Multimedia Tools and Applications (2020) 79:30069–30085
https://doi.org/10.1007/s11042-020-09534-w

A fault diagnosis method of rolling bearing based on VMD Tsallis entropy and FCM clustering

Xing Ting-ting[1,2] · Zeng Yan[2] · Meng Zong[1] · Guo Xiao-lin[1]

Received: 11 November 2019 / Revised: 29 July 2020 / Accepted: 4 August 2020 /
Published online: 13 August 2020

Abstract
A new fault diagnosis method of rolling bearings was presented based on variational mode decomposition (VMD), Tsallis entropy and Fuzzy C-means clustering (FCM) algorithm. Firstly, the measured vibration signals were decomposed with VMD in different scales to obtain a series of band-limited intrinsic modal function (BIMF). The VMD parameters were determined according to the change of the BIMF center frequency. Then, the Tsallis entropy of BIMF components were calculated and used as the signal features. Finally, the features were put into FCM classifier to recognize different fault types. It is proved by experiments that this method is feasible and the proposed approach could obtain better result compared with the method based on mode decomposition (EMD) and local mean decomposition (LMD).

Keywords Variational mode decomposition (VMD) · Tsallis entropy · Fuzzy C-means clustering (FCM) algorithm · Fault diagnosis · Rolling bearing

1 Introduction

Rolling bearing is an important component of rotation machinery, its operation directly affects the working condition of the whole mechanical equipment. The bearing failure will cause huge security risks in the manufacturing process. Therefore, it has a very import significant for on-line monitoring and fault-diagnosis of rolling bearings [2, 28]. This is why fault diagnosis of rolling bearing becomes a research focus, and many the vibration analysis methodologies have been proposed. When a rolling bearing fails, the collision occurs between the faulty part and other components, and non-stationary, non-linear shock signals can be obtained from the sensor installed on the device. This is also the basic principle of these analysis methodologies.

✉ Zeng Yan
zengyan3925@163.com

[1] Key Laboratory of Measurement Technology and Instrumention of HeBei Province, Yanshan University, Qinhuangdao 066004 Hebei, China

[2] Tangshan Polytechnic College, Tangshan 063299 Hebei, China

Most methods include two typical steps: feature extraction and selection, condition classification. In the first step, time domain, frequency domain, and time-frequency domain analysis are often applied [21]. The extracted time domain features like peak-to-peak value, root mean square value, kurtosis indicator, etc. obtained from the raw signals can be used, but some information are not easily observed. Frequency domain analysis could solve this problem which conduct FFT on the raw vibration signal, then analyzing the power spectrum, kurtosis spectrum, order cepstrum, envelope spectrum, etc. for diagnosis [5, 11, 32]. However, frequency domain analysis has limited analytical capabilities for non-stationary signals and this is why time-frequency domain analysis have been used for non-stationary signals diagnosis [8]. No matter which domain features are adopted, most of the features extracted are redundant. It is need to choose the typical information used the method like PCA [36], Kullback Leibler (K–L) divergence [3], distance evaluation technique [18], feature discriminant analysis, and compressed sensing [1, 12].

In addition, there are other transform domain analysis which are proved to be effective. Especially, Empirical Mode Decomposition (EMD) and Local Mean Decomposition (LMD) are widely used in feature extraction [22, 23, 31, 41], but the decomposition error is lager and decomposition result is susceptible to sampling frequency in these two feature extraction methods. Because EMD and LMD both are the recursive modal decomposition, modal aliasing is existed which make it difficult to separate the components with similar frequency, and the end effect has also appeared. Compared with EMD and LMD, VMD shows great advantages in bearing fault diagnosis [37, 42] . VMD determines the frequency center and bandwidth of each decomposition mode by iteratively searching for the optimal model. It is the non-recursive and variational modal decomposition, which could avoid modal aliasing and successfully separate two pure harmonic signals with similar frequencies. It has the characteristics of high precision and fast convergence and shows good robustness to noise. Therefore, VMD is used in this paper, and the Tsallis entropy of multiple modalities are calculated as the signal features.

In the second step, some artifificial intelligence algorithms have been proposed, like support vector machine(SVM) [38], artificial neural network(ANN) [16], random forest [40], etc. In addition, there are some algorithms constructed based on the objects [43]. Most of the algorithms can be used for rolling bearing diagnosis, but they are heavily dependent on the features extraction and much signal prior knowledge.

With the development of deep learning, the concept of building a network is applied in many aspects, which includes Image-Text Matching [10, 15], Social Multimedia [13, 14, 35], fault diagnosis [6, 7, 39]. At the same time, there are many deep learning-based bearing diagnosis methods proposed [29, 30]. Especially CNN has excellent feature extraction ability. Attention mechanism is introduced which could assist the deep network to extract the discriminative features and visualize the learned diagnosis knowledge effectively under the condition that there is only a small data set [19]. In addition, The multi-layers is also utilized as the main architecture of the fault diagnosis, and is proved efficient [20]. However, most of method based on deep learning have a long training time for the learning model, and can't check run process which is not flexible. Therefore, two steps: feature extraction and selection, condition classification are adopted considering the actual operation of bearing signals. FCM, which has been proved convergent [24–27], is used for classification after the feature extraction .

Specifically, the method proposed in this paper is: 1) adopt VMD to decompose the obtained raw signal. 2) The obtained Tsallis entropy after signals decomposition with VMD

are used as the signal features considering Tsallis entropy can solve the non-extensive problem of the system. 3) Fuzzy c-means clustering algorithm (FCM) is applied for better diagnosis. Next, the remainder of this paper is organized as follows. In Section 2, the feature extraction is presented. Then the FCM is described in Section 3. The Overall framework is presented in Section 4. The experimental results are shown and discussed in Section 5, and the conclusion is then presented in Section 6.

2 Feature extraction

In the feature extraction, the obtained vibration signal is decomposed with VMD firstly to get a series of band-limited intrinsic modal function (BIMF). Then, the Tsallis entropy of BIMF components were calculated and used as the signal features. Specially, the VMD and Tsallis entropy calculation are described as follows.

2.1 VMD

VMD is a non-recursive decomposition method, which can decompose multi-component signals of complex signals into amplitude-frequency modulation (AM-FM) component signals. The basic process of this algorithm are: assuming that each eigenmode function has a limited bandwidth with different center frequencies firstly, then the variational problem is solved by conversion, and each eigenmode function is demodulated to Corresponding base frequency band in order to minimize the sum of the estimated bandwidth of each eigenmode function, finally extract each eigenmode function and corresponding center frequency.

Decompose a real signal $f(t)$ into K sparse and independent sub-signals, its AM and FM signal form can be defined as:

$$u_k(t) = A_k(t)\cos[\varphi_k(t)] \tag{1}$$

$u_k(t)$ is the K IMF components obtained by VMD decomposition of signal $f(t)$, $\{u_k(t)\} = \{u_1(t), u_2(t), \cdots, u_K(t)\}$, $(k = 1, 2, \cdots K)$. $\varphi_k(t)$ is a non-monotonically decreasing phase function and $\varphi_k^{'}(t) \geq 0$,$A_k(t)$ is the instantaneous amplitude of $u_k(t)$ (envelope) which satisfies$A_k(t) \geq 0$.

The instantaneous frequency of $u_k(t)$ is

$$\omega_k(t) = \varphi_k^{'}(t) = \frac{d\varphi_k(t)}{d(t)} \tag{2}$$

Obviously, $A_k(t)$ and $\omega_k(t)$ are gradually changing relative to $\varphi_k(t)$, that is, within the interval of $[t-\delta, t+\delta]$ (where $\delta = 2\pi/\varphi'(t)$), $u_k(t)$ can be regarded as a harmonic signal with amplitude $A_k(t)$ and frequency $\omega_k(t)$.

Here, assume that each mode of the signal has a limited bandwidth with a center frequency, variational problems can be described as seeking k modal functions $u_k(t)$ so that the sum of the estimated bandwidth of each mode is the smallest and the constraint is the sum of each mode is the original input signal $f(t)$.

Specifically, the analytical signal of each modal function $u_k(t)$ is obtained through the Hilbert transform, and then its unilateral frequency spectrum can be obtained:

$$\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t) \tag{3}$$

Where $n \leftarrow 0$ is a unit pulse function, j is an imaginary unit, and * is convolution.

Then the analysis signal of each mode is added an estimated center frequency $d_{ij} = \|x_j - v_i\|$, the spectrum of each mode is modulated to the corresponding base band:

$$\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t} \tag{4}$$

Where $\{\omega_k\} = \{\omega_1, \omega_2, \cdots, \omega_K\}$, $(k = 1, 2, \cdots K)$ is the center frequency of each $\{u_k(t)\}$.

Calculate the squared L^2 norm of the demodulated signal gradient to estimate the bandwidth of each modal signal. The variation problem is expressed as follows:

$$\min_{\{u_k\}\cdot\{\omega_k\}}\left\{\sum_{k=1}^{K}\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}\right\|_2^2\right\}, \qquad s.t.\ \sum_{k=1}^{K}u_k(t) = f(t) \tag{5}$$

To find the optimal solution of the above constrained variational model, transform the constrained variational problem to be solved into a non-constrained variational problem by introducing quadratic penalty factor and Lagrange operator. And the extended Lagrangian function is:

$$L(\{u_k\},\{\omega_k\},\lambda) = \alpha\sum_{k=1}^{K}\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}\right\|_2^2 + \left\|f(t)-\sum_{k=1}^{K}u_k(t)\right\|_2^2 + \left\langle\lambda(t), f(t)-\sum_{k=1}^{K}u_k(t)\right\rangle \tag{6}$$

Where α is the second penalty factor, it guarantees the reconstruction accuracy of the signal in the presence of Gaussian noise, $\lambda(t)$ is Lagrange operator and keeps the constraints strictly, $\langle\cdot, \cdot\rangle$ represents inner product.

Next, Alternate Direction Method of Multipliers(ADMM) is adopted. Seek the "saddle point" of the Lagrange expression by alternately updating u_k^{n+1}, ω_k^{n+1}, andλ^{n+1}.

The problem of solving u_k^{n+1} can be expressed as:

$$u_k^{n+1} = \underset{u_k\in X}{\operatorname{argmin}}\left\{\alpha\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}\right\|_2^2 + \left\|f(t)-\sum_{i=1}^{K}u_i(t)+\frac{\lambda(t)}{2}\right\|_2^2\right\} \tag{7}$$

Where X is the solution space of u_k. Using the Parseval/Plancherel Fourier isometric method to solve this problem in the frequency domain.

$$u_k^{n+1}(\omega) = \underset{u_k,u_k\in X}{\operatorname{argmin}}\left\{\alpha\|j\omega[(1+\operatorname{sgn}(\omega+\omega_k))u_k(\omega+\omega_k)]\|_2^2 + \left\|f(\omega)-\sum_{i=1}^{K}u_i(\omega)+\frac{\lambda(\omega)}{2}\right\|_2^2\right\} \tag{8}$$

Where u, f, λ is the Fourier transform of the corresponding time signal respectively. In the first term of Eq. (8), the variable$\omega \leftarrow \omega - \omega_k$.

$$u_k^{n+1}(\omega) = \underset{u_k, u_k \in X}{\operatorname{argmin}} \left\{ \alpha \| j(\omega\text{-}\omega_k)[(1+\operatorname{sgn}(\omega))u_k(\omega)] \|_2^2 + \left\| f(\omega) - \sum_{i=1}^{K} u_i(\omega) + \frac{\lambda(\omega)}{2} \right\|_2^2 \right\} \quad (9)$$

Using Hermitian symmetry of the real signal in the reconstructed fidelity term, these two terms can be written as half-space integrals at non-negative frequencies.

$$u_k^{n+1}(\omega) = \underset{u_k, u_k \in X}{\operatorname{argmin}} \left\{ \int_0^{\infty} \left[4\alpha(\omega\text{-}\omega_k)^2 |u_k(\omega)|^2 + 2 \left| f(\omega) - \sum_{i=1}^{K} u_i(\omega) + \frac{\lambda(\omega)}{2} \right|^2 \right] \mathrm{d}\omega \right\} \quad (10)$$

The solution to this quadratic optimization problem is:

$$u_k^{n+1}(\omega) = \frac{f(\omega) - \sum_{i \neq k} u_i(\omega) + \frac{\lambda(\omega)}{2}}{1 + 2\alpha(\omega\text{-}\omega_k)^2} \quad (11)$$

In addition, because the center frequency ω_k only appears in the low-frequency bandwidth. It can be expressed:

$$\omega_k^{n+1} = \underset{\omega_k}{\operatorname{argmin}} \left\{ \left\| \partial_t \left[\left(\delta(t) \frac{j}{\pi t} \right) * u_k(t) \right] e^{-jw_k t} \right\|_2^2 \right\} \quad (12)$$

And get

$$\omega_k^{n+1} = \frac{\int_0^{\infty} \omega |u_k(\omega)|^2 \mathrm{d}\omega}{\int_0^{\infty} |u_k(\omega)|^2 \mathrm{d}\omega} \quad (13)$$

Obviously, ω_k is at the center of gravity of the corresponding modal power spectrum.

Iterate $u_k(\omega)$ and ω_k using Eqs. 11 and 14 to get the optimal solution. Generally, the termination criterion of iteration number n satisfies:

$$\frac{\sum_{k=1}^{K} \left\| u_k^{n-1} - u_k^n \right\|_2^2}{\sum_{k=1}^{K} \left\| u_k^n \right\|_2^2} < e \quad (14)$$

Where $e(e > 0)$ is the convergence constraint of fixed precision.

According to the above description, the specific process of VMD algorithm is as follows

(1) Initialize $\{u_k^1(\omega)\}$, $\{\omega_k^1\}$, $\lambda^1(\omega)$, $\widehat{f}$;
(2) Repeat

$\widehat{\lambda}(\omega)$ for $\widehat{\lambda}^{n+1}(\omega) \leftarrow \widehat{\lambda}^n(\omega) + \gamma \left(\widehat{f}(\omega) - \sum_{k=1}^{K} \widehat{u}_k^{n+1}(\omega) \right)$ do

(1) Update γ for all $\omega \geq 0$:

$$u_k^{n+1}(\omega) \leftarrow \frac{f(\omega) - \sum_{i<k} u_i^{n+1}(\omega) - \sum_{i>k} u_i^n(\omega) + \frac{\lambda^n(\omega)}{2}}{1 + 2\alpha\left(\omega - \omega_k^n\right)^2} \tag{15}$$

The quadratic penalty factor α improves convergence. Especially when the signal contains noise, the Lagrange multiplier using the quadratic penalty function effectively approximates the precise reconstruction of the signal.

(2) Update $\sum_{k=1}^{K} \left(\left\| \widehat{u}_k^{n+1} - \widehat{u}_k^n \right\|_2^2 / \left\| \widehat{u}_k^n \right\|_2^2 \right) < e$:

$$\omega_k^{n+1} \leftarrow \frac{\int_0^\infty \omega \left| u_k^{n+1}(\omega) \right|^2 d\omega}{\int_0^\infty \left| u_k^{n+1}(\omega) \right|^2 d\omega} \tag{16}$$

Where $\widehat{u}_k$ is the modal function in the frequency domain; $\widehat{\lambda}$ represents the Lagrange multiplier operator in frequency domain and plays a mandatory role; $\widehat{f}$ represents the original signal in frequency domain.

(3) Dual ascent for all $\omega \geq 0$:

$$\lambda^{n+1}(\omega) \leftarrow \lambda^n(\omega) + \gamma\left(f(\omega) - \sum_{k=1}^{K} u_k^{n+1}(\omega) \right) \tag{17}$$

In this formula, γ represents noise tolerance coefficient. To achieve good de-noising effect, it can be set: $\gamma = 0$.

until convergence: $\sum_{k=1}^{K} \left(\left\| u_k^{n+1} - u_k^n \right\|_2^2 / \left\| u_k^n \right\|_2^2 \right) < e$. At the end of the iteration, K components are output.

It is need to be noted that the decomposition layers K has an effect on the decomposition results, the specific impact has been explained in Reference [17] and the optimal number of decomposition layers for rolling bearing diagnosis has been proved in the experiment of this paper.

2.2 Tsallis entropy

The concept of Shannon entropy was first proposed by American scholar C.E. Shannon in 1948. The theory states that if an event has multiple possible outcomes and the probability of each outcome is $p_i (i = 1, 2, \cdots, N)$, the information obtained by a certain result can be expressed by $I_i = \log_\alpha(1/p_i)$, and the information entropy defined for time series is

$$S_{BG}^{(d)} = -k\sum_{i=1}^{N} p_i \ln p_i \tag{18}$$

Where $k = 1$. Obviously, Shannon entropy is based on thermodynamic B-G entropy, and it is extensive.

Tsallis entropy introduces non-extensive parameter q on the basis of Shannon entropy and constructs a new form of entropy function. It can be expressed:

$$S_q^{(d)} = \frac{k}{q-1}\left(1-\int f(x)^q \mathrm{d}x\right), q \in R \tag{19}$$

Where $f(x)$is the probability density distribution function which satisfies $\int f(x)\mathrm{d}x = 1$, and q is the non-extensive parameter.

In addition, Tsallis entropy can be expressed discretely:

$$S_q^{(d)} = \frac{k}{q-1}\left(1-\sum_{i=1}^{n}\left(p_i^q\right)\right), q \in R \tag{20}$$

Where p_i is the probability density distribution function of random variables i,k is a constant. In this paper $k = 1, \sum_{i=1}^{n}\left(p_i^q\right) = 1$.

The selection of the non-extensive coefficient q of different tested systems has a great significance to the calculation of Tsallis entropy, q can describe the non-extensive degree of the test system, and make system entropy meets the following pseudo-additivity:

$$\frac{S_s(A+B)}{k} = \frac{S_s(A)}{k} + \frac{S_s(B)}{k} + (1-q)\frac{S_s(A)S_s(B)}{k^2} \tag{21}$$

Therefore, q makes information measurement more targeted and flexible. $q < 1$ and $q > 1$ denote the system's specific super-extendability and sub-extensibility, respectively. Especially $q \to 1$, Tsallis entropy is equivalent to Shannon entropy which be proved in below formula. Therefore, Tsallis entropy which is the extension of Shannon entropy also can describe systems with extensive characteristics, and it is often used in the analysis of random complex signals.

$$\lim S_q^{(d)} = \lim_{q\to 1}\frac{k}{q-1}\left(1-\sum_{i=1}^{n} p(i)^q\right) = \lim_{q\to 1}\frac{k}{q-1}\left(\sum_{i=1}^{n} p(i)\left(1-p(i)^{q-1}\right)\right) = -k\sum_{i=1}^{n} p(i)\ln p(i) = S_{BG}^{d} \tag{22}$$

In this paper, Tsallis entropy is suitable due to the randomness of vibration signal from the rolling bearing fault. After the vibration signal is decomposed by VMD, k eigenmode functions are obtained. Then choose the appropriate non-extended parameter q to calculate the Tsallis entropy of each eigenmode function. The features of the fault information of the signal can be distinguished according to the change of entropy value [9, 33, 34].

