

## Research Article

# Title Fault Diagnosis Method of Gearbox Multifeature Fusion Based on Quadratic Filter and QPSO-KELM

Shuo Meng <sup>1</sup>, Jianshe Kang <sup>1</sup>, Xupeng Die,<sup>2</sup> Xiaohan Wu,<sup>3</sup> Kuo Chi <sup>4</sup>, Zhipeng Dong,<sup>1</sup> and Weiyi Wu<sup>1</sup>

<sup>1</sup>Shijiazhuang Branch, Army Engineering University of PLA, Shijiazhuang 050000, China

<sup>2</sup>31690 Troops, Jilin 132000, China

<sup>3</sup>Dingzhou Yangjiazhuang Junior High School, Dingzhou 073000, China

<sup>4</sup>China Satellite Maritime Tracking and Control Department, Jiangyin 214431, China

Correspondence should be addressed to Shuo Meng; 864748906@qq.com

Received 15 July 2020; Accepted 4 August 2020; Published 26 September 2020

Academic Editor: Xiangyu Meng

Copyright © 2020 Shuo Meng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Effective filtering and noise reduction, feature extraction and fault diagnosis, and prognostics technology are important to Prognostics and Health Management (PHM) of equipment. Therefore, a multifeature fusion fault diagnosis method based on the combination of quadratic filtering and QPSO-KELM algorithm is proposed. In the quadratic filtering, stable filtering can reduce the impact of noise and fast-kurtogram can filtrate fault frequency bands with rich fault information. Then, the time-domain, frequency-domain, and time-frequency parameters of the secondary filter signal are extracted. MSSST was used to analyze the filtered signal, and the time-frequency image was obtained. The time-frequency parameter was extracted from the time-frequency image by 2DPCA, and all the extracted parameters are taken as the fusion fault feature of the gearbox. Finally, the fault feature parameters are taken as the training sample and testing sample of QPSO-KELM for training and testing to achieve the purpose of fault diagnosis. The experimental results show that the proposed method can effectively filter the noise, complete the fault mode identification of gearbox, and improve the fault diagnosis accuracy better than other methods.

## 1. Introduction

As a key component of mechanical transmission system, gearbox failure will lead to equipment shutdown, economic losses, and even casualties. The earlier the gearbox fault is found and repaired in time, the less the loss will be caused. Gearbox is mainly composed of rotating shaft, bearing, and gear. Its common fault types are gear fault and bearing fault. The fault will produce an impact signal, but the background noise interference and complex propagation path make it difficult to judge the fault state of gearbox timely and accurately. In order to filter out the noise, extract the effective state information, and better monitor and diagnose the state of gearbox, the related research is being carried out gradually.

As the signal is inevitably disturbed by noise in the process of acquisition, filtering technology as a signal

denoising method plays an important role in the follow-up research. At present, the commonly used methods of filtering and denoising can be divided into linear filtering and nonlinear filtering. Linear noise reduction algorithm is not suitable for processing complex vibration signals. Therefore, the nonlinear filtering noise reduction method is the mainstream noise reduction method in the field of mechanical equipment fault diagnosis. The commonly used technologies include noise reduction method based on empirical mode decomposition (EMD) [1–4], noise reduction method based on wavelet transform [5–7], and noise reduction method based on manifold learning [8–11]. For example, Boudraa and Cexus [1, 2] proposed a noise reduction method based on EMD threshold, Gu et al. [7] used wavelet analysis to denoise data, and Song et al. [11] and others used the manifold learning method to denoise. Although the above methods can achieve some effective results,

the mechanical transmission system is often disturbed by impact noise in the motion process, and the impact noise can be better filtered by a stable filter.

Feature extraction technology is the core of fault diagnosis. The commonly used feature parameters include time-domain feature [12, 13], frequency-domain feature [14], and time-frequency feature [15]. Time-domain feature is the feature information directly extracted from the time-domain signal, which can intuitively obtain the signal feature. The frequency-domain feature is extracted from the frequency spectrum obtained by Fourier transform. The time-frequency feature contains more information. The commonly used time-frequency analysis methods include short-time Fourier transform (STFT) [16, 17], continuous wavelet transform (CWT) [18, 19], and S-transform (ST) [20–22]. The S-transform inherits the advantages of STFT and CWT and develops on the basis of STFT and CWT. However, there are still some problems such as low resolution and energy leakage of the time-frequency images obtained by using ST technology.

Intelligent classification algorithm is widely used in the field of fault diagnosis, and its effect is good. Common intelligent algorithms include neural network [23–25], support vector machine (SVM) [26, 27], extreme learning machine (ELM) [28, 29], kernel extreme learning machine (KELM) [30], deep learning [31, 32], and other methods. Among them, the KELM algorithm has stronger comprehensive advantages in sample demand, network generalization, calculation speed, and accuracy and is more suitable for solving gearbox fault diagnosis problems with high calculation speed and accuracy requirements. However, the classification ability of KELM network is affected by the value of kernel parameters and penalty coefficient, so it is necessary to optimize the above structural parameters. Therefore, a method, using the quantum particle swarm optimization algorithm [33–35] to optimize the structural parameters of KELM (QPSO-KELM), is proposed in this paper. The algorithm based on QPSO can optimize the structural parameters of KELM with strong global search ability, so as to improve the learning speed and classification accuracy of KELM and improve the accuracy of gearbox fault diagnosis.

At present, most of the methods are not perfect in terms of diagnosis process and structure, and main problems are as follows:

- (1) The effect of signal filtering and noise reduction is not obvious enough, leading to the fault feature extraction which is not obvious enough.
- (2) The common time-frequency analysis methods also have the problems of energy leakage and low resolution, which cannot extract effective two-dimensional features.
- (3) The fault diagnosis accuracy of intelligent classification algorithm is not enough. To a large extent, the classification accuracy can be improved by improving the parameters.

Therefore, filtering noise reduction, feature extraction, and fault diagnosis technology are the focus of this paper [36].

In order to solve the above problems, denoising, feature extraction, and intelligent fault diagnosis technology are listed as the research priorities in this paper, and the following researches are carried out. The related steps involved are shown in Figure 1.

The working process is summarized as follows:

- (1) Quadratic filtering: collecting the signals of gearbox under different fault conditions, as the signals are susceptible to noise, the stable filter is first used for primary filtering, and then, the fast-kurtogram is used to filter frequency bands with rich fault information [37–39] to complete quadratic filtering.
- (2) Feature extraction: one-dimensional features and two-dimensional features of the filtered signal are extracted. One-dimensional features are characteristic parameters in the time domain and frequency domain. The extraction process of two-dimensional features is that MSSST is first used for time-frequency analysis of the filtered signal, and then, 2DPCA [40, 41] is used to extract the two-dimensional characteristic parameters of the time-frequency image.
- (3) Fault diagnosis: the extracted features are divided into training sample and test sample, which are used as the input of the QPSO-KELM model. The test sample is input into the trained QPSO-KELM, and then, the result of fault diagnosis can be obtained.

## 2. Basic Theory of the Proposed Method

**2.1. Stable Filter and Fast-Kurtogram.** Materials and Methods should be described with sufficient details to allow others to replicate and build on published results. Please note that publication of your manuscript implicates that you must make all materials, data, computer code, and protocols associated with the publication available to readers. Please disclose at the submission stage any restrictions on the availability of materials or information. New methods and protocols should be described in detail, while well-established methods can be briefly described and appropriately cited.

Research manuscripts reporting large datasets that are deposited in a publicly available database should specify where the data have been deposited and provide the relevant accession numbers. If the accession numbers have not yet been obtained at the time of submission, please state that they will be provided during review. They must be provided prior to publication.

Interventional studies involving animals or humans and other studies requiring ethical approval must list the authority that provided approval and the corresponding ethical approval code.

Conventional signal processing methods usually filter noise by linear algorithms. But when the noise is impact and the trailing is serious, the linear filter cannot deal with it effectively. Stable filter is better for this kind of impact signal.

The standard additive noise model of a signal is as follows:

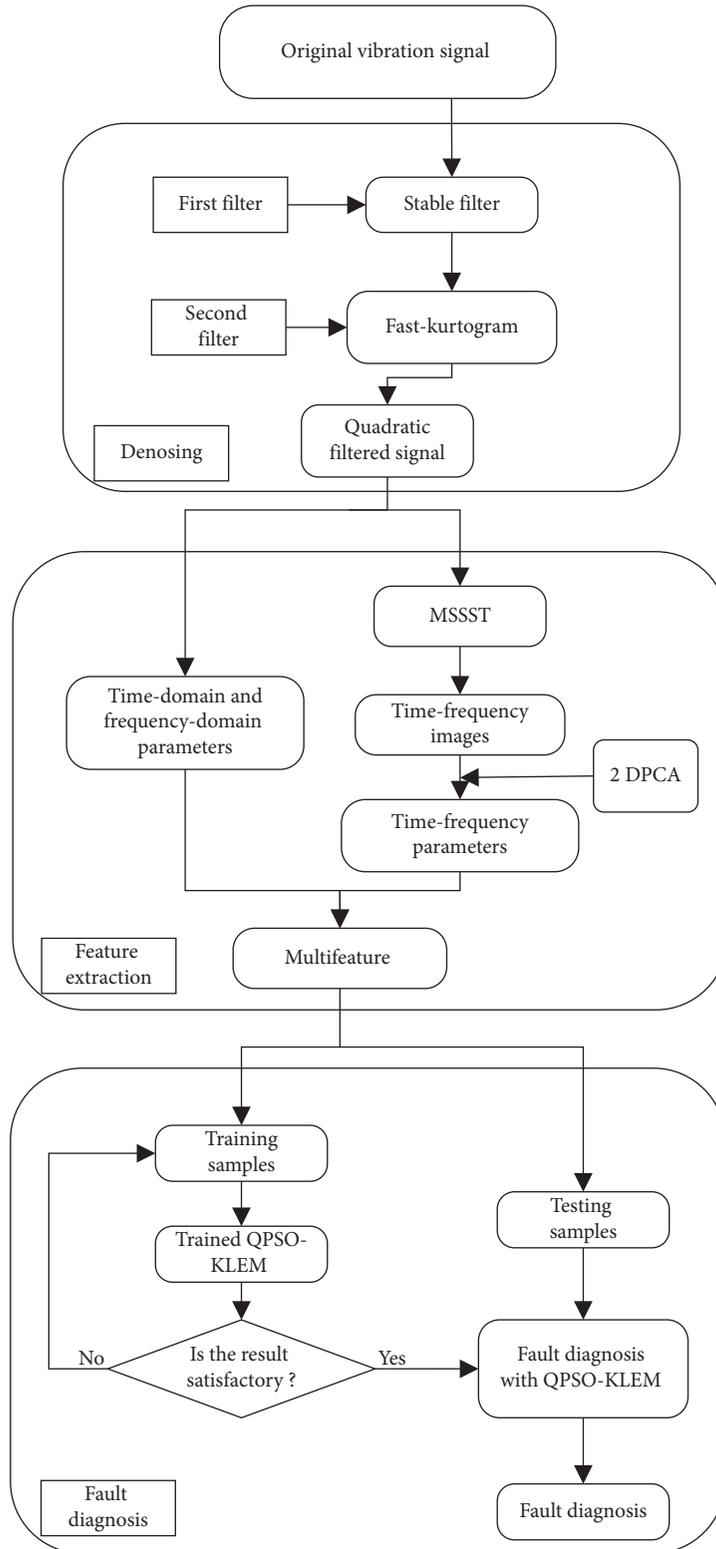


FIGURE 1: The flow of the proposed method. The structure consists of three parts: denoising, feature extraction, and fault diagnosis.

$$x_t = s_t + n_t, \quad t = 1, 2, 3, \dots, \quad (1)$$

where  $s_t$  is the signal of interest and  $n_t$  is the noise. Noise signals come from a wide range of sources, such as equipment operation, resonance, and environmental noise.

