Multi-fault Diagnosis of Rolling Bearing using Two-dimensional Feature Vector of WP-VMD and PSO-KELM Algorithm

tingyu jiang (✉ 1969000240@qq.com)

Beihang University

Sheng Hong

Beihang University

Hao Liu

Beihang University

Research Article

Keywords: rolling bearing, two-dimensional feature vector, fault diagnosis, WP-VMD, PSO-KELM

Posted Date: July 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-714508/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Multi-fault diagnosis of rolling bearing using two-dimensional feature vector of
WP-VMD and PSO-KELM algorithm**

Tingyu Jiang^a, Sheng Hong^b, Hao Liu^{*c}

^aSchool of Cyber Science and Technology, Beihang University, Beijing, 100191,
P.R.China. 1969000240@qq.com

^bSchool of Cyber Science and Technology, Beihang University, Beijing 100191,
P.R.China. jty@buaa.edu.cn

^cSchool of Astronautics, Beihang University, Beijing, 100191, P.R.China.
liuhao13@buaa.edu.cn

^{*}Corresponding author Tel:010-82339496 E-mail addresses: 1969000240@qq.com

Abstract

In order to achieve accurate fault diagnosis of rolling bearing under random noise, a new fault diagnosis method based on wavelet packet-variational mode decomposition (WP-VMD) and kernel extreme learning machine (KELM) optimized by particle swarm optimization (PSO) is proposed in this paper. Firstly, the time-frequency domain feature vectors of the original rolling bearing fault signals are effectively obtained by preprocessing of WMD and decomposition and reconstruction of VMD. Then, the extracted two-dimensional feature vector is input into the KELM neural network for fault identification, and combined with PSO, KELM parameters were optimized. The experimental results show that the proposed method can effectively diagnose the rolling bearing under random noise, with the features of fast speed, stable performance and high accuracy. By comparison, this paper obtains better accuracy and real-time performance with fewer features, which provides a simple and efficient solution for fault diagnosis of rolling bearings.

Key words: rolling bearing; two-dimensional feature vector; fault diagnosis; WP-VMD; PSO-KELM

1. Introduction

Rolling bearing has long been one of the most critical and vulnerable parts of rotating equipment, whose function largely affects the operating state of the entire mechanical system (Huang WT et al. 2019). Failure of rolling bearings can lead to machinery malfunctions or even catastrophic accidents. Therefore, it is of great significance to achieve reliable health monitoring and fault diagnosis for rolling element bearings (Zhang et al. 2018).

Since the rolling bearing vibration signals have the characteristics of non-stationary, nonlinear, multi-modulation, and multi-component, accompanied by random noise interference, it has been a challenging issue that how to effectively acquire the fault feature of the rolling bearing vibration signal in the signal processing field. At present, common extraction methods include wavelet transform (WT) (Zhang et al. 2019; Gilles 2013; Li et al. 2015), short-time Fourier transform (STFT) (He et al. 2014), empirical mode decomposition (EMD) and so on (Liang et al. 2005). However, there still exists some imperfections. For example, the WT can not subdivide well the high-frequency of the signal, inducing information loss and energy leakage (Meng et al. 2016), and the analysis window of STFT needs to be further optimized. These shortcomings make these classic methods not be fully self-adaptive in essential (Yao et al. 2018). As a self-adaptive time-frequency analysis method, EMD is able to handle non-stationary and non-linear signals better, but it has disadvantages such as over-envelope, under-envelope and end-effect, what's more, its modal aliasing

property will lead to the distortion of the decomposed IMFs (Yao et al. 2016). In view of the defects mentioned above, Dragomiretskiy and Zosso (2014) put forward a completely non-recursive self-adaptive signal decomposition method, named variational mode decomposition (VMD) method in 2014, which has a solid mathematical foundation (Li et al. 2019), because it iteratively searches for the optimal solution of the mathematical model on the variational problem. In this way, the precision of IMFs component is high without these problems like over-envelope and under-envelope, furthermore, the phenomenon of mode aliasing is effectively suppressed (Xu et al. 2019). And the superior performance of VMD was proved in the literature (Dragomiretskiy and Zosso 2014; Wang et al. 2015). Gu et al. (2017) decomposed the vibration signal of rolling bearing by VMD, Li et al. (2017) applied VMD to bearing fault diagnosis.