3 FCM

FCM algorithm is a kind of partition-based clustering algorithm. It is an improvement of the classic C-means algorithm. The principle of FCM algorithm is to maximize the similarity between the objects that are divided into the same cluster and minimize the similarity between

the objects of different cluster. In this process, it is need to minimize the Euclidean distance between all data points and each cluster center, and the weighted sum of fuzzy membership firstly, then correct the fuzzy classification matrices and cluster centers continuously until the convergence constraints for a given precision are met. Lastly, clustering the data points with similarity.

Assume the sample set is $X=\{x_1, x_2, \cdots x_n\}$, where n is the number of samples. The cluster center vector $V=[v_1, v_2, \cdots, v_c]^T$, where c is the number of cluster centers. The fuzzy classification matrix is $U=[u_{ij}]_{c\times n}$, where u_{ij} is the membership degree of the data point $x_j$relative to the cluster center v_i. The clustering objective function is

$$J_{fcm}(U,V)=\sum_{j=1}^{n}\sum_{i=1}^{c}u_{ij}^{m}d_{ij}^{2} \tag{23}$$

Where d_{ij} is the Euclidean distance from the data point x_j to the cluster center v_i, it can be expressed as $d_{ij}=\|x_j-v_i\|$. The parameter m is a fuzzy weighted index, generally $m=2$. In addition, introducing the following constraints in FCM algorithm so as to find the smallest partition of the objective function though calculating U and V iteratively under the constraints.

$$\begin{cases} 0\le u_{ij}\le 1 \\ \sum_{i=1}^{c}u_{ij}=1 \\ \sum_{j=1}^{n}u_{ij}>0 \end{cases} \quad 1\le i\le c,\quad 1\le j\le n \tag{24}$$

Specific steps:

1) Set the number of cluster centers c, precision$\varepsilon(\varepsilon>0)$ and the fuzzy weighted index m, initialize the fuzzy classification matrix, and set the iteration number $l=0$.

(Update)v_i

$$v_i=\sum_{j=1}^{n}u_{ij}^{m}x_j/\sum_{j=1}^{n}u_{ij}^{m} \tag{25}$$

(Update)U

$$u_{ij}=1/\sum_{k=1}^{c}\left(\frac{d_{ij}}{d_{kj}}\right)^{2/(m-1)} \tag{26}$$

2) Determine whether U satisfies the constraint:

$$\left\|U^{l+1}-U^{l}\right\|<\varepsilon \tag{27}$$

If the constraint is satisfied, stop iteration, otherwise repeat the step (2) and (3) to get the optimal result.

In addition, the effect of clustering can be evaluated by the classification coefficient F and the average fuzzy entropy H. The more the classification coefficient Ftends to 1, the more the average fuzzy entropy$H$tends to 0, the better the clustering effect.

$$F = \frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{c} u_{ij}^2 \tag{28}$$

$$H = -\frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{c} u_{ij}\ln u_{ij} \tag{29}$$

4 Our methodology

The fault diagnosis method of rolling bearing in this paper is described in Fig. 1: (1) Collect the vibration signal, set the second penalty factor α, the decomposition level$K$, and perform VMD decomposition on the vibration signal. (2) Through continuous optimization iterations, when the parameters meet the convergence constraint of a given precision $e(e>0)$, K BIMF components are output.(3) Set the non-extensive parameter q, find the Tsallis entropy of each BIMF function, and get the feature entropy value. (4) Perform FCM cluster analysis on the entropy value to determine the fault type of the vibration signal.

5 Experiment

In this paper, the Western Reserve University bearing test bench data are used for experiments [4]. The bearing test bench is shown in Fig. 2. The platform consists of a 1.5W motor, a torque sensor/decoder, a power test meter and an electronic controller.

Specifically, the experimental data come from the drive end bearing whose model is 6205-2RS JEM SKF deep groove ball bearing. The bearing inner ring diameter is 25mm, the outer ring diameter is 52mm, the thickness is 15mm, the rolling element diameter is 7.94mm and the pitch diameter It is 39.04mm. The rolling bearing fault is caused by artificial damage to the bearing by EDM. Then acquiring the vibration signal through accelerometers which are mounted in the motor housing to get the vibration signals under different faults, different speeds, and different load conditions. Lastly, analyzing the vibration signal to get whether there is one or more faults on the bearing. In addition, the vibration signal is collected by a 16-channel data recorder, and the power and speed are measured by a torque sensor/decoder.

In order to verify the effectiveness of this paper's method, there were two cases in the experiment (1) Research on different types of fault diagnosis for the bearing with same shaft diameter; (2) Research on the same type of fault diagnosis for the bearing with different shaft diameter.