The distribution of such noise can be expressed by symmetrical stable distribution. Stable filter can filter out the noise to the maximum extent and recover the signal  $s_t$ .

For the noise which conforms to the stable distribution, the nonlinear filtering technology based on the same steady-

state distribution model has the best processing effect. These filters have good robustness and can also effectively deal with the situation of serious signal tailing distribution. Suppose  $\rho(x) = -\log f(x)$  is the negative value of logarithmic density, and the cost function is defined as follows:

$$C(\theta; x_1, \dots, x_m) = \sum_{i=1}^m \rho(x_i - \theta). \quad (2)$$

The stable filter is defined as the minimum  $\theta$  value of the cost function:

$$\begin{cases} \hat{s}_t = \hat{\theta}_{\text{STABLE}}(x_{t-k_1}, x_{t-k_1+1}, \dots, x_{t-k_2}), \\ \hat{\theta}_{\text{STABLE}}(x_1, \dots, x_m) = \arg \min_{\theta} C(\theta; x_1, \dots, x_m). \end{cases} \quad (3)$$

When  $\rho(x) = -\log f(x)$ , the minimum value of cost function is the maximum likelihood estimation of parameter  $\theta$ . When the distribution is the Gaussian distribution of  $\alpha = 2$ ,  $\rho(x) = -x^2/2$ , and the minimum value  $\hat{\theta}_{\text{LINEAR}}$  can be obviously got. When  $0 < \alpha < 2$ , filtering is nonlinear, and the minimum value of equation (3) must be found by numerical method.

A variety of filtering forms can be obtained by replacing the cost function equation (2). The simplest extension is to add a nonnegative weight  $\omega_i$ :

$$C_{\text{weighted}}(\theta; x_1, \dots, x_m) = \sum_{i=1}^m \rho(\omega_i(x_i - \theta)). \quad (4)$$

In the detection problem, the known signal  $s_1, \dots, s_n$  is extracted from the signal interfered by stable distributed noise, which can be expressed as follows:

$$x_t = \theta s_t + n_t, \quad t = 1, 2, 3, \dots \quad (5)$$

If there is no known signal in the signal,  $\theta = 0$ , if there is a known mode in the signal  $\theta \neq 0$ . The filtering process can be realized by constructing a real filter function:

$$C_{\text{matched}}(\theta; x_1, \dots, x_m) = \sum_{i=1}^m \rho(x_i - \theta s_i). \quad (6)$$

By the minimization equation (6),  $\hat{\theta}$  can be obtained, and  $\hat{\theta}$  is the estimated value of the extracted known signal.

After getting the stable-filtered signal, the fast-kurtogram is used to select the frequency band with rich fault information. The basic theory of fast-kurtogram is as follows.

Spectral kurtosis is a performance index reflecting the instantaneous impact strength. Its basic theory is to calculate the kurtosis value of each spectral line, so as to find the frequency band of impact [42]. The calculation equation is as follows:

$$K_k^i = \frac{\langle |c_k^i(n)|^4 \rangle}{\langle |c_k^i(n)|^2 \rangle^2} - 2, \quad (i = 0, \dots, 2^k - 1, k = 0, \dots, K - 1), \quad (7)$$

where  $c_k^i(n)$  is the complex envelope of signal  $x(n)$  at the corresponding central frequency  $f_i = (i + 2^{-1})2^{-k-1}$  and bandwidth  $(\Delta f)_k = 2^{-k-1}$ .

The spectral kurtosis is proposed by Antoni et al. But the calculation of kurtosis is cumbersome and not conducive to engineering practice. Therefore, Antoni further proposed the concept of fast spectral kurtosis. The fast spectral kurtosis is an inverted pyramid structure. The abscissa represents the frequency, the ordinate represents the number of decomposition layers, and the color depth of the image represents the spectral kurtosis under different central frequencies and bandwidths. According to the color depth of fast spectral kurtosis, the optimal bandwidth and center frequency can be obtained. Then, the fault frequency band is extracted by band-pass filter, and then, the square envelope order spectrum of fault signal is obtained by square envelope analysis. The inverted pyramid structure of the fast-kurtogram is shown in Figure 2.

2.2. MSSST-2DPCA. For signal  $x(t)$ , the S-transform of  $x(t)$  can be defined as follows:

$$S(t, f) = \int_{-\infty}^{+\infty} x(\tau)g(\tau - t, f)e^{-i2\pi f\tau} d\tau, \quad (8)$$

$$g(t - \tau, f) = \frac{|f|}{\sqrt{2\pi}} e^{-(f^2(t-\tau)^2/2)}, \quad (9)$$

where the result of  $S(t, f)$  represents the time spectrum of the signal;  $g(t - \tau, f)$  is the Gaussian window function;  $t$  and  $f$  represent time and frequency, respectively; and  $\tau$  is a time parameter that represents the position of the Gaussian window on the timeline.

According to equation (9), synchronous squeezing S-transformation (SSST) is performed on signal  $x(t)$ , and the calculation steps are as follows.

The coefficients of S-transform are  $S_x(t, f)$ , which can be obtained by equation (8).

Calculate the time partial derivative of the S-transform, and the equation is as follows:

$$\begin{aligned} \frac{\partial S(t, f)}{\partial t} &= i\omega \cdot \int_{-\infty}^{+\infty} x(\tau)g(\tau - t, f)e^{-i2\pi f\tau} d\tau \\ &= i\omega S(t, f). \end{aligned} \quad (10)$$

The instantaneous frequency  $\omega_x$  can be obtained by transforming equation (10), and the calculated result is

$$\omega_x(t, \omega) = \frac{\partial_t S(t, f)}{iS(t, f)}. \quad (11)$$

By compressing the value of the interval  $[\omega_x - (1/2)\Delta\omega, \omega_x + (1/2)\Delta\omega]$  to  $\omega_x$ , we can get the synchronous compression transformation value SSST. The formula is as follows:

$$\text{SSST}(t, f) = \int_{-\infty}^{+\infty} S(t, \xi)\delta(\pi f - \omega_x(t, \xi))d\xi. \quad (12)$$

The time-frequency analysis method of MSSST can be obtained by multiple iterations of the SSST, and the iterative equation of MSSST is as follows:

$$\begin{aligned}
\text{SSST}^{[1]} &= \text{SSST}, \\
\text{SSST}^{[2]}(t, f) &= \int_{-\infty}^{+\infty} \text{SSST}^{[1]}(t, \xi) \delta(2\pi f - \omega_x(t, \xi)) d\xi \\
&= \int_{-\infty}^{+\infty} S(t, \xi) \delta(2\pi f - \delta(2\pi f - \omega_x(t, \xi))) d\xi, \\
\text{SSST}^{[3]}(t, f) &= \int_{-\infty}^{+\infty} \text{SSST}^{[2]}(t, \xi) \delta(2\pi f - \omega_x(t, \xi)) d\xi \\
&= \int_{-\infty}^{+\infty} S(t, \xi) \delta(2\pi f - \delta(2\pi f - \delta(2\pi f \\
&\quad - \omega_x(t, \xi)))) d\xi, \\
&\vdots \\
\text{SSST}^{[N]}(t, f) &= \int_{-\infty}^{+\infty} \text{SSST}^{[N-1]}(t, \xi) \delta(2\pi f - \omega_x(t, \xi)) d\xi \\
&= \int_{-\infty}^{+\infty} S(t, \xi) (\delta(\dots 2\pi f - \delta(2\pi f \\
&\quad - \delta(2\pi f - \omega_x(t, \xi)) \dots)) d\xi,
\end{aligned} \tag{13}$$

where  $\text{SSST}^{[1]}$  is the SSST and  $N$  is the number of iterations ( $N \geq 2$ ). Substituting  $\text{SSST}^{[k-1]}$  into  $\text{SSST}^{[k]}$  ( $k = 2, 3, \dots, N$ ), after  $N - 1$  calculations, the final result can be simplified to

$$\begin{aligned}
\text{MSSST} &= \text{SSST}^{[N]}(t, f) \\
&= \int_{-\infty}^{+\infty} S(t, \xi) \delta(2\pi f - \omega_x^{[N]}(t, \xi)) d\xi.
\end{aligned} \tag{14}$$

For MSSST, it is only required to calculate  $\omega_x^{[N]}(t, \xi)$  and substitute it into the first-order SSST, and the calculated result is the same as equation (13). Therefore, the computational complexity of MSSST is similar to that of the first-order SSST, and it is beneficial to improve the calculation speed. The proposed MSSST is effective to solve energy leakage and low resolution of the spectrum.

After getting the time-frequency images, 2DPCA is used to extract the feature matrix. The theory of two-dimensional principal component analysis (2DPCA) is as follows.

Suppose  $X$  is the  $n$ -dimensional column vector,  $A$  is the image matrix of  $m \times n$ , and  $Y$  is the  $m$ -dimensional projection vector of  $A$  on  $X$  after linear transformation. The process can be expressed as follows:

$$Y = AX. \tag{15}$$

The trace of covariance matrix  $Y$  of  $S_x$  is defined as the total divergence:

$$J(X) = \text{tr}(S_x). \tag{16}$$

In maximum total divergence criterion, when the divergence of the projected vector  $Y$  is the largest, the direction of vector  $X$  is the optimal projection direction, where the equation of  $S_x$  is

$$\begin{aligned}
S_x &= E(Y - EY)(Y - EY)^T \\
&= E[AX - E(AX)][AX - E(AX)]^T \\
&= E[(A - EA)X][(A - EA)X]^T.
\end{aligned} \tag{17}$$

The total divergence of  $S_x$  is

$$\text{tr}(S_x) = X^T [E(A - EA)^T (A - EA)] X. \tag{18}$$

The covariance matrix expression of the image is as follows:

$$G_t = E(A - EA)^T (A - EA). \tag{19}$$

According to equation (19),  $G_t$  is a nonnegative definite  $n \times n$ -dimensional matrix. Suppose there are  $M$  training images, the  $j$ -th image is represented as  $A_j$ , and the mean value of all training images is recorded as  $\bar{A}$ , then the divergence matrix expression is as follows:

$$G_t = \frac{1}{M} \sum_{j=1}^M (A_j - \bar{A})^T (A_j - \bar{A}). \tag{20}$$

The normalized expression is as follows:

$$J(X) = X^T G_t X. \tag{21}$$

The  $X$  obtained from the criterion formula of the maximum divergence is called the optimal projection axis. The optimal projection axis  $X$  is the eigenvector corresponding to the maximum eigenvalue of  $G_t$ . In general, one optimal projection axis is not enough. Therefore, the first  $d$  orthogonal unit eigenvectors with the largest eigenvalues are selected as the optimal projection axis, which can be expressed by the following equation:

$$\begin{cases} \{X_1, \dots, X_d\} = \arg \max J(X), \\ X_i^T X_j = 0, \quad i \neq j, j = 1, \dots, d. \end{cases} \tag{22}$$

The optimal projection vector  $X_1, \dots, X_d$  of 2DPCA is used to extract the two-dimensional features of the image. Suppose sample image  $A$ , according to the above equations,  $Y_1, \dots, Y_d$  can be obtained, which is called the principal component vector of sample image  $A$ .

The matrix  $B = [Y_1, \dots, Y_d]$  of  $m \times d$  composed of principal component vectors is the feature matrix (feature image) of sample image  $A$ .

**2.3. QPSO-KELM.** KELM algorithm is adding kernel function into ELM, and the input weights and offsets in ELM are replaced by kernel function mapping, which makes the output of the ELM more stable and solves the over learning problems.

ELM is a single hidden layer feedforward neural network, and its hidden layer weight does not need to be adjusted by feedback regulation. The structure of the ELM is shown in Figure 3.