In fact, rolling bearing fault diagnosis can be regarded as a model recognition problem. In general, support vector machines (SVM) (Liu et al. 2015) and artificial neural networks (ANNs) are commonly used to implement pattern recognition (Jia et al. 2015). Nevertheless, they belong to the shallow learning model, lacking of strong characterization efficiency and having troubles in expressing the complex relationship in fault diagnosis problems in an efficient manner (Shao et al. 2017). Based on the disadvantages of SVM and ANNs, Huang et al. stated an available solution, the extreme learning machine (ELM) (Huang et al. 2016), equipped with the advantage of fast running speed. Whereas, there still remains deficiencies such as unstable results

and confusions on determining the number of hidden neurons in practice. To improve it, the kernel function is introduced into the ELM, and KELM is proposed by Huang et al. (2014). By integrating kernel functions, it maps the linear non-separable mode to the high-dimensional feature space, making it linearly separable and improving the accuracy of the algorithm. Compared with ELM, KELM only needs to select the kernel function and its related parameters to obtain the output weight, which has fewer adjustable parameters, better generalization performance, etc. (Qin et al. 2017; Li et al. 2017) applied it to the diagnosis of rolling bearings then gained better accuracy. On the other hand, kernel function makes the KELM be very sensitive to parameter settings. Consequently, it is the key to optimize the KELM parameters for improving the accuracy of state recognition.

Aiming at the above analysis, a new method for fault diagnosis of rolling bearings based on WP-VMD and PSO-KELM is voiced in this paper. The following contents of this paper is organized as follows: In section 2, the fault feature extraction method based on WP-VMD model is introduced. In section 3, the fault diagnosis method based on PSO-KELM model is introduced. In section 4, the test platform is presented and the proposed method of WP-VMD and PSO-KELM was verified by experiments. In section 5, the method proposed in this paper is analyzed and discussed. Finally, the conclusion is summarized in section 6.

2. Introduction of WP-VMD model

2.1. Wavelet Packet Decomposition

Since the WT is incapable to subdivide the high-frequency signal well, easily causing information loss. The more detailed signal decomposition algorithm, WPD (Wang et al. 2015) is proposed to solve this defect, it hierarchically divides the analyzed signals, and improves the resolution of the frequency domain analysis, according to various characteristics of the signal to self-adaptively match the spectrums (Learned et al. 2015; Yan et al. 2018).

Taking the three-layer wavelet packet analysis as an example, the WPD diagram is shown in Fig. 1. It can be seen that WPD does not neglect the further decomposition of the high-frequency of the signal, performing fine decomposition on each node of each layer as well, which improves the resolution of frequency domain analysis.

Fig. 1 Wavelet packet decomposition diagram

The decomposition equation of WPD algorithm is given as

$$d_l^{j,2n} = \sum_k h_{k-2l} d_k^{j+1,n} \quad (1)$$

$$d_l^{j,2n+1} = \sum_k g_{k-2l} d_k^{j+1,n} \quad (2)$$

Where h_{k-2l} and g_{k-2l} are the high-pass and low-pass filter coefficients, $d_l^{j,2n}$ and $d_l^{j,2n+1}$ are wavelet packet coefficients. The reconstruction equation is given as

$$d_l^{j+1,n} = \sum_k (h_{l-2k} d_k^{j,2n} + g_{l-2k} d_k^{j,2n+1}) \quad (3)$$

2.2. Variational Mode Decomposition

The VMD is mainly divided into the construction and solution of variational problems. The corresponding specific steps are as follows:

(1) Construction of constrained variational problem

The VMD constructs a variational problem with the sum of the K IMFs being equal to the input signal f as a constraint and the sum of the estimated bandwidths of the K IMFs as the minimum:

① Each IMF $u_k(t)$ is converted into a corresponding analytical signal by a Hilbert transform, and a unilateral spectrum of the signal is obtained:

$$\left[\delta(t) + \frac{j}{\pi t} \right] * u_k(t) \quad (4)$$

② The pre-estimated center frequency $e^{-j\omega_k t}$ of all modally resolved signals is combined and the spectrum of each IMF is modulated to the corresponding baseband:

$$\left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \quad (5)$$

③ The square root of the time gradient L^2 norm of the demodulated signal is calculated, and the bandwidth of each modal classification is estimated. The representation of the constrained variational model is:

$$\begin{cases} \min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ s.t. \sum_{k=1}^K u_k = f \end{cases} \quad (6)$$

Where $\delta(t)$ is the Dirichlet function, $*$ is the convolution operation. $\{u_k\} = \{u_1, \dots, u_K\}$ represents the set of K IMFs components after VMD decomposition, $\{\omega_k\} = \{\omega_1, \dots, \omega_K\}$ represents the set of K IMFs component center frequencies and f is the input signal.