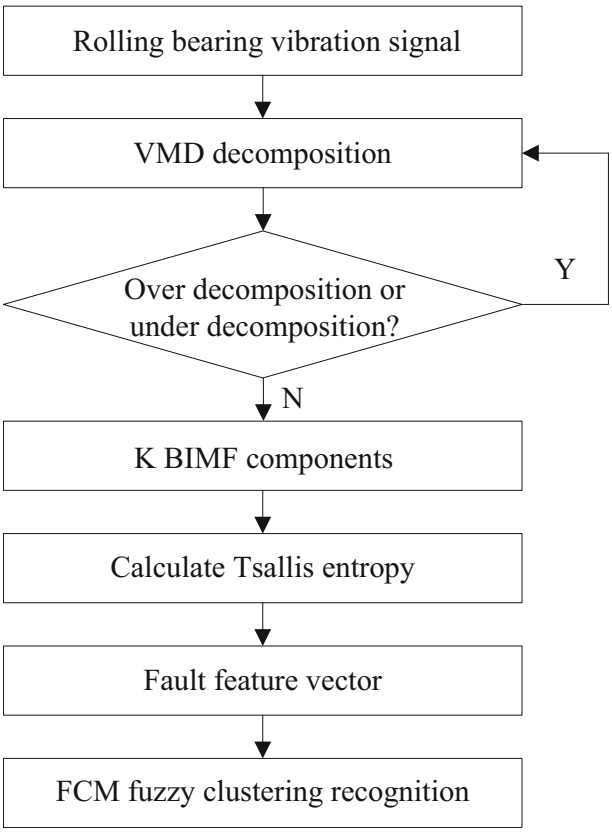

Fig. 1 Flow diagram of rolling bearing fault diagnosis method in this paper

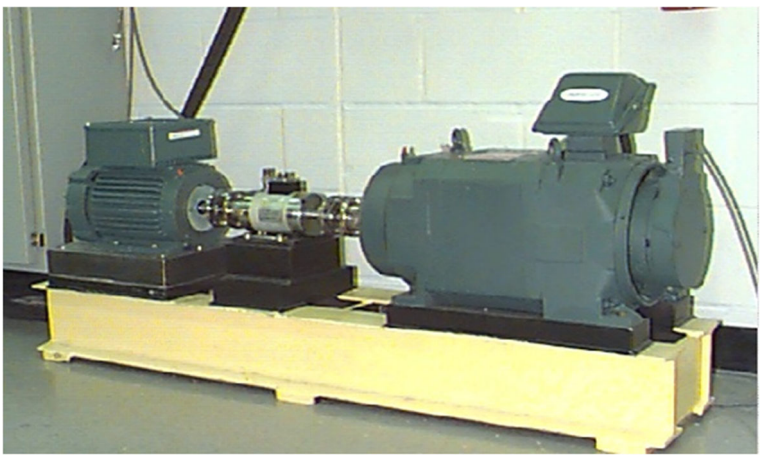

Fig. 2 Bearing experiment platform of Western Reserve University

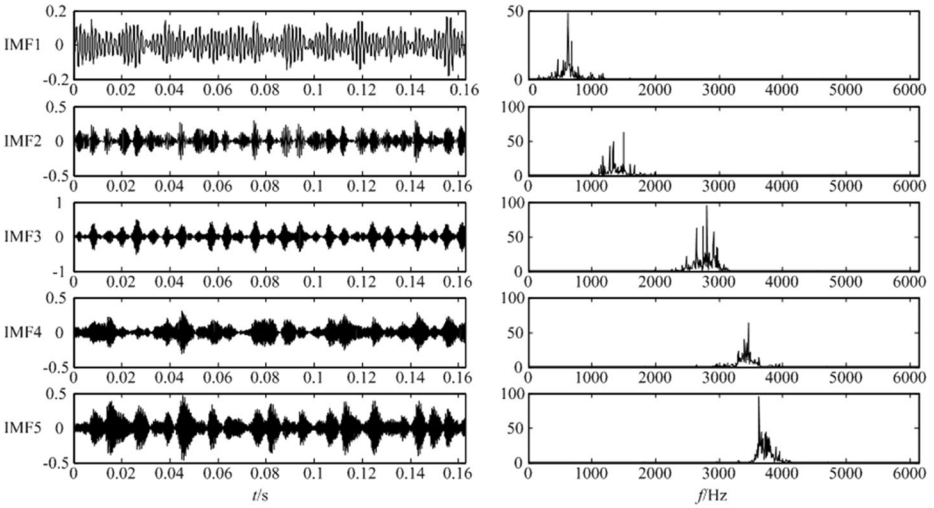

Fig. 3 The VMD decomposition result of signal with inner race fault when $K = 5$

5.1 Different types of fault diagnosis for the bearing with same shaft diameter

In this case, the chosen shaft damaged diameter is 0.1778mm, the speed is 1772r/min, and the sampling frequency is 12kHz. In addition, The bearing has four status which are normal(NO), inner race(IR), outer race(OR) and rolling element (RE). In order to obtain a better diagnostic effect, the VMD parameters are determined experimentally in this section firstly, and then the effectiveness of the proposed diagnosis method is verified.

(1) Parameter determination

When the original signal is decomposed based on VMD, the scale value K needs to be preset. The scale value K will affect the decomposition result, which in turn affects the feature extraction result and diagnosis result. Therefore, it is necessary to set the appropriate K value and prevent under-decomposition or over-decomposition. Here, a set of sample data with inner race faults is tested, and the length of sample data is 4096. Set the appropriate K by observing the center frequency of the signal decomposed at different K values. When $K = 5$, the sequence

Table 1 The BIMF components center frequencies of signal with inner race fault at different K values

Scale	BIMF1	BIMF2	BIMF3	BIMF4
$K = 3$	620.54	2807.8	3625	–
$K = 4$	620.54	1499.1	2807.8	3625
$K = 5$	620.54	1499.1	2807.8	3459.1
$K = 6$	620.54	1499.1	2641.9	2912.3
$K = 7$	620.54	1278	1499.1	2807.8
$K = 8$	620.54	1278	1499.1	2641.9
Scale	BIMF5	BIMF6	BIMF7	BIMF8
$K = 5$	3625	–	–	–
$K = 6$	3459.1	3625	–	–
$K = 7$	3459.1	3625	3784.7	–
$K = 8$	2912.3	3459.1	3625	3784.7

Table 2 The BIMF components center frequencies of signal with three different faults at different K

Fault type	Scale	BIMF1	BIMF2	BIMF3	BIMF4	BIMF5
NO	$K=3$	92.16	2150.4	1093.6		
	$K=4$	92.16	1093.6	2150.4	5007.4	
	$K=5$	92.16	1093.6	2150.4	3072	5007.4
RE	$K=3$	546.82	2967.6	3446.8		
	$K=4$	368.64	1148.9	2967.6	3446.8	
	$K=5$	368.64	1148.9	2746.4	3299.3	3563.5
OR	$K=3$	1118.2	2936.8	3483.6		
	$K=4$	1118.2	2936.8	3483.6	5363.7	
	$K=5$	737.28	1118.2	2936.8	3483.6	5363.7

diagram and the spectrogram of each component are shown in Fig. 3. When K is taken different values, the center frequency of each BIMF component are shown in Table 1.

It can been seen from Table 1 that when $K>4$, the center frequencies of different BIMF components change little. Especially when $K=5$, the center frequencies of BIMF4 and BIMF5 are similar. This shows that when $K>4$, the signal is over-decomposed. At the same time, when $K<4$, the signal is under-decomposition. The frequency 1499.1Hz signal is missing when $K=3$.Therefore, the best VMD decomposed scale $K=4$ for the signal with the inner race fault .