Suppose  $Z = \{x_i, t_i\}_{i=1}^N$  are  $N$  training sample sets, where  $x_i = [x_{i1}, x_{i2}, \dots, x_{im}]^T \in R^m$  is the input characteristic

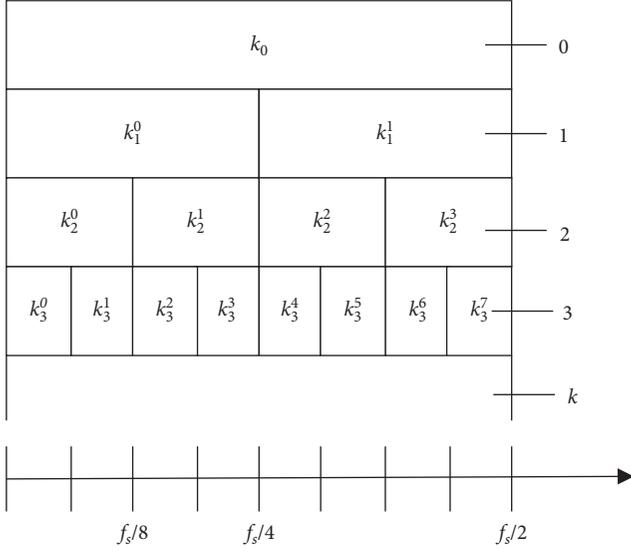


FIGURE 2: Fast-kurtogram. The grid in the diagram represents the frequency band.

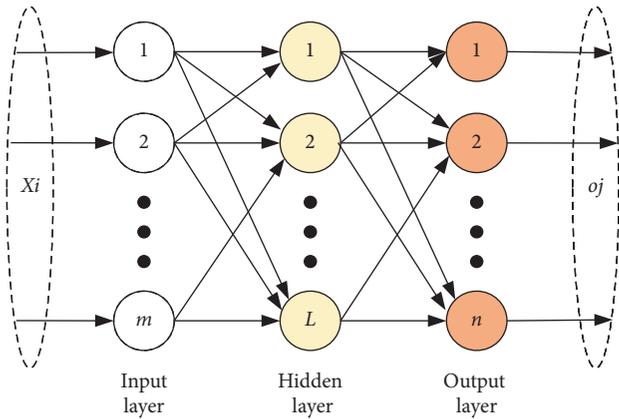


FIGURE 3: ELM model. ELM is made up of three parts: input layer, hidden layer, and output layer.

parameter and  $t_i = [t_{i1}, t_{i2}, \dots, t_{in}]^T \in R^n$  is the sample label, then the ELM training model can be expressed as

$$\sum_{i=1}^M \beta_i g(w_i \cdot x_i + b_i) = o_i, \quad i = 1, 2, \dots, N, \quad (23)$$

where  $w_i = [w_{1i}, w_{2i}, \dots, w_{mi}]$  and  $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{in}]^T$  represent the input and output weight matrix of the  $i$ th hidden layer node, respectively;  $o_j = [o_{j1}, o_{j2}, \dots, o_{jn}]^T$  represents the output layer vector of the ELM.

The goal of ELM training is to make the error between the actual output of the training sample and the sample label close to 0, and the equation is

$$\sum_{i=1}^N \|o_i - t_i\| = 0 \beta_i w_i b_i \sum_{i=1}^M \beta_i g(w_i \cdot x_i + b_i) = t_i, \quad (24)$$

$$i = 1, 2, \dots, NH \beta = T.$$

So  $\beta_i$ ,  $w_i$ , and  $b_i$  can be used to obtain equation (25), as follows:

$$\sum_{i=1}^M \beta_i g(w_i \cdot x_i + b_i) = t_i, \quad i = 1, 2, \dots, N. \quad (25)$$

Equation (23) can also be expressed as a matrix:

$$H\beta = T, \quad (26)$$

where  $H$  is the output matrix of the hidden layer, and the equation of  $\beta$  is as follows:

$$\beta' = H^+ T, \quad (27)$$

where  $H^+$  represents the inverse matrix of  $H$ , and  $\beta'$  is the optimal quadratic solution to  $\beta$ .

According to the structural optimization and ERM criteria, the output weight  $\beta$  of ELM can be determined, that is,

$$\text{minimize } L_{P_{ELM}} = \frac{1}{2} \|\beta\|^2 + C \frac{1}{2} \sum_{i=1}^N \|\xi_i\|^2, \quad (28)$$

$$\text{subject to } h(x_i)\beta = t_i^T - \xi_i^T, \quad i = 1, \dots, N,$$

where  $\xi_i = [\xi_{i1}, \dots, \xi_{im}]^T$  represents the network calculation error,  $(1/2)\|\beta\|^2$  is the structural error,  $(1/2)\sum_{p=1}^N \|\xi_p\|^2$  is the empirical error, and  $C$  is the penalty factor.

Through the optimization of the solution, the optimal solution of  $\beta$  can be obtained as follows:

$$\beta = H^T \left( \frac{I}{C} + HH^T \right)^{-1} T. \quad (29)$$

Therefore, the output model function of ELM is as follows:

$$o_i = h_i H^T \left( \frac{I}{C} + HH^T \right)^{-1} T. \quad (30)$$

By using kernel operation to replace matrix operation  $H$  in ELM, the KELM classification model can be established. The kernel matrix is defined as follows:

$$\Omega = HH^T: \Omega_{i,j} = h(x_i) \cdot h(x_j) \\ = K(x_i, x_j), \quad (31)$$

where  $\Omega$  is the symmetric matrix of  $N \times N$  and  $K(x, y)$  is the kernel function.

Equation (30) can be expressed by kernel function as follows:

$$o_i = [K(x_i, x_j), \dots, K(x_i, x_j)] \lambda, \quad (32)$$

where  $\lambda$  represents the output weight matrix of KELM.

Different support vector machines can be constructed by selecting different kernel functions. Among the commonly used kernel functions, compared with the polynomial kernel function and *sigmoid* kernel function, the radial basis kernel function has the advantages of simple parameters and strong adaptability to randomly distributed samples. Therefore, a

SVM based on radial basis kernel function is established in this paper, and its expression is shown as follows:

$$K(x, x_i) = \exp\left(\frac{-\|x - x_i\|^2}{\sigma^2}\right). \quad (33)$$

According to equations (28) and (33), it can be seen that KELM contains kernel parameter and penalty coefficient  $C$ , and it is necessary to find the optimal parameter to make the classification effect of KELM the best.

By quantizing the iterative updating process of particles of PSO algorithm, the QPSO can reduce the algorithm complexity and improve algorithm convergence speed and global search ability. The basic principles of QPSO are as follows.

Assume that  $\Omega$  is the  $d$ -dimensional search space, and the population number of particles in the space is  $b$ , then the position of the  $i$ th particle can be expressed as

$$V_i = (v_{i1}, v_{i2}, \dots, v_{id}). \quad (34)$$

Suppose the individual optimal position of the particle is  $p_{ibest}$ , and the global optimal position of the particle is  $p_{gbest}$ ; the  $p_{ibest}$  and  $p_{gbest}$  are as follows:

$$\begin{aligned} p_{ibest} &= (p_{i1}, p_{i2}, \dots, p_{id}), \\ p_{gbest} &= (p_{g1}, p_{g2}, \dots, p_{gd}). \end{aligned} \quad (35)$$

The particle can find and update its individual optimal position  $p_i$  and population optimal position  $p_g$  through iterative operation, and the average optimal position  $mbest$  can be introduced as the population optimal center. Then, the particle optimization process can be expressed as

$$\begin{aligned} mbest &= \frac{1}{M} \sum_{i=1}^M p_{ibest} \\ &= \left[ \frac{1}{M} \sum_{i=1}^M p_{i1}, \frac{1}{M} \sum_{i=1}^M p_{i2}, \dots, \frac{1}{M} \sum_{i=1}^M p_{id} \right], \end{aligned} \quad (36)$$

$$p_{id} = \varphi p_{id} + (1 - \varphi) p_{gd}, \quad \varphi \in (0, 1), \quad (37)$$

$$v_{id} = p_{id} \pm \alpha |mbest - v_{id}| \ln\left(\frac{1}{u}\right), \quad u \in (0, 1), \quad (38)$$

where  $\alpha$  is the contraction expansion factor, which is dynamically adjusted in the iterative operation according to the following equation:

$$\alpha = (\alpha_1 - \alpha_2) \times \frac{N - t}{N} + \alpha_2, \quad (39)$$

where  $N$  represents the maximum number of iterations.  $\alpha_1$  and  $\alpha_2$  are the initial and final values of  $\alpha$ , respectively; usually, the two parameters are set as  $\alpha_1 = 0.5$  and  $\alpha_2 = 1$ .

According to the basic principle of KELM and QPSO, the network training process which used QPSO to optimize the network structure parameters of KELM is as follows:

*Step 1.* Initialize the position of population particles, and set parameters such as particle swarm size, iteration step size, and termination conditions.

*Step 2.* Initialize the current position of each particle as  $p_{ibest}$ , define the classification accuracy of KELM as fitness function, and calculate the fitness value of each particle. The position of the particle with the maximum fitness value was initialized as  $p_{gbest}$ .

*Step 3.* Update the particle position according to equations (36)–(38).

*Step 4.* The fitness value of each particle is calculated, and the individual optimal position  $p_{best}$ , the group optimal position  $p_{gbest}$ , and the group optimal center  $mbest$  are updated based on the optimal fitness value.

*Step 5.* Judge whether the termination conditions are met. If so, stop the calculation and output the result; if not, return to step 3.

According to the above steps, the modeling flow chart of QPSO-KELM can be drawn, as shown in Figure 4.

### 3. Experimental Verification and Discussion

To verify whether the proposed method is available, three preset fault tests of the gearbox are conducted. The fault simulation test rig is shown in Figure 5. The test rig consists of four parts: the power and control part, the bearing fault simulation part (not used), the gear fault simulation part, and the data acquisition part (not shown). This section mainly uses the gearbox fault simulation part. The perspective and internal structure of the gearbox are shown in Figure 6. The teeth number in the gearbox from high-speed shaft to low-speed shaft is 41, 79, 36, and 90. The fault gear is gear 3 with 36 teeth. The fault types include wear, break, and miss teeth. The fault bearing is located in the deep groove ball bearing at the end cover of the gear 3 side of the medium-speed shaft (Figure 6(b)). The deep groove ball bearing type is ER-16k, and its fault types include inner race fault and outer race fault.

In the test, two vibration acceleration sensors are set in vertical and horizontal directions, respectively (Figure 6(c)). The data acquisition parameters are set as follows: sampling frequency  $fs = 20.48$  kHz, sampling time  $t = 48$  s, and motor speed 20 r/s. The main dimension parameters of the bearing ER-16K are shown in Table 1. According to the speed of the motor and the parameters of the gearbox and the bearing ER-16K, the relevant frequency can be calculated. The calculation results are shown in Table 2.

$f_{r1}$  is the rotating frequency of high-speed axis,  $f_{r2}$  is the rotating frequency of medium-speed axis, and  $f_{r3}$  is the rotating frequency of low-speed axis;  $f_{m1}$  is the meshing frequency of gear 1 and gear 2, and  $f_{m2}$  is the meshing frequency of gear 3 and gear 4;  $f_{BPFO}$  is the fault characteristic frequency of the bearing outer race,  $f_{BPFI}$  is the fault characteristic frequency of the bearing inner race, and  $f_{BSF}$  is the fault characteristic frequency of the bearing ball spin.

*3.1. Denoising with Quadratic Filter.* The quadratic filtering method can effectively filter out the interference noise in the signal and select the fault frequency band with rich fault information. Two cases of bearing inner race fault and gear

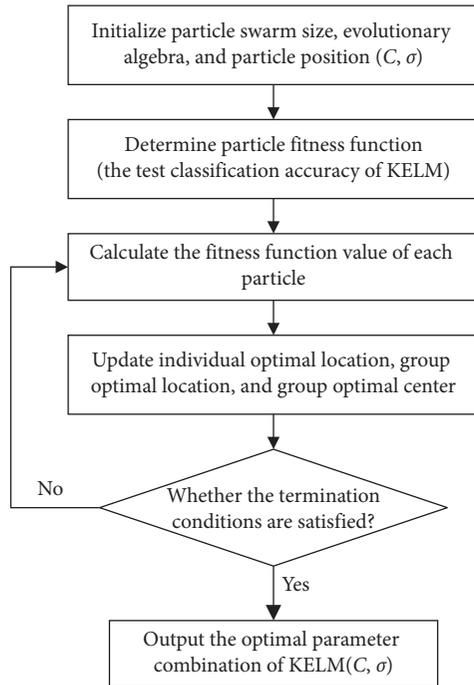


FIGURE 4: Flow chart of QPSO-KELM. QPSO is used to search global optimal parameters of KELM.

broken tooth fault are selected as fault cases in the quadratic filtering part.

Firstly, the stable filter is used to filter out the noise, and the detailed process is as follows.

The fault signal of bearing inner race and gear broken tooth are filtered with stable filter, and the filtering parameters are set as  $\alpha = 1.5$ ,  $\beta = 0$ , and  $\gamma = 1$ ; then, the filtered spectrum can be obtained, as shown in Figure 7.

Comparing with the spectrum before and after filtering, the signal-to-noise ratio (SNR) is improved with a stable filter. The SNR of bearing inner race fault before filtering is  $-9.3491$ , the SNR after filtering is increased to  $-8.2371$ , the SNR before filtering is  $-8.7745$ , and the SNR after filtering is  $-8.1016$ . Comparing with the spectrum before and after filtering, it can be seen that the trailing part of the signal is reduced, which indicates that the stable filter is effective to reduce noise.

Secondly, the fast-kurtogram technology is used to process the filtered signals, and the fault frequency band with rich fault information can be obtained. The fast-kurtogram diagram and square envelope results of bearing inner race fault and tooth break fault are shown in Figure 8.

Figure 9 is the square envelope spectrum of bearing inner ring fault and gear broken tooth after quadratic filtering. It can be seen from Figure 9 that the feature frequency of bearing inner race fault is obvious, and the bearing inner race fault can be judged according to the fault feature frequency. In the square envelope spectrum of tooth broken

fault, the meshing frequency, frequency doubling, and the surrounding side band can be observed, so as to judge the gear tooth broken fault.

The above results show that the fault information of gearbox can be reflected to a great extent through the signal analysis after quadratic filtering.

**3.2. Feature Parameters of Time Domain and Frequency Domain.** Time-domain feature parameters after quadratic filtering are as follows: mean square value ( $P1$ ), root mean square ( $P2$ ), variance ( $P3$ ), standard deviation ( $P4$ ), energy ( $P5$ ), root square amplitude ( $P6$ ), mean square amplitude ( $P7$ ), mean square amplitude ( $P8$ ), kurtosis ( $P9$ ), skewness ( $P10$ ), waveform index ( $P11$ ), peak index ( $P12$ ), pulse index ( $P13$ ), margin index ( $P14$ ), and clearance coefficient ( $P15$ ). These characteristic parameters are extracted from the original signal. Frequency-domain characteristic parameters include mean frequency (MF,  $P16$ ), frequency center (FC,  $P17$ ), root mean square frequency (RMSF,  $P18$ ), and standard deviation frequency (STDF,  $P19$ ). In order to facilitate calculation, the extracted feature parameters are normalized, and the normalized parameters are expressed by NP, as shown in Table 3.

**3.3. Feature Images Extraction with MSSST-2DPCA.** Time-frequency analysis is carried out on the filtered inner race fault signal of bearing and tooth break fault signal, and the time-frequency analysis results of STFT, CWT, ST, and MSSST are compared. The time-frequency analysis results are shown in Figures 10 and 11.

From the time-frequency analysis images, it can be seen that the order of time-frequency resolution and energy aggregation is  $MSSST > ST > CWT > STFT$ , both in bearing inner race fault or gear broken tooth fault.

After getting the time-frequency images, 2DPCA is used to extract the six cases of normal, tooth miss, tooth break, tooth wear, and bearing inner race fault and outer race fault. The parameter settings are as follows: the original image size is  $258 * 330$ , and the image size after 2DPCA feature extraction is  $72 * 99$ . By using the MSSST method, 960 images are generated for each case, and a total of 5760 images can be obtained. Then, the 2DPCA algorithm is used to extract the feature matrix of the image. Comparing STFT and MSSST, and the results are as follows.

The images in Figure 12 and 13 are the time-frequency features of the filtered fault signal. Comparing the results of 2DPCA-STFT and 2DPCA-MSSST, it can be seen that the feature extracted by 2DPCA-MSSST contains more abundant fault information.

**3.4. Feature Fusion.** In the previous paper, 19 feature parameters are extracted by calculation, and the feature matrix corresponding to the time-frequency image is obtained by using MSSST-2DPCA technology. The fusion features are

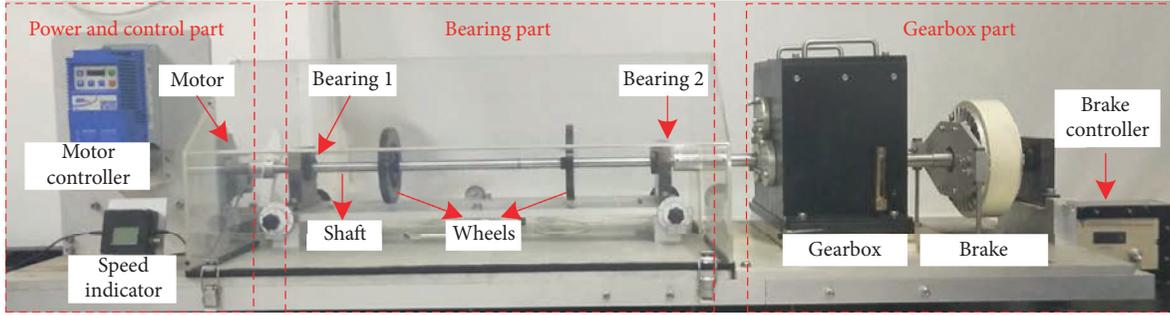


FIGURE 5: The mechanical fault simulation test bench.

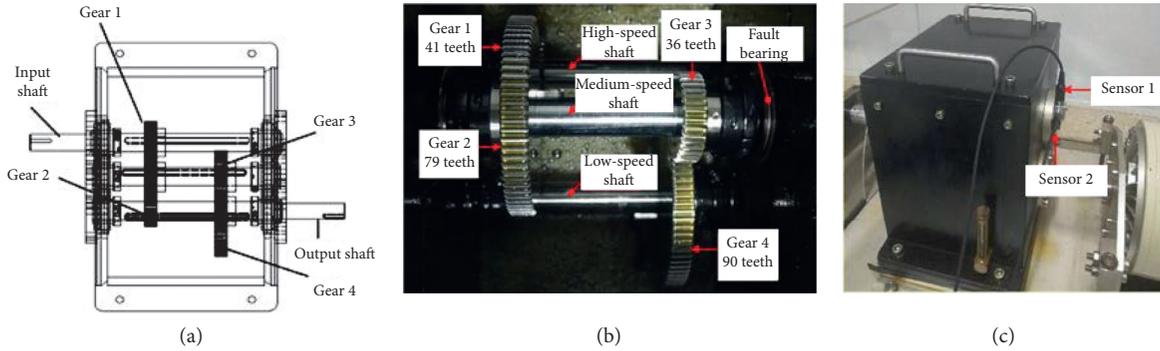


FIGURE 6: The structure of gearbox: (a) perspective view; (b) the inner structure; (c) the outer structure.

TABLE 1: Main dimensions of bearing ER-16K.

Number of rolling elements $N_b$	Ball diameter $d$ (inch)	Pitch diameter $D$ (inch)	Contact angle $\theta$ ( $^\circ$ )
9	0.3125	1.516	0

TABLE 2: Relevant frequencies of the gearbox.

Rotating frequency ( $r \cdot s^{-1}$ )			Meshing frequency (Hz)		Fault feature frequency of bearing (Hz)		
$f_{r1}$	$f_{r2}$	$f_{r3}$	$f_{m1}$	$f_{m2}$	$f_{BPFO}$	$f_{BPF1}$	$f_{BSF}$
20	10	4	820	371.67	36.588	54.965	23.787

obtained by fusing the feature parameters and the feature matrix. The specific steps are as follows:

- (1) Since the size of the feature image is  $72 * 99$ , in order to ensure the same length of the feature vector, 99 segments of the same length signal are intercepted from the collected signals, 19 feature parameters are extracted, respectively, and the feature parameter matrix of  $19 * 99$  can be obtained.
- (2) For the two sensors (Figure 6(c)), the matrix of feature parameters is extracted according to the method in step 1, and the eigenvector of the feature matrix is calculated.
- (3) Calculate the feature vectors of feature image.
- (4) By combining the feature vectors corresponding to the two feature parameters and the eigenvectors

corresponding to the feature images, fusion feature matrix can be obtained.

The corresponding image is shown in Figure 14.

**3.5. Fault Diagnosis Based on QPSO-KELM.** The time-domain, frequency-domain, and time-frequency feature parameters obtained above are summarized; then, 960 groups of features can be obtained, and the features are randomly divided into training samples and testing samples according to the ratio of 3:1; the training samples number is 720, and the testing samples number is 240. The 720 training samples are used to train the QPSO-KELM. After training, the 240 testing samples are input into the QPSO-KELM, and then, the test results are obtained, as shown in Figure 15. There is no error point in Figure 15, so the classification accuracy is 100%.

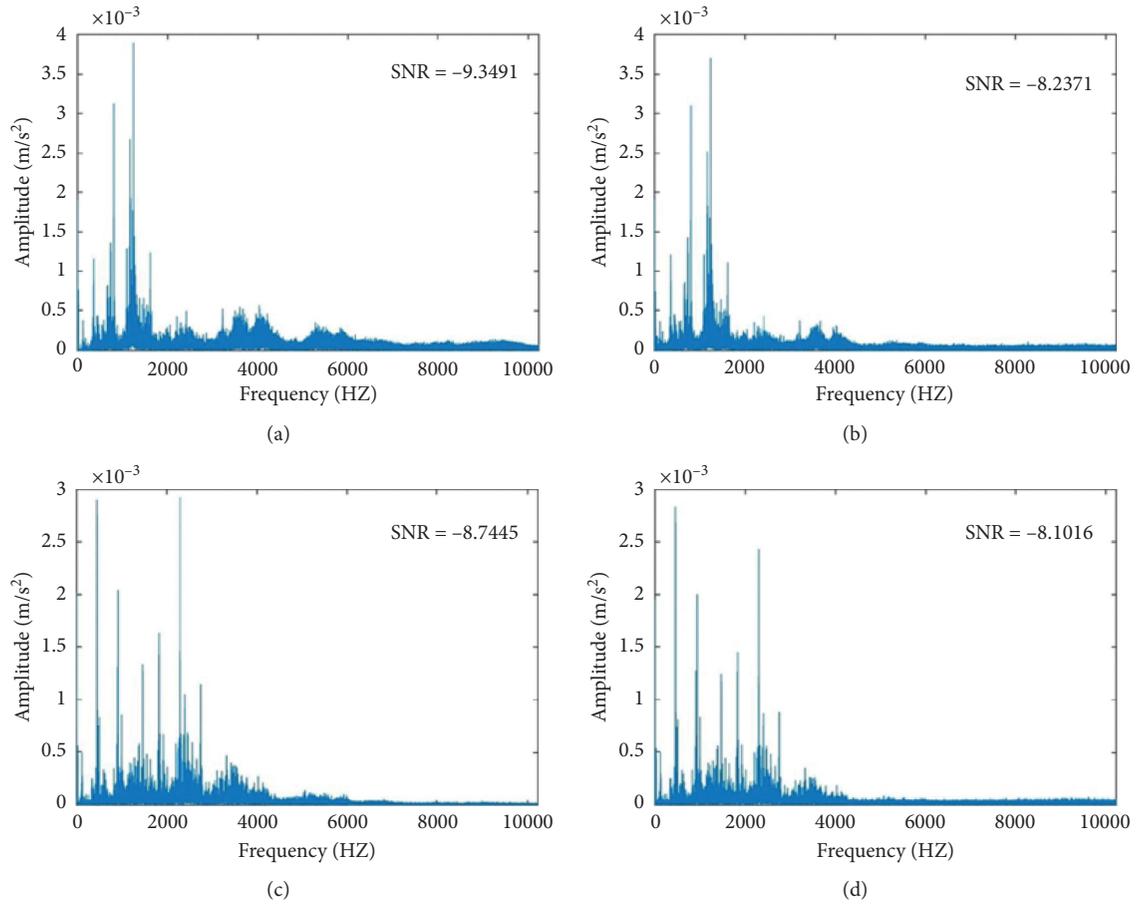


FIGURE 7: The spectrum of fault signal before and after filtering: (a) inner race fault-without filter; (b) inner race fault-stable filter; (c) tooth break-without filter; (d) tooth break-stable filter.

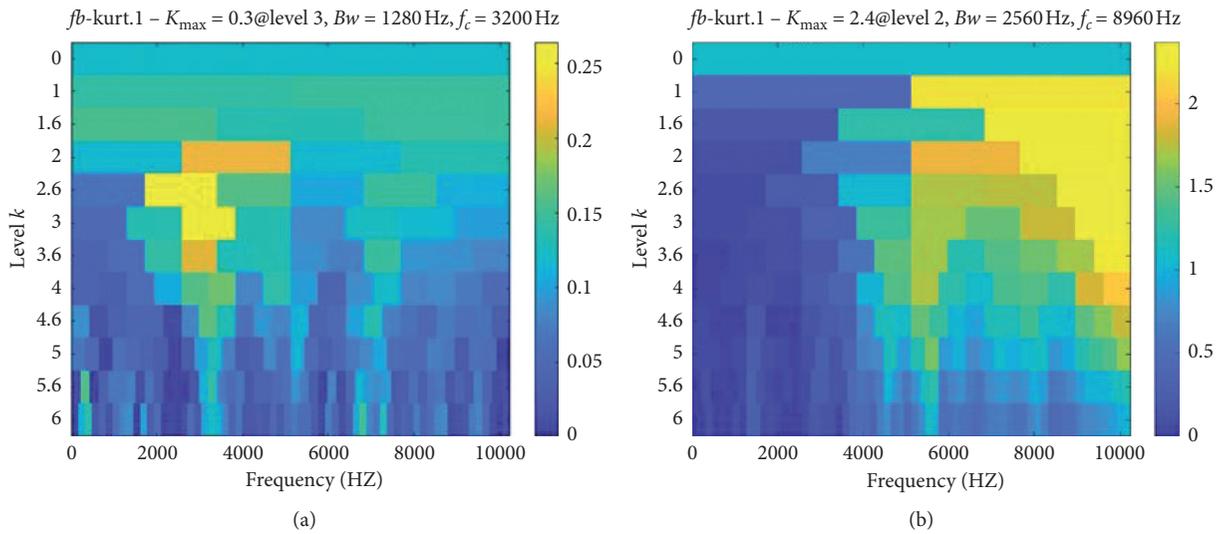


FIGURE 8: The fast-kurtogram of fault signal: (a) inner race fault; (b) tooth break fault.

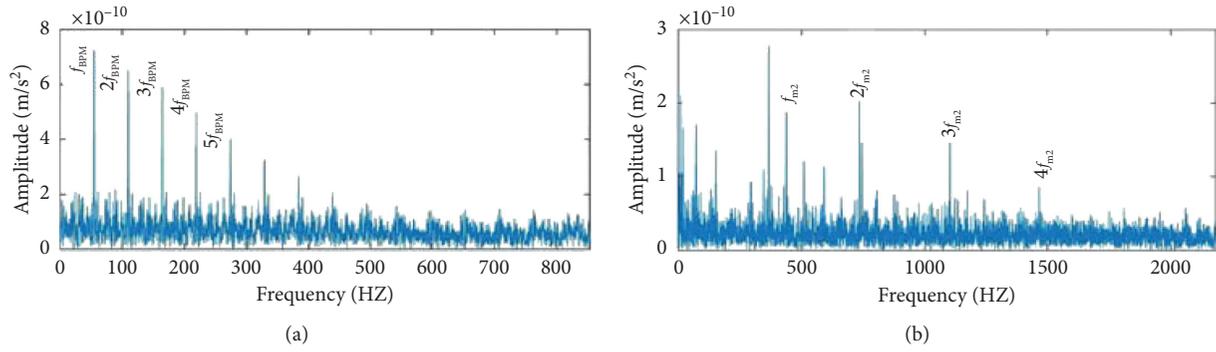


FIGURE 9: The square envelope spectrum of fault signal: (a) inner race fault; (b) tooth break fault.

TABLE 3: Normalized fault characteristic parameters.

Parameters	Normal	Tooth miss	Tooth break	Tooth wear	Bearing inner race	Bearing outer race
NP1	-0.1722	-0.2070	-0.5214	-0.4619	0.2612	-0.0560
NP2	0.2163	0.0154	-0.2039	0.5636	-0.0097	-0.2371
NP3	0.3545	-0.0780	0.1330	-0.0686	-0.6341	-0.0873
NP4	-0.3568	-0.7051	-0.7467	0.3346	-0.5576	-0.0312
NP5	-0.2626	-0.3410	0.0753	0.4721	0.4040	-0.1450
NP6	0.0315	0.6561	-0.3575	-0.5123	0.3392	0.3301
NP7	0.2004	0.1586	0.1160	0.1654	-0.1836	-0.1569
NP8	-0.0779	0.2327	0.0570	-0.4184	-0.0005	-0.5677
NP9	-0.0356	-0.0050	-0.0466	-0.0359	-0.0488	-0.1385
NP10	0.2045	0.2679	0.2693	-0.3857	0.1252	0.2183
NP11	-0.3601	0.0658	0.0311	0.0880	-0.1048	-0.2282
NP12	-0.1763	-0.1525	-0.2617	-0.4034	-0.2318	0.1652
NP13	-0.5131	-0.0078	0.3056	-0.3098	-0.0876	-0.6876
NP14	-0.4827	0.3851	-0.1396	0.4718	0.0795	-0.3908
NP15	-0.1136	0.1951	-0.3124	-0.5820	-0.1669	0.1302
NP16	-0.2533	-0.0725	0.4445	0.3444	-0.0275	-0.3765
NP17	-0.3963	-0.3155	-0.2906	0.0142	-0.2101	-0.2109
NP18	-0.6192	-0.0573	0.1311	0.0116	0.1430	0.1435
NP19	-0.2871	0.2230	-0.0180	0.5400	0.1824	0.1822

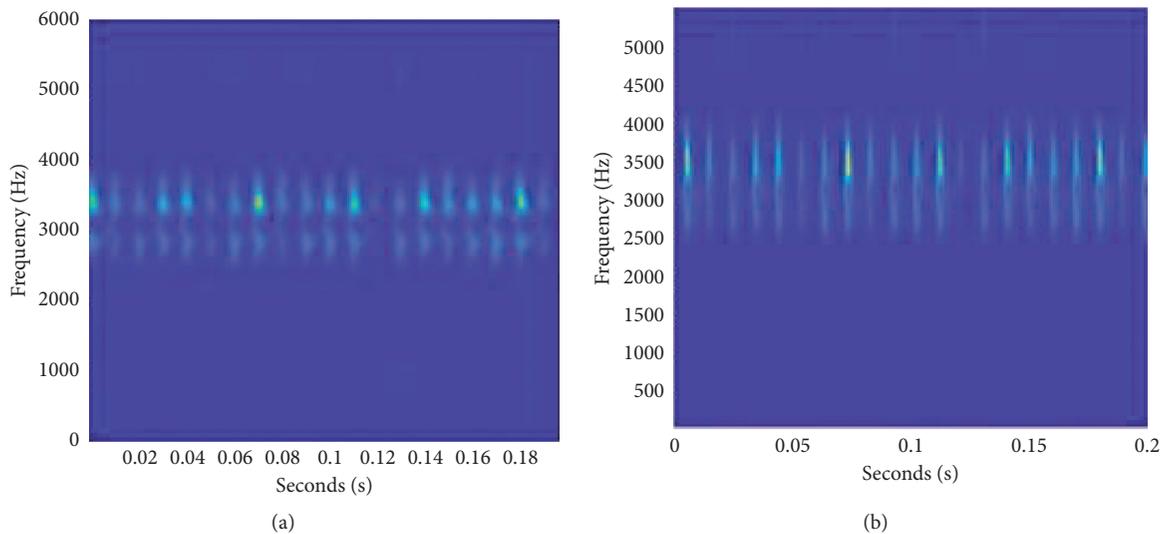


FIGURE 10: Continued.

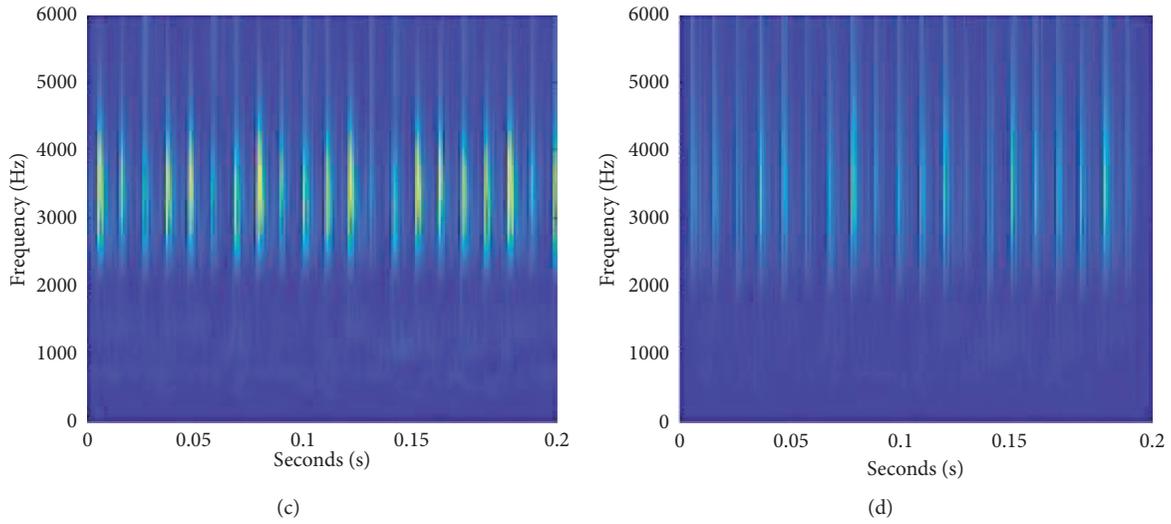


FIGURE 10: Time-frequency analysis of inner race fault: (a) STFT; (b) CWT; (c) ST; (d) MSSST.