(2) Solution of constrained variational problem

VMD transforms the problem of constructed constrained variational solution into an unconstrained variational solution by means of importing the quadratic penalty factor α and the Lagrange operator $\lambda(t)$. In the Gaussian noise environment, α can ensure the reconstruction accuracy of the signal, meanwhile $\lambda(t)$ can strengthen the rigor of the constraint, and extended Lagrange expressions are as follows:

$$\begin{aligned} L(\{u_k\}, \{\omega_k\}, \lambda) = & \alpha \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \\ & + \left\| f(t) - \sum_{k=1}^K u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^K u_k(t) \right\rangle \end{aligned} \quad (7)$$

The saddle point of the extended Lagrange expression is found by the alternating direction multiplier method (ADMM). Specific steps are as follows:

- ① Initialize $\{\hat{u}_k^1\}$, $\{\omega_k^1\}$, $\hat{\lambda}^1$, $n=0$.
- ② $n=n+1$, for all $\omega \geq 0$, update $\hat{u}_k$.

$$\hat{u}_k^{n+1}(\omega) \leftarrow \frac{\hat{f}(\omega) - \sum_{i < k} \hat{u}_i^{n+1}(\omega) - \sum_{i > k} \hat{u}_i^n(\omega) + \frac{\hat{\lambda}^n(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2}, \quad k \in \{1, K\} \quad (8)$$

③ Update ω_k :

$$\omega_k^{n+1} \leftarrow \frac{\int_0^\infty \omega \left| \hat{u}_k^{n+1}(\omega) \right|^2 d\omega}{\int_0^\infty \left| \hat{u}_k^{n+1}(\omega) \right|^2 d\omega}, \quad k \in \{1, K\} \quad (9)$$

④ Update λ :

$$\hat{\lambda}^{n+1}(\omega) \leftarrow \hat{\lambda}^n(\omega) + \gamma \left[\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega) \right] \quad (10)$$

Where γ is the noise tolerance parameter and is generally taken as 0.

⑤ Repeat steps ②-④ until the iteration constraint is met to stop the loop.

$$\sum_{k=1}^K \left\| \hat{u}_k^{n+1} - \hat{u}_k^n \right\|_2^2 / \left\| \hat{u}_k^n \right\|_2^2 < \varepsilon \quad (11)$$

Before applying the VMD to signal decomposition, the number of modes K should be set ahead, which will lead to disparate results with different values of the parameter. When K is too small, the key information will lose, while too large, the frequency aliasing will come up. Aiming at this situation, it is known that if there is a component similar to the center frequency during the decomposition process, it indicates that over-decomposition has occurred. This paper sets forth a method to determine the number of modes K by using this feature. Firstly, the number of initialization modes is $K = 2$, and then it is determined whether the center frequency of the IMFs overlap. If there is no overlap, $K = K + 1$ is repeated. Otherwise, we get the final mode number as $K = K - 1$.

2.3. The model of WP-VMD

Fig. 2 shows the fault feature vector extraction process of rolling bearing based on WP-VMD method. Firstly, the original vibration signal is denoised by the WPD method, and the denoised signal is decomposed into three layers to obtain signals of eight groups of different frequency bands. The signal energy of each group is calculated, and the frequency bands with larger energy is selected to be reconstructed. Then, using the VMD method to decompose the reconstructed signal into multiple IMFs and taking advantage of the kurtosis-correlation criteria to determine the IMFs proportion as the reconstruction coefficient. Finally, feature extraction is exerted on the reconstructed signal, and the best feature vectors in the time-frequency domain is pick up to construct a two-dimensional fault feature vector.

Fig. 2 The procedure of rolling bearing fault feature vector extraction

3. Introduction of PSO-KELM model

Through detailed analysis of the neural network model in fault classification, choose the KELM that introduces the kernel function framework in SVM under the ELM model to fault classification. Since radial basis function (RBF) has strong generalization ability and smoothness (Lian et al. 2013), this paper takes RBF as the kernel function of KELM. In order to achieve the best search performance of KELM, it is essential to single out the best penalty coefficient C and kernel function parameter γ , afterwards, a fault diagnosis model based on PSO-KELM is proposed.

The detailed modeling processes are as follows:

- ① Initialize PSO parameters and determine the coding form as $\theta = (C, \gamma)$.
- ② The KELM training is performed on the training subset using the parameters, obtained by the initial population decoding, and the average error rate obtained by the 5-CV cross-validation method is recognized as the fitness function. The average error rate is calculated as:

$$AER = 1 - \frac{\sum_{i=1}^m T(i)}{m} \quad (12)$$

Among them, $T(i)$ represents the test accuracy of the i -th training sample obtained by the 5-CV method, and m represents the number of cross-validation, where $m = 5$.