Next, VMD decomposition on the signals with the other three types fault are performed at different K, the obtained center frequencies are shown in Table 2. Obviously, when $K=4$, neither over-decomposition nor under-decomposition exists in VMD decomposition. Hence, K is set to 4 in the following experiments.

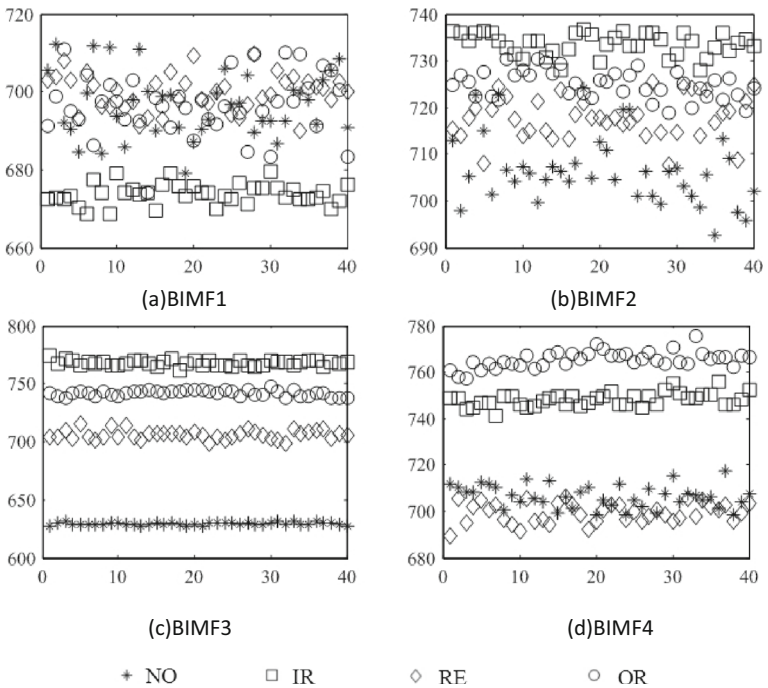

Fig. 4 The obtained Tsallis entropy with VMD decompsition of the signals with different faults

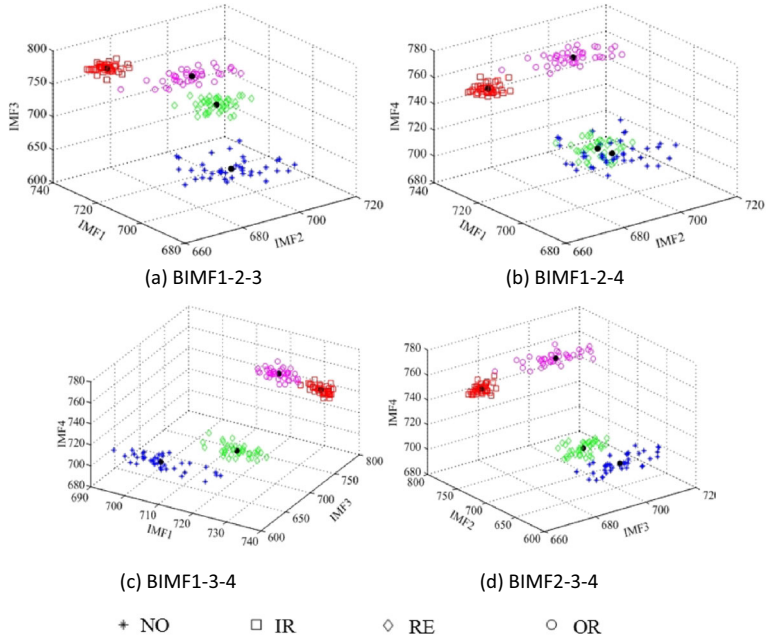

Fig. 5 The clustering results of different fault signals

(2) Fault Diagnosis

In this part, choosing 40 sets of signals as samples, and there are 2048 data per group. Set the decomposition scale K to 4 and perform VMD decomposition on the signal with different types of fault. Then calculate the Tsallis entropy of each decomposed components. The results are shown in Fig. 4.

According the obtained Tsallis entropy, a 160×4matrix can be constructed, which can be used as the feature in diagnosis. The FCM clustering results are shown in Fig. 5, and the center coordinates are shown in Table 3. Specially, the clustering center number $c = 4$, fuzzy weighted index $m = 2$, convergence precision $e = 0.001$.

To further test the diagnosis effect, calculate the classification coefficient and the average fuzzy entropy, and get $F = 0.95466$ and $H = 0.14579$. Obviously, F tends to 1 and H tends to 0. These results indicate that the FCM clustering result has good effect, and the proposed method that combining VMD, Tsallis entropy and FCM clustering is feasible in rolling bearing diagnosis .

Table 3 The clustering center coordinates of different fault signals

Fault type	v_1	v_2	v_3	v_4
NO	673.76	733.55	767.57	748.28
IR	696.49	706.49	629.27	706.52
RE	699.9	717.73	705.97	699.3
OR	697.01	724.9	741.74	765.27

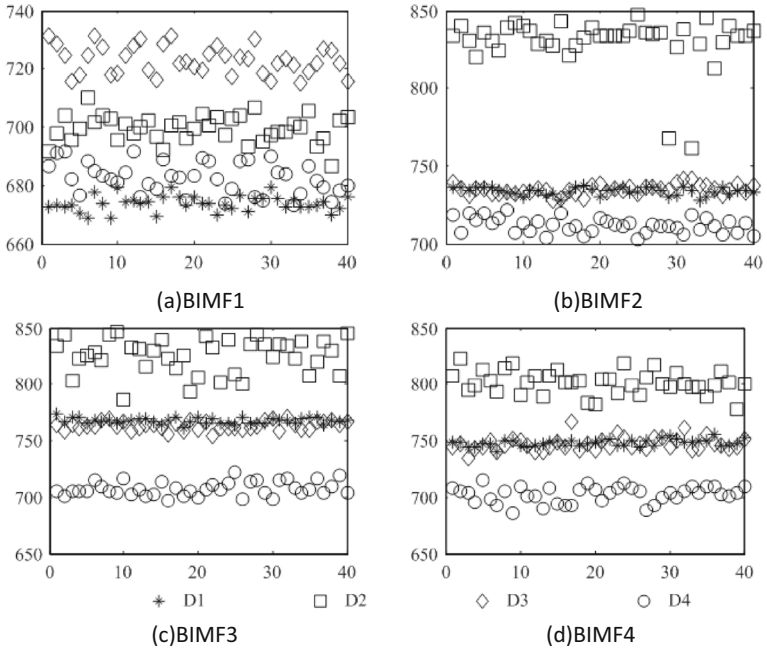

Fig. 6 The Tsallis entropy with VMD decompsition of the inner fault signals with different shaft diameter

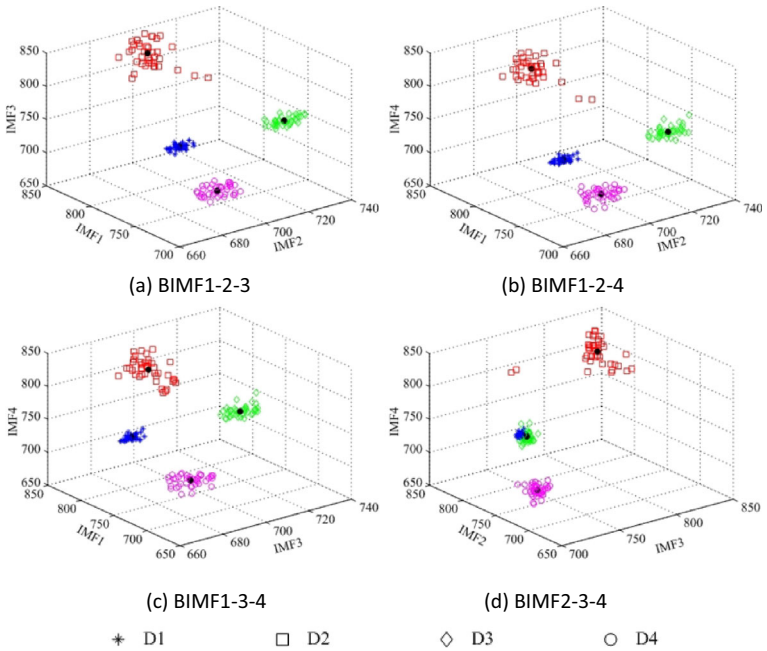

Fig. 7 The clustering results of inner fault signals with different shaft diameter

Table 4 The clustering center coordinates of inner fault signals with different shaft diameter

Diameter	v_1	v_2	v_3	v_4
D1	722.88	734.75	763.36	747.02
D2	699.49	832.52	825.92	802.23
D3	673.87	733.8	767.96	748.36
D4	682.58	711.73	707.46	702.33

5.2 Same type of fault diagnosis for the bearing with different shaft diameter

In this part, the rolling bearings with four different shaft damaged diameters were used for experiment, which are D1 = 0.1778mm, D2 = 0.3556mm, D3 = 0.5334mm, D4 = 0.7112mm respectively. The speed was 1772r/min, the sampling frequency was 12kHz, and only the inner race fault of rolling baring is tested. Similar to the first part, 40 sets of signals are chosen as samples, and there are 2048 data per group. K is still set to 4. The Tsallis entropy of each decomposed components shown in Fig. 6. The cluster result is shown in Fig. 7 and the cluster center coordinates are shown in Table 4(The parameters of FCM are the same as the first part).

Calculate the classification coefficient and the average fuzzy entropy, get $F = 0.94598$ and $H = 0.15506$. It is clear that the proposed method in this paper is also applicable for the fault bearing diagnosis with different shaft diameter.

In addition, for proving the superiority of the method, comparing this paper's method with another two methods, which are almost same to the above process, except the VMD is replaced by EMD (EMD + FCM) or LMD(LMD + FCM). The classification coefficient F and the average fuzzy entropy H obtained are shown in Table 5.

It can be seen that the classification coefficient F is greatest in the above both cases, and the average fuzzy entropy H is smallest. Obviously, the method proposed in this paper is more advantageous in fault diagnosis of rolling bearing .

6 Conclusion

In this paper, a new method for rolling bearing fault diagnosis is proposed, which apply VMD in signals decomposition, then use Tsallis entropy as the signal feature, lastly, combine FCM algorithm to diagnose. To verify the feasibility of the method, a series of experiments are preformed, the results are optimistic. Further, comparing with another methods which are EMD + FCM and LMD + FCM, it turns out that the method proposed in this paper is the best.

Table 5 The clustering effect of different methods

Condition	Parameter	EMD + FCM	LMD + FCM	this paper's method
Same shaft diameter, different faults	F	0.8988	0.9106	0.95466
	H	0.3608	0.2506	0.14579
Same fault, Different shaft diameter	F	0.8807	0.9024	0.94598
	H	0.3974	0.3016	0.15506

References

1. Ahmed HOA, Nandi AK (2019) Three-stage hybrid fault diagnosis for rolling bearings with compressively sampled data and subspace learning techniques. IEEE Trans Ind Electron 66(7):5516–5524
2. Akhand R, Upadhyay SH (2016) A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings. Tribology International 96:289–306
3. Brkovic A, Gajic D, Gligorijevic J, Savic-Gajic I, Georgieva O, Gennaro SD (2017) Early fault detection and diagnosis in bearings for more efficient operation of rotating machinery. Energy 136:63–71
4. Case Western Reserve University Bearing Data Center n.d.. [Online]. Avail-able: http://csegroup.case.edu/bearingdatacenter/home
5. Cerrada M, Sanchez RV, Li C, Pacheco F, Cabrera D, Oliveira JVD, Rafael EV (2018) A review on data-driven fault severity assessment in rolling bearings. Mechanical Systems & Signal Processing 99:169–196
6. Chellamuthu S, Sekaran EC (2019) Fault detection in electrical equipment's images by using optimal features with deep learning classifier. Multimed Tools Appl 78:27333–27350
7. Chen F, Fu Z, Zhen L (2019) Thermal power generation fault diagnosis and prediction model based on deep learning and multimedia systems. Multimed Tools Appl 78(4):4673–4692
8. Ding X, Li Q, Lin L, He Q, Shao Y (2019) Fast time-frequency manifold learning and its reconstruction for transient feature extraction in rotating machinery fault diagnosis. Measurement 141:380–395
9. Furuichi S, Yanagi K, Kuriyama K (2004) Fundamental properties of Tsallis relative entropy. J Math Phys 45(12):4868–4877
10. Garg S, Kaur K, Kumar N, Rodrigues JJPC (2019) Hybrid deep-learning-based anomaly detection scheme for suspicious flow detection in SDN: a social multimedia perspective. IEEE Transactions on Multimedia 21(3):566–578
11. Gu X, Yang S, Liu Y, Hao R (2016) Rolling element bearing faults diagnosis based on kurtogram and frequency domain correlated kurtosis. Meas Sci Technol 27(12):125019
12. Hu Z, Wang Y, Ge MF, Liu J (2020) Data-driven fault diagnosis method based on compressed sensing and improved multi-scale network. IEEE Trans Ind Electron 67(4):3216–3225
13. Huang F, Zhang X, Xu J, Zhao Z, Li Z (2019) Multimodal learning of social image representation by exploiting social relations. In IEEE Transactions on Cybernetics 99:1–13
14. Huang F, Zhang X, Zhao Z, Li Z (2019) Bi-directional spatial-semantic attention networks for image-text matching. IEEE Trans Image Process 28(4):2008–2020
15. Huang F, Zhang X, Zhao Z, Xu J, Li Z (2019) Image-text sentiment analysis via deep multimodal attentive fusion. Knowl-Based Syst 167:26–37
16. Kanai RA, Desavale R, Chavan SP (2016) Experimental-based fault diagnosis of rolling bearings using artificial neural network. Journal of Tribology 138(3):031103
17. Konstantin D, Dominique Z (2014) Variational mode decomposition. IEEE Trans Signal Process 62(3): 531–544
18. Li H, Wang W, Huang P, Li Q (2019) Fault diagnosis of rolling bearing using symmetrized dot pattern and density-based clustering. Measurement 152:107293
19. Li X, Zhang W, Ding Q (2019) Understanding and improving deep learning-based rolling bearing fault diagnosis with attention mechanism. Signal Process 161:136–154
20. Li X, Zhang W, Ding Q, Sun JQ (2019) Multi-layer domain adaptation method for rolling bearing fault diagnosis. Signal Process 157:180–197
21. Lu W, Jiawei X, Yi L (2019) Time-frequency-based maximum correlated kurtosis deconvolutionapproach for detecting bearing faults under variable speed conditions. Meas Sci Technol 30(12):125005
22. Meng Z, Gu W, Hu M, Xiong J (2016) Early weak fault feature extraction of rolling bearings based on improved singular value decomposition and empirical mode decomposition. Acta Metrologica Sinica 37(4): 406–410
23. Meng Z, Li S, Wang Y (2015) Rotating machinery fault diagnosis method based on LMD and local time-frequency entropy. Acta Metrologica Sinica 36(1):77–81
24. Omar AA (2015) Adaptation of reproducing kernel algorithm for solving fuzzy Fredholm-Volterra integrodifferential equations. Neural Comput & Applic 28:1–20
25. Omar AA, Mohammd AS (2020) Fuzzy conformable fractional differential equations: novel extended approach and new numerical solutions. Soft Comput 24:12501–12522
26. Omar AA, Mohammd AS, Momani S, Hayat T (2016) Numerical solutions of fuzzy differential equations using reproducing kernel Hilbert space method. Soft Comput 20(8):3283–3302
27. Omar AA, Mohammd AS, Momani S, Hayat T (2017) Application of reproducing kernel algorithm for solving second-order, two-point fuzzy boundary value problems. Soft Comput 21(23):7191–7206
28. Robert BR, Jwrôme A (2011) Rolling element bearing diagnostics - a tutorial. Mech Syst Signal Process 25(2):485–520

29. Shao H, Jiang H, Zhang H, Duan W, Liang T, Wu S (2018) Rolling bearing fault feature learning using improved convolutional deep belief network with compressed sensing. Mechanical systems and signal processing 100(FEB.1):743–765
30. Shao H, Jiang H, Zhang X, Niu M (2015) Rolling bearing fault diagnosis using an optimization deep belief network. Meas Sci Technol 26(11):115002
31. Shi P, Wang J, Wen J, Tian G (2016) Study on rotating machinery fault diagnosis method based on envelopes fitting algorithms EMD. Acta Metrologica Sinica 37(1):62–66
32. Tian J, Morillo C, Azarian MH, Pecht M (2016) Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with k-nearest neighbor distance analysis. IEEE Trans Ind Electron 63(3):1793–1803
33. Tsallis C (1988) Possible generalization of Boltzmann-Gibbs statistics. J Stat Phys 52(1–2):479–487
34. Tsallis C, Mendes RS, Plastino AR (1998) The role of constraints within generalized nonextensive statistics. Physica A 261(3):534–554
35. Wang S, Hu X, Yu PS, Li Z (2014). MMRate: Inferring multi-aspect diffusion networks with multi-pattern cascades. KDD '14: Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. Page:1246–1255
36. Wang F, Sun J, Yan D, Zhang S, Cui L, Xu Y (2015) A feature extraction method for fault classification of rolling bearing based on PCA. Journal of Physics Conference 628:012079
37. Wei D, Jiang H, Shao H, Li X, Lin Y (2019) An optimal variational mode decomposition for rolling bearing fault feature extraction. Meas Sci Technol 30(5):055004
38. Wu C, Chen T, Jiang R (2017) Bearing fault diagnosis via kernel matrix construction based support vector machine.Journal of. Vibroengineering 19(5):3445–3461
39. Xia M (2019) Multimedia based multi-fault diagnosis of satellite sensor based on gauss Bayesian algorithm. Multimed Tools Appl 78:22601–22611
40. Xu G, Liu M, Jiang Z, Söffker D, Shen W (2019) Bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning. Sensors 19(5):1088
41. Xu Y, Zhang K, Ma C, Li S, Zhang H (2019) Optimized LMD method and its applications in rolling bearing fault diagnosis. Meas Sci Technol 30(12):125017
42. Zan T, Pang Z , Wang M, Gao X (2018). Research on early fault diagnosis of rolling bearing based on VMD. 2018 6th international conference on mechanical, automotive and materials engineering (CMAME): pp. 41-45
43. Zhang X, Zhang Y, Wang S, Yao Y, Fang B, Yu PS (2018) Improving stock market prediction via heterogeneous information fusion. Knowl-Based Syst 143:236–247