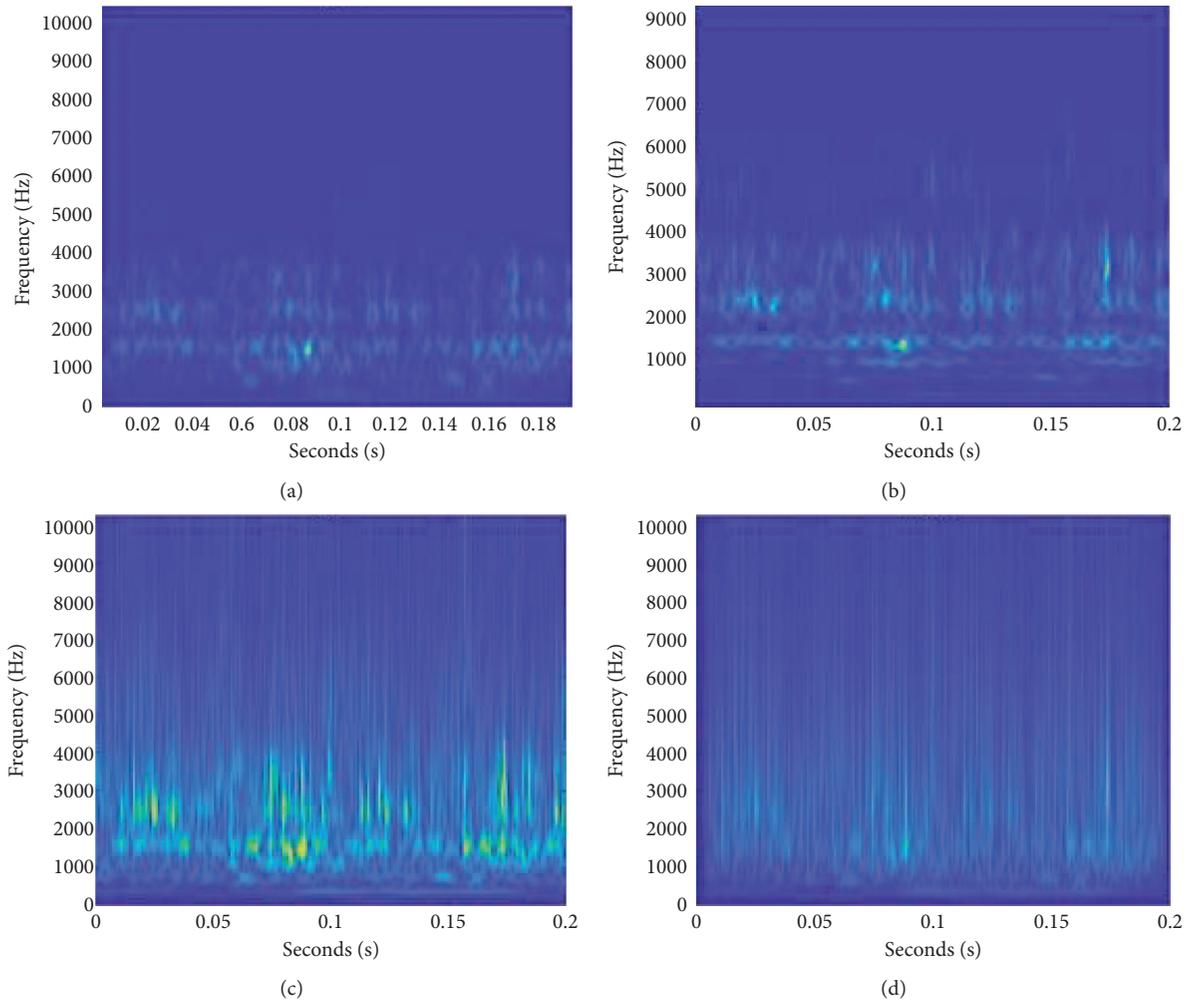


FIGURE 11: Time-frequency analysis of tooth break fault: (a) STFT; (b) CWT; (c) ST; (d) MSSST.

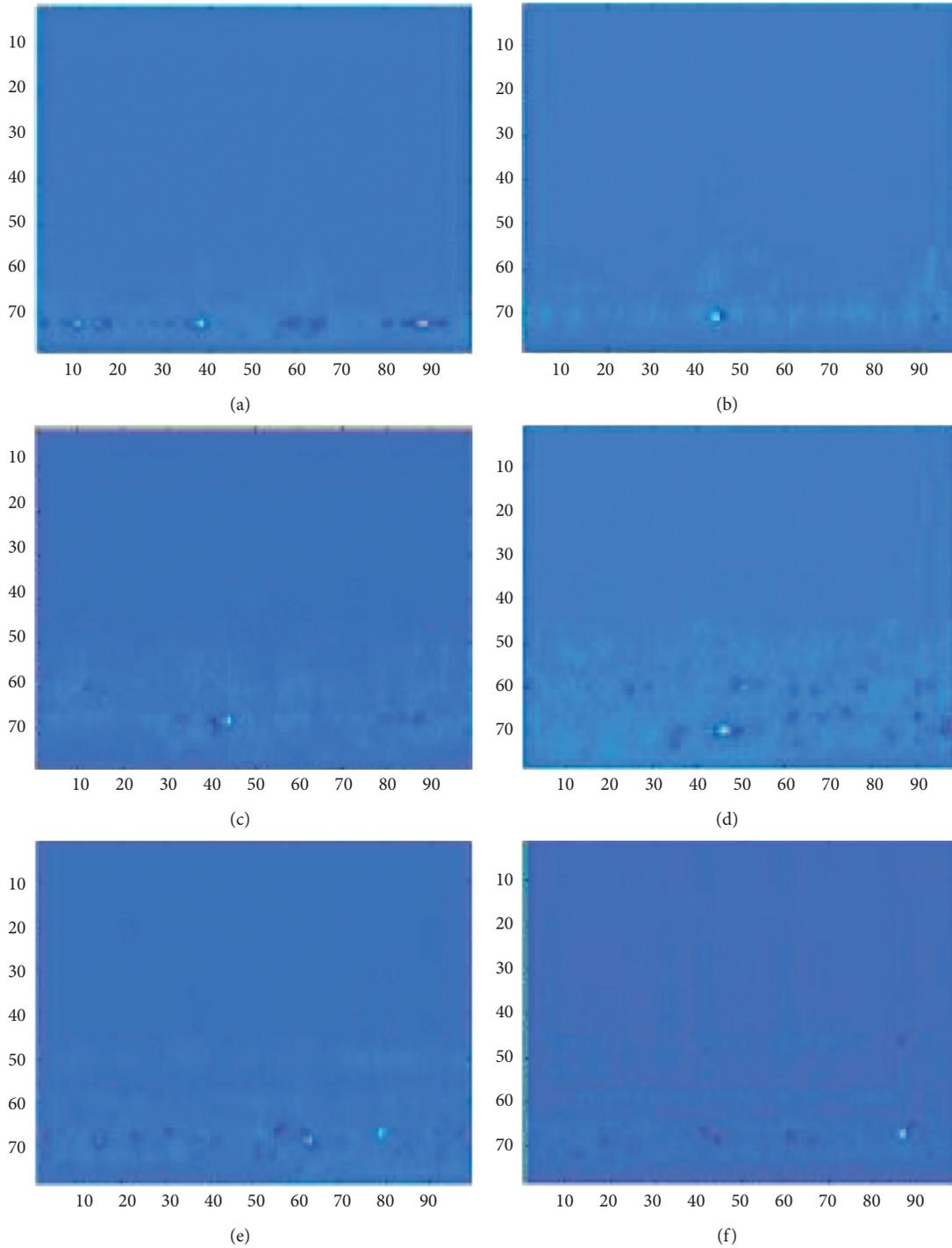


FIGURE 12: Time-frequency feature based on 2DPCA-STFT: (a) normal; (b) miss tooth fault; (c) break tooth fault; (d) wear tooth fault; (e) bearing inner race fault; (f) bearing outer race fault.

#### 4. Comparison and Discussion

In order to verify the effectiveness of the proposed method, ELM and PSO-KELM algorithms are used to compare with the QPSO-KELM algorithm. Meanwhile, in order to verify

the noise filtering effect of quadratic filter, the fault classification accuracy before filtering is also analyzed based on the above three methods, which is compared with the fault diagnosis effect after filtering. The experimental results are shown in Table 4 and Figure 16.

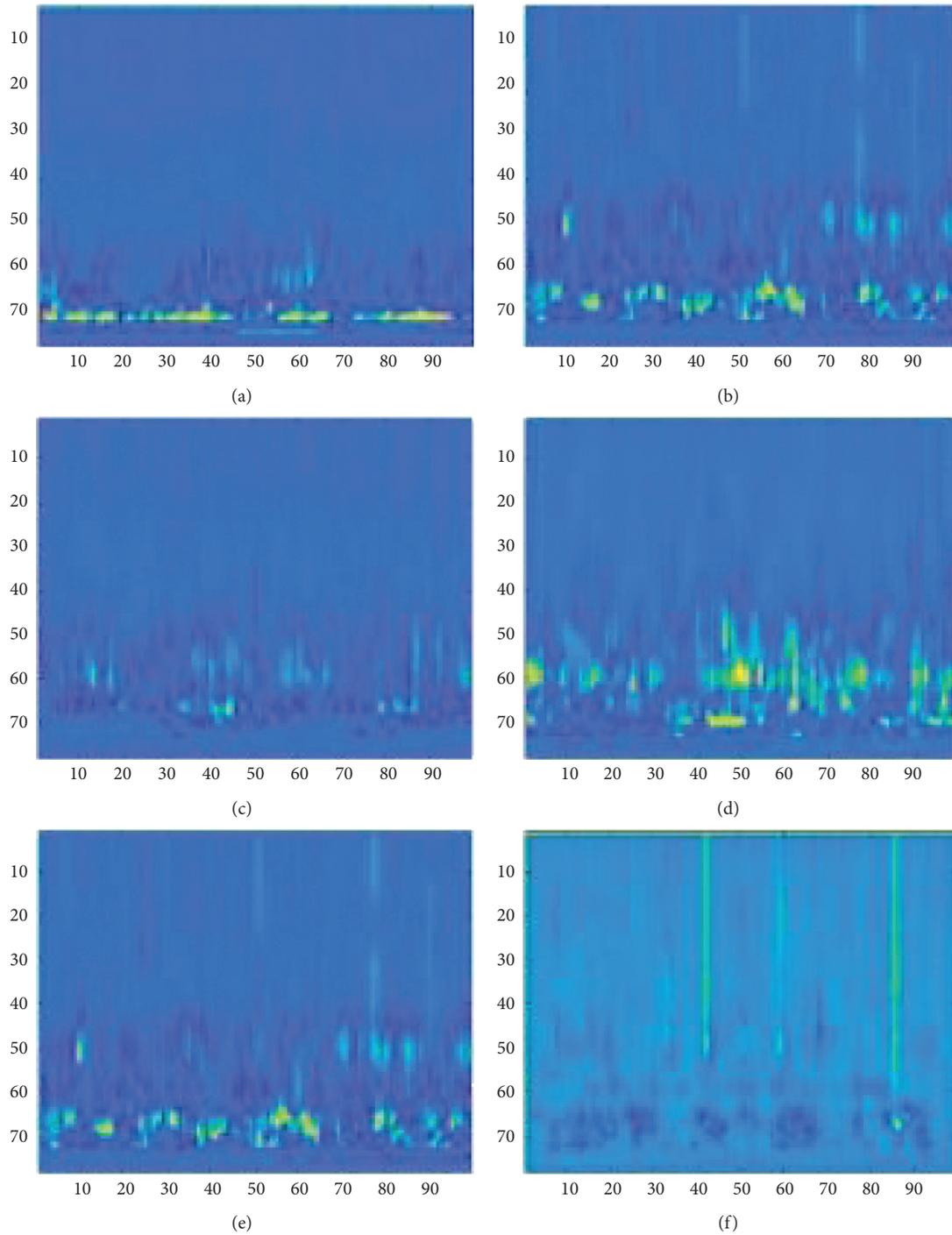


FIGURE 13: Time-frequency feature based on 2DPCA-MSSST: (a) normal; (b) miss tooth fault; (c) break tooth fault; (d) wear tooth fault; (e) bearing inner race fault; (f) bearing outer race fault.

