A multi-scale feature extraction and fusion method for diagnosing bearing faults

Zhixiang Chen[1,#], Hang Wang[1,3,#,*], Yuanyuan Zhou[1], Yang Yang[2], Yongbin Liu[1,3]

1. School of Electrical Engineering and Automation, Anhui University, Hefei, China

2. China North Vehicle Research Institute, Beijing, China

3. Smart Grid Digital Collaborative Technology Joint Laboratory of Anhui Province, Hefei, China

# Zhixiang Chen and Hang Wang contributed equally to this work.

*Corresponding author: hangwang@ahu.edu.cn

Received Month X, XXXX | Accepted Month X, XXXX | Posted Online Month X, XXXX

Abstract: Bearing fault diagnosis is vital to safeguard the heath of rotating machinery. It can help to avoid economic losses and safe accidents in time. Effective feature extraction is the premise of diagnosing bearing faults. However, effective features characterizing the health status of bearings are difficult to extract from the raw bearing vibration signals. Furthermore, inefficient feature extraction results in substantial time wastage, making it hard to apply in real time monitoring. A novel feature extraction method for diagnosing bearing faults using multi-scale improved envelope spectrum entropy (MIESE) is proposed in this work. First, bearing vibration signals are analyzed across multiple scales, and improved envelope spectