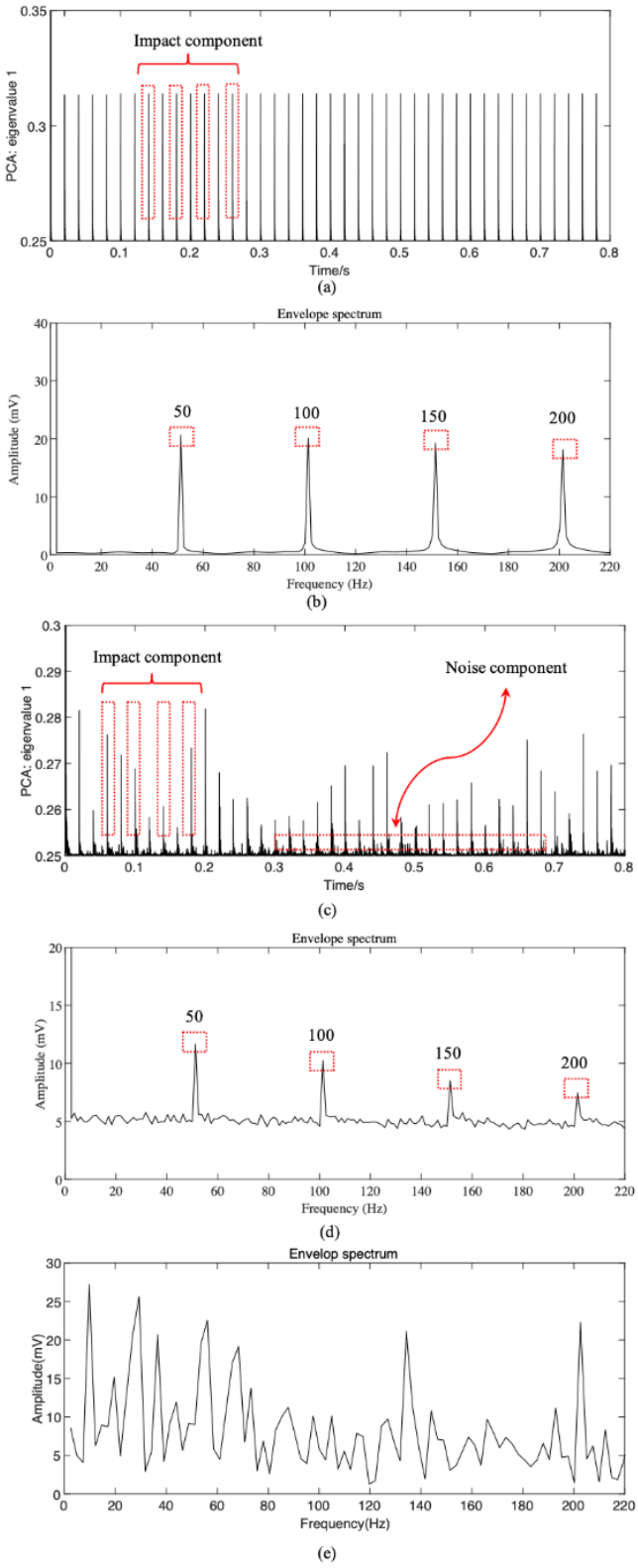


Fig.2-10 The results shown stepwise, as per the proposed method. (a) The principal component eigenvalue of fault signals without noise; (b) The envelope analysis of principal component value; (c) The principal component eigenvalue of fault signals

(noise); (d) The envelope analysis of principal component value. (e) The envelope analysis of noise signals.

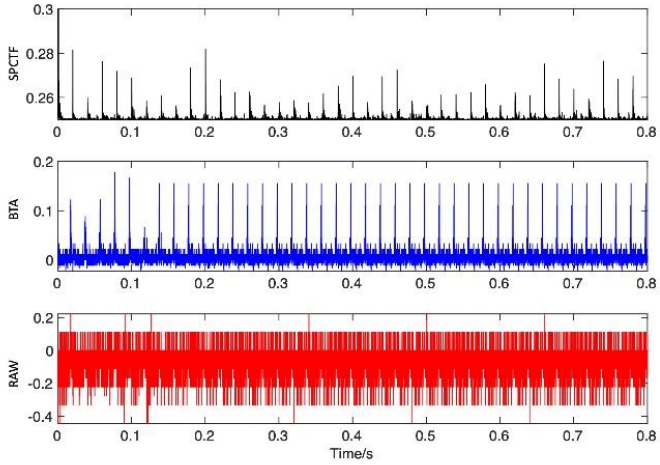


Fig.2-11 Comparative experiment using the reduced-order method.

2.6 Case study

2.6.1 Experiment platforms

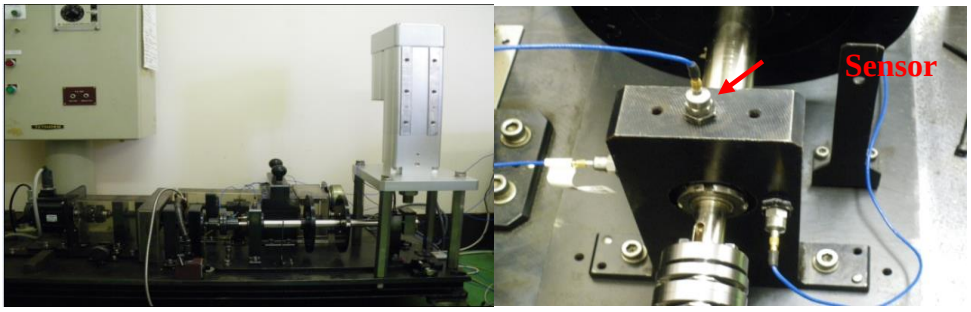
In order to verify the effectiveness of SPCTF, raw signals of bearings in different states (normal, outer race fault, inner race fault, and roller fault) were collected from the experimental platform, as shown in Fig.2-12 (a). The experimental device is composed of a servo motor, load equipment, coupling and a rotor system. There are three kinds of fault in the bearing (NSK 205), as shown in Fig.2-12(b). The bearing failure is due to manual cutting, the depth of damage is 0.25 mm, and the width is 0.3 mm.

Vibration signals of bearings due to different types of faults are measured by the acceleration sensor (PCB ma352a60, PCB Piezotronics Inc., New York, NY, USA). The signals sampling frequency is 100 kHz， and the sampling time was approximately 3s. The characteristic fault frequencies are calculated according to Apana et al. The fault characteristic frequencies of the inner race fault are shown in Table 2-4. There are three kinds of rotating speeds in this experimental design, 500, 1000, and 1500 rpm.

Table 2-4

Fault characteristic frequency (Inner 500 RPM)

Fault	Outer race	Inner race	Roller
Characteristic frequency	36.59Hz	55.07Hz	39.67Hz



(a)

(b)

Fig.2-12. Experimental platform. (a) The experimental devices; (b) Faulty bearings.

2.6.2 Intelligent diagnosis based on proposed method

Both two dimensional and three dimensionless parameters, including maximum value, peak-to-peak value, kurtosis, skewness and crest factors, are selected to be used as evaluation metrics (inner race fault signals is treated as an example). The comparative results for SPCTF and BTA are listed in Table 2-5. Although BTA is outstanding in the large-scale converters field, the increase of fault signal parameters characteristic index is more resilient using SPCTF.

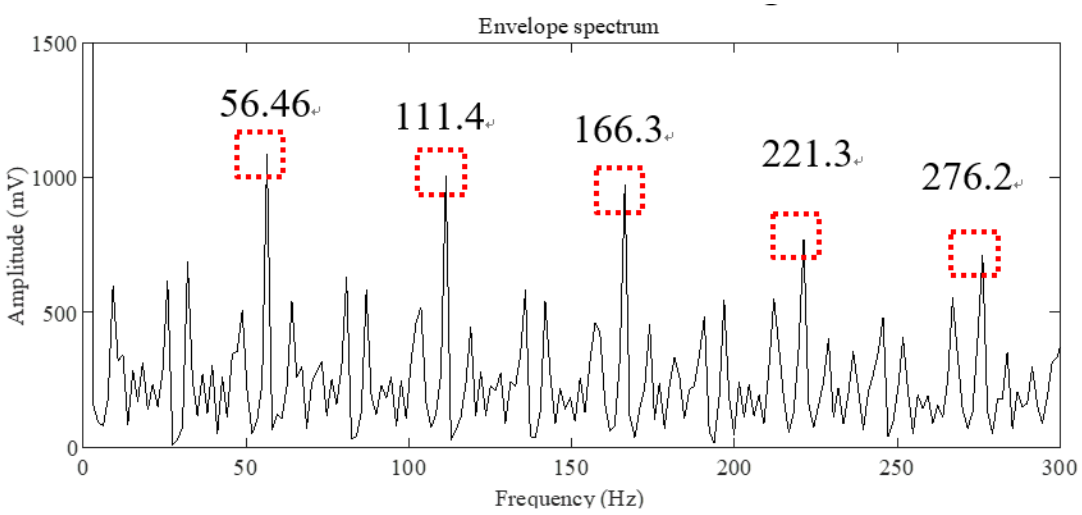


Fig.2-13. Inner race fault envelope spectrum.

Table 2-5

Dimensionless parameters of inner race fault signals.

	K	S	C	Max	Ptp
RAW	1.0558	0.7709	7.1191	11.988	29.392
SPCTF	4.9474↑	2.4499↑	9.8833↑	14.281↑	33.275↑
BTA	2.339	1.358	8.369	13.279	31.254

Moreover, the envelope analysis of bearing inner race fault characteristic signals are shown in Fig. 13 and Table 2-6. The calculated fault frequency of the inner race fault signal is 55.1 Hz, 110.1Hz, 165.2 Hz, 220.3 Hz and 275.4 Hz. The characteristic fault frequencies obtained with SPCTF are 56.5 Hz, 111.4 Hz, 166.3 Hz, 221.3 Hz, and 276.2 Hz, in line with the calculated failure frequency of bearing inner race. As a result of the low signal-to-noise ratio of the raw bearing inner race fault signal, the envelope spectrum of the original signal is incapable of quickly identifying the fault characteristics. The effectiveness of the proposed method is validated using this simple system.

Table 2-6

Characteristic frequency comparison (Inner 500 RPM)

Fault bearings	Characteristic frequency	Signal	Double	Triple	Quadruple	Quintuple
Inner race fault	Theoretical value/Hz	55.1	110.1	165.2	220.3	275.4
	SPCTF/Hz	56.5	111.4	166.3	221.3	276.2
	Envelop/Hz	79.6	135.7	149.8	183.8	235.9

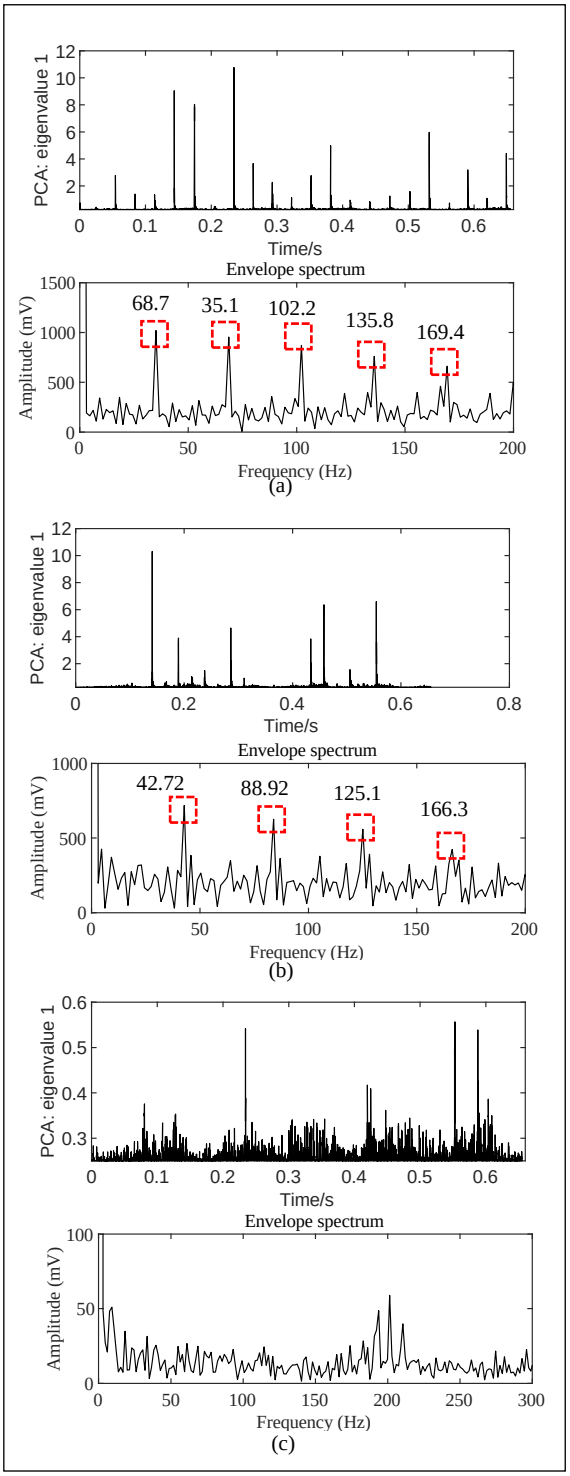


Fig.2-14 Results of the experiment. (a) Principal component characteristics of the outer race and its envelope spectrum; (b) Principal component characteristics of the roller and its envelope spectrum. (c) Principal component characteristics of typical signals and their envelope spectrum.

Table 2-7

Characteristic frequency comparison (500 RPM)

Fault bearings	Characteristic frequency	Signal	Double	Triple	Quadruple	Quintuple
Outer race fault	Theoretical value/Hz	36.59	73.18	109.77	146.36	182.95
	Experimental value/Hz	35.1	68.66	102.2	135.8	169.4
Roller race fault	Theoretical value/Hz	39.67	79.34	119.01	158.68	198.38
	Experimental value/Hz	42.72	83.92	125.1	166.3	209

The outer race fault, roller fault and typical rolling bearings signals are analyzed, and the results are shown in Fig.2-14 and Table 2-7. The experimental results show that the bearing fault types can be clearly distinguished by analyzing the envelope spectrum of the fault characteristic principal component matrix extracted from SPCTF, which again validates the effectiveness of the proposed method for experimental and practical bearing signal analyses.

2.6.3 Comparative experiment

To verify the effectiveness and feasibility of the proposed method, Hilbert spectrum and high pass filtering methods were used for comparative experiments. The Hilbert spectrum performs a Hilbert spectrum analysis on the raw signal filtered by Kalman and the proposed SPCTF. Moreover, the high pass filtering diagnoses and analyzes different fault signals types (outer race, inner race, roller) at three speeds (500, 1000, and 1500rpm).

1) Hilbert spectrum experiment.

The Hilbert spectrum decomposes the original signal using EMD and then calculates its Hilbert and Hilbert marginal spectra. The Hilbert spectrum emphasizes the local nature of the signal, which has a more intuitive physical meaning. The specific algorithm is as follows：

- Perform EMD decomposition on the original signal to obtain the eigenmode components；
- Hilbert transform for each eigenmode component c_i；

$$\hat{c}_i(t) = \frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{c_i(\tau)}{t-\tau}d\tau \qquad (2\text{-}29)$$

- Construct an analytical signal $z_i(t) = c_i(t) + j\hat{c}_i(t) = a_i(t)e^{j\varphi_i(t)}$, where $a_i(t)$

is the instantaneous amplitude and $\varphi_i(t)$ is the instantaneous phase function. Ignoring the residuals r_n , we get:

$$x(t) = \mathrm{Re}\sum_{i=1}^{n} a_i(t)e^{j\varphi_i(t)} \quad (2\text{-}30)$$

- Calculate the instantaneous frequency:

$$f_i(t) = \frac{1}{2\pi} w_i(t) = \frac{1}{2\pi}\frac{d\varphi_i(t)}{dt} \quad (2\text{-}31)$$

- Then we get the Hilbert spectrum

$$H(w,t) = \mathrm{Re}\sum_{i=1}^{n} a_i(t)e^{j\int w_i(t)dt} \quad (2\text{-}32)$$

- Calculate the integral of the Hilbert spectrum on the time axis to get the Hilbert marginal spectrum:

$$H(w) = \int_0^T H(w,t)dt \quad (2\text{-}33)$$

The comparative experiment, based on the Hilbert spectrum, is divided into three groups. The data used in the experiment are all under 500 rpm, and each sample length is 16384 points, of which the first 8192 data are the outer race fault signals, and the last 8192 data are the inner race fault signals. Notably, Label 1 represents the raw signals, Label 2 represents the signals processed using traditional Kalman filtering, and Label 3 represents the signal processing by SPCTF. It intuitively identifies the fault type through the Hilbert contour distribution of different kinds of fault signals. The experimental results are shown in Fig.2-15, 2-16, 2-17. As shown in Fig.2-15, there are many false components in the spectrum of Hilbert signals due to the influence of system noise, which makes the distribution of the spectrum contour map scattered. The focus of the fault frequency is extremely poor, and the outer race and inner ring faults cannot be effectively identified.

As shown in Fig.2-16, the main frequency band of the contour map of bearing outer race fault frequency is concentrated at 0.5-0.8 Hz, and the primary frequency band of bearing inner race fault frequency is focused at 0.2-0.8 Hz. Compared with the raw signal, the fault type can be differentiated using the traditional Kalman filter, but a mixed area still causes the fault type to be inaccurately determined. The Kalman filter cannot realize the dynamic tracking system state since it is easily affected by the time step. As shown in Fig.2-17, the main frequency band of the Hilbert contour map of the outer race fault is 0.5 Hz-0.8 Hz, while that of the inner race fault is 0.2 Hz-0.5 Hz, clearly identifying the fault type. Therefore, the proposed SPCTF method can

effectively remove the extensive noise interference and enable high-precision fault diagnosis.

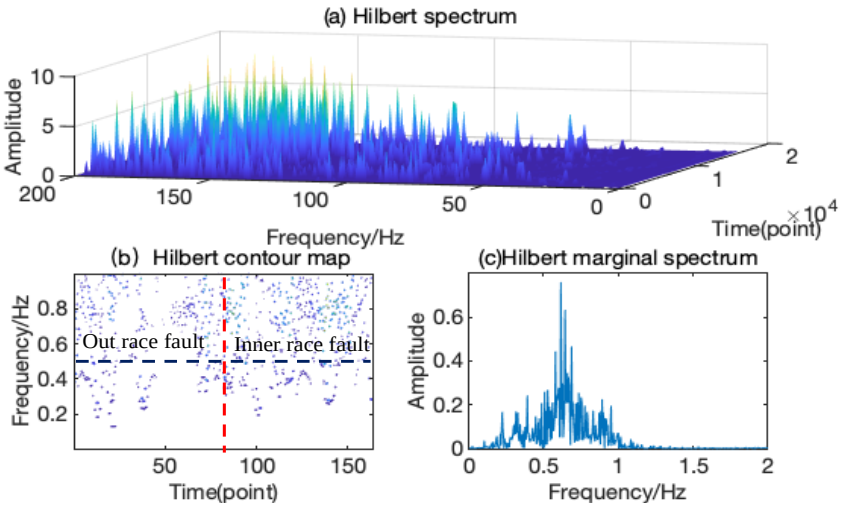


Fig.2-15. The Hilbert spectrum result shows raw bearing signals. (a)Hilbert spectrum; (b) Hilbert contour map; (c) Hilbert marginal spectrum (label 1).

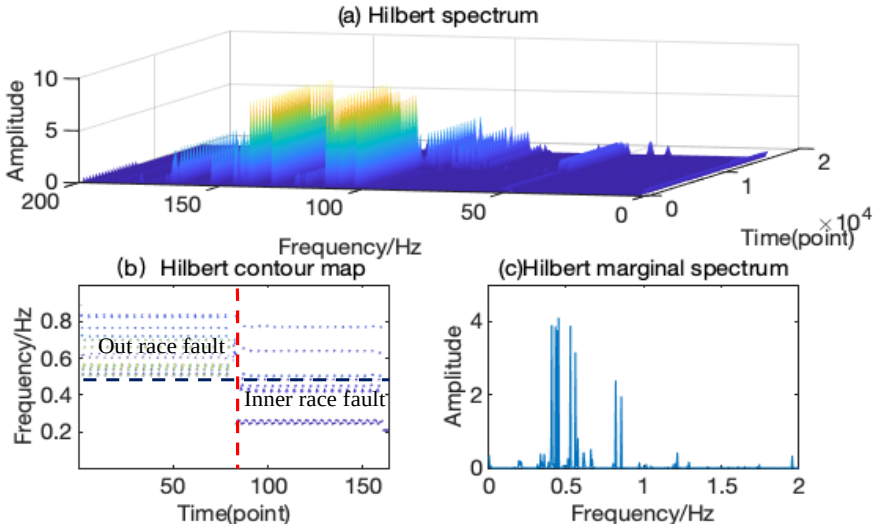


Fig.2-16. The Hilbert spectrum result of Kalman processed signals. (a)Hilbert spectrum; (b) Hilbert contour map; (c) Hilbert marginal spectrum (label 2).

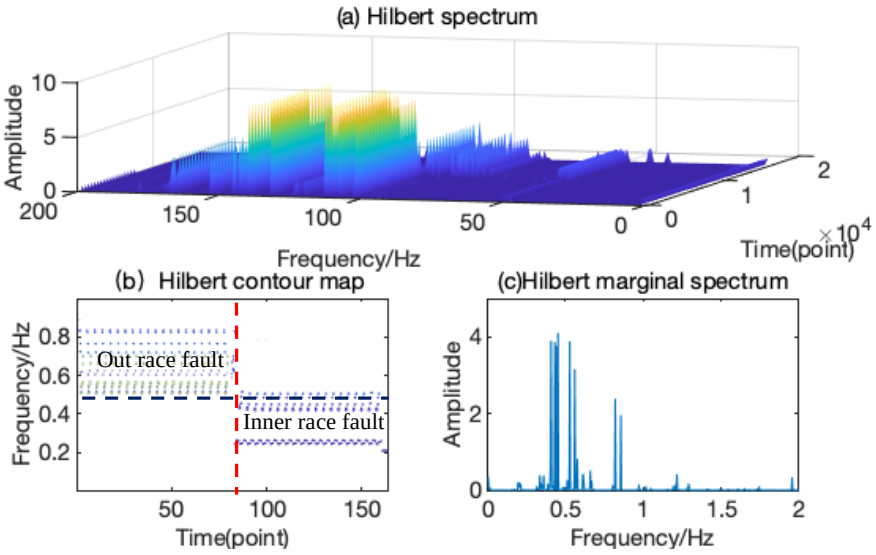


Fig.2-17. The Hilbert spectrum result of SPCTF processed signals. (a)Hilbert spectrum; (b) Hilbert contour map; (c) Hilbert marginal spectrum (label 3).

2) High pass filter experiment

The high pass filter is used for comparative experiments since the proposed algorithm was derived from the filtering method. The experimental data include various bearing fault states: the outer race, inner race, and roller. Generalization is an essential indicator of any diagnosis when changing the working conditions. It is difficult at low speed to perform the bearing fault diagnosis. Moreover, in the case of varying speeds, the operation affects the vibration signal from the bearings, which makes diagnoses difficult [34]. Therefore, speed variation is treated as an experimental condition change for verifying the generalized ability of the proposed method, including low speed (500rpm) and higher speeds (1000 and 1500rpm). The fault characteristic frequency corresponding to the bearing fault state under various speeds is calculated using empirical equations and are listed in Table 2-8. In this experiment, the data sample length of the bearing signal is 16384, and the high pass filter frequency is 2000Hz. The experimental results are shown in Fig2-18, 2-19.

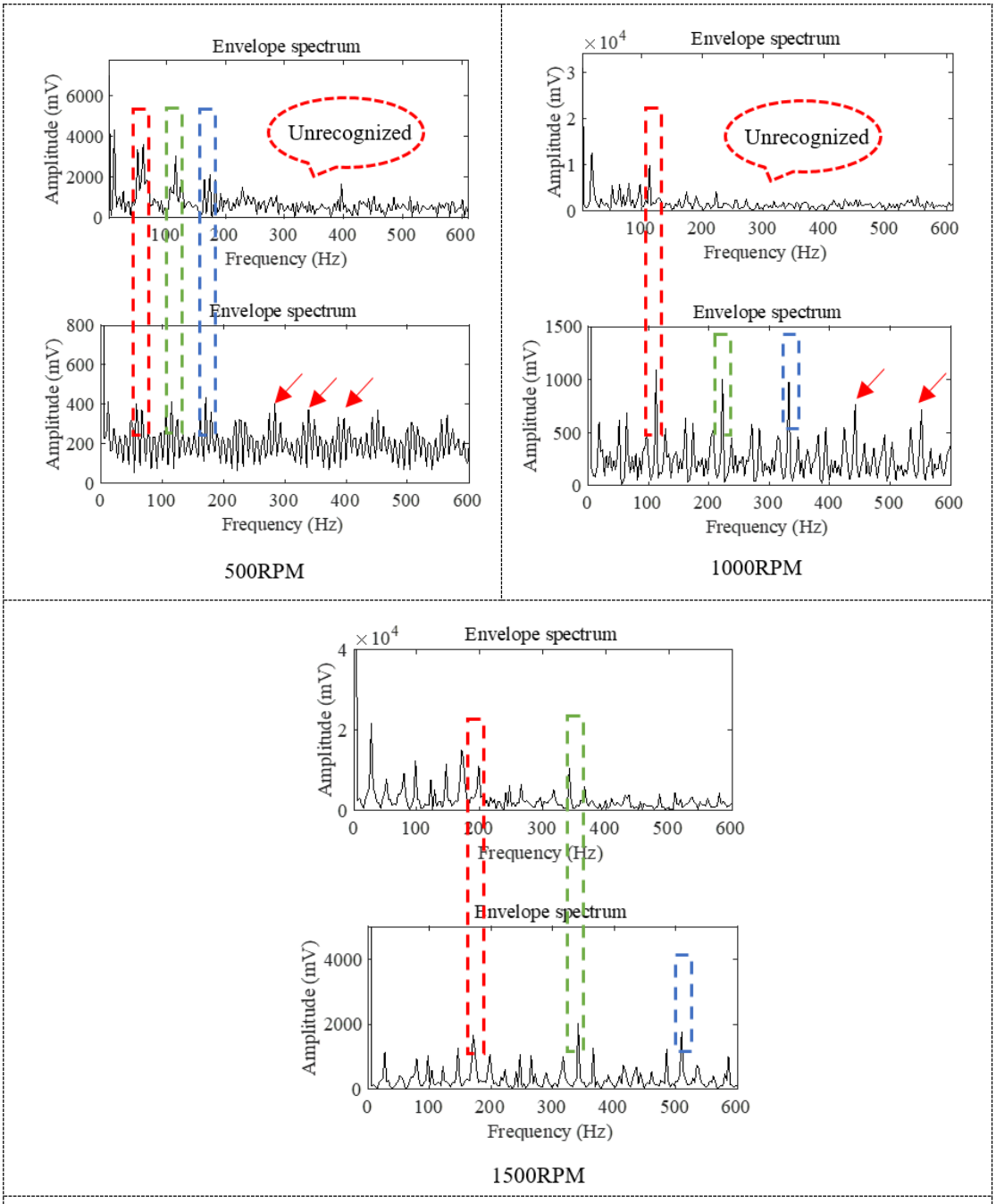


Fig.2-18. Fault characteristic envelope spectrum of inner race fault at different speeds

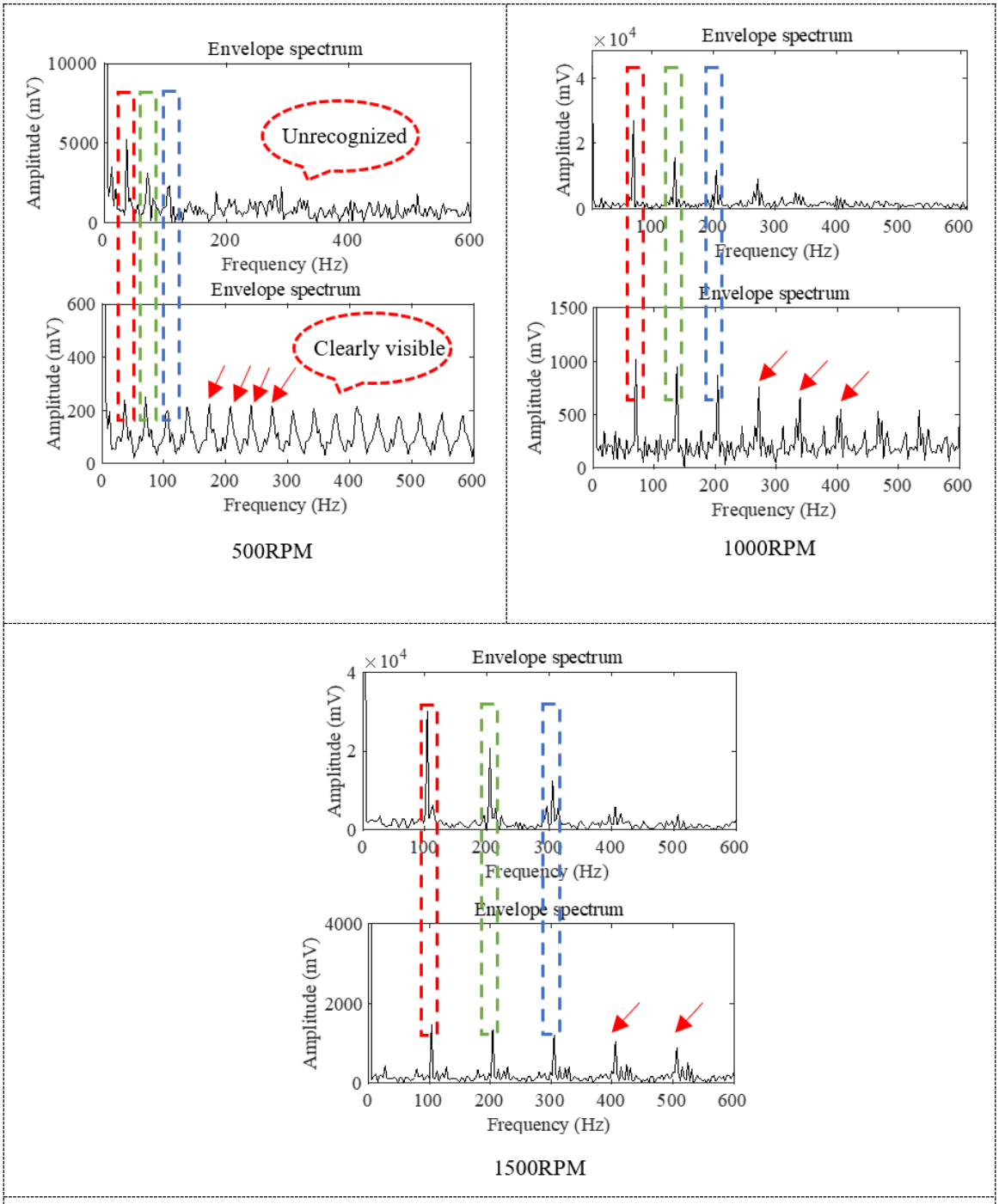


Fig.2-19. Fault characteristic envelope spectrum of outer race fault at different speeds

Table 2-8.

Fault characteristic frequency.

RPM	Outer race fault	Inner race fault	Roller fault
1000	73.18 Hz	110.15 Hz	79.28 Hz
1500	109.77 Hz	165.23 Hz	118.92 Hz

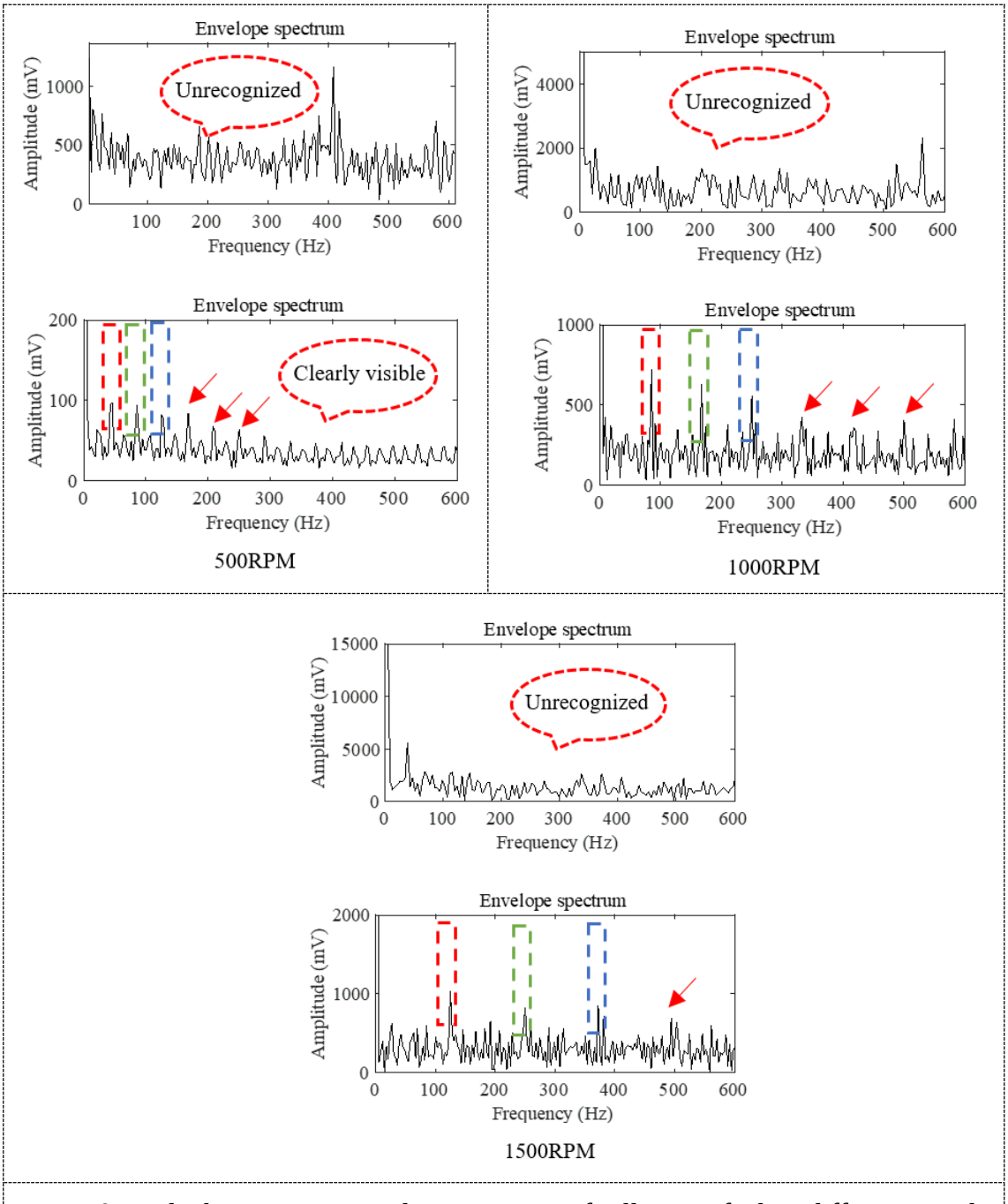


Fig.2-20. Fault characteristic envelope spectrum of roller race fault at different speeds

As shown in Fig.2-18, the signal envelope spectrum (after HPF processing) can identify the characteristic fault frequency of the outer race fault, as well as its double and triple frequency. However, after the third harmonic generation, the fault characteristic frequency cannot be clearly identified, most apparent under the 500rpm. Relatively, the envelope spectrum of the proposed SPCTF method is very obvious at

approximately five times the frequency and clearly identifies the fault characteristics of the bearing outer race. The fault diagnosis result of inner race fault signals is shown in Fig.2-19. Less than 1000rpm, the HPF only has one peak at the fault frequency, so it cannot identify the fault type. The proposed SPCTF can effectively identify the inner race fault types under 500, 1000 and 1500rpm. Significantly, the proposed method outperforms the HPF methodology in the roller fault experiment, and the experimental results are shown in Fig.2-20. Therefore the proposed SPCTF process outperforms in comparison with HPF, although the interference of complex background noise is substantial.

2.7 Summary

In this section a novel method based on state-space principal component tracking filtering (SPCTF) is proposed for fault diagnoses of bearings, removing noise interference and extracting fault features effectively to realize high-precision bearing fault diagnosis. The proposed method is applied to simulation signals and actual vibration signals, demonstrating the effectiveness and feasibility of the developed process. Moreover, the comparative experiments performed with the Hilbert spectrum method and high pass filter validate the superiority of the SPCTF. The experimental results show that the fault types can be identified effectively by comparing the distribution of different fault signals in the Hilbert spectrum contour map. In addition, the proposed SPCTF bearing fault diagnosis method can effectively identify various bearing fault states under different speed conditions compared with the high pass filter, which verifies the generalized applicability of SPCTF.

3. Fault diagnosis for low-speed bearings

3.1 Introduction

In the condition monitoring and fault diagnosis technology of rolling bearings, because the vibration signal is closely related to the mechanical structure, the diagnosis technology based on the vibration signal has been widely used. The principal element state filtering model proposed in Chapter 2 can effectively realize fault feature extraction and fault type discrimination of high-speed rolling bearings. In general, when the bearing speed is lower or the degree of failure is less severe, the periodic impact caused by bearing failure is weaker. The characteristics of the vibration signal of rolling bearing running at low speed are that the fault feature is not obvious, the fault signal is weak and the signal-to-noise ratio is low. The key to realize the fault diagnosis of rolling bearing is how to accurately extract more fault information from the strong background noise. Because the fault signal of mechanical equipment, especially the bearing fault signal will generate periodic pulse impact force, the modulation phenomenon of vibration signal will be generated, and amplitude modulation or frequency modulation signal will be formed. Therefore, signal modulation technology is widely used in fault diagnosis and analysis.

TEO has the ability to denoising and enhance the fault signal. TEO is a nonlinear difference operator, it estimates the total energy of the signal through transient signal and signal differential. TEO can enhance the transient features of the vibration signal, effectively detects shock components. Which has been widely applied in the fields of fault diagnosis to demodulate signal and extract the fault feature. However two inherent drawbacks limit its effect of signal processing: 1) It is sensitive to noise and vibration interference; 2) The energy spectrum may have negative values. For improving that, we proposed an improved Teager energy operator (ITEO) algorithm: the traditional forward difference is replaced by central-finite difference to smooth the raw signal, and the Hilbert transform is performed to redefine the analytic energy operator. The ITEO enhances the robustness of the traditional TEO, and improves the feature extraction capability.

ITEO can reduce the noise of the raw signals and improve the signal to noise ratio (SNR), which provides a good foundation for effective feature extraction. However, currently the features extract methods are mainly depend on artificial trial and error, which may bring the uncertainty for rolling bearings vibration signals under different

operating conditions. With the development of artificial intelligence in recent years, deep learning methods have been applied in many fields and achieved good results. It shows better generalization performance than traditional methods. Auto-encoding (AE) as one of the deep learning method, can adaptive extract the fault features by auto encoding operation [114-118]. Based on this characteristic, this paper propose an ITEO-AE-PCA method for low speed rolling bearing fault diagnosis: ITEO is employed to enhance the weak transient shock component of the raw signal. The processed signals then input into the AE to adaptive extract the signal fault features. The ITEO–AE model combines the capability of signal process and feature extraction, which do not require complicated prior knowledge and reduce the computational complexity. Finally the adaptive extracted fault feature parameters are employed to build the diagnosis model based on PCA method for diagnosing faults of low speed rolling bearing.

3.2 Principle of the proposed algorithm

3.2.1 Teager energy operator

Traditional TEO was first proposed by Herbert M. Teager and Shushan M. Teager. The method to compute the energy of the generator named "Teager's energy algorithm" was proposed by J. F. Kaiser and the Teager Operator has since been defined for continuous signals, both real and complex ones. Kaiser used the following differential equation as a starting point for the operator:

$$\frac{d^2x}{dt^2} + \frac{k}{m}x = 0 \tag{3-1}$$

This second order differential equation describes an object with mass m suspended by a string with constant k. We can regard this as a simple (but incomplete) model of a mechanical-acoustical system, where the object may oscillate, thus creating pressure waves in the surrounding medium. None of the medium's characteristics are included in the model, though.

The solution to equation 3-1 is a periodic oscillation given by $x(t) = A\cos(\omega t + \varphi)$, where $x(t)$ is the position of the object at time t, A is the amplitude of the oscillation, $\omega = \sqrt{k/m}$ is the frequency of the oscillation, and φ is the initial phase. If $\varphi = 0$, we have that the object is not initially in equilibrium.

The total energy of the object is in Newtonian physics given as the sum of the potential energy of the spring and the kinetic energy of the object, given by

$$E = \frac{1}{2}kx^2 + \frac{1}{2}mv^2 \tag{3-2}$$

By substituting $v = {dx}/{dt}$, and $x = A\cos(\omega t + \varphi)$, we get

$$\mathrm{E} = \frac{1}{2} m\omega^2 A^2 \tag{3-3}$$

From this, we immediately see that the energy of the object is proportional to both A and ω. Note that the energy E is implicitly a function of time.

3.2.2 The continuous Teager energy operator

Historically, the discrete version of the Teager Energy Operator was defined first. However, it is natural to begin with the continuous operator here. In continuous time, we define the Teager Energy Operator to be

$$\Psi(\mathrm{x(t)}) = (\dot{x}^2(\mathrm{t})) - \mathrm{x(t)}\ddot{x}(t) \tag{3-4}$$

Inserting $\mathrm{x(t)} = \mathrm{A}\cos(\omega t)$ into the equation above yields,

$$\Psi(\mathrm{x(t)}) = A^2\omega^2 \tag{3-5}$$

Which is the amplitude and frequency product squared.

We now look at a digital signal x_n given by

$$\mathrm{x}_n = A\cos(\Omega n + \varphi) \tag{3-6}$$

Where Ω is the digital frequency **Ω** = 2πf /Fs, where φ is an arbitrary phase, f is the analog frequency, and Fs is the sampling frequency.

The equation above has three parameters, which means that we in principle should be able to set up three different instances of this formula, and solve it to determine the three unknowns. To do this, we choose x_n, x_{n-1} and x_{n+1} as unknowns:

$$\mathrm{x}_n = A\cos(\Omega n + \varphi) \tag{3-7}$$

$$\mathrm{x}_{n-1} = A\cos(\Omega(n-1) + \varphi) \tag{3-8}$$

$$\mathrm{x}_{n+1} = A\cos(\Omega(n+1) + \varphi) \tag{3-9}$$

By using the trigonometric identities we get

$$\mathrm{x}_{n-1}\mathrm{x}_{n+1} = A^2\cos^2(\Omega n + \varphi) - A^2\sin^2\Omega \tag{3-10}$$

From this, we see that the first term of the sum is x_n^2; we can then write just

$$A^2\sin^2\Omega = \mathrm{x}_n^2 - \mathrm{x}_{n-1}\mathrm{x}_{n+1} \tag{3-11}$$

If we restrict **Ω** to be positive and less than π/2, the solution above is exact and unique. In practice, this is when **Ω** is less than one quarter of the sampling frequency.

The formula above can be approximated further if we notice that $\sin(\mathbf{\Omega}) \approx \mathbf{\Omega}$, for small **Ω**s. The approximation error is less than 11% if **Ω** < π/4. From Fig.3-1, we see that the largest deviation from sin(x) is at x = π/4. The ratio between the two functions is 1.11 at that point, which explains the 11% approximation error. We thus

end up with the following basic definition of the discrete Teager Energy Operator:

$$\Psi[x_n] = x_n^2 - x_{n-1}x_{n+1} \tag{3-12}$$

So, for a given signal x[n], we can estimate the instantaneous energy of the signal by substituting for x in the formula above.

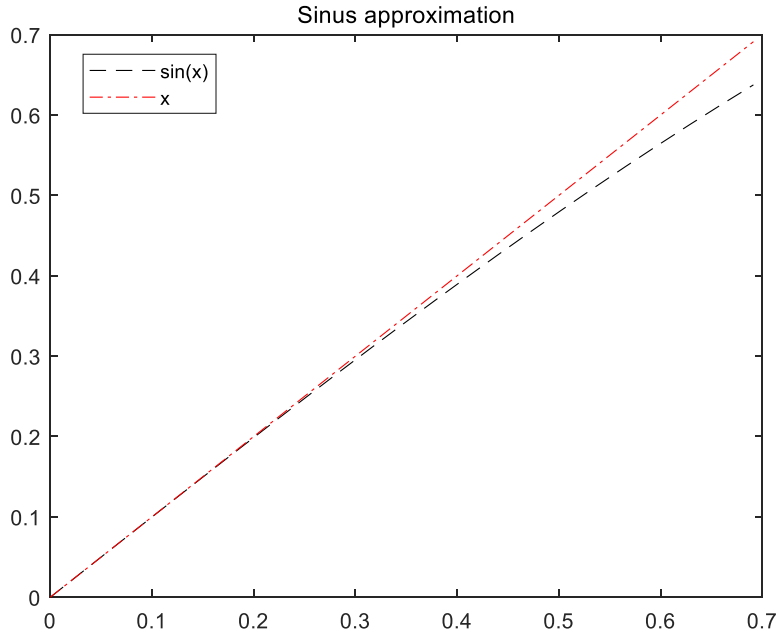


Fig.3-1 Linear approximation of the sine function in the interval $[0,\frac{\pi}{4}]$

Own to only three samples are required for the energy computation at each time instant, TEO hold the ability to capture the transient energy fluctuations. But the forward difference method lead to high demodulation error, sensitive to noise and vibration interference. Therefore, to smooth the original discrete signal and improve the accuracy of demodulation, instead of forward difference the central finite difference sequence is adopted in the symmetrical difference analytic energy operator, which is expressed as:

$$\dot{x}_n = \frac{x_{n+1} - x_{n-1}}{2} \tag{3-13}$$

On the basis of analytic signal theory, amplitude demodulation and transient frequency can represent the characteristics of a signal. Combine the central-finite difference and analytic signal theory.

Define the analytic energy operator as follows:

$$\Gamma[x(t)] = x'(t)\hat{x}(t) - x(t)\hat{x}'(t) \tag{3-14}$$

where $x'(t)$ is the first derivative of x(t), $\hat{x}'(t)$ is the first derivative of the $\hat{x}(t)$, $\hat{x}(t)$ is the HT transform of x(t).

The discrete form of Eq.(3-13) as shown in Eq. (3-15):

$$\Gamma[\mathrm{x(n)}] = x'(n)\hat{x}(n) - x(n)\hat{x}'(n) \tag{3-15}$$

Use the central-finite difference replace the forward difference, substitute Eq. (3-13) into Wq. (3-15), we get the central-finite difference energy operator:

$$\Gamma[x_n)] = \frac{[x_{n+1}-x_{n-1}]\cdot h[x_n]}{2} - \frac{h[x_{n+1}-x_{n-1}]\cdot x_n}{2} \tag{3-16}$$

The result of TEO, DTEO, and ITEO signal demodulate show in Fig. 3-2. It shows that the ITEO can effectively reduce the noise, enhance the faint transient shock component of the raw vibration signals, and keep the symmetry and original amplitude of the raw signals. The enhanced signals are prone to extract the abnormal features.

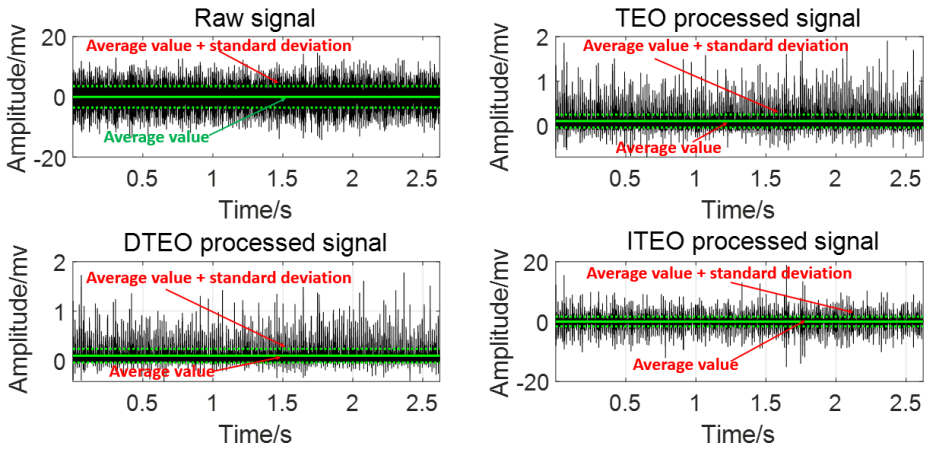


Fig.3-2. TEO, DTEO , ITEO signal processing

3.2.3 The deep Learning Auto-encoder Algorithms

An autoencoder is a special type of neural network, it is consist of three layers: the encoder layer, the hidden layer and the decoder layer as shown in Fig.3-3. The encoder maps the input data from a high-dimensional space into codes in a low-dimensional space, and the decoder reconstruct the input data from the corresponding codes.

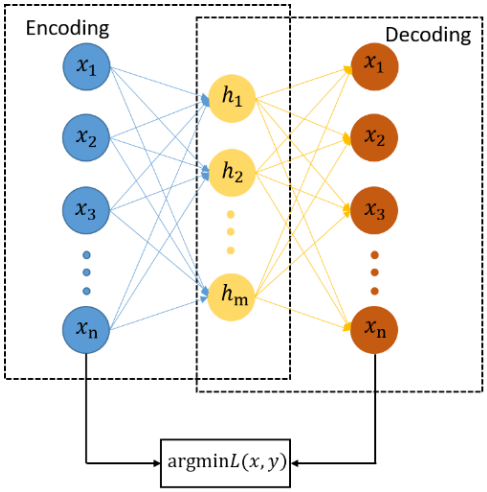


Fig.3-3 Auto-encoder network structure

Encoding: Given unlabeled fault sample data x, the encoder transforms the input vector x into a hidden representation h through the sigmoid function $f_\theta(x)$.

$$h = f_\theta(x) = S(W^{(1)}x + b^{(1)}) \tag{3-17}$$

Decoding: the vector h is transformed back to a reconstruction vector y through the $g_{\theta'}(h)$.

$$y = g_{\theta'}(h) = S(W^{(2)}x + b^{(2)}) \tag{3-18}$$

Where S is an activation function:$S(t) = {}^{1}/_{(1-e^{-t})}$, $\theta = \{W^{(1)}, b^{(1)}\}$ is the encoder parameter set. $\theta' = \{W^{(2)}, b^{(2)}\}$ is the decoder parameter set.

The autoencoder training aims to optimize the parameter set θ and, θ' minimizing the reconstruction error.

$$\arg\min_{\theta,\theta'} \min L(x, g_{\theta'}(h)) \tag{3-19}$$

The cross entropy function is usually used to represent the reconstruction error

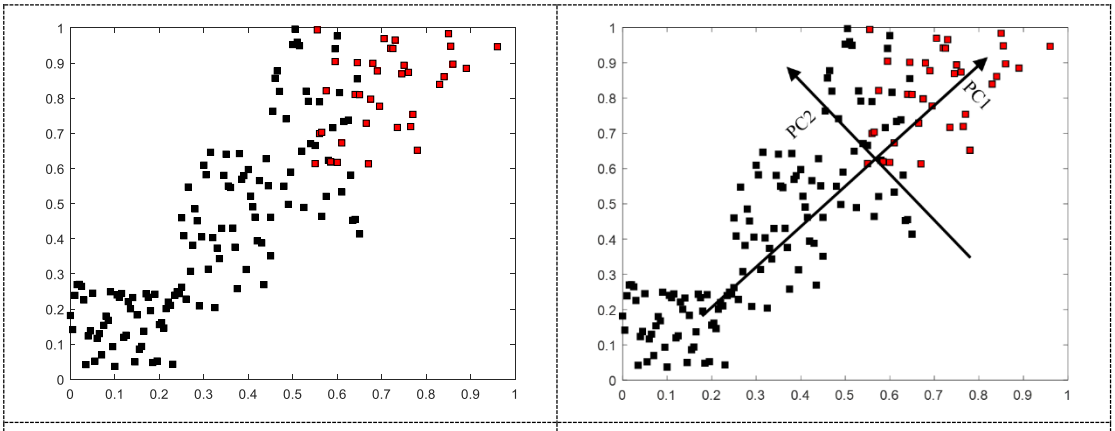
$$L(x,y) = \frac{1}{m}\sum(x\ln y + (1-x)\ln(1-y)) \tag{3-20}$$

3.2.4 Principal component analysis

Principal component analysis (PCA) is a mathematical algorithm that reduces the dimensionality of the data while retaining most of the variation in the data set1. It accomplishes this reduction by identifying directions, called principal components, along which the variation in the data is maximal. By using a few components, each sample can be represented by relatively few numbers instead of by values for thousands of variables. Samples can then be plotted, making it possible to visually assess similarities and differences between samples and determine whether samples can be

grouped.

The basics of PCA can be explained with simple eometrical interpretations of the data. To allow for such interpretations, the normalized vibration signal of the faulty bearing is shown in Figure 3-4(a). PCA identifies new variables, the principal components, which are linear combinations of the original variables. The two principal components for our two-dimensional vibration signal are shown in Figure 3-4(b).



(a) (b)

Fig.3-4 Principal component analysis (PCA) of discrete data

It is easy to see that the first principal component is the direction along which the samples show the largest variation. The second principal component is the direction uncorrelated to the first component along which the samples show the largest variation. If data are standardized such that each signal is centered to zero average expression level, the principal components are normalized eigenvectors of the covariance matrix of the signals and ordered according to how much of the variation present in the data they contain. Each component can then be interpreted as the direction, uncorrelated to previous components, which maximizes the variance of the samples when projected onto the component.

PCA provides a means to compress the data dimension and eliminate the correlation among different statistical features. Given a set of n dimension feature vectors $\mathrm{X} = (x_1, x_2, \cdots x_m,)$, where n and m can be considered as the number of observations and measured variables. The principal component decomposition of X can be represented as:

$$X = QP^T = \sum_{i=1}^{m} q_i p_i^T \qquad (3\text{-}21)$$

where $q_i \in R^n$, $p_i \in R^m$ are respectively the score vector and loading vector. And

the Q is the score matrix and P is the loading matrix.

Calculate the covariance $\sum(\lambda_1, \lambda_2, \cdots, \lambda_m)$ of matrix X, where $\lambda_1, \lambda_2, \cdots, \lambda_m$ are the eigenvalues and they are sorted in descending order, $P = (p_1, p_2, \cdots, p_m,)$are the corresponding eigenvectors. According to the matrix decomposition theorem：

$$\sum = P\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{bmatrix} P^T = \sum_{i=1}^{m} \lambda_i \mathrm{p}_i \, \mathrm{p}_i^T \qquad (3\text{-}22)$$

where P is the unit orthogonal matrix

$$p_i^T p_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \qquad (3\text{-}23)$$

Simultaneous right-multiplying p_i at both ends of Eq. (3-19):

$$Xp_i = q_1 p_1^T p_i + q_2 p_2^T p_i + \cdots + q_m p_m^T p_i \qquad (3\text{-}24)$$

$$q_i = Xp_i \qquad (3\text{-}25)$$

Eq. (3-21) could be rewritten as:

$$X = \hat{X} + E \qquad (3\text{-}26)$$

$$\hat{X} = \sum_{i=1}^{l} q_i p_i^T = \hat{Q}\hat{P}^T, E = \sum_{i=l+1}^{m} q_i p_i^T = \tilde{Q}\tilde{P}^T \qquad (3\text{-}27)$$

where $\hat{Q}$ is principal score matrix $\hat{Q} \in R^{n \times l}, \hat{Q} = X\hat{P}$, $\hat{P}$ is principal loading matrix, $\hat{P} \in R^{n \times m}$, l is the number of principal components and it is far smaller than n. $\hat{X}$ represents the principal element model of X, E is the error matrix, $\widetilde{Q}$ is residual score matrix, $\widetilde{P}$ is residual loading matrix. Generally, the cumulative variance contribution rate method is used to determine the number of retained principals.

3.3 ITEO-AE based fault diagnosis

A flow diagram of the proposed ITEO-AE-PCA based fault diagnosis scheme is shown in Fig. 3-5. The diagnosis process of the low speed rolling bearings is divided into two stage: training stage and testing stage.

Training stage:

- The normal signals and single abnormal state fault signals is demodulated by ITEO to improve the signal to noise ratio(SNR).
- Construct N samples with the ITEO processed signal. Then put samples into the AE model to adaptive extract fault extract.
- Build the diagnosis model by PCA method based on the AE extracted features.

Testing stage:

- Demodulate the test signals by ITEO.
- Construct the N samples with ITEO processed test signal and adaptive extract the features by AE.
- Map the extractd features to training stage PCA space and judge the condition of bearing.

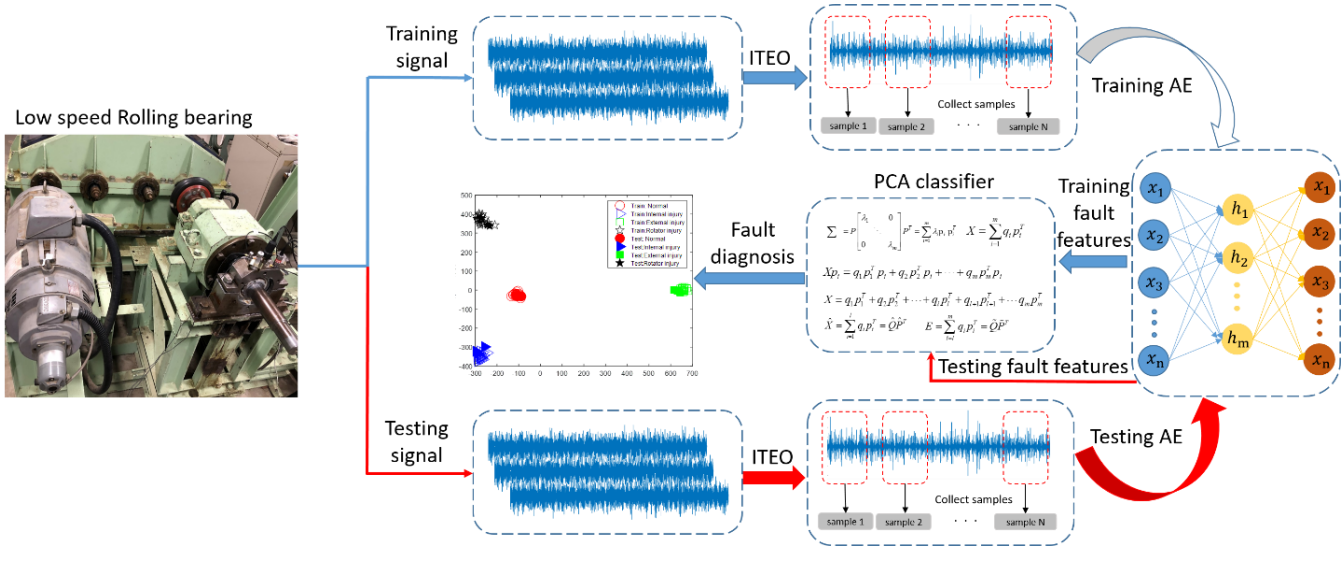


Fig.3-5 Flowchart of the proposed method

3.4 Case study

3.4.1 Experiment platforms

In order to verify the effectiveness of the proposed AE-ITEO, raw signals of bearings in different states (normal, outer race fault, inner race fault, and roller fault) were collected from the experimental platform, as shown in Fig.3-6. There are three kinds of fault in the bearing. The bearing failure is due to manual cutting, the depth of damage is 0.3 mm, 5.0mm in width. In addition, we have added the measurement of various fault signals when the lubrication is poor. The signals sampling frequency is 100 kHz，and the sampling time was approximately 20s. In order to verify the effect of low-speed running bearing fault diagnosis，three kinds of rotating speeds in this experimental design, 50 rpm, 70 rpm and 100 rpm.

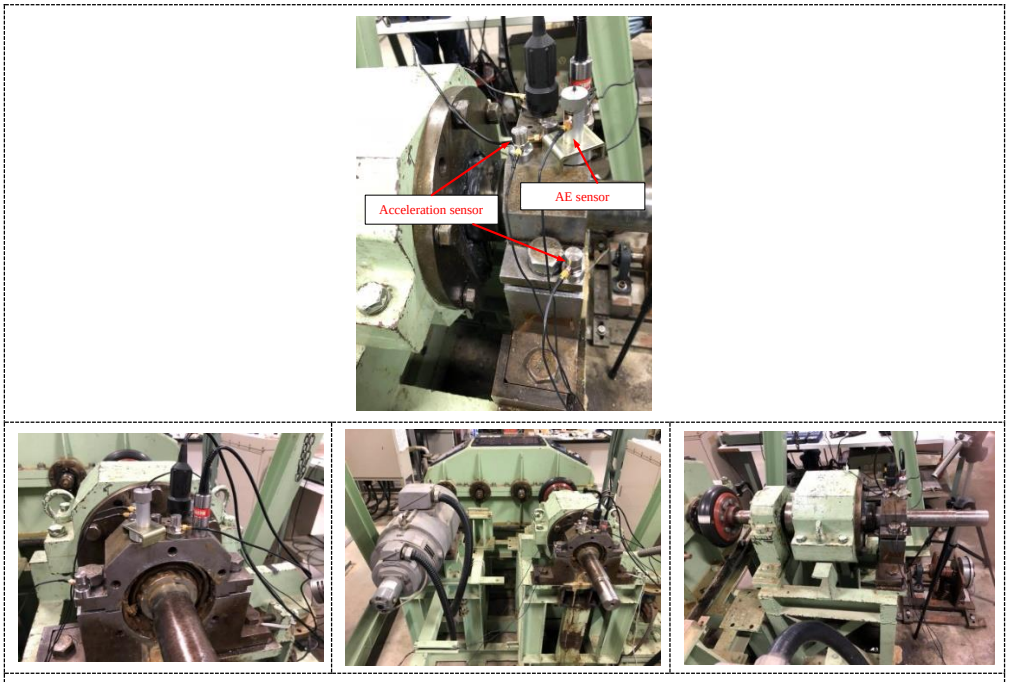


Fig.3-6 Experimental platform for low speed rolling bearings

3.4.2 Intelligent diagnosis based on proposed method

1) ITEO fault features extract process

In this section, bearing signal of normal condition, inner race fault condition, outer race fault condition and roller fault are selected to verify the effectiveness of proposed method. The data of the faulty bearing at 100 RPM is selected. And the wound size is 5.0W * 0.3d. The signals sampling frequency is 10 kHz to ensure that the length of each sample contains the information more than one turn. Each sample length is 16384. In this experiment, 20 samples are selected for training and 3 samples are selected for testing. So, the signal length of each state is 23*16384. The signals of four types rolling bearings is processed by ITEO respectively. The result of the inner race fault signals processed by ITEO is shown in Fig.3-7. Compared to the raw signals, the noise of the signals are reduced by ITEO, and the weak transient shock component of the raw signals are enhanced.

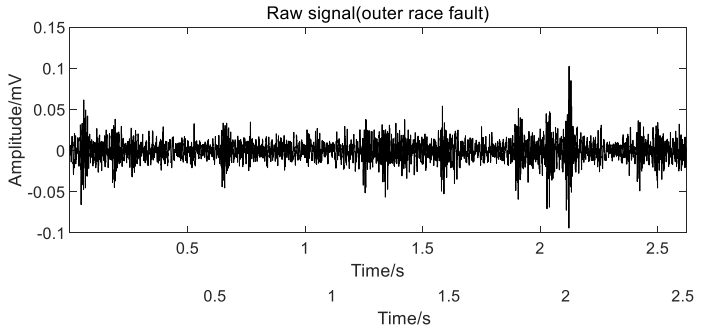


Fig.3-7 Signal processed by ITEO

2) Adaptive extract fault features based on AE

First divide the ITEO processed signal into 23 samples, the signal length of each sample is 16384. Among them, 20 samples are selected for training and 3 samples are selected for testing. The encoding layer, hidden layer and decoding layer dimensions of the AE features extract model are 16384, 200 and 16384, respectively. Manual experiment shows, when set the learning rate to 0.9 and the number of iterations to 300, it can get better diagnosis result. Therefore, it use the above parameters to initialize the AE model. After continuous iterative training, it can get the value of the AE hidden layer and retain the features parameters. Four types of rolling bearing correspond to 80 training samples. Visualize the features result show in Fig.3-8:

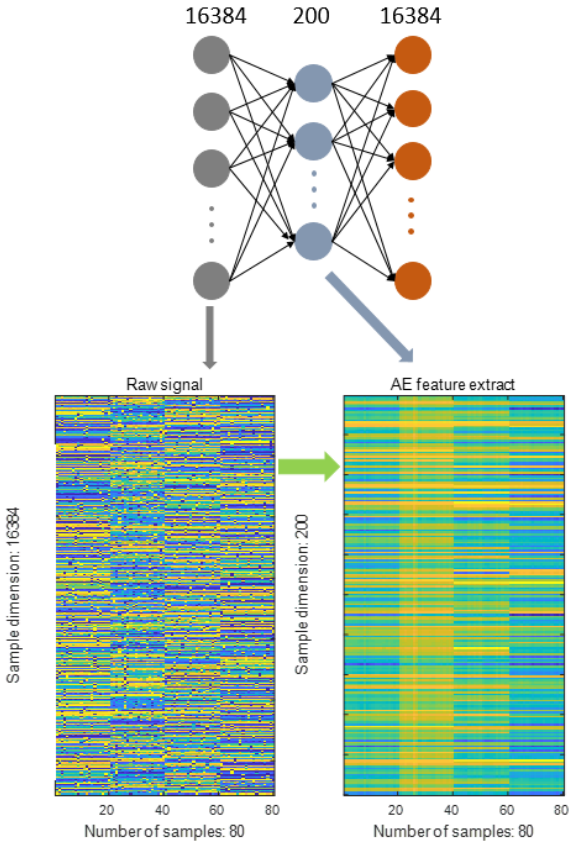


Fig.3-8 AE fault features extract

Fig.3-8 shows that the raw signal is rather messy and the features extracted by AE are relatively orderly, which provides a good basis for further PCA analysis.

3) Diagnosis model building by PCA.

The features adaptive extracted by AE are substituted into the PCA model. It takes

the 80 (20*4) samples for training, and the 12 (3*4) samples for testing. The cumulative contribution rate of different principal elements is shown in Fig.3-9.

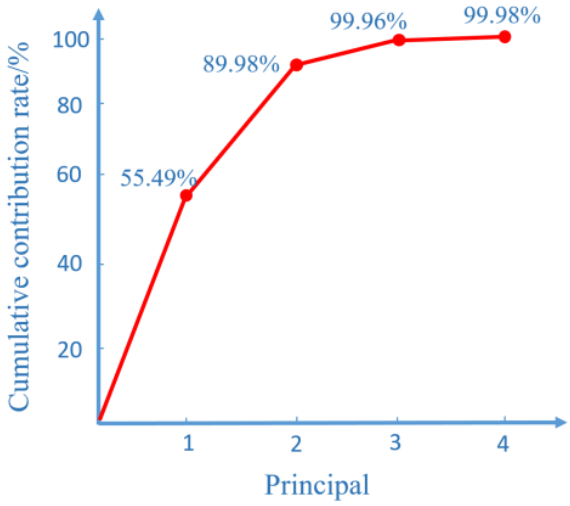


Fig.3-9 Cumulative contribution rate

It can be seen from the Fig. 3-9, the cumulative contribution rate of the first principal element and the second principal element over 89%. So, it select the coefficient of the first principal component and the second principal component to build the model. The coefficients are multiplied to obtain the principal component scores, and then map the scores of each sample on a two-dimensional plane. The testing samples component scores are been mapped on the same PCA model. The results of the PCA fault identification is shown in Fig.3-10.

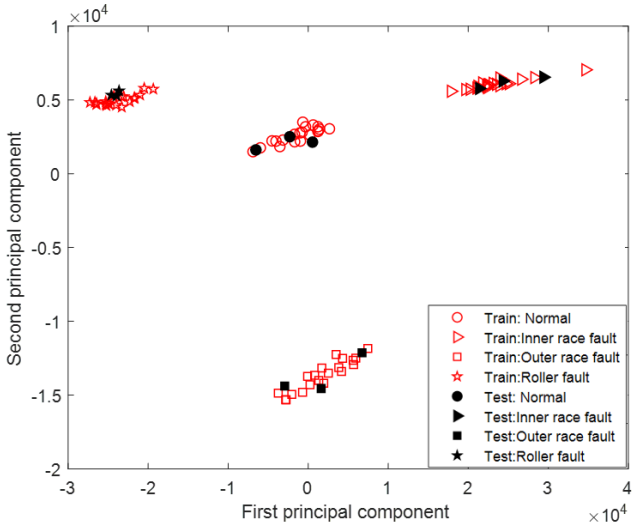


Fig.3-10 The diagnostic result （RBM100）

Circles, triangles, squares and five-pointed stars represent the four types of normal

bearing, out race fault, inner race fault, and roller fault. And the red hollow shape represents the training data, and the black solid shape represents the test data. Under the 100 RPM operating condition, the training data and testing data are coincident, the proposed method can accurate distinguish and classify the four states of the bearings. It proved the effectiveness of the proposed method.

3.5 Comparative Experiment

3.5.1 Traditional artificially designed characteristic parameters

In order to validate the ITEO-AE-PCA diagnosis effect, the same rolling speed bearings signals (70RBM) is processed by the traditional artificially designed characteristic parameters to extract the features. After the ITEO signal processing, the characteristics of the signal are calculated by the following characteristic formula.

$$p_1=\sqrt{\sum_{i=N/3}^{N} f_i^2 S(f_i) \Big/ \sum_{i=N/3}^{N} S(f_i)} \tag{3-28}$$

$$p_2=\sqrt{\sum_{i=N/3}^{N} f_i^4 S(f_i) \Big/ \sum_{i=N/3}^{N} f_i^2 S(f_i)} \tag{3-29}$$

$$p_3=\sum_{i=N/3}^{N} f_i^2 S(f_i) \Big/ \sqrt{\sum_{i=N/3}^{N} f_i^2 S(f_i) \sum_{i=N/3}^{N} S(f_i)} \tag{3-30}$$

$$p_4=\frac{\sigma}{\bar{f}} \tag{3-31}$$

$$p_5=\sum_{i=N/3}^{N} \left(f-\bar{f}\right)^3 S(f_i) \Big/ \sigma^3 N \tag{3-32}$$

$$p_6=\sum_{i=N/3}^{N} \left(f-\bar{f}\right)^4 S(f_i) \Big/ \sigma^4 N \tag{3-33}$$

$$p_7=\sum_{i=N/3}^{N} \sqrt{\left|\left(f-\bar{f}\right)\right|} S(f_i) \Big/ \sqrt{\sigma} N \tag{3-34}$$

However，

$$\bar{f}=\sum_{i=N/3}^{N} f_i S(f_i) \Big/ \sum_{i=N/3}^{N} S(f_i), \quad \sigma=\sqrt{\sum_{i=N/3}^{N} \left(f-\bar{f}\right)^2 S(f_i) \Big/ N} \tag{3-35}$$

where N is the dimension of the signal, f is frequency, $s(f)$ is transient spectrum, $i=1,2,\cdots N$ 。

In addition, in order to eliminate the influence of the amplitude variation caused by the fluctuation of the amplifier voltage, before calculating the characteristic parameters, normalize the transient spectrum by Eq. (3-34) and normalize the feature parameters by Eq.(3-35).

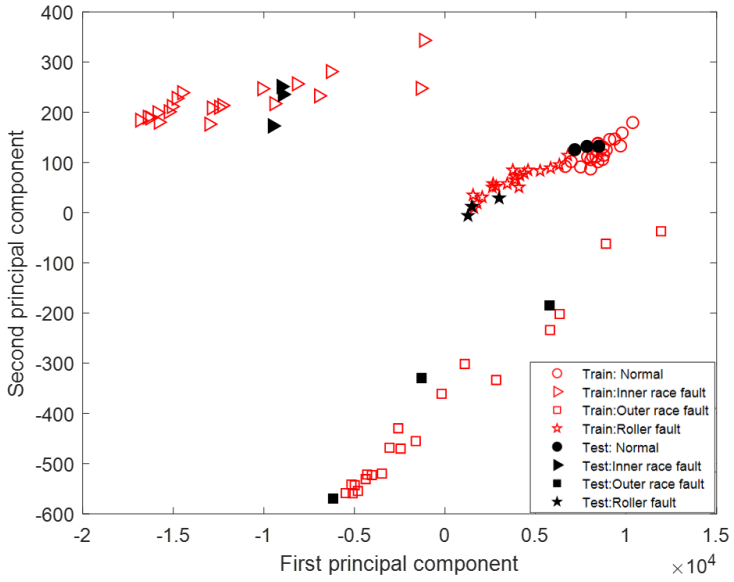
$$S(f_i) = S'(f_i) \Big/ \sum_{i=1}^{N} S'(f_i) \tag{3-36}$$

where $S'(f_i)$ and $S(f_i)$ is the transient spectrum before and after normalization, respectively.

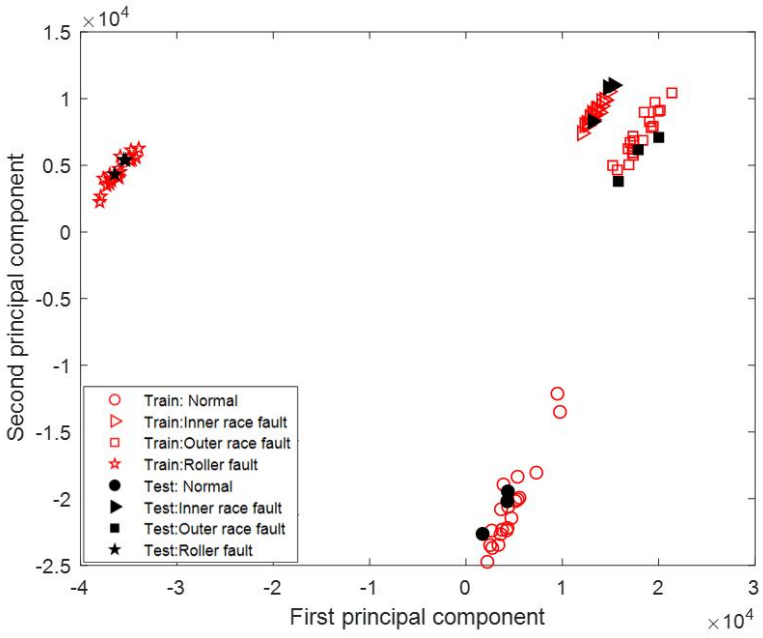
$$p_i = \left(p_i' - \bar{p}_{ni} \right) \Big/ Sp_{ni} \tag{3-37}$$

where p_i and p_i' is the characteristic parameters before and after normalization, respectively. $\bar{p}_{ni}$ and Sp_{ni} are the mean and standard deviation of the characteristic parameters.

The signals of four types rolling bearings (normal, outer race fault, inner race fault, roller fault) with 70RBM is calculated by the above formula to obtain the characteristic parameters. Substitutie the characteristic parameters into the PCA model, and map the principal component scores on the two-dimensional plane. The distribution result is shown in Fig.3-11.



a. Traditional method



b. Proposed method

Fig.3-11 The diagnostic result

Due to low vibration energy, fault features of rolling bearings under low speed operating conditions is prone to be submerged and distorted by the background noise interference, the artificial feature extraction method lead to the fault classification aliasing phenomenon(As shown in Fig.10.a). The proposed ITEO-AE-PCA model could accurate realize the fault diagnosis of low speed rolling bearings.

3.5.2 No processed by ITEO

Construct N samples with the raw signals(without processing by ITEO), put the samples into the AE model to adaptive extract fault feature. Map the extractd features to trained PCA space and the result show in Fig.3-12

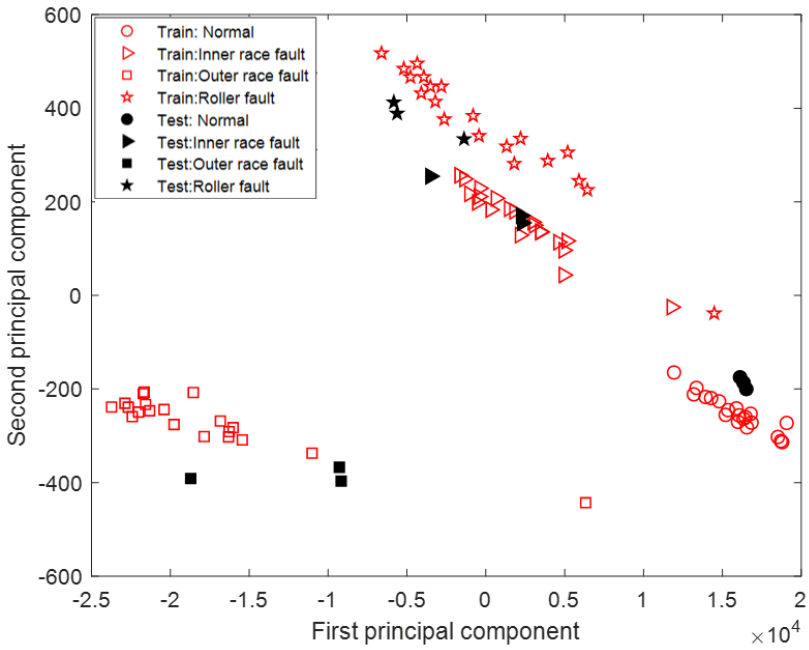


Fig.3-12 The diagnostic result (without ITEO)

The low speed rolling bearings vibration signals is prone to be interference, it is difficult to extract the effective fault feature from low signal to noise ratio samples which led to the poor diagnostic result.

3.6 Summary

In this section, an ITEO-AE-PCA based low speed rolling bearing features extract and accurate fault diagnosis is proposed. ITEO is used to reduce the noise and enhance the transient features of the raw signal under the complex operating environments. AE can adaptive extract the fault features without depending on human experience and prior knowledge. PCA map the fault features signals extracted by AE into a low-dimensional subspace. The ITEO-AE-PCA model is performed well in the diagnosis of real low speed rolling bearings with different faults. The effectiveness of the proposed method is further verified by comparison with conventional feature parameter methods.

4. Diagnosis for early fault bearing

4.1 Introduction

The bearing fault diagnosis method based on state-space principal component tracking filtering (SPCTF) proposed in Chapter 2 can effectively extract the fault characteristics of high-speed (500RPM-1500RPM) running bearings, and analyze their fault types through envelope spectrum. For low-speed running bearing fault diagnosis, the traditional method of judging fault types by fault characteristic frequency is difficult to diagnose bearing faults [119-123]. For this reason, a fault diagnosis model based on the ITEO-AE method is established. Experiments show that this method can realize low-speed running bearing fault online diagnosis. ITEO can effectively extract low-speed bearing fault features, but due to the poor robustness of the deep learning AE method, especially when the probability distribution of test samples and training samples is quite different, the diagnosis effect is not good. For the early low-speed bearing fault, due to the small size of the wound, the fault characteristic signal is easily submerged, and this method cannot effectively identify the fault type of the rolling bearing. Restricted Boltzmann machine (RBM) is one of the deep learning methods, which can extract the characteristics of self-adaptive signals compared to the RBM method and the conventional feature parameter method, and relies on human experience and expertise at the time of equipment diagnosis. Therefore, it is suitable for extracting characteristic parameters of abnormal signals in early fault bearings for which precise diagnosis is difficult. Furthermore, we propose a method for discriminating the type of anomaly by integrating the characteristic parameters extracted from RBM by the principal component analysis method [124-127].

As a deep learning classifier, RBM is widely used in the field of data processing. In this context, two approaches are usually followed. First, an RBM is trained in an unsuper vised manner to model the distribution of the inputs. Then, the RBM is used in one of twoways: either its hidden layer is used to preprocess the input data by replacing it with the representation given by the hidden layer, or the parameters of the RBM are used to initialize a feedforward neural network. In both cases, the RBM is paired with some other learning algorithm (the classifier using the preprocessed inputs or the neural network) to solve the supervised learning problem at hand. This approach unfortunately requires one to tune both sets of hyper-parameters (those of the RBM and of the other learning algorithm) at the same time. Moreover, since the RBM is trained in an

unsupervised manner, it is blind to the nature of the supervised task that needs to be solved and provides no guarantees that the information extracted by its hidden layer will be useful. Based on the Classification Restricted Boltzmann Machine(CRBM) approach and learning algorithms address both aforementioned issues[128].

This section proposed a low speed rolling bearing fault diagnosis based on improved Teager energy operator (HDTEO) and classification Restricted Boltzmann machine (CRBM) model adaptive symptom parameters extraction. The combination of HDTEO and CRBM is performed to adaptive extract the symptom parameters of fault diagnosis. Based on the principal component analysis (PCA) diagnosis model, it can discriminate the fault condition of low speed rolling bearing. The effectiveness of the proposed method is validated by low speed rolling bearing fault diagnosis experiment.

4.2 Principle of the proposed algorithm

4.2.1 The restricted Boltzmann machines

The restricted Boltzmann machine (RBM) is a probabilistic model that uses a layer of hidden binary variables or units to model the distribution of a visible layer of variables as shown in Fig.4-1. It has been successfully applied to problems involving high dimensional data such as images and text [129-132]. For a given state vector h and v, the current energy function of the RBM can be expressed as:

$$E(\mathbf{v},\mathbf{h}) = -a^T v - b^T h - h^T W v \tag{4-1}$$

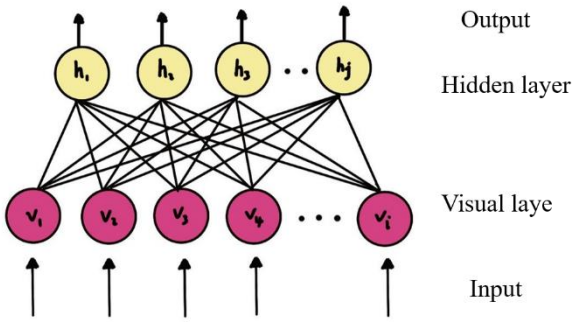


Fig. 4-1 The restricted Boltzmann machines model

With the energy function, we can define the state of the RBM as the probability distribution of a given v, h as:

$$p(\mathbf{v},\mathbf{h}) = \frac{1}{Z} e^{-E(\mathbf{v},\mathbf{h})} \tag{4-2}$$

The conditional distribution *P(h|v)* can be expressed as:

$$p(\mathbf{v},\mathbf{h}) = \frac{1}{P(v)}\frac{1}{Z}\exp(\mathrm{a}^{\mathrm{T}}\mathrm{v} + b^T h + h^T \mathrm{W}\mathrm{v})$$

$$= \frac{1}{Z'}\exp\{\sum_{j=1}^{nh}(b_j^T h_j + h_j^T W_{j,:} v)\}$$

$$= \frac{1}{Z'}\prod_{j=1}^{nh}\exp\{b_j^T h_j + h_j^T W_{j,:} v\} \tag{4-3}$$

where Z' is the new normalization coefficient, and the expression is:

$$\frac{1}{Z'} = \frac{1}{P(v)}\frac{1}{Z}\exp(a^T v) \tag{4-4}$$

As described above, since there is no coupling between the elements in the layer, the conditional probabilities of the visible element and the hidden element are expressed in equations (4-5) and (4-6).

$$p(v_i = 1|\boldsymbol{h}) = \sigma\left(a_i + \sum_j \omega_{ij} h_j\right) \tag{4-5}$$

$$p(h_j = 1|\boldsymbol{v}) = \sigma\left(b_j + \sum_i \omega_{ij} v_i\right) \tag{4-6}$$

Where σ is a sigmoid function.

4.2.2 The Classification Restricted Boltzmann Machines

The Classification Restricted Boltzmann Machine (CRBM) models the joint distribution of an input $\mathrm{x} = (x_1, x_2, \cdots, x_D)$ and target class $\mathrm{y} \in \{1,2, \cdots \mathrm{C}\}$ using a hidden layer of binary stochastic units $\mathrm{h} = (h_1, h_2, \cdots, x_H)$. This is done by first defining an energy function:

$$\mathrm{E}(\mathrm{y}, \mathrm{x}, \mathrm{h}) = -h^T W x - b^T x - c^T h - d^T e_y - h^T U e_y \tag{4-7}$$

with parameters Θ = (W, b, c,d, U) and where $e_y = (1_{i=y})_{i=1}^{C}$ is the "one out of C" representation of y. From the energy function, we assign probabilities to values of y, x and h as follows:

$$\mathrm{p}(\mathrm{y}, \mathrm{x}, \mathrm{h}) = \frac{\exp(-E(y,x,h))}{Z} \tag{4-8}$$

Where Z is a normalization constant (also called partition function) which ensures that Equation 1 is a valid probability distribution. We will assume that the elements of x are binary, but extensions to real-valued units on bounded or unbounded intervals are

straightforward. An illustration of the Class RBM is given in Fig.4-2.

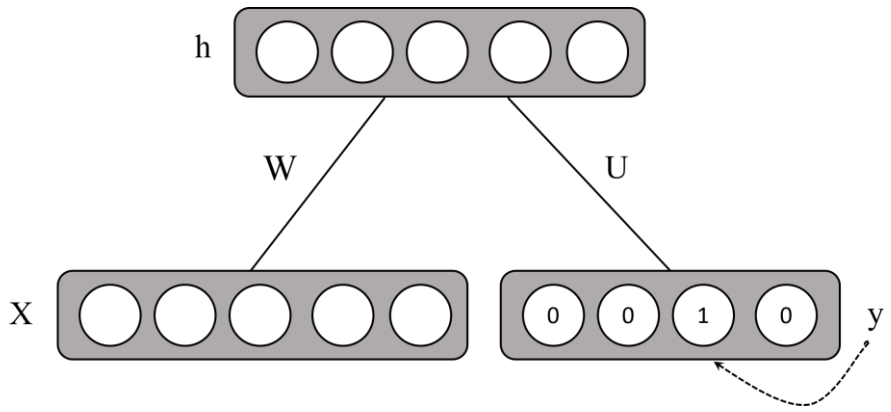


Fig.4-2 Illustration of a Classification Restricted Boltzmann Machine

Unfortunately, computing $\mathrm{p}(\mathrm{y},\mathrm{x},\mathrm{h})$ or $\mathrm{p}(\mathrm{y},\mathrm{x})$ is typically intractable. However, it is possible to sample from the ClassRBM, using Gibbs sampling, that is, alternating between sampling a value for the hidden layer given the current value of the visible layer (made of variables x and the e_y representation of y), and vice versa. All the required conditional distributions are very simple.

When conditioning on the visible layer, we have

$\mathrm{p}(h|y,x) = \prod_j p(h_j|y,x)$, with $p(h_j = 1|y,x) = \mathrm{sigm}(c_j + U_{jy} + \sum_j W_{ji}\, x_i)$（4-9）

where sigm(a) = 1/(1 + exp(−a)) is the logistic sigmoid function. When conditioning on the hidden layer, we have

$$\mathrm{p}(x|h) = \prod_j p(x_i|h), \text{ with } p(x_i = 1|h) = \mathrm{sigm}(b_i + \sum_j W_{ji}\, h_j) \qquad (4\text{-}10)$$

$$\mathrm{p}(y|h) = \frac{\exp(d_y + \sum_j U_{jy} h_j)}{\sum_{y^*} \exp(d_{y^*} + \sum_j U_{jy} h_j)} \qquad (4\text{-}11)$$

It is also possible to compute p(y|x) exactly and hence perform classification. Indeed, noticing that:

$$\sum_{h_1 \in \{0,1\}} \cdots \sum_{h_H \in \{0,1\}} \exp(h^T W_x + b^T x + c^T h + d^T e_y + h^T U_{ey})$$

$$= \exp(d_y) \sum_{h_1 \in \{0,1\}} \exp(h_1(c_1 + U_{1y} + \sum_i W_{1i} x_i)) \cdots \sum_{h_H \in \{0,1\}} \exp(h_H(c_H + U_{Hy} + \sum_i W_{Hi} x_i))$$

$$= \exp(d_y)(1+\exp(c_1 + U_{1y} + \sum_i W_{1i}x_i)) \cdots (1+\exp(c_n + U_{Hy} + \sum_i W_{ni}x_i))$$

$$= \exp(d_y + \sum_j \log(1 + \exp(c_j + U_{jy} + \sum_i W_{ji}x_i)))$$

$$= \exp(d_y + \sum_j \text{softolus}(c_j + U_{jy} + \sum_i W_{ji}x_i)) \quad (4\text{-}12)$$

where softolus (a) = log(1+exp(a)), then we can write

$$\text{p}(y|x) = \frac{\exp(-F(y,x))}{\sum_{y^* \in \{1,\cdots,C\}} \exp(-F(y,x))} \quad (4\text{-}13)$$

Where $F(y,x)$ is referred to as the free energy. Precomputing the terms $c_j + \sum_i W_{ji}x_i$ and reusing them when computing $\text{softolus}(c_j + U_{jy^*} + \sum_i W_{ji}x_i)$ for all classes y^* yields a procedure for computing this conditional distribution in time *O(HD+HC).*

One way of interpreting Equation 2 is that, when assigning probabilities to a particular class y for some input x, the ClassRBM looks at how well the input x fits or aligns with the different filters associated with the rows W_j of W. These filters are shared across the different classes, but different classes will make comparisons with different filters by controlling the class-dependent biases U_{jy} in the $\text{softolus}(c_j + U_{jy} + \sum_i W_{ji}x_i)$ terms. Notice also that two similar classes could share some of the filters in W, that is, both could simultaneously have large positive values of U_{jy} for some of the rows W_j.

4.2.3 Feature extraction by CRBM

Fig.4-3 shows the flow of the self-adaptive feature extraction method by CRBM.

According to Fig.4-3, the measured vibration signals of each single state are first reconstructed for the training sample (the reconstructed sample is X). Divide X into a training sample and a validation sample. The training sample is input to the CRBM model and trained.

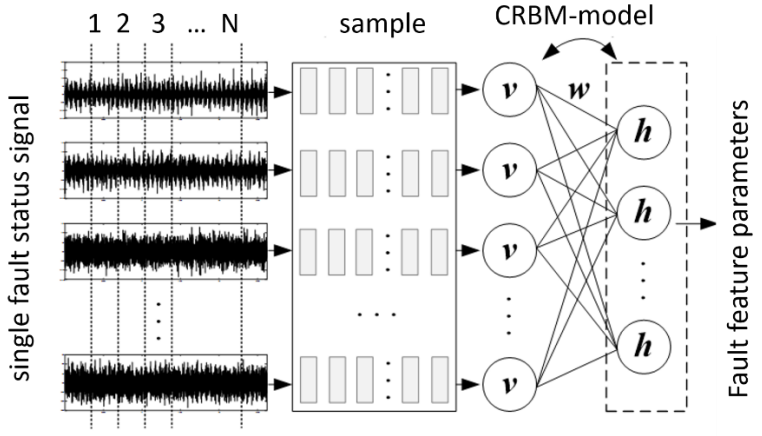


Fig.4-3 Flow of signal self-adaptive feature extraction method by CRBM

The input dimension of the CRBM model is the data length of the sample. For example, if the data length of the sample is 16384, the input dimension is also 16384. The number of elements in the CRBM hidden layer is the number of feature parameters. The hidden layer of the trained CRBM model corresponds to the feature parameter. By this method, it is possible to extract the characteristic parameters of each state signal by self-adaptation.

In order to integrate the structural feature parameters by principal component analysis, first, the calculated structural feature parameter values (P) are normalized.

Equation (3-14) is used for m main components (z_1 to z_m).

$$\begin{Bmatrix} z_1 \\ \vdots \\ z_m \end{Bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \begin{Bmatrix} p_1 \\ \vdots \\ p_m \end{Bmatrix} = AP \tag{4-14}$$

Here, *P* is a feature parameter sequence, *A* is called a coefficient matrix, and the coefficient a_{ij} is obtained from the data in the normal state (or reference state). Further, assuming that the eigenvalues of the covariance matrix of the structural feature parameter p_{si}'obtained in the normal state (or the reference state) are λ_1 to λ_m, λ_i is the standard deviation of the i-th principal component.

First, the signal of the reference state, which is the single abnormal state (outer race fault, inner race fault, roller fault) is HDTEO processed, the signal is divided into samples, and the characteristic parameters are extracted by CRBM. Using the feature parameter value as training data, the discriminant function (main component) for identifying a single state is obtained by Eq. (3-14). The first and second principal components obtained by the reference state discrimination function are represented by ,. Here, O (outer ring wound), I (inner ring wound), R (rolling body injury).

Then, let the first and second principal components at the time of diagnosis obtained by substituting the CRBM feature parameter values extracted from the data measured at the time of diagnosis into the principal component coefficients of each reference state. The contents described above are summarized in the table below.

Table 4-1 principal component coefficients

Training principal component z_{ik}	Diagnosis principal component z_{ix}
$z_{1k}\ (k=O,I,R) = \sum_{j=1}^{M} a_{1j} p_{jk}$	$z_{1x} = \sum_{j=1}^{M} a_{1j} p_{jx}$
$z_{2k}\ (k=O,I,R) = \sum_{j=1}^{M} a_{2j} p_{jk}$	$z_{2x} = \sum_{j=1}^{M} a_{2j} p_{jx}$

4.3 The early fault diagnosis based on HDTEO-CRBM

The diagnosis processing method proposed in this section is described. The flow of the diagnostic method is shown in Fig.4-4. In the offine diagnosis stage, the HDTEO of the vibration signal measured in each reference state, that is, each single abnormal state (outer race fault, inner race fault, roller fault) is calculated, and different kinds of fault features are adaptively extracted by CRBM.

The principal component analysis of the characteristic parameters of each abnormal component extracted by CRBM is performed, and the first principal component coefficient, the second principal component coefficient, and the respective principal component scores are obtained. At the time of diagnosis, the HDTEO of the measured vibration signal is first obtained as in the learning stage, and the self-adaptive characteristic parameters are calculated using the good model learned in the learning stage. Next, the first principal component score and the second principal component score are matched with the principal component section of each reference state, and the abnormality type is specified by inference based on the matching score.

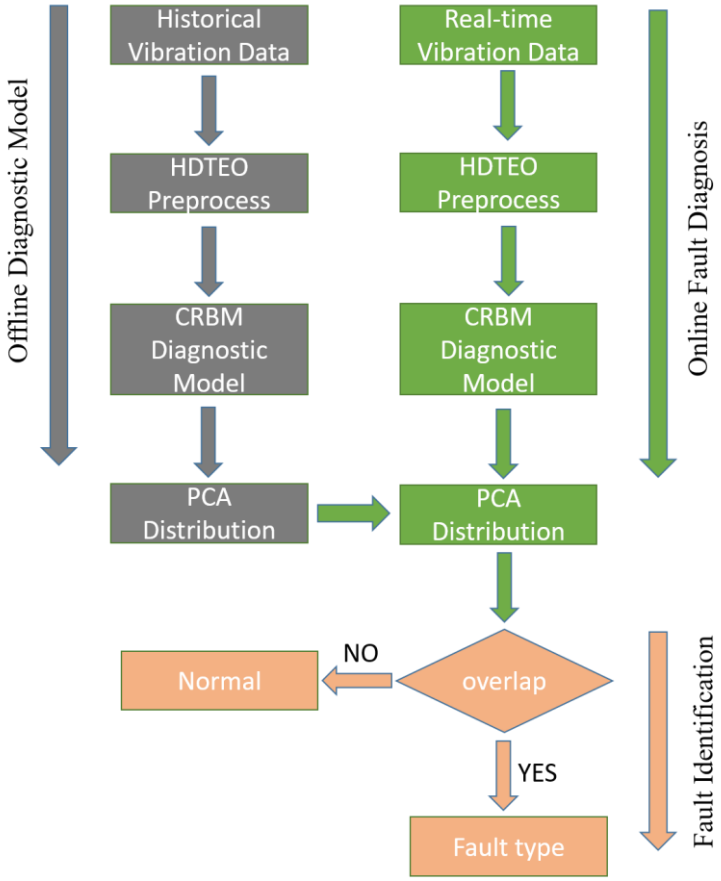


Fig.4-4 The faut diagnosis flowchart of the proposed method

4.4 Case study

4.4.1 Experiment platforms

In order to verify the effectiveness of the proposed method, raw signals of bearings in different states (normal, outer race fault, inner race fault, and roller fault) were collected from the experimental platform, as shown in Fig.4-5. There are three kinds of beatings in the bearing as shown in Fig.4-5, Fig. 4-6, and Fig.4-7. In order to more realistically simulate the early failure data, the bearing failure is due to manual cutting, the depth of damage is 0.3 mm, and four different sizes of faults in total from 0.6mm to 5.0mm in width as shown in Table 4-2. In addition, we have added the measurement of various fault signals when the lubrication is poor. The signals sampling frequency is 100 kHz， and the sampling time was approximately 20s. As shown in Table 4-3, there are eight kinds of rotating speeds in this experimental design, 10 rpm, 20 rpm, 30 rpm, 40 rpm, 50 rpm, 60 rpm, 70 rpm and 100 rpm. The naming methods of different vibration signals are shown in the Table 4-3, Table 4-4 and Table 4-5.

W: 0.6 D: 0.3

W: 1.2 D: 0.3

W: 2.0 D: 0.3

W: 5.0 D: 0.3

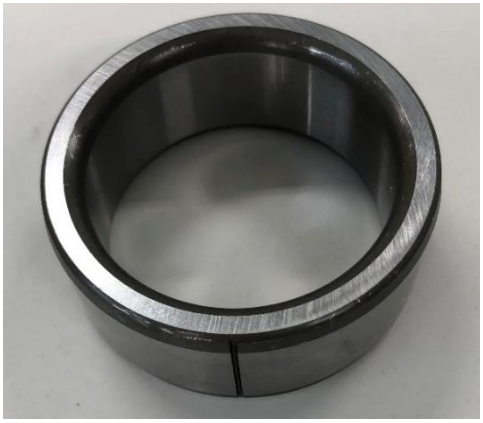
Fig.4-5 The outer race fault bearings

W: 0.6 D: 0.3

W: 1.2 D: 0.3

W: 2.0 D: 0.3	W: 5.0 D: 0.3

Fig.4-6 The roller fault bearings



W: 0.6 D: 0.3	W: 1.2 D: 0.3

W: 2.0 D: 0.3	W: 5.0 D: 0.3

Fig.4-7 The inner race fault bearings

Table 4-2 Low Speed Bearing Failure Types

Type of failure	Lubrication	Model	Wound size
Outer race	Lubrication normal	N312	0.6W＊0.3d
	Lubrication normal	N312	1.2W＊0.3d
	Lubrication normal	N312	2.0W＊0.3d
	Lubrication normal	N312	5.0W＊0.3d
	Poor lubrication	1312	2.0W＊0.3d

Type of failure	Lubrication	Model	Wound size
Inner race	Lubrication normal	NU312	0.6W＊0.3d
	Lubrication normal	NU312	1.2W＊0.3d
	Lubrication normal	NU312	2.0W＊0.3d
	Lubrication normal	NU312	5.0W＊0.3d
	Poor lubrication	1312	2.0W＊0.3d

Type of failure	Lubrication	Model	Wound size
Roller race	Lubrication normal	NU312	0.6W＊0.3d
	Lubrication normal	NU312	1.2W＊0.3d
	Lubrication normal	NU312	2.0W＊0.3d
	Lubrication normal	NU312	5.0W＊0.3d

Table 4-3 Outer race fault data naming

Type	Outer race fault									
Model	N312								1312	
RPM	CH7	CH8	CH7	CH8	CH7	CH8	CH7	CH8	CH7	CH8
10	O_0603_10_A	O_0603_10_E	O_1203_10_A	O_1203_10_E	O_2003_10_A	O_2003_10_E	O_5003_10_A	O_5003_10_E	O_1_10_A	O_132003_10_E

20	O_0603_20_A	O_0603_20_E	O_1203_20_A	O_1203_20_E	O_2003_20_A	O_2003_20_E	O_5003_20_A	O_5003_20_E	O_12003_20_A	O_12003_20_E
30	O_0603_30_A	O_0603_30_E	O_1203_30_A	O_1203_30_E	O_2003_30_A	O_2003_30_E	O_5003_30_A	O_5003_30_E	O_12003_30_A	O_1003_30_E
40	O_0603_40_A	O_0603_40_E	O_1203_40_A	O_1203_40_E	O_2003_40_A	O_2003_40_E	O_5003_40_A	O_5003_40_E	O_12003_40_A	O_12003_40_E
50	O_0603_50_A	O_0603_50_E	O_1203_50_A	O_1203_50_E	O_2003_50_A	O_2003_50_E	O_5003_50_A	O_5003_50_E	O_12003_50_A	O_2003_50_E
60	O_0603_60_A	O_0603_60_E	O_1203_60_A	O_203_60_E	O_2003_60_A	O_2003_60_E	O_5003_60_A	O_5003_60_E	O_12003_60_A	O_12003_60_E
70	O_0603_70_A	O_0603_70_E	O_1203_70_A	O_1203_70_E	O_2003_70_A	O_2003_70_E	O_5003_70_A	O_5003_70_E	O_12003_70_A	O_12003_70_E
100	O_0603_100_A	O_0603_100_E	O_1203_100_A	O_1203_100_E	O_2003_100_A	O_2003_100_E	O_5003_100_A	O_5003_100_E	O_12003_100_A	O_12003_100_E

Table 4-4 Roller fault data naming

Type	Roller fault							
Model	NU312							
RPM	CH7	CH8	CH7	CH8	CH7	CH8	CH7	CH8
10	E_0603_10_A	E_0603_10_E	E_1203_10_A	E_1203_10_E	E_2003_10_A	E_2003_10_E	E_5003_10_A	E_5003_10_E
20	E_0603_20_A	E_0603_20_E	E_1203_20_A	E_1203_20_E	E_2003_20_A	E_2003_20_E	E_5003_20_A	E_5003_20_E
30	E_0603_30_A	E_0603_30_E	E_1203_30_A	E_1203_30_E	E_2003_30_A	E_2003_30_E	E_5003_30_A	E_5003_30_E
40	E_0603_40_A	E_0603_40_E	E_1203_40_A	E_1203_40_E	E_2003_40_A	E_2003_40_E	E_5003_40_A	E_5003_40_E
50	E_0603_50_A	E_0603_50_E	E_1203_50_A	E_1203_50_E	E_2003_50_A	E_2003_50_E	E_5003_50_A	E_5003_50_E

60	E_0603_60_A	E_0603_60_E	E_1203_60_A	E_1203_60_E	E_2003_60_A	E_2003_60_E	E_5003_60_A	E_5003_60_E
70	E_0603_70_A	E_0603_70_E	E_1203_70_A	E_1203_70_E	E_2003_70_A	E_2003_70_E	E_5003_70_A	E_E_5003_70_E
100	E_0603_100_A	E_0603_100_E	E_1203_100_A	E_1203_100_E	E_2003_100_A	E_2003_100_E	E_5003_100_A	E_5003_100_E

Table 4-5 Inner race fault data naming

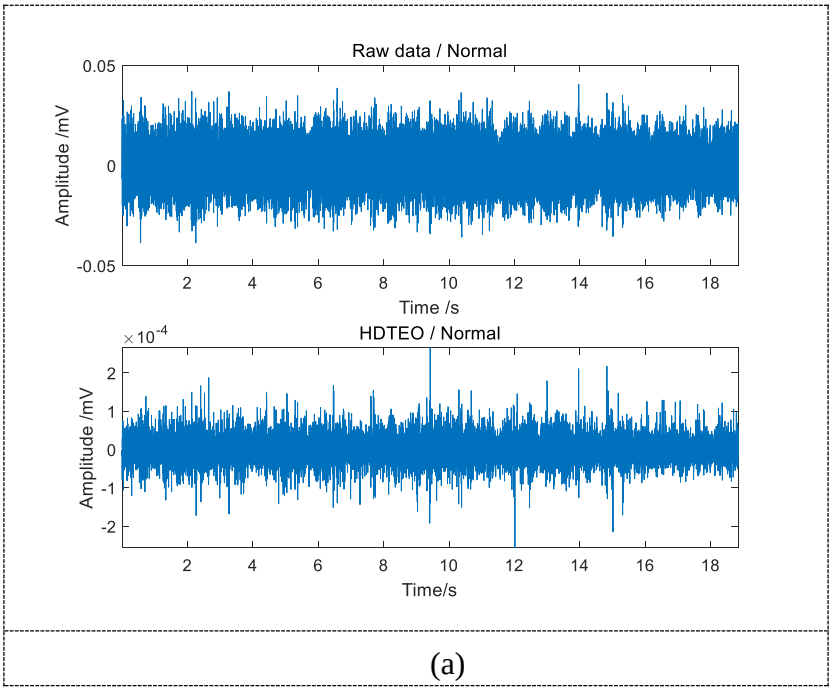
Type	Inner race fault									
Model	NU312								1312	
RPM	CH7	CH8	CH7	CH8	CH7	CH8	CH7	CH8	CH7	CH8
10	I_0603_10_A	I_0603_10_E	I_1203_10_A	I_1203_10_E	I_2003_10_A	I_2003_10_E	I_5003_10_A	I_5003_10_E	I_12003_10_A	I_12003_10_E
20	I_0603_20_A	I_0603_20_E	I_1203_20_A	I_1203_20_E	I_2003_20_A	I_2003_20_E	I_5003_20_A	I_5003_20_E	I_12003_20_A	I_12003_20_E
30	I_0603_30_A	I_0603_30_E	I_1203_30_A	I_1203_30_E	I_2003_30_A	I_2003_30_E	I_5003_30_A	I_5003_30_E	I_12003_30_A	I_12003_30_E
40	I_0603_40_A	I_0603_40_E	I_1203_40_A	I_1203_40_E	I_2003_40_A	I_2003_40_E	I_5003_40_A	I_5003_40_E	I_12003_40_A	I_12003_40_E
50	I_0603_50_A	I_0603_50_E	I_1203_50_A	I_1203_50_E	I_2003_50_A	I_2003_50_E	I_5003_50_A	I_5003_50_E	I_12003_50_A	I_12003_50_E
60	I_0603_60_A	I_0603_60_E	I_1203_60_A	I_1203_60_E	I_2003_60_A	I_2003_60_E	I_5003_60_A	I_5003_60_E	I_12003_60_A	I_12003_60_E
70	I_0603_70_A	I_0603_70_E	I_1203_70_A	I_1203_70_E	I_2003_70_A	I_2003_70_E	I_5003_70_A	I_5003_70_E	I_12003_70_A	I_12003_70_E
100	I_0603_100_A	I_0603_100_E	I_1203_100_A	I_1203_100_E	I_2003_100_A	I_2003_100_E	I_5003_100_A	I_5003_100_E	I_12003_100_A	I_12003_100_E

4.4.2 Intelligent diagnosis based on proposed method

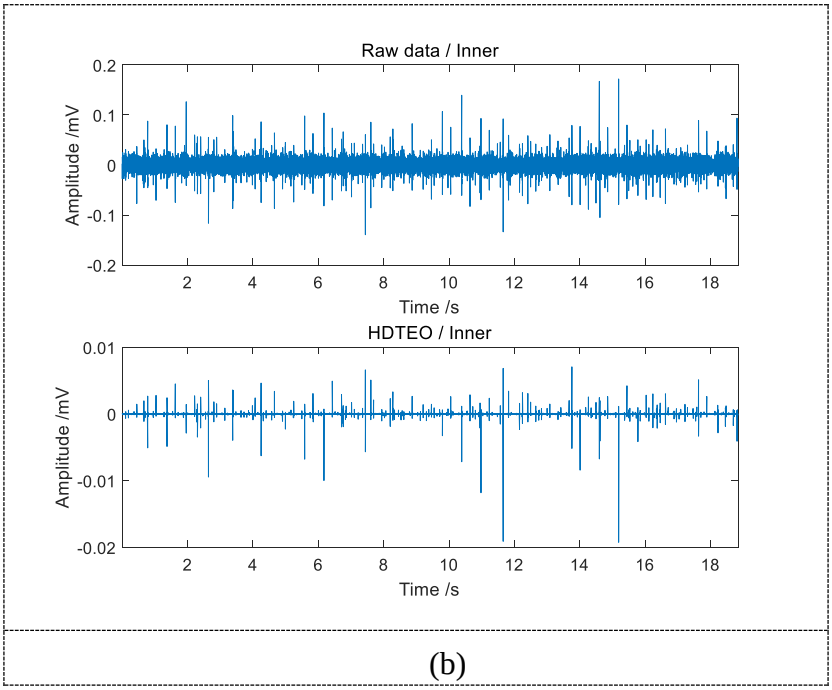
In order to verify the effectiveness of the method proposed in this study, a bearing with a single artificial scratch (outer race fault, inner race fault, and roller fault) was used as an experimental installed in the rotating machine. It was attached and a verification experiment was conducted. The size of the bearing scratches used in the verification content of this report is outer ring scratches, inner ring scratches. In order to better simulate the early failure of real bearings, in this experiment, the rolling bearings with scratch size 5 0.6 mm * 0.3 mm (W * D) are selected.

1) The extraction of the early low-speed rolling bearings

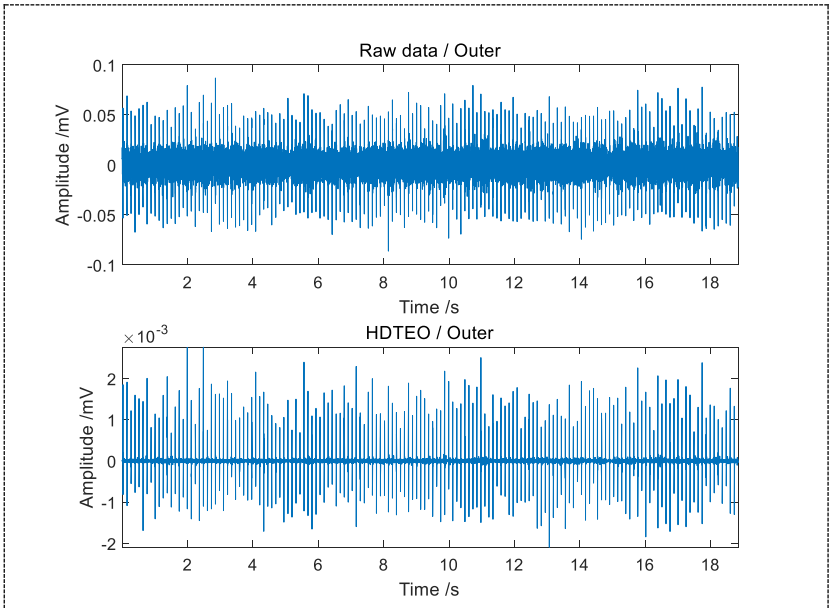
The HDTEO method is used to process signals measured in a single bearing state (normal state, outer ring scratches, inner ring scratches and rolling element scratches). The length (number) of discrete signals in each state is 376832. The results at 20 RPM will be used below. Fig.4-8 shows the result of HDTEO processing. From Fig.4-8, it can be seen that the instantaneous signal can be strengthened and the impact component in the signal can be detected by HDTEO processing, and the characteristics of the processed signal can be clearly confirmed.



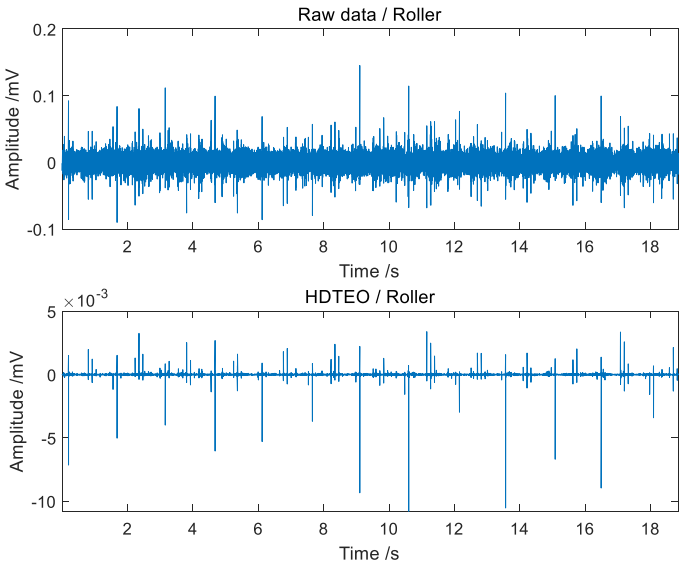
(a)


Raw data / Inner
Amplitude /mV
0.2
0.1
0
-0.1
-0.2
2
4
6
8
10
12
14
16
18
Time /s
HDTEO / Inner
Amplitude /mV
0.01
0
-0.01
-0.02
2
4
6
8
10
12
14
16
18
Time /s

(b)


Raw data / Outer
Amplitude /mV
0.1
0.05
0
-0.05
-0.1
2
4
6
8
10
12
14
16
18
Time /s
HDTEO / Outer
×10-3
Amplitude /mV
2
1
0
-1
-2
2
4
6
8
10
12
14
16
18
Time /s

(c)



(d)

Fig.4-8 Comparison before and after fault feature extraction

2) Adaptive feature parameter extraction based on CRBM

Feature parameters are self-adaptively extracted by CRBM from the signal processed by HDTEO. The rotation speed is 10-100 RPM. The data length of the sample is 16384 to include information for one or more rotations. The number of samples in this paper is 23. Among them, the number of learning samples is 20, and the number of verification samples is 3. That is, the signal length of each state is set to 376832 (23 * 16384). The experimental results at 10 RPM are shown below as an example.

The number of hidden layers of CRBM is 50, and the number of repeated learnings is 300. Set the Batch size to 80 and the learning rate to 0.01. The number of dimensions input to CRBM is 16384, which is the same as the sample data length. The number of input samples in the learning stage is 80 (20X4, 4 types), and Fig. 7 shows the input data in the learning stage and the data that has been trained by CRBM. As shown in Fig. 7, it can be seen that the raw data is distorted before processing, but it becomes clean after learning by CRBM. It should be noted that there is no method for theoretically determining the optimum network structure in deep learning, and the current situation is that it is determined by trial and error.

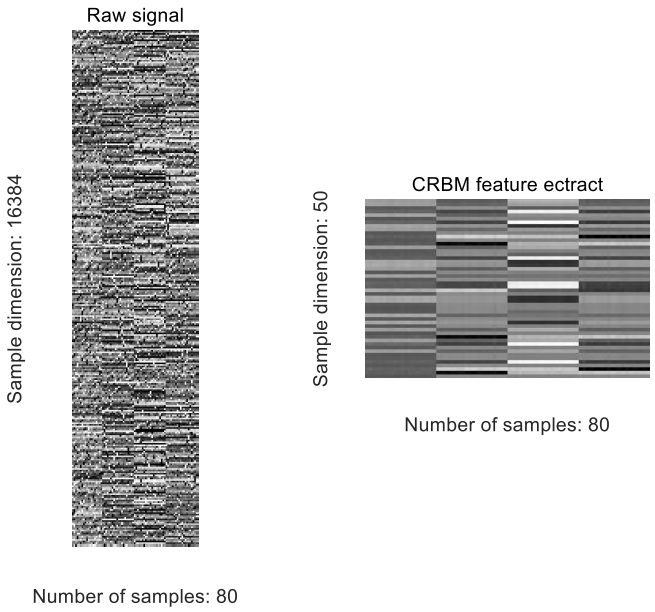


Fig.4-9 CRBM based fault feature extraction

3) The low-speed early fault diagnosis based on proposed method

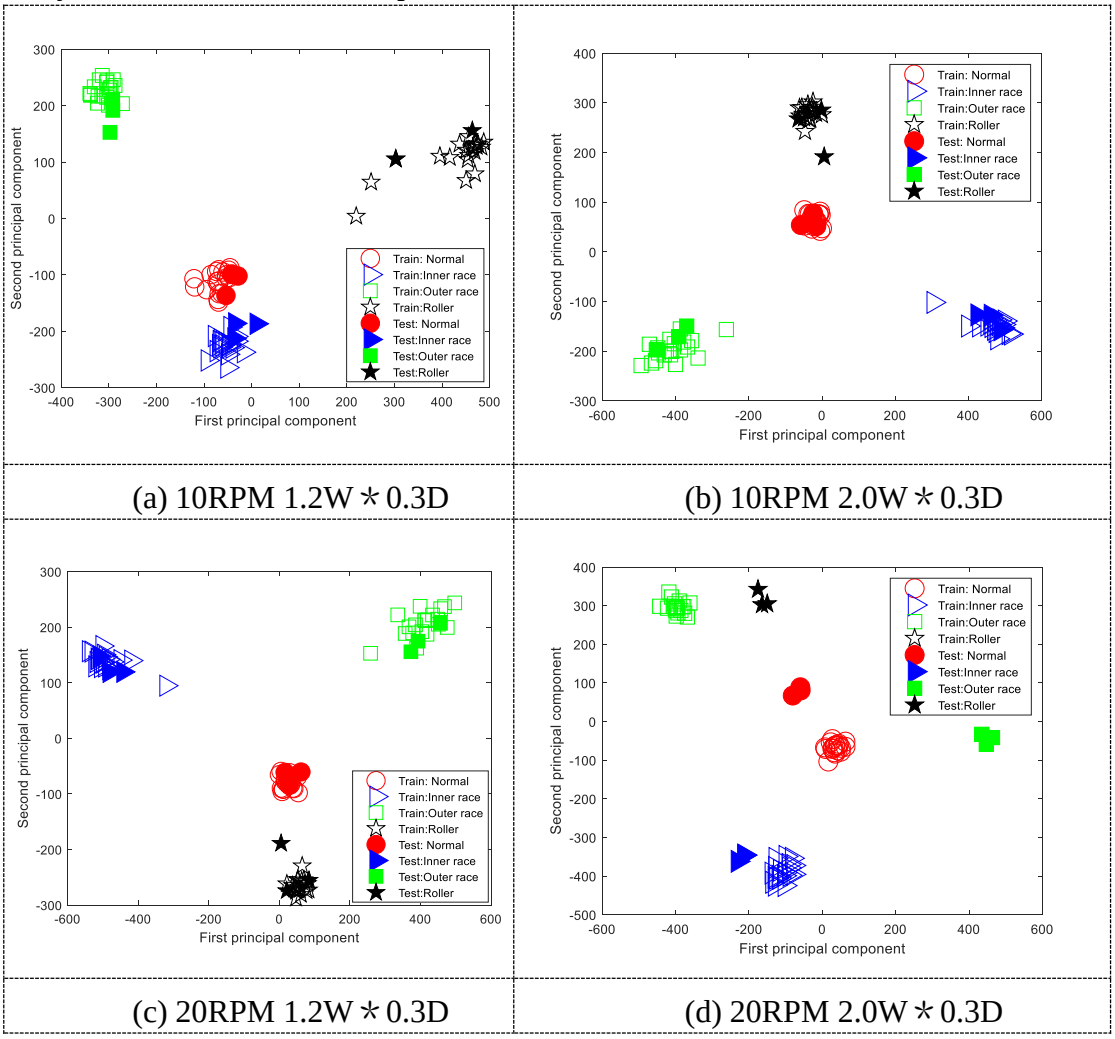
The feature parameters extracted self-adaptive by CRBM are analyzed by the PCA method to obtain eigenvalues, contribution rate and cumulative contribution rate. Table 2 shows the eigenvalues, contributions, and cumulative contributions obtained by PCA analysis at 10 RPM.

From Table 4-6, the cumulative contribution rate of the first and second eigenvalues in the learning stage is 80% or more. Then, bearing abnormality diagnosis is performed using the first and second principal components. Plot the principal components on the axis system. When the scores of the first and second principal components of each state are plotted in coordinates, a crowd of each state in the learning stage appears. That is, points in different fault states appear at distant places, and points in the same fault state appear at places close to each other.

Table 4-6 Eigenvalues, contributions and cumulative contributions

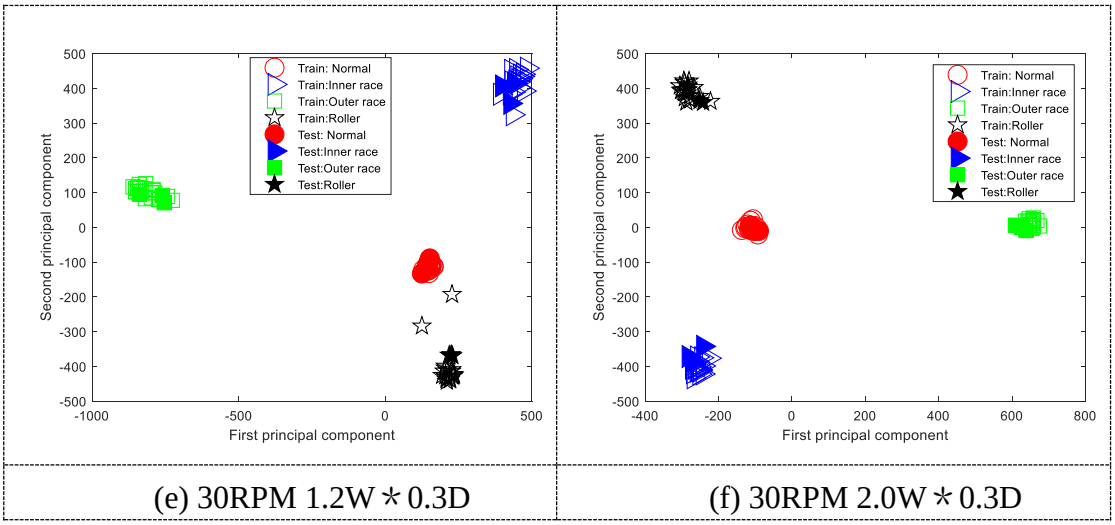
	Eigenvalues	contributions（%）	cumulative contributions（%）
1	87389.54	55.4950	55.4950
2	54313.86	34.4909	89.9859
3	15718.12	9.9815	99.9673
…	…	…	…

The scores of the first principal component and the second principal component of each state in the verification stage can be discriminated from each other by the intersection of the scores in the learning stage. Fig.4-10 shows the results of PCA analysis for different rotation speeds.

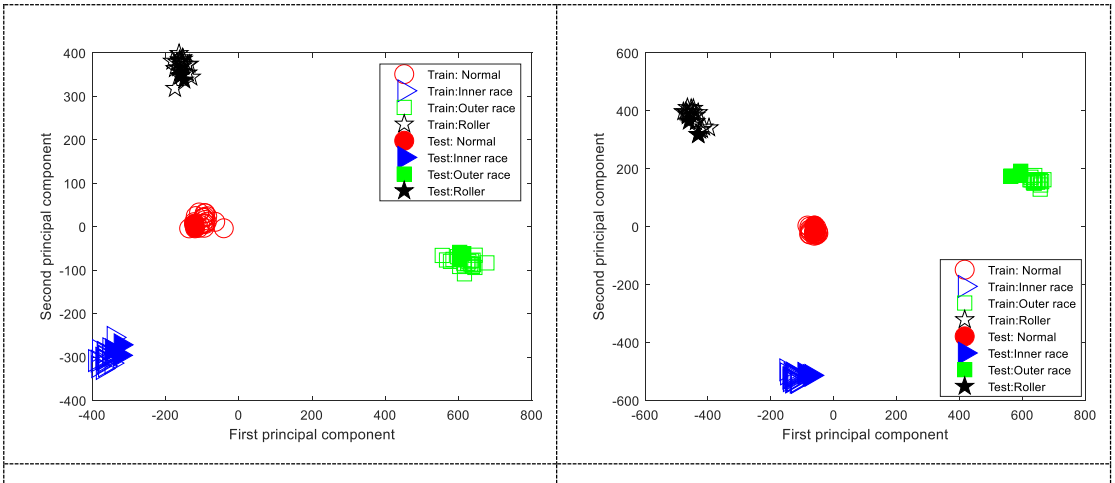


(a) 10RPM 1.2W＊0.3D　(b) 10RPM 2.0W＊0.3D

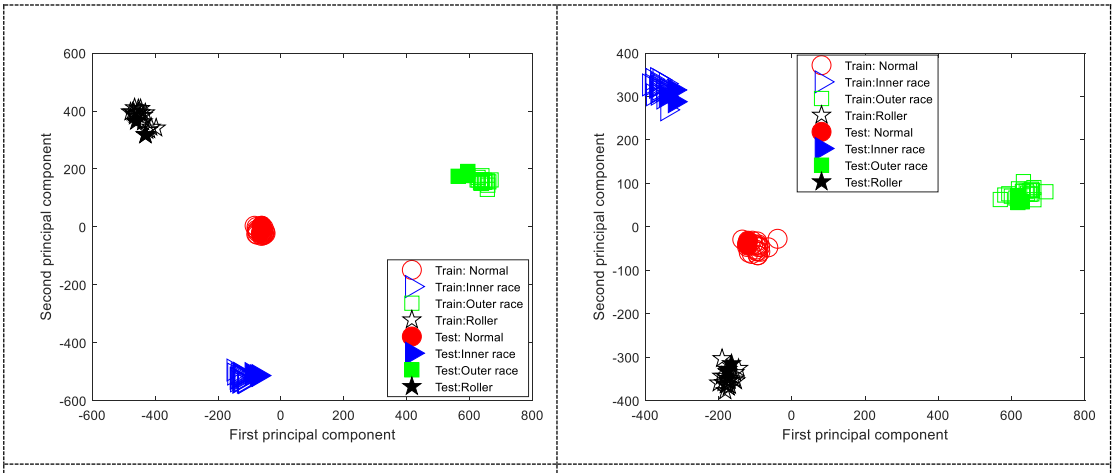
(c) 20RPM 1.2W＊0.3D　(d) 20RPM 2.0W＊0.3D



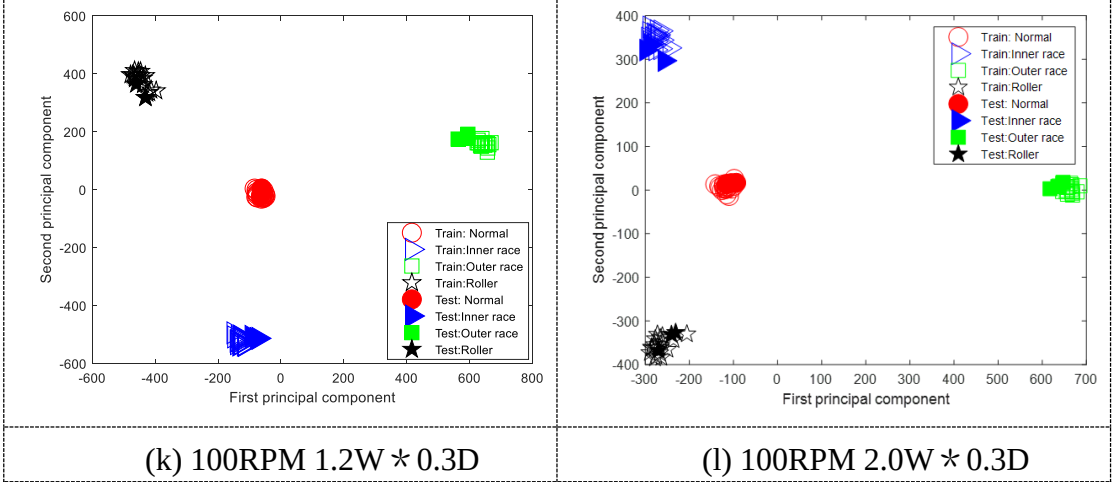
(e) 30RPM 1.2W＊0.3D　(f) 30RPM 2.0W＊0.3D



(g) 50RPM 1.2W＊0.3D　　(h) 50RPM 2.0W＊0.3D



(i) 70RPM 1.2W＊0.3D　　(j) 70RPM 2.0W＊0.3D



(k) 100RPM 1.2W＊0.3D　　(l) 100RPM 2.0W＊0.3D

Fig.4-10 Part of the experimental results

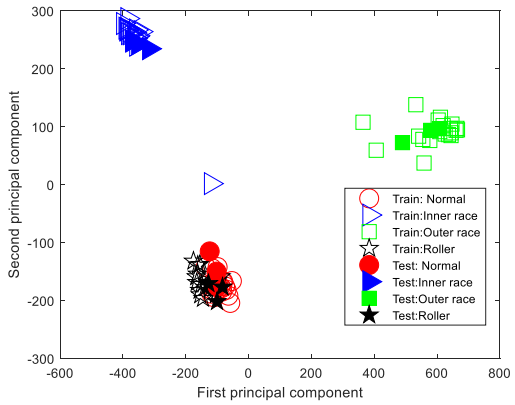
Figure 4-10 shows some results of low-speed shaft fault diagnosis based on the

proposed method. It can be seen from the figure that the proposed method can effectively extract fault features and classify and identify different fault types. Fig4-10(a),(b). Respectively represent the fault diagnosis results of the bearing with a fault size of 1.2W*0.3D and 2.0W*0.3D when the bearing rotates at an ultra-low speed of 10RPM. The smaller size of the damage can be closer to early bearing failure in real life. It can be seen from the experimental results that although the roller fault diagnosis results are somewhat biased, the overall diagnosis effect is good, which can prove the effectiveness of the proposed method for early fault diagnosis. Fig4-10(c), (d). Respectively represent the fault diagnosis results of the bearing with a flaw size of 1.2W*0.3D and 2.0W*0.3D when the bearing rotates at an ultra-low speed of 20RPM. For the faulty bearing with a damage size of 1.2W*0.3D, Although there is a small amount of deviation in the roller fault diagnosis, it is still within the allowable range of the error, and the fault type can still be identified. For the faulty bearing with a damage size of 2.0W*0.3D, the proposed method can accurately identify three different types of bearing faults, which proves the effectiveness of the proposed method. Fig4-10(e), (f). Respectively represent the fault diagnosis results of the bearing with a flaw size of 1.2W*0.3D and 2.0W*0.3D when the bearing rotates at an ultra-low speed of 30RPM, For the flaw size is 1.2W*0.3D the two states of roller fault and normal are closely distributed, but the deviation is still within the allowable range, which can effectively identify the fault type. For bearings with a damage size of 2.0W*0.3D, different fault types can be accurately identified. For the fault diagnosis of bearing speed above 50RPM, it can be seen from Fig4-10(g), (h), (i), (j), (k), (l) that the proposed method can effectively identify For different types of bearing faults with dimensions of 1.2W*0.3D and 2.0W*0.3D, the experiments demonstrate the effectiveness of the proposed method for early fault diagnosis of bearings operating at low speeds.

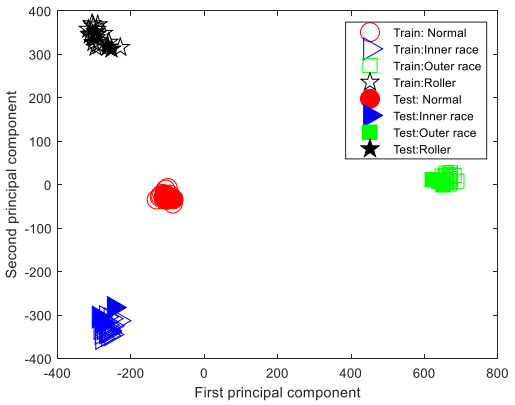
4.5 Comparative experiment

In order to realize the fault diagnosis of low-speed bearings, the third chapter of this paper proposes a fault feature extraction method based on the improved teagar energy operator, and analyzes and diagnoses three faulty bearing data with a damage size of 2.0 W*0.3D. The results show that the method can identify 70RPM and 100RPM low speed bearing faults. However, for early fault bearings, the fault characteristics are relatively weak. In order to identify the early fault diagnosis of low-speed bearings, this section proposes a bearing fault diagnosis method based on deep learning CRBM. The

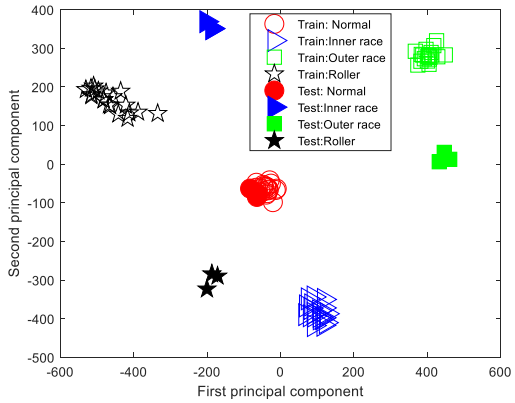
characteristic of the early bearing failure is that the damage size is small. Therefore, the fault bearing signal with the damage size of 0.6W*0.3D at ultra-low speed of 10RPM and 20RPM is used. The comparison results of the two methods are shown in Fig.4-11:



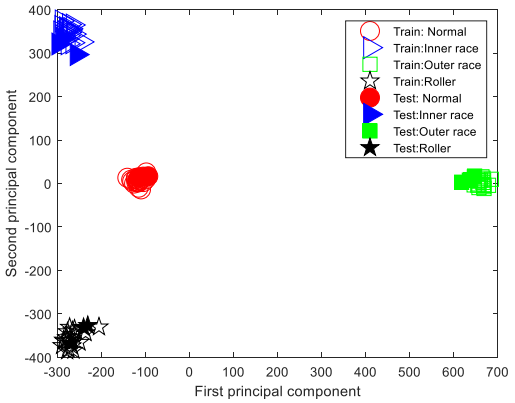
(a) 10RPM 0.6W＊0.3D AE



(b) 10RPM 0.6W＊0.3D CRBM



(c) 20RPM 0.6W＊0.3D AE



(d) 20RPM 0.6W＊0.3D CRBM

Fig. 4-11Comparative experimental results

As can be seen from the figure, for bearing fault diagnosis with ultra-low speed of 10RPM and damage size of 0.6W*0.3D, although deep learning AE can identify the fault type of outer race fault, the two states of normal and roller fault overlap, and some parts of the distribution of inner race faults is scattered. The proposed method performs better in this case. For bearing fault diagnosis with a low speed of 20RPM and a damage size of 0.6W*0.3D, deep learning AE can only identify the normal state, and cannot determine the type of bearing fault. The proposed method can still clearly identify different faults under this condition, which further proves the effectiveness of the proposed method.

4.6 Summary

In this section, in order to diagnose early fault of rotary machinery, we proposed a method to accurately identify abnormal types by HDTEO-CRBM-PCA and applied it to rolling bearings early fault diagnosis. The Teager energy operator has a clear advantage in the amount of calculation and the efficiency of algorithm execution, which can detect the impact component in the signal more strongly.

Compared to the conventional feature parameter extraction method, the method can automatically extract signal features in a self-adaptive manner, to realize automatic diagnosis of rolling bearings early fault. In addition, the three kind of rolling bearing fault signals is used in the experimental, a total of 7 different rotational speeds from 10RPM to 100RPM were used for the experiment. And the effectiveness of the proposed method was proven. In the future, the signal processing method and identification method proposed in this research will be introduced into the actual equipment diagnosis system, and will be applied and improved to on-site rotating machines under various operating conditions.

5. Conclusions and Future works

5.1 Conclusions

5.1.1 Result and discussion

In this thesis, a variety of bearing fault feature extraction and diagnosis methods have been proposed. These proposed algorithms and theories have been verified by various simulation experiments and case studies. Firstly, the bearing structure and characteristics are explained, and the calculation method of bearing fault characteristic frequency is described in detail. Existing bearing fault diagnosis methods based on vibration signals are listed, and the advantages and disadvantages of these methods are briefly analyzed. In order to better track the degradation of bearing fault state, a bearing fault feature extraction and intelligent diagnosis method based on state-space principal component tracking filtering (SPCTF) is proposed in chapter 2. The proposed method can dynamically track the bearing fault status and detect bearing faults in real time. Experiments on bearing data of three different speeds (500RPM, 1000RPM, and 1500RPM) and three fault states (inner race fault, outer race fault and roller fault) show the effectiveness of the proposed method. Fault feature extraction is the key to bearing fault diagnosis, low-speed running bearing fault signals are easily overwhelmed by noise, Ordinary fault diagnosis methods based on envelope spectrum analysis of fault characteristic frequencies cannot effectively detect bearing faults. In order to effectively extract low-speed bearing fault features, an improved Teager energy operator (ITEO) is proposed in chapter 3. ITEO overcomes the shortcomings of sensitive to noise and vibration interference of the traditional TEO, through the method of amplifying the shock energy of the vibration signal, the fault features of the low-speed bearing are extracted. Then, an intelligent diagnosis model of low-speed bearings based on ITEO-AE is proposed, which can realize the fault diagnosis of low-speed bearings. The effectiveness of the proposed method is proved by the diagnostic experiments of bearing vibration signals for two different speeds (70RPM and 100RPM) and three fault states (inner race fault, outer race fault and roller fault). Once the rolling bearing fails, it will often cause equipment damage, factory shutdown, and serious economic losses. Therefore, early diagnosis is very important for rolling bearing fault diagnosis. In this thesis, searing signals of small size faults are used in experiments. In order to more realistically simulate the early failure of bearings, three different sizes of faulty bearings (0.6W*0.3D, 1.2W*0.3D and 2.0W*0.3D) were used in the experiments. In Chapter 4,

the early fault diagnosis model of rolling bearing based on CRBM-PCA is proposed，which can adaptively extract the early faults of rolling bearings， excellent for diagnosing faulty bearings with small dimensions.

5.1.2 Experiments and data sampling

To better verify the validity of the proposed method, the experimental data used are all obtained by the operation of laboratory mechanical equipment and measured by vibration sensors. We obtain fault bearing data at 500 RPM-1500 RPM through the high-speed platform, bearings with different failure types are used for data acquisition. The low-speed experimental platform played an important role to better study low speed and early bearing failures. Faulty bearings of different sizes were used in the experiments. These data are very valuable and important for the research of bearing fault diagnosis method.

5.2 Future works

This thesis corresponding solutions are proposed for bearing fault diagnosis at different speeds. These methods work well for a single type of bearing fault, but in real industry, there are often multiple fault fusion bearings. In the future, it is necessary to find suitable methods to identify compound failures of rolling bearings.

References

[1] Li, Y. Wang, X. Liu, Z. Liang, X. Si, S. The entropy algorithm and its variants in the fault diagnosis of rotating machinery: A review. IEEE Access 2018, 6, 66723–66741.

[2] Y. Wang, H. Zhang, S. Wei, D. Zhou, and B. Huang, Control performance assessment for ILC-controlled batch processes in a 2-D system framework, IEEE Trans. Syst., Man, Cybern., Syst., 2018, 48, 1493–1504.

[3] A.K. Jardine. D. Lin. D. Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance, Mech. Syst. Sign. Process. 2006, 20, 1483–1510.

[4] M. Cerrada et al., 'A review on data-driven fault severity assessment in rolling bearings, Mech. Syst. Signal Process., 2018, 99, 169–196.

[5] P. Baraldi. L. Podofillini. L. Mkrtchyan. E. Zio, V.N. Dang, Comparing the treatment of uncertainty in bayesian networks and fuzzy expert systems used for a human reliability analysis application, Reliab. Eng. Syst. Saf. 2015, 138, 176–193

[6] A. Yunusa-Kaltungo. J.K. Sinha. A.D. Nembhard. A novel fault diagnosis technique for enhancing maintenance and reliability of rotating machines, Struct. Health Monit. 2015, 14 , 604–621.

[7] Osman S and Wang W. An enhanced Hilbert–Huang transform technique for bearing condition monitoring Meas. Sci. Technol. 2013, 24, 085004.

[8] M. M. Tahir, A. Q. Khan, N. Iqbal, A. Hussain, and S. Badshah, Enhancing fault classification accuracy of ball bearing using central tendency based time domain features, IEEE Access, 2017, 5, 72–83.

[9] Q. B. He, Z. H. Feng, and F. R. Kong, Principal component representations for machine noise monitoring, Guangxue Jingmi Gongcheng/Optics and Precision Engineering. 2006, 14, 1093-1099.

[10] R.B. Randall, J. Antoni, Rolling element bearing diagnostics-a tutorial, Mech. Syst. Signal Process. 2011, 25, 485–520.

[11] J. Tian, C. Morillo. M.H. Azarian. M. Pecht. Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with k-nearest neighbor distance analysis, IEEE Trans. Ind. Electron. 2016, 63, 1793–1803.

[12] H. Huang, N. Baddour, M. Liang, Auto-OBSD: automatic parameter selection for

reliable oscillatory behavior-based signal decomposition with an application to bearing fault signature extraction, Mech. Syst. Signal Process. 2017, 86, 237–259.

[13] X. Wan. D. Wang. W.T. Peter, G. Xu. Q. Zhang, A critical study of different dimensionality reduction methods for gear crack degradation assessment under different operating conditions, Measurement 2016, 78, 138–150.

[14] Y. Lei, J. Lin, Z. He, and M. J. Zuo, A review on empirical mode decom-position in fault diagnosis of rotating machinery, Mech. Syst. Signal Process., 2013, 35, 108–126.

[15] S. Yin, O. Kaynak. Big data for modern industry: challenges and trends, Proc. IEEE 2015, 103, 143–146.

[16] F. Piltan. J. M. Kim. Bearing fault diagnosis using an extended variable structure feedback linearization observer. Sensors, Dec. 2018, 18, 4359.

[17] L. Bo. G. Xu. X. Liu. J. Lin. Bearing fault diagnosis based on sub band time-frequency texture tensor, IEEE Access, 2019, 7, 37611-37619.

[18] D. T. Hoang. H. J. Kang. Rolling element bearing fault diagnosis using convolutional neural network and vibration image, Cognitive Systems Research, 2018, 53, 42-50.

[19] V. Purushotham. S. Narayanan. S. A. Prasad. Multi-fault diagnosis of rolling bearing elements using wavelet analysis and hidden Markov model based fault recognition, Ndt & E International, 2005, 38, 654-664.

[20] HD. Shao. HK. Jiang. XQ. Li, SP. Wu. Intelligent fault diagnosis of rolling bearing using deep wavelet auto-encoder with extreme learning machine. Knowledge-Based Systems, 2017, 140, 1-14.

[21] Y. Lei. J. Lin. Z. He. M. J. Zuo. A review on empirical mode decomposition in fault diagnosis of rotating machinery. Mechanical systems and signal processing, 2012, 35, 108-126.

[22] X. Zhang. Y. Liang. J. Zhou. A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM, Measurement, 2015, 69, 164-179.

[23] M. Kedadouche. M. Thomas. A. J. M. S. Tahan. A comparative study between Empirical Wavelet Transforms and Empirical Mode Decomposition Methods: Application to bearing defect diagnosis, Mechanical Systems and Signal Processing, 2016, 81, 88-107.

[24] R. Abdelkader. A. Kaddour. Z. Derouiche. Enhancement of rolling bearing fault

diagnosis based on improvement of empirical mode decomposition denoising method, The International Journal of Advanced Manufacturing Technology, 2018, 97, 3099-3117.

[25] X. Wang. Z. Yang. X. Yan. Novel particle swarm optimization-based variational mode decomposition method for the fault diagnosis of complex rotating machinery, IEEE/ASME Transactions on Mechatronics, 2018, 23, 68-79.

[26] C. Yi. Y., Lv. Z. Dang. A fault diagnosis scheme for rolling bearing based on particle swarm optimization in variational mode decomposition, Shock and Vibration, 2016, 9372691.

[27] S. Zhang. Y. Wang. S. He, Z. Jiang. Bearing fault diagnosis based on variational mode decomposition and total variation denoising, Measurement Science and Technology, 2016, 27, 075101.

[28] L. Wang. Z. Liu. Q. Miao. X. Zhang. Time–frequency analysis based on ensemble local mean decomposition and fast kurtogram for rotating machinery fault diagnosis, Mechanical Systems and Signal Processing, 2017, 103, 60-75.

[29] P. Shi. X. Ma. D. Han. A weak fault diagnosis method for rotating machinery based on compressed sensing and stochastic resonance. Journal of Vibroengineering, 2019, 654-664.

[30] G. Li. J. Li. S. Wang, X. Chen. Quantitative evaluation on the performance and feature enhancement of stochastic resonance for bearing fault diagnosis, Mechanical Systems and Signal Processing, 2016, 81, 108-125.

[31] Z. Li. X. Liu. X. Wang. T. He. Y. Shan. A multi-parameter constrained potential underdamped stochastic resonance method and its application for weak fault diagnosis, Journal of Sound and Vibration, 2019, 459, 114862.

[32] H. Wang. P. Chen. Fault diagnosis for a rolling bearing used in a reciprocating machine by adaptive filtering technique and fuzzy neural network, WSEAS Transactions on Systems, 2008, 7, 1-6.

[33] Z. Liao. L. Song. P. Chen. S. Zuo. An automatic filtering method based on an improved genetic algorithm—with application to rolling bearing fault signal extraction, IEEE Sensors Journal, 2017, 17, 19, 6340-6349.

[34] J. Long. H. Wang. D. Zha. P. Li. H. Xie. L. Mao. Applications of fractional lower order S transform time frequency filtering algorithm to machine fault diagnosis, PloS one, 2017,12, 4, e0175202.