- ③ Increase the number of iterations.

④ Update the position and velocity of the particles, and the fitness value of each population.

⑤ The combination of parameters obtained by decoding the new particles of the previous step, is trained on KELM, and the fitness value is calculated according to equation (12).

⑥ If it has not reached the maximum population, go back to the fourth step to execute, otherwise, the execution will continue.

⑦ Compare the current fitness value with the global optimal fitness value and determine the global optimal fitness value.

⑧ If the maximum number of iterations has not been reached, go back to the third step, otherwise, continue.

⑨ Output global optimal parameter combination (C, γ) .

⑩ The optimal parameter combination (C, γ) is combined with the training set to train on KELM, and then the fault diagnosis result is obtained.

4. Test and Verification

4.1. Experiment platform

The experimental data used in this paper was collected by the electro-discharge machining (EDM) experimental platform built by Case Western Reserve University. Fig. 3 shows the fault simulation experimental platform. The rolling bearing model in the experiment is SKF6205, which is a deep groove ball bearing with an inner diameter of 25 mm, an outer diameter of 52 mm, and a thickness of 15 mm. The

experimental platform employs EDM technology, to introduce a single point of failure to the test bearing, and records the vibration data corresponding to the bearing fault condition. The sampling frequency of the experimental platform is 12k, and a total of about 10 seconds of data is collected. Fig. 4 shows four states of the rolling bearing. Under the situation when the motor speed is 1772/min, the drive end data is collected. Considering the sampling frequency and the characteristic frequency, the 2400 sample length is enough to be on behalf of the features of the bearing. In the four states of normal, inner race fault, outer race fault and rolling fault, 50 sets of data are obtained, thus a total of 200 sets of data are obtained. Fig. 5 shows the waveforms of original time domain vibration signal in each state of the rolling bearing.

Fig. 3 Rolling bearing experimental platform

Fig. 4 Bearing conditions [9]

Fig.5 Vibration signals of each state: a) Normal state; b) Inner race fault; c) Outer race fault; d) Rolling fault

4.2. Fault feature extraction based on WP-VMD

Taking the data of the inner race fault as an example, the wavelet packet denoising is performed on the signal, and the denoised signal is decomposed by WPD into three layers to obtain signals of eight groups of different frequency bands, which are recorded as $x_3^1(t), x_3^2(t), \dots, x_3^8(t)$, and the energy of each group of signals is calculated, and recorded as $E_3^1(t), E_3^2(t), \dots, E_3^8(t)$, as shown in Fig. 6. The main energy bands are selected for reconstruction, and the reconstructed signal is used as the VMD decomposition object, where the mode number is $K=3$. The original reconstructed signal and the decomposition results are shown in Fig. 7.

Fig. 6 Energy of each wavelet packet sub-band

Fig. 7 The reconstructed signal and the decomposition results of inner race fault

In order to extract the fault signal characteristic more precisely, the IMFs components are comprehensively analyzed in accordance with the kurtosis-correlation coefficient criterion. For the sake of paying more attention to the IMFs components with a larger amount of information, the normalized proportion of the two indicators

of the k IMFs components are calculated to the range[0, 1], recorded as q_i and w_i respectively, the reconstruction coefficient r_i of the corresponding IMF component is

$$r_i = \frac{q_i + w_i}{2} \quad (13)$$

Where $i = 1, 2, \dots, k$, and the calculation results are shown in Table 1.

Table 1 Kurtosis $Kurt$, correlation coefficient $Corrcoef$, and normalized proportion of IMFs components

IMFs component	IMF1	IMF2	IMF3
$Kurt$	2.2134	4.4649	3.0613
$Corrcoef$	0.5607	0.6845	0.5642
r	26.86%	41.83%	31.31%

The appropriate feature quantity selection is the key to fault diagnosis. In the time domain, the Multi-Scale Entropy (MSE) of the signal is deemed to the feature quantity. Owing to the MSE is sensitive to the scale setting, the selection of the best scale is constantly too subjective. On the basis of the idea of statistics, the maximum value of the scale was set as 30, and the peak, variance, and root mean square value of the MSE at 1-30 scale in the statistics are figured out. In the frequency domain, the frequency point of energy concentration and its corresponding energy can be installed as parameters to signify the energy distribution trend in the spectrum, which is a crucial feature reflecting the bearing condition. Then calculating the center of gravity frequency f_c and the center of gravity amplitude a_c of the signal to reveal these

features. The kurtosis $Kurt$, the correlation coefficient $Corrcoef$, and the MSE are introduced as follows:

$$Kurt = \frac{1}{N} \sum_{i=1}^N \left(\frac{x(i) - \bar{x}}{\sigma^4} \right)^2 \quad (14)$$

$$Corrcoef(x, y) = \frac{\sum_{i=1}^N x(i) y(i)}{\left[\sum_{i=1}^N x(i)^2 \sum_{i=1}^N y(i)^2 \right]^{\frac{1}{2}}} \quad (15)$$

Among them, N is the length of the signal sequence, $x(i)$ is the value of the i -th signal sequence, $\bar{x}$ is the average value of the signal, σ is the standard deviation of the signal.

Costa et al. brings in the idea of coarse granulation and elaborate the MSE based on sample entropy (Madalena et al. 2005; Madalena et al. 2007), on the reason that the sample entropy can only reflect the time series complexity on one scale. The calculation procedures are as follows:

(1) Let the time series be $\{x(1), x(2), \dots, x(N)\}$ and the length be N , establish a coarse vector

$$p_i(\tau) = \frac{1}{\tau} \sum_{j=(i-1)\tau+1}^{i\tau} x_j, 1 \leq i \leq N/\tau \quad (16)$$

Where τ is the scale factor, and when $\tau=1$, the sequence is not subjected to the coarse granulation process, and $p_i(1)$ is the original sequence.

(2) Give the dimension m and threshold r of the model, and obtain the sample entropy of each coarse-grained sequence (Richman and Moorman 2000). The accessible sample entropy $SampEn(m, r, N)$ at each scale is described as MSE.

The five mathematical characteristics of peak p , variance s^2 , root mean square value RMS , center of gravity frequency f_c , and center of gravity amplitude a_c are

$$p = \frac{1}{2}(x_{\max} - x_{\min}) \quad (17) \quad s^2 = \frac{\sum_{i=1}^N (x(i) - \bar{x})^2}{N} \quad (18) \quad RMS = \sqrt{\frac{\sum_{i=1}^N (x(i) - \bar{x})^2}{N-1}} \quad (19)$$

$$f_c = \frac{\sum_{i=1}^N f_i P(f_i)}{\sum_{i=1}^N P(f_i)} \quad (20) \quad a_c = \frac{\sum_{i=1}^N f_i P(f_i)}{\sum_{i=1}^N f_i} \quad (21)$$

Where f is the spectrum of the signal sequence and P is the power spectrum of the signal sequence.

The results of the five feature vectors of the reconstructed signal are shown in Fig. 8-12. Each picture implies the corresponding distinctions of the four states of the bearing vibration signal. The larger the difference of the feature vectors in the four states, the better the effect of fault classification. Fig. 8 and Fig. 9 symbolize the peak and variance characteristics of the MSE respectively. In Fig. 8, there are aliasing phenomena in the four state feature vector curves. In Fig. 9, the outer race and the rolling feature vector curves are aliased. It is overt that both of them fail in distinguishing well the state of rolling bearings. Compared with Fig. 8 and Fig. 9, the root mean square value feature vector curves of Fig. 10 has no aliasing phenomenon, the feature vector curves of the four states are distinctly different as well, in this way, the root mean square value of the MSE of the signal is selected in the time domain as feature vector. Similarly, from the comparison of Fig. 11 and Fig. 12, the feature

vector curves aliasing phenomenon of Fig. 11 is serious. The feature vector curves of Fig. 12 have a large spacing, which can distinguish four states well. Therefore, the center of gravity amplitude of the signal is recognized as the feature vector in the frequency domain.

Consequently, taking the signal's root mean square value of the MSE as well as center of gravity amplitude as two-dimensional feature vector in the time-frequency domain for fault diagnosis. Compared with other feature extraction methods, not only does this method comprehensively take into account the time-frequency domain characteristics of the signal, but also its fewer dimensions on construction of signal eigenvectors.

Fig. 8 Peak of MSE of the reconstructed signal

Fig. 9 Variance of MSE of the reconstructed signal

Fig. 10 Root mean square value of MSE of the reconstructed signal

Fig. 11 Center of gravity frequency of the reconstructed signal

Fig. 12 Center of gravity amplitude of the reconstructed signal

4.3. Fault diagnosis based on PSO-KELM

Avoiding the impact of various magnitude as well as dimension feature data on convergence speed and classification accuracy of the neural network model, all sample data is normalized to $[-1, 1]$. The formula is:

$$y = 2 \times \frac{x - x_{\min}}{x_{\max} - x_{\min}} + (-1) \quad (22)$$

Where x is the true value of a sample of a single data, $x_{\min}$, $x_{\max}$ are respectively the minimum and maximum values of the sample data, and the y is a standardized output data.

Among 200 groups of experimental data, 160 groups were taken as training samples and 40 groups as testing samples. The minimum average error rate obtained under the 5-CV method was taken as the fitness evaluation function, after 100 iterations adopting PSO algorithm, simultaneously, the optimal parameter combination (C, γ) was reached. The experiment was repeated 50 times to verify the effectiveness of this method. The fault diagnosis results of the first 20 experiments are shown in Table 2. From the results in Table 2, the fault diagnosis model based on PSO-KELM classifies accurately the fault state, providing preferable diagnostic effect.

Table 2 The results of first 20 experiments of PSO-KELM

Test times	Testing Accuracy				(C, γ)
	Normal state	Inner race state	Outer race state	Rolling fault	
1	100%	100%	100%	100%	(9562.46, 17.04)
2	100%	97.5%	100%	100%	(9133.86, 86.59)

3	100%	100%	100%	100%	(1874.53, 49.78)
4	100%	100%	100%	100%	(3584.95, 26.98)
5	100%	100%	100%	100%	(4411.90, 32.98)
6	100%	100%	100%	100%	(6692.76, 1.78)
7	100%	100%	100%	100%	(6933.55, 46.27)
8	100%	100%	100%	100%	(3554.06, 15.13)
9	100%	97.5%	100%	100%	(8529.92, 54.15)
10	100%	97.5%	100%	100%	(4552.70, 75.69)
11	100%	100%	100%	100%	(3430.79, 18.44)
12	100%	100%	100%	100%	(7906.51, 61.83)
13	100%	100%	100%	100%	(1884.71, 2.56)
14	100%	100%	100%	100%	(368.36, 82.05)
15	100%	100%	100%	100%	(1155.89, 3.45)
16	100%	100%	100%	100%	(2625.39, 21.90)
17	100%	100%	100%	100%	(6210.61, 12.96)
18	100%	100%	100%	100%	(7019.02, 11.65)
19	100%	97.5%	100%	100%	(514.31, 80.22)
20	100%	97.5%	100%	100%	(6426.57, 98.49)

5. Analysis and discussion

5.1. Experimental data analysis

Since the experimental data processed in this article is collected by an ordinary EDM experimental platform, the data collected by the experimental platform cannot accurately reflect the state of the bearing itself. This is because a large amount of noise is easily introduced during the collection process and cannot be describe the sample-and-hold behavior of the measurement output well. At present, documents

(Qiu et al. 2020a; Qiu et al. 2020b) propose more effective methods for data acquisition of random nonlinear systems to describe the sample-and-hold behavior of the measurement output of the data and to filter the data. Therefore, the use of better methods to collect data will further improve our experimental results.

5.2. Performance comparison

The extracted two-dimensional feature vector is input into the BP and SVM for fault diagnosis. In the 50 experiments, the average accuracy of PSO-KELM proposed in this paper is 99.65%, BP and SVM were 86.45% and 96.88% accuracy. The specific classification results of the three methods are shown in Table 3.

Table 3 Comparisons among BP, SVM and PSO-KELM

Algorithm	Accuracy (%)			
	Normal state	Inner race fault	Outer race fault	Rolling fault
BP	97%	92.6%	63%	93.2%
SVM	100%	97.5%	97.5%	92.5%
PSO-KELM	100%	98.6%	100%	100%

5.3. Time analysis

To illustrate the high-efficiency of the WP-VMD method proposed in this paper, the comparison of the real-time performance between the WP-VMD and the VMD, by calculating the total running time of 50 samples in each state of the rolling bearing. Averting the contingency of the results, a total of 50 experiments were performed to figure out the average total running time of two algorithms. Under the same test

conditions, the WP-VMD costs less time and has better real-time performance than the VMD. By summarizing the four state feature extraction times of the rolling bearing under the VMD and WP-VMD, the efficiency of the WP-VMD method can be improved by about 27.82%. The experimental results are shown in Fig. 13.

Fig. 13 Comparisons of WP-VMD and VMD

6. Conclusions

In order to solve the problem of fault diagnosis of rolling bearing, this paper proposes a new method based on WP-VMD and PSO-KELM rolling bearing fault diagnosis. The validity of the proposed method is verified by the experimental data of Case Western Reserve University. Furthermore, the accuracy as well as real-time performance of the proposed method are analyzed and discussed.

1) The WP-VMD method is utilized to preprocess the signal and decompose it into multiple IMFs, then the kurtosis-correlation criteria is employed to determine the IMFs proportion as the reconstruction coefficient.