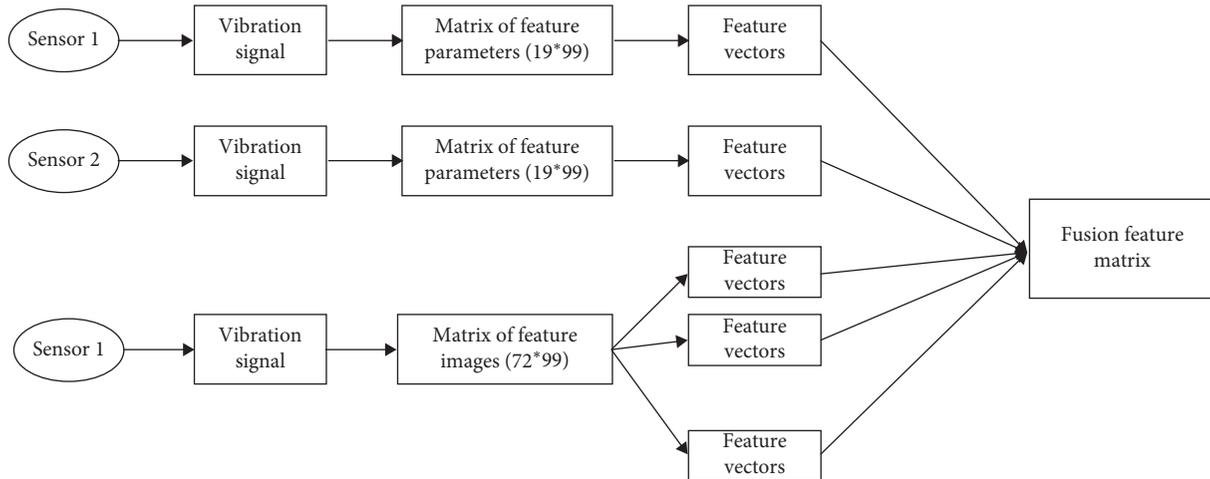


FIGURE 14: Fusion feature flow. Feature vectors are extracted from parameter matrix and image matrix, respectively, for feature fusion.

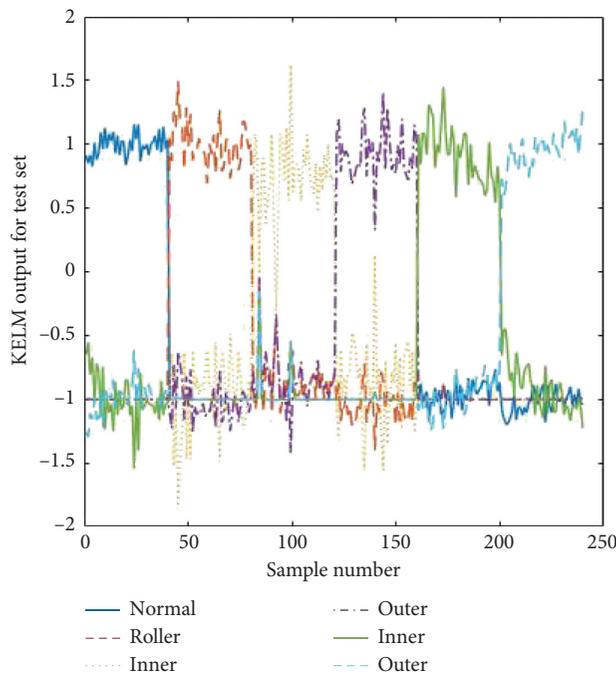


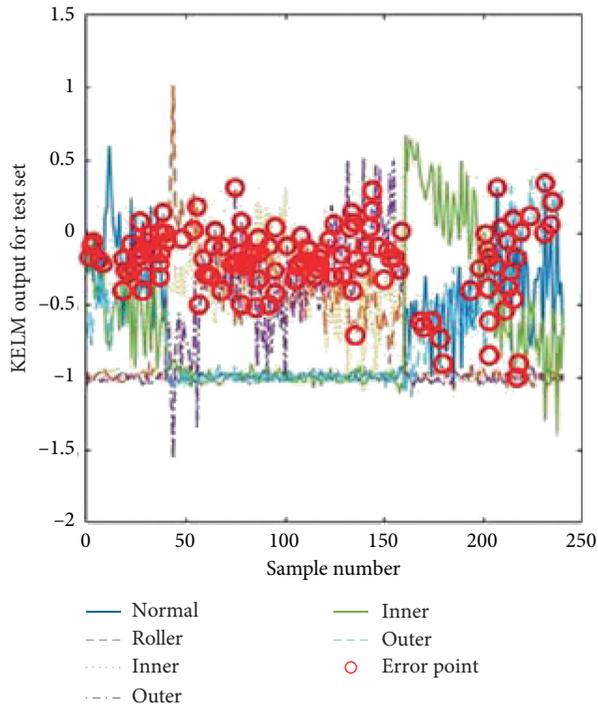
FIGURE 15: QPSO-KELM classification results with quadratic filter. The horizontal axis is the number of samples, and the vertical axis is the output of QPSO-KELM.

TABLE 4: Comparison of fault diagnosis results.

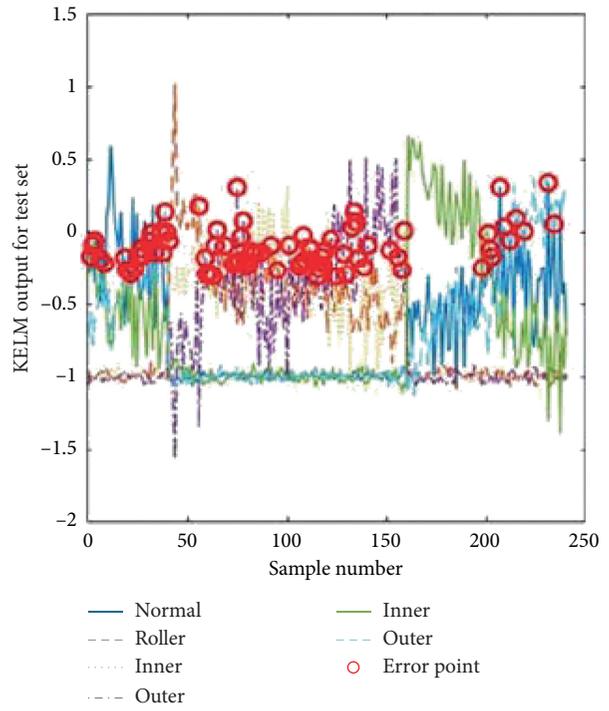
	ELM		PSO-KELM		QPSO-KELM	
	Without filter	Filter	Without filter	Filter	Without filter (%)	Filter (%)
Training accuracy	53.34	75.67	72.67	98.17	79.83	100
Testing accuracy	53.34	75.67	66.67	98.75	70.83	100

In Figure 16 and Table 4, it can be seen that, among all the classification methods, QPSO-KELM has the best classification effect and the highest fault diagnosis accuracy. By comparing the fault diagnosis accuracy before and after

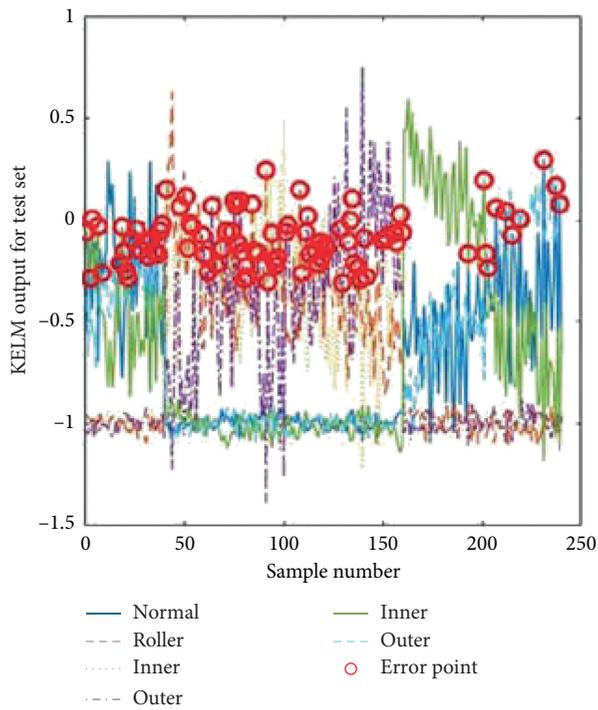
filtering, it can be concluded that no matter which method is used, the fault diagnosis accuracy after filtering has been improved. Therefore, the effectiveness of the proposed method is proved.



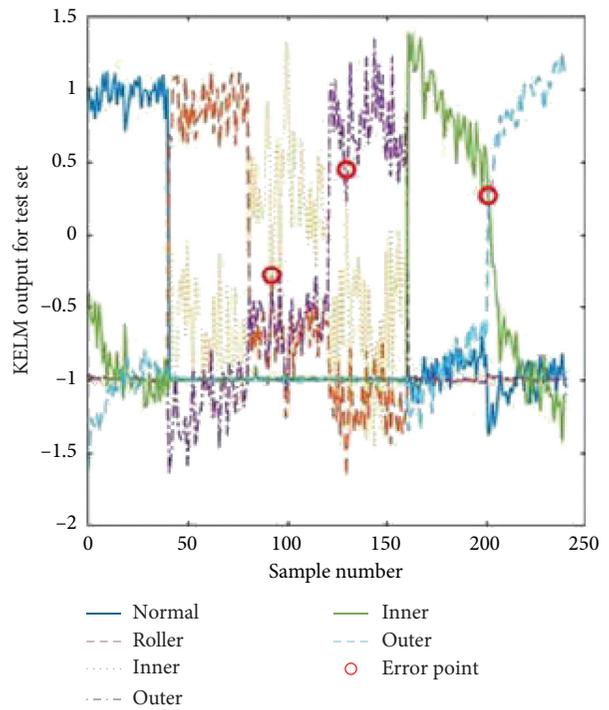
(a)



(b)



(c)



(d)

FIGURE 16: Continued.

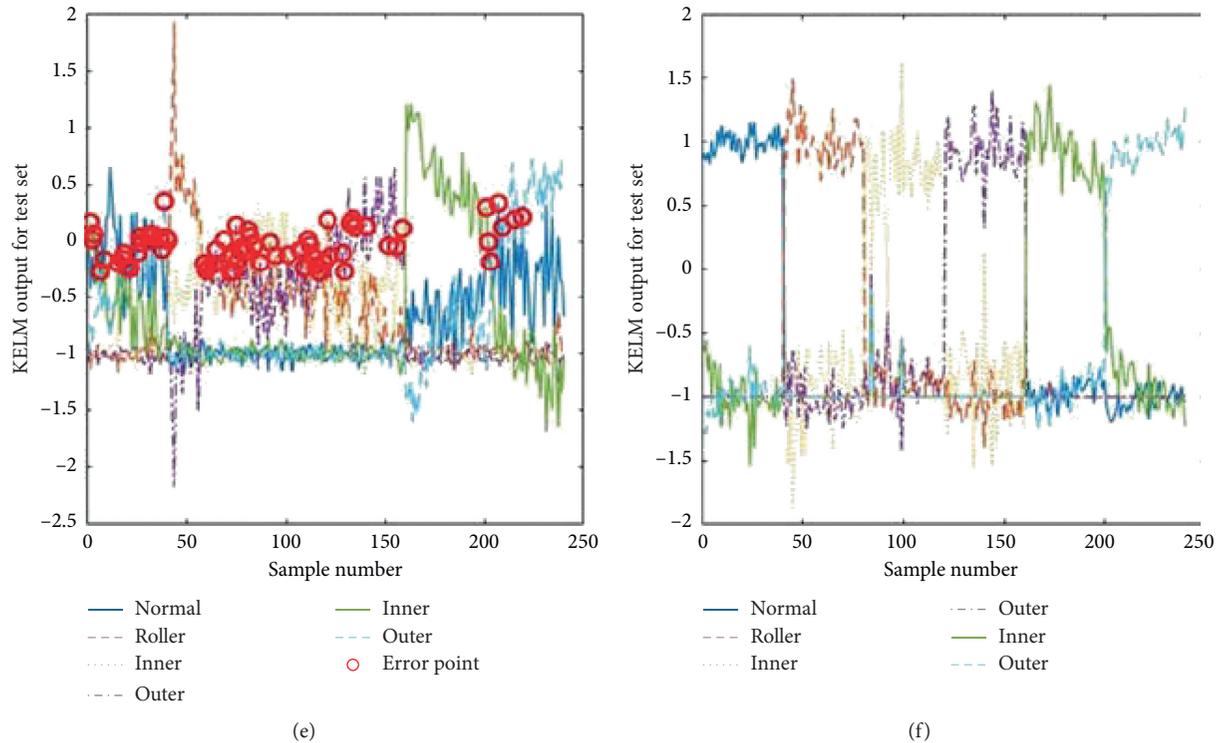


FIGURE 16: Comparison of fault diagnosis results: (a) ELM classification results without filter; (b) ELM classification results with quadratic filtering; (c) PSO-KELM classification results without filter; (d) PSO-KELM classification results with quadratic filter; (e) QPSO-ELM classification results without filter; (f) QPSO-KELM classification results with quadratic filter.

## 5. Conclusion

In this paper, a fault diagnosis method based on gearbox multifeature fusion, quadratic filter, and QPSO-KELM algorithm is proposed. Through the analysis of the experimental results, the following conclusions can be drawn: after stable filtering of the collected signal, the noise interference can be effectively filtered out, and the SNR can be improved. Combined with the method of fast-kurtogram, the fault frequency band with rich fault characteristic signals can be screened out, which is also conducive to the improvement of fault diagnosis accuracy. By extracting the time-domain, frequency-domain, and time-frequency characteristic parameters, more fault characteristic information can be extracted, which is beneficial to improve the accuracy of fault diagnosis. The MSSST method is proposed to improve the time-frequency analysis method, which can effectively solve the problems of energy leakage and low resolution. The KELM algorithm optimized by QPSO, to a large extent, improves the learning ability and classification accuracy of the KELM algorithm. Compared with other methods, the classification effect is better.

Although the method proposed in this paper is better than the traditional fault diagnosis method, it still needs to continue to explore in some aspects, looking for some new fault characteristic parameters to make it more conducive to monitor the current state of the equipment. In this paper, only some single fault modes of the gearbox are analyzed, but in the actual situation, the compound fault of gearbox is more common. So, the next step of the study is to extract

more sensitive and effective fault feature parameters and try to analyze more complex gearbox failure.

## Data Availability

The experiment data are obtained through the laboratory bench of our unit, which is not open to the public.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## References

- [1] A.-O. Boudraa and J.-C. Cexus, "EMD-based signal filtering," *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 6, pp. 2196–2202, 2007.
- [2] A. O. Boudraa and J. C. Cexus, "Denosing via empirical mode decomposition," *International Symposium on Communications, Control and Signal Processing*, vol. 6, 2006.
- [3] R. J. Hao, X. J. An, and Y. L. Shi, "Single-channel blind source separation based on EMD and CICA and its application to gearbox multi-fault diagnosis," *Journal of Vibration and Shock*, vol. 38, no. 8, pp. 230–236, 2019.
- [4] Y. Shen, Z. Zhao, and S. Yang, "Improved EMD based denoising method," *Journal of Vibration and Shock*, vol. 28, no. 12, pp. 35–37, 2009.
- [5] M. Y. Wang, "The fault monitor of planetary gear grinding based on wavelet and Hilbert transform," *Machine Design and Manufacturing Engineering*, vol. 48, no. 5, pp. 105–108, 2019.

- [6] Z. L. Wang, H. L. Ren, and J. G. Ning, "Acoustic emission source location based on wavelet transform de-noising," *Journal of Vibration and Shock*, vol. 37, no. 4, pp. 226–232, 2018.
- [7] H. Gu, J. Zhao, and X. Zhang, "Hybrid methodology of degradation feature extraction for bearing prognostics," *Eksploracja I Niezawodnosc-Maintenance and Reliability*, vol. 15, no. 2, pp. 195–201, 2013.
- [8] H. S. Seung and D. D. Lee, "COGNITION: the manifold ways of perception," *Science*, vol. 290, no. 5500, pp. 2268–2269, 2000.
- [9] S. T. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
- [10] J. B. Tenenbaum, V. Silva, and J. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.
- [11] T. Song, B. P. Tang, and F. Li, "Fault diagnosis method for rotating machinery based on manifold learning and K-nearest neighbor classifier," *Journal of Vibration and Shock*, vol. 32, no. 5, pp. 149–153, 2013.
- [12] J. F. Wang and F. C. Li, "Signal processing methods in fault diagnosis of machinery-analyses in time domain," *Noise & Vibration Control*, vol. 4, no. 2, pp. 198–202, 2013.
- [13] C. Li and M. Liang, "Continuous-scale mathematical morphology-based optimal scale band demodulation of impulsive feature for bearing defect diagnosis," *Journal of Sound and Vibration*, vol. 331, no. 26, pp. 5864–5879, 2012.
- [14] J. F. Wang and F. C. Li, "Signal processing methods in fault diagnosis of machinery-analyses in frequency domain," *Noise & Vibration Control*, vol. 2, no. 1, pp. 173–180, 2013.
- [15] J. F. Wang and F. C. Li, "Review of signal processing methods in fault diagnosis for machinery using time-frequency analysis," *Noise and Vibration Control*, vol. 3, no. 2, pp. 202–206, 2013.
- [16] L. H. Wang, X. P. Zhao, J. X. Wu et al., "Motor fault diagnosis based on short-time fourier transform and convolutional neural network," *Chinese Journal of Mechanical Engineering*, vol. 30, no. 6, pp. 272–277, 2017.
- [17] M. T. Zhu, X. J. Wu, G. F. Zheng et al., "Load signal edition method based on the short-time fourier transform to durability test of vehicle component," *Journal of Mechanical Engineering*, vol. 55, no. 4, pp. 143–151, 2019.
- [18] J. D. Jia, C. Z. Wu, X. Y. Jia et al., "A time-frequency analysis method suitable for engine vibration signals," *Automotive Engineering*, vol. 39, no. 1, pp. 97–101, 2017.
- [19] J. D. Zheng, H. Y. Pan, S. B. Yang et al., "Generalized variational mode decomposition and its applications to gearbox fault diagnosis under variable conditions," *Journal of Vibration Engineering*, vol. 30, no. 3, pp. 502–509, 2017.
- [20] Q. Wang, J. Gao, N. Liu, and X. Jiang, "High-resolution seismic time-frequency analysis using the synchrosqueezing generalized S-transform," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 3, pp. 374–378, 2018.
- [21] Z. Huang and J. Zhang, "Synchrosqueezing S-transform," *Scientia Sinica Informationis*, vol. 46, no. 5, pp. 643–650, 2016.
- [22] D.-H. Pham and S. Meignen, "High-order synchrosqueezing transform for multicomponent signals analysis-with an application to gravitational-wave signal," *IEEE Transactions on Signal Processing*, vol. 65, no. 12, pp. 3168–3178, 2017.
- [23] J. Li, B. Li, and M. Q. Wang, "Expert system of gear's fault diagnosis based on neural network," *Journal of Mechanical Transmission*, vol. 31, no. 5, pp. 81–83, 2007.
- [24] J. L. Qu, L. Yu, T. Yuan et al., "Adaptive fault diagnosis algorithm for rolling bearings based on one-dimensional convolutional neural network," *Chinese Journal of Scientific Instrument*, vol. 39, no. 7, pp. 134–143, 2018.
- [25] Y. K. Wang, "A diagnosis method for electrical submersible pump based on RSGWPT-MSE and PNN," *Mechanical Engineering*, vol. 1, pp. 167–170, 2018.
- [26] H. L. Tian, H. R. Li, and B. H. Xu, "Fault prediction for hydraulic pump based on EEMD and SVM," *China Mechanical Engineering*, vol. 24, no. 20, pp. 926–931, 2013.
- [27] Y. Yu, R. Bai, and P. Yang, "Application of support vector machine based on genetic algorithm optimization on acoustic emission detection of gear fault," *Journal of Mechanical Transmission*, vol. 42, no. 1, pp. 163–166, 2018.
- [28] G. B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 2, pp. 513–529, 2012.
- [29] F. Gao, H. Li, and B. Xu, "Applications of extreme learning machine optimized by ICPSO in fault diagnosis," *China Mechanical Engineering*, vol. 24, no. 20, pp. 2753–2757, 2013.
- [30] A. Iosifidis, A. Tefas, and I. Pitas, "On the kernel extreme learning machine classifier," *Pattern Recognition Letters*, vol. 54, no. 1, pp. 11–17, 2015.
- [31] P. M. Shi, K. Liang, and N. Zhao, "An S.J. Intelligent fault diagnosis for gears based on deep learning feature extraction and particle swarm optimization SVM state identification," *China Mechanical Engineering*, vol. 28, no. 9, pp. 1056–1061, 2017.
- [32] J. T. Wen, C. H. Yan, J. D. Sun et al., "Bearing fault diagnosis method based on compression acquisition and deep learning," *Chinese Journal of Scientific Instrument*, vol. 39, no. 1, pp. 171–179, 2018.
- [33] J. Xu, Y. Wang, and Z. Ji, "Fault diagnosis method of rolling bearing based on WKELM optimized by whale optimization algorithm," *Journal of System Simulation*, vol. 29, no. 9, pp. 2189–2197, 2017.
- [34] H. B. Duan, D. B. Wu, and X. F. Yu, "Ant colony algorithm: survey and prospect," *Engineering Science*, vol. 9, no. 2, pp. 98–102, 2007.
- [35] H. S. Wu, F. M. Zhang, and L. S. Wu, "New swarm intelligence algorithm-wolf pack algorithm," *Systems Engineering and Electronics*, vol. 35, no. 11, pp. 2430–2438, 2013.
- [36] K. L. Tsui, N. Chen, Q. Zhou, Y. Hai, and W. Wang, "Prognostics and Health management: a review on data driven approaches," *Mathematical Problems in Engineering*, vol. 2015, no. 1, 17 pages, Article ID 793161, 2015.
- [37] C. Jiang, S. L. Liu, and R. H. Jiang, "IMF of EEMD selection based on fast spectral kurtosis," *Journal of Vibration, Measurement & Diagnosis*, vol. 35, no. 170, pp. 1173–1178, 2015.
- [38] R. Xueping, L. Pan, and W. Chao, "Rolling bearing early fault diagnosis based on VMD and FSK," *Bearing*, vol. 11, no. 12, pp. 39–43, 2017.
- [39] Z. Liu, L. Tang, and L. Cao, "Fault diagnosis for rolling bearings based on kalman filter and spectral kurtosis," *Bearing*, vol. 12, no. 8, pp. 39–43, 2014.
- [40] W. H. Li, L. Lin, and W. P. Shan, "Fault diagnosis for rolling bearings based on GST and bidirectional 2DPCA," *Journal of Vibration, Measurement & Diagnosis*, vol. 16, no. 3, pp. 499–506, 2015.
- [41] S. Zhang, Y. Cai, and W. Mu, "Coding feature extraction and diagnosis of I.C.engine vibration time-frequency images," *Journal of Mechanical Science and Technology*, vol. 37, no. 3, pp. 391–395, 2018.
- [42] P. Lian, J. W. Yang, and D. C. Yao, "Fault diagnosis method of gear based on adaptive demodulated resonance," *Journal of Mechanical Transmission*, vol. 42, no. 3, pp. 169–174, 2018.