[35] Z. Liu. L. Tang. L. Cao. Feature Extraction Method for Rolling Bearing's Week

Fault Based on Kalman Filter and FSK, Applied Mechanics and Materials, 2014, 574, 684–689.

[36] Xinghong. L, Tao. C, Liangsheng. Q, Zhenwu. L, Lizhu. Y, Energy operator demodulation method and its application in mechanical signal demodulation, Mechanical Engineering, 1998, 34, 85-90.

[37] Jie,.T, Yilun. L, Zhuo. F, Dalian. Y, Fang. T,Fault damage degrees diagnosis for rolling bearing based on Teager energy operator and deep belief network, Central South University (Science and Technology), 2017, 48, 61-68.

[38] R. K. Singleton. E. G. Strangas. S. Aviyente. Extended Kalman filtering for remaining-useful-life estimation of bearings, IEEE Transactions on Industrial Electronics, 2015, 62, 1781-1790.

[39] C. Anger. R. Schrader. U. Klingauf. Unscented Kalman Filter with Gaussian Process Degradation Model for Bearing Fault Prognosis, PHM Society European Conference, 2012, 1, 1-12.

[40] L. Cui. X. Wang. Y. Xu. H. Jiang. J. Zhou. A novel switching unscented Kalman filter method for remaining useful life prediction of rolling bearing, Measurement, 2018, 135, 678-684.

[41] C. K. R. Lim. D. Mba. Switching Kalman filter for failure prognostic, Mechanical Systems and Signal Processing, 2014, 52, 426-435.

[42] R. Wang. Q. Sun. P. Tu. J. Xiao. Y. Gui. P. Wang. Reduced-Order aggregate model for large-scale converters with inhomogeneous initial conditions in DC microgrids, IEEE Transactions on Energy Conversion, 2021.

[43] X. Hu. H. Zhang. D. Ma. R. Wang. A tnGAN-Based leak detection method for Pipeline network considering incomplete sensor data, Access IEEE, 2021, 9, 21735-21745.

[44] V. Manfredi. S. Mahadevan. J. Kurose. Switching kalman filters for prediction and tracking in an adaptive meteorological sensing network, IEEE SECON. 2005, 197-206.

[45] K. Li. Y. Zhang. Z. Li. Application research of Kalman filter and svm applied to condition monitoring and fault diagnosis, Applied Mechanics and Materials, 2011, 121, 268-272.

[46] D. Zhao. W. Cheng. R. X. Gao. R. Yan. P. Wang. Generalized Vold–Kalman Filtering for Nonstationary Compound Faults Feature Extraction of Bearing and Gear, IEEE Transactions on Instrumentation and Measurement, 2020, 69, 401-410.

[47] X. Wang. V. Makis. Autoregressive model-based gear shaft fault diagnosis using

the Kolmogorov-Smirnov test, Journal of sound and vibration, 2009, 327, 413-423.

[48] Y. Zhan. V. Makis. K. Jardine. Adaptive state detection of gearboxes under varying load conditions based on parametric modelling, mechanical systems and signal processing, 2006, 20, 188-221.

[49] L. Jiang. X. Fu, J. Cui, Z. Li. Fault detection of rolling element bearing based on principal component analysis, Chinese Control and Decision Conference (CCDC) IEEE. 2012, 24, 2944-2948.

[50] D. Appana. A. Prosvirin. J. Kim. Reliable fault diagnosis of bearings with varying rotational speeds using envelope spectrum and convolution neural networks, Soft Computing, 2018, 22, 6719-6729.

[51] M. Hamadache. D. Lee, K. Veluvolu. Rotor speed-based bearing fault diagnosis (RSB-BFD) under variable speed and constant load, IEEE transactions on industrial electronics, 2015, 62, 6486-6495.

[52] H. Sun. Z. He. Y. Zi. J. Yuan, X. wang. J. Chen and S. He, Multiwavelet transform and its applications in mechanical fault diagnosis-a review, Mech Syst Signal Process 2014, 43, 1–24.

[53] Xiaoxi. D, Quanchang. L, Lun. L, Qingbo. H, Yimin. S, Fast time-frequency manifold learning and its reconstruction fortransient feature extraction in rotating machinery fault diagnosis, Measurement, V2019, 141, 380-395.

[54] Zhixing. L, Xiangdong. L, Xiaoruan. W, Tian. H, Yingchun. S, A multi-parameter constrained potential underdamped stochastic resonance method and its application for weakfault diagnosis, Sound and vibration, 2019, 459, 114862

[55] Tongle. X, Zhaojie. Y, Daoyong. C, Diankun. Z. Fault diagnosis for rotating machinery based on local mean decomposition morphology filtering and least square support vector machine, Intelligent & Fuzzy Systems, 2017, 32, 2061-2070.

[56] Xianbo. W, Zhixin. Y., Xiaoan. Y, Novel particle swarm optimization-based variational mode decomposition method for the fault diagnosis of complex rotating machinery, IEEE/ASME Transactions on Mechatronics, 2017, 23, 68-79.

[57] Zong. M, Yuyang. Z, Jing. L, Zuozhou. P, An enhancement denoising autoencoder for rolling bearing faultdiagnosis, Measurement, 2018,130, 448-454.

[58] Y. Lei and Z. He, Advances in applications of hybrid intelligent fault diagnosis and prognosis technique, Journal of Vibration and Shock 2011, 30, 129–135.

[59] Y. Bengio. A. Courville. P. Vincent. Representation learning: a review and new perspectives, IEEE Trans. Pattern Anal. Mach. Intell. 2013, 35, 1798–1828.

[60] H. Shao. H. Jiang. X. Zhang. M. Niu. Rolling bearing fault diagnosis using an optimization deep belief network, Meas. Sci. Technol. 2015, 26, 115002.

[61] S. Hong. Z. Zhou. E. Zio. W. Wang. An adaptive method for health trend prediction

of rotating bearings, Digit. Sign. Process. 2014, 35, 117–123.

[62] Y. Shatnawi, M. Al-khassaweneh, Fault diagnosis in internal combustion engines using extension neural network, IEEE Trans. Ind. Electron. 2014, 6, 11434–11443.

[63] B. Samanta. K.R. Al-Balushi. S.A. Al-Araimi. Artificial neural networks and genetic algorithm for bearing fault detection, Soft Comput. 2006, 10, 264–271.

[64] H. Yaghobi. H.R. Mashhadi. K. Ansari. Artificial neural network approach for locating internal faults in salient-pole synchronous generator, Expert Syst. Appl. 2011, 38, 13328–13341.

[65] Z. Chen, W. Li, Multisensor feature fusion for bearing fault diagnosis using sparse auto encoder and deep belief network, IEEE Trans. Instrum. Meas. 2017, 66, 1693–1702.

[66] Liu ZW, He ZJ, Guo W, et al. A hybrid fault diagnosis method based on second generation wavelet de-noising and local mean decomposition for rotating machinery. ISA Trans, 2016, 61, 211–220.

[67] Y. Lei, F. Jia, J. Lin, S. Xing, S.X. Ding, An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data, IEEE Trans.Ind. Electron. 2016, 63, 3137–3147.

[68] Wang L, Liu ZW, Miao Q, et al. Time–frequency analysis based on ensemble local mean decomposition and fast kurtogram for rotating machinery fault diagnosis. Mech Syst Signal Process 2018, 103, 60–75.

[69] Jiao JF, Zhao N, Wang G, et al. A nonlinear quality-related fault detection approach based on modified kernel partial least squares. ISA Trans, 2017, 66, 275–283.

[70] L. Wen, X.Y. Li, L. Gao, Y.Y. Zhang, A new convolutional neural network-based data-driven fault diagnosis method, IEEE Trans. Ind. Electron. 2018, 65, 5990–5998.

[71] Sun RB, Yang ZB, Chen XF, et al. Gear fault diagnosis based on the structured sparsity time-frequency analysis. MechSyst Signal Process, 2018, 102, 346–363.

[72] Ciabattoni, L. Ferracuti, F. Freddi, A. Monteriù, A. Statistical spectral analysis for fault diagnosis of rotating machines. IEEE Trans. Ind. Electron. 2018, 65, 4301–4310.

[73] Yang ZB, Radzienski M, Kudela P, et al. Fourier spectral-based modal curvature analysis and its application to damagedetection in beams. Mech Syst Signal Process, 2017, 84, 763–781.

[74] S.Z. Su, Z.H. Liu, S.P. Xu, S.Z. Li, Sparse auto-encoder based feature learning for

human body detection in depth image, Signal Process. 2015, 112, 43–52.

[75] Li X, Yu DJ and Zhang DC. Fault diagnosis of rolling bearings based on the resonance-based sparse signal decomposition with optimal Q-factor. J Vib Eng 2015, 28, 998–1005.

[76] X. Jia, M. Zhao, Y. Di, P. Li, J. Lee, Sparse filtering with the generalized lp/lq norm and its applications to the condition monitoring of rotating machinery, Mech. Syst. Signal Process. 2017, 102, 198–213

[77] Cao, H.; Fan, F.; Zhou, K.; He, Z. Wheel-bearing fault diagnosis of trains using empirical wavelet transform. Measurement, 2016, 82, 439–449.

[78] Sun, W.; Shao, S.; Yan, R. Induction motor fault diagnosis based on deep neural network of sparse auto-encoder. J. Mech. Eng. 2016, 52, 65–71.

[79] Song LY, Wang HQ and Chen P. Vibration-based intelligent fault diagnosis for roller bearings in low-speed rotating machinery. IEEE Trans Instrum Meas. 2018, 67, 1887–1899.

[80] F. Jia, Y.G. Lei, L. Guo, J. Lin, S. Xing, A neural network constructed by deep learning technique and its application to intelligent fault diagnosis of machines, Neuro computing. 2018, 272, 619–628.

[81] Dalvand, F. Dalvand, S. Sharafi, F. Pecht, M. Current noise cancellation for bearing fault diagnosis using time shifting. IEEE Trans. Ind. Electron. 2017, 64, 8138–8147.

[82] Qin, F.W. Bai, J. Yuan, W.Q. Research on intelligent fault diagnosis of mechanical equipment based on sparse deep neural networks. J. Vibroeng. 2017, 19, 2439–2455.

[83] Liang, X. Zuo, M.J. Liu, L. A windowing and mapping strategy for gear tooth fault detection of a planetary gearbox. Mech. Syst. Signal Process. 2016, 80, 445–459.

[84] Frosini, L.; Harlisca, C.; Szabo, L. Induction machine bearing fault detection by means of statistical processing of the stray flux measurement. IEEE Trans. Ind. Electron. 2015, 62, 1846–1854.

[85] Feng, Z.; Chen, X.; Wang, T. Time-varying demodulation analysis for rolling bearing fault diagnosis under variable speed conditions. J. Sound Vib. 2017, 400, 71–85.

[86] Donoho DL, Tsaig Y and Drori I. Sparse solution of underdetermined systems of linear equations by stage wise orthogonal matching pursuit. IEEE Trans Inform Theory, 2012, 58, 1094–1121.

[87] Xue, H. Wang, H. Chen, P. Li, K. Song, L. Automatic diagnosis method for

structural fault ofrotating machinery based on distinctive frequency components and support vector machines under varied operating conditions. Neuro computing 2013, 116, 326–335.

[88] Yi, C. Lv, Y. Ge, M. Xiao, H. Yu, X. Tensor Singular Spectrum Decomposition Algorithm Based on Permutation Entropy for Rolling Bearing Fault Diagnosis. Entropy. 2017, 19, 139.

[89] Wang, X. Zheng, Y. Zhao, Z. Wang, J. Bearing fault diagnosis based on statistical locally linear embedding. Sensors 2015, 15, 16225–16247.

[90] Lv, Y. Yuan, R. Song, G. Multivariate empirical mode decomposition and its application to fault diagnosis of rolling bearing. Mech. Syst. Signal Process. 2016, 81, 219–234.

[91] Zhao, M. Lin, J. Xu, X. Lei, Y. Tacholess envelope order analysis and its application to fault detection of rolling element bearings with varying speeds. Sensors 2013, 13, 10856–10875.

[92] Roopa S and Narasimhan SV. S-transform based on analytic discrete cosine transform for time–frequency analysis. Signal Process 2014, 105, 207–215.

[93] Zvokelj, M. Zupan, S. Prebil, I. EEMD-based multiscale ICA method for slewing bearing fault detection and diagnosis. J. Sound Vib. 2016, 370, 394–423.

[94] Li, Z. He, Z.J. Zi, Y.Y. Chen, X.F. Bearing condition monitoring based on shock pulse method and improved redundant lifting scheme. Math. Comput. Simul. 2008, 79, 318–338.

[95] Lei, Y. Li, N. Lin, J. Wang, S. Fault diagnosis of rotating machinery based on an adaptive ensemble empirical mode decomposition. Sensors 2013, 13, 16950–16964.

[96] Imaouchen, Y. Kedadouche, M. Alkama, R. Thomas, M. A frequency-weighted energy operator and complementary ensemble empirical mode decomposition for bearing fault detection. Mech. Syst. Signal Process. 2017, 82, 103–116.

[97] Ng, S.S.Y. Tse, P.W. Tsui, K.L. A One-Versus-All Class Binarization Strategy for Bearing Diagnostics of Concurrent Defects. Sensors. 2014, 14, 1295–1321.

[98] Zhu, L. Chen, L. Zhao, D. Zhou, J. Zhang, W. Emotion recognition from Chinese speech for smart affective services using a combination of SVM and DBN. Sensors 2017, 17, 1694.

[99] Chen, J. Pan, J. Li, Z. Zi, Y. Generator bearing fault diagnosis for wind turbine via empirical wavelet transform using measured vibration signals. Renew. Energy.

2016, 89, 80–92.

[100] Li, C. Sánchez, R.-V. Zurita, G. Cerrada, M. Cabrera, D. Fault diagnosis for rotating machinery using vibration measurement deep statistical feature learning. Sensors. 2016, 16, 895.

[101] Yan BK and Zhou FX. Initial fault identification of bearing based on coherent cumulant stage wise orthogonal matching pursuit. J Mech Eng, 2015, 50, 88–96.

[102] Osman S and Wang W. A normalized Hilbert–Huang transform technique for bearing fault detection. J Vib Control 2016, 22, 2771–2787.

[103] Xie, J. Du, G. Shen, C. Chen, N. Chen, L.; Zhu, Z. An End-to-End Model Based on Improved Adaptive Deep Belief Network and Its Application to Bearing Fault Diagnosis. IEEE Access 2018, 6, 63584–63596.

[104] Zhang, N. Che, L.Z. Wu, X.J. Present Situation and Prospect of Data-driven Based Fault Diagnosis Technique. Comput. Sci. 2017, 44, 37–43.

[105] Elasha, F. Mba, D. Ruiz-Carcel, C. Effectiveness of adaptive filter algorithms and spectral kurtosis in bearing faults detection in a gearbox. Vib. Eng. Technol. Mach. 2015, 219–229.

[106] Li, K. Zhang, Q. Wang, K. Chen, P. Wang, H. Intelligent condition diagnosis method based on adaptive statistic test filter and diagnostic bayesian network. Sensors 2016, 16, 76.

[107] Wen, L. Li, X. Gao, L.; Zhang, Y. A New Convolutional Neural Network Based Data-Driven Fault Diagnosis Method. IEEE Trans. Ind. Electron. 2018, 65, 5990–5998.

[108] Fan, W. Zhou, Q. Li, J. Zhu, Z. A wavelet-based statistical approach for monitoring and diagnosis of compound faults with application to rolling bearings. IEEE Trans. Autom. Sci. Eng. 2018, 15, 1563–1572.

[109] Sohaib, M.; Kim, C.-H.; Kim, J.-M. A Hybrid Feature Model and Deep-Learning-Based Bearing Fault Diagnosis. Sensors 2017, 17, 2876.

[110] Ma JP and Jiang J. Analysis design of modified window shapes for S-transform to improve time–frequency localization. Mech Syst Signal Process, 2015, 58–59: 271–284.

[111] Lu, C. Wang, Z.Y. Qin, W.L. Ma, J. Fault diagnosis of rotary machinery components using a stacked denoising auto encoder-based health state identification. Signal Process. 2017, 130, 377–388.

[112] Hou FT, Chen J and Dong GM. Weak fault feature extraction of rolling

bearings based on globally optimized sparse coding and approximate SVD. Mech Syst Signal Process 2018, 111, 234–250.

[113] G.E. Hinton, S. Osindero and Y. Teh, A fast learning algorithm for deep belief nets, Neural Comput, 2006, 18, 1527–1554.

[114] X. Zhang and J. Wu, Denoising deep neural networks based voice activity detection, in: Proceedings of the International Conference on Acoustics Speech and Singnal Processing ICASSP, 2013, 853–857.

[115] X. Ding, R. Zhu and J. Li, JPEG2000 standard and its novel technologies, Journal of Computer Application, 2003, 23, 1–4.

[116] D. Taubman, High performance scalable image compression with EBCOT, IEEE Transactions on Image Processing, 2009, 9, 1158–1170.

[117] J. Liu, Y. Liu and X. Luo, Research and development on deep learning, Application Research of Computers, 2014, 31, 1921–1930.

[118] D. Erhan, Y. Bengio, A. Courville, et al., Why does unsupervised pre-training help deep learning, The Journal of Machine Learning Research, 2010, 11, 625–660.

[119] G.E. Hinton, N. Srivastava, A. Krizhevsky, et al. Improving neural network by preventing co-adaptation of feature detectors, Research Gate, 2012, 3, 212–223.

[120] C. Li, R. Sa´nchez, et al., Fault diagnosis for rotating machinery using vibration measurement deep statistical feature learning , Sensors, 2016, 16, 895.

[121] L.J.P. Maaten and G.E. Hinton, Visualizing high–dimensional data using t-sne, Journal of Machine Learning Research, 2008, 9, 2579–2605.

[122] Liu J. Shao Y. Overview of dynamic modelling and analysis of rolling element bearings with localized and distributed faults. Nonlinear Dynamics, 2018, 93, 1765-1798.

[123] Naiquan S. U, Jianbin X, Qinghua Z, et al. Research methods of the rotating machinery fault diagnosis. Machine Tool and Hydraulics, 2018, 46, 133-139.

[124] Ferguson D, Catterson V. Big data techniques for wind turbine condition monitoring. European Wind Energy Association Annual Event, 2014.

[125] Candè E. J, Wakin M. B. An introduction to compressive sampling. IEEE Signal Processing Magazine, 2008, 25, 21-30.

[126] Tang G, Hou W, Wang H, et al. Compressive sensing of roller bearing faults via harmonic detection from under-sampled vibration signals. Sensors, 2015, 15, 25648.

[127] Benzi R. Stochastic resonance: from climate to biology. Nonlinear Processes

in Geophysics, 2007, 17, 431-441.

[128] Gammaitoni L, Hanggi P, Jung P, et al. Stochastic resonance: a remarkable idea that changed our perception of noise. European Physical Journal B, 2009, 69, 1-3.

[129] Li J, Chen X, He Z. Multi-stable stochastic resonance and its application research on mechanical fault diagnosis. Journal of Sound and Vibration, 2013, 332, 5999-6015.

[130] Zhang X, Hu N, Hu L, et al. Enhanced detection of bearing faults based in signal cepstrum pre-whitening and stochastic resonance. Journal of Mechanical Engineering, 2012, 48, 83-89.

[131] Candes E. J, Romberg J. Errata for quantitative robust uncertainty principles and optimally sparse decompositions. Foundations of Computational Mathematics, 2007, 7, 529-531.

[132] Hugo L. Michael M. Razvan P. Yoshua B. Learning Algorithms for the Classification Restricted Boltzmann Machine. Journal of Machine Learning Research. 2012, 13, 643-669.

Acknowledgements

Time flies, I have spent almost three years in Japan. Looking back now, the excitement and novelty when I first arrived in Japan seem to be vivid in my mind. I would like to express my gratitude to all those who helped me during my PhD program.

My deepest gratitude goes first and foremost to Prof. Jinyama Hou, my supervisor professor during my PhD program, for his constant encouragement and guidance. Prof. Jinyama Hou has given me insightful suggestions and constant encouragement both in my study and in my life. Without his consistent and illuminating instruction, this thesis could not have reached its present form.

My sincere thanks also go to all the teachers and staffs in bio-resources Mie University, who gave me a lot of help and encouragement. Also, I own my thanks to all the friends in Japan and China. They helped me out when I was lost and helpless, it was their support that gave me the courage to persevere.

Furthermore, thanks to the school, the International Exchange Center, and the Japanese government for their financial support. Thanks to help me to finish my studies despite the severe epidemic.

Finally, my sincere thanks to my parents and family. Studying in Japan means more expenses and less reunion for families. Even so, they still support me without hesitation. Because of the epidemic, I haven't been home for three years, and I feel very guilty. I think that graduating well, working hard, and living hard every day is the best reward for them. Sincerely thank you for your support and encouragement.

Tangshao Duan

April 2022