2) In the time-frequency domain, the root mean square value of the MSE as well as the center of gravity amplitude of the reconstructed signal are respectively selected as two-dimensional feature vector, which is input into the PSO-KELM algorithm for training and diagnosis. Due to the comprehensive consideration of the time-frequency domain indicators and the use of fewer features, the accuracy of fault diagnosis and the speed of fault diagnosis are improved.

3) Considering that the KELM is sensitive to parameter settings, whose classification performance is unstable. The PSO-KELM is proposed, which single out the best penalty coefficient C and kernel function parameter γ of KELM, with better and stable fault diagnosis performance.

4) The accuracy and real-time performance of the method in this paper are analyzed and discussed. On the one hand, compared with the traditional neural network BP and SVM, the PSO-KELM is equipped with higher fault accuracy. On the other hand, compared with the VMD, the proposed WP-VMD method costs less time and has preferable real-time performance.

Although the method proposed in this paper has a good performance in the multi-fault diagnosis of rolling bearing, it cannot identify faults of unknown fault types and cannot diagnose two or more mixed faults. Therefore, the main research

problems in the future will focus on identifying unknown fault types and diagnosing mixed faults of rolling bearing.

7. Declarations

Funding: This study was funded by National Natural Science Foundation of China (61773001 and 61873012), Beijing Natural Science Foundation-Rail Traffic Control Science and Technology Joint Fund (L161003), Beijing Natural Science Foundation (4172023).

Conflict of Interest: Author Tingyu Jiang declares that he has no conflict of interest. Author Sheng Hong declares that he has no conflict of interest. Author Hao Liu declares that he has no conflict of interest.

Availability of data and material: All data generated or analysed during this study are included in this published article.

Code availability: The code that support the findings of this study are available on request from the corresponding author. The code are not publicly due to privacy or ethical restrictions.

Authors' contributions: All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Tingyu Jiang, Sheng Hong and Hao Liu. The first draft of the manuscript was written by Tingyu Jiang and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

References

- Dragomiretskiy K, Zosso D (2014) Variational mode decomposition. *IEEE Transactions on Signal Processing* 62(3):531-544
- Gu XJ, Chen CZ (2017) Rolling Bearing Fault Signal Extraction Based on Stochastic Resonance-Based Denoising and VMD. *International Journal of Rotating Machinery* 6:1-12 <https://doi.org/10.1155/2017/3595871>
- Gilles J (2013) Empirical Wavelet Transform. *IEEE Transactions on Signal Processing* 61(16):3999-4010
- Huang WT, Sun HJ, Luo NJ, Wang WJ (2019) Periodic feature oriented adapted dictionary free OMP for rolling element bearing incipient fault diagnosis. *Mechanical Systems and Signal Processing* 126:137-160
- Huang BG, Zhu QY, Siew CK (2016) Extreme learning machine: Theory and applications. *Neurocomputing* 70(1):489-501 <https://doi.org/10.1016/j.neucom.2005.12.126>
- Huang BG (2014) An insight into extreme learning machines: random neurons, random features and kernels. *Cognitive Computation* 6(3):376-390
- He Q, Wang XX, Zhou Q (2014) Vibration sensor data denoising using a time-frequency manifold for machinery fault diagnosis. *Sensors* 14(1):382-402
- Jia HC, Pan DH, Yuan Y, Zhang WC (2015) Using a BP neural network for rapid assessment of populations with difficulties accessing drinking water because of drought. *Human and Ecological Risk Assessment: An International Journal* 21(1):100-116 <https://doi.org/10.1080/10807039.2013.879025>
- Li JM, Yao XF, Wang H, Zhang JF (2019) Periodic impulses extraction based on improved adaptive VMD and sparse code shrinkage denoising and its application in rotating machinery fault diagnosis. *Mechanical Systems and Signal Processing* 126:568-589
- Li K, Su L, Wu JJ, Wang HQ, Chen P (2017a) A rolling bearing fault diagnosis method based on variational mode decomposition and an improved kernel

- nel extreme learning machine. *Applied Sciences-Basel* 7(10):1004 <https://doi.org/10.3390/app7101004>
- Li K, Wu J, Wang HQ, Chen P (2017b) A rolling bearing fault diagnosis method based on variational mode decomposition and an improved kernel extreme learning machine. *Applied Sciences* 7(10):1004
- Li YB, Xu MQ, Zhao HY, Wei Y, Huang WH (2015) A new rotating machinery fault diagnosis method based on improved local mean decomposition. *Digital Signal Processing* 46:201-214 <https://doi.org/10.1016/j.dsp.2015.07.001>
- Liu XF, B Lin, Luo HL (2015) Bearing faults diagnostics based on hybrid LS-SVM and EMD method. *Measurement* 59:145-166 <https://doi.org/10.1016/j.measurement.2014.09.037>
- Liang HL, Bressler SL, Desimone R, Fries P (2005) Empirical mode decomposition: a method for analyzing neural data. *Neurocomputing* 65-66:801-807 <https://doi.org/10.1016/j.neucom.2004.10.077>
- Lian RJ (2013) Adaptive Self-Organizing Fuzzy Sliding-Mode Radial Basis-Function Neural-Network Controller for Robotic Systems. *IEEE Transactions on Industrial Electronics* 61(3):1493-1503
- Learned RE (1992) Wavelet packet based transient signal classification. *Time-frequency & Time-scale Analysis*, IEEE-SP International Symposium.
- Meng L, Li CS, Zhang XY, Li RH, An XL (2016) Compound feature selection and parameter optimization of ELM for fault diagnosis of rolling element bearings. *Isa Transactions* 65:556-566 <https://doi.org/10.1016/j.isatra.2016.08.022>
- Madalena C, Goldberger AL, Peng CK (2007) Multiscale entropy analysis of complex physiologic time series. *Physical Review Letters* 89(6):705-708
- Madalena C, Goldberger AL, Peng CK (2005) Multiscale entropy analysis of biological signals. *Physical Review E Statistical Nonlinear & Soft Matter Physics* 71:021906

- Qiu JB, Ji WQ, Rudas IJ, Gao HJ (2020a) Asynchronous Sampled-Data Filtering Design for Fuzzy-Affine-Model-Based Stochastic. *IEEE Transactions on Cybernetics*:1-11
- Qiu JB, Ji WQ, Lam HK, Wang M (2020b) Fuzzy-Affine-Model Based Sampled-Data Filtering Design for Stochastic Nonlinear Systems. *IEEE Transactions on Fuzzy Systems*:1-13
- Qin B, Sun GD, Wang JG, Hu J (2017) Fault Features Extraction and Identification based Rolling Bearing Fault Diagnosis. 12th International Conference on Damage Assessment of Structures (pp. 1-14)
- Richman JS, Moorman JR (2000) Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology* 278(6):2039-2049
- Shao HD, Jiang JK, Wang F, Wang YN (2017) Rolling bearing fault diagnosis using adaptive deep belief network with dual-tree complex wavelet packet. *ISA Transactions* 69:187-201 <https://doi.org/10.1016/j.isatra.2017.03.017>
- Wang Y, Xu GH, Liang L, Jiang KS (2015) Detection of weak transient signals based on wavelet packet transform and manifold learning for rolling element bearing fault diagnosis. *Mechanical Systems and Signal Processing* 54-55:259-276
- Wang YX, Markert R, Xiang JW, Zheng WG (2015) Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system. *Mechanical Systems and Signal Processing* 60-61:243-251 <https://doi.org/10.1016/j.ymssp.2015.02.020>
- Xu B, Zhou FX, Li HP, Yan BK (2019) Early fault feature extraction of bearings based on Teager energy operator and optimal VMD. *ISA Transactions* 86:249-265

- Yan XA, Jia MP, Zhang W, Zhu L (2018) Fault diagnosis of rolling element bearing using a new optimal scale morphology analysis method. *Isa Transactions* 73:165-180 <https://doi.org/10.1016/j.isatra.2018.01.004>
- Yao DC, Yang JW, Pang ZF, Nie CM, Wen Fang (2018) Railway axle box bearing fault identification using LCD-MPE and ELM-AdaBoost. *Journal of Vibration Engineering* 20(1):165-174 <https://doi.org/10.21595/jve.2017.18502>
- Yao DC, Yang JW, Bai YL (2016) Railway rolling bearing fault diagnosis based on multi-scale intrinsic mode function permutation entropy and extreme learning machine classifier. *Advances in Mechanical Engineering* 54:168-176 <https://doi.org/10.1177/1687814016676157>
- Zhang XX, Zhou GW, Li DD (2019) Application of variational mode decomposition and permutation entropy for rolling bearing fault diagnosis. *International journal of acoustics and vibration* 24:303-311
- Zhang W, Li CH, Peng GL, Chen YH, Zhang ZJ (2018) A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load. *Mechanical Systems and Signal Processing* 100:439-453 <https://doi.org/10.1016/j.ymssp.2017.06.022>

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Thadataoffeatureextraction.xlsx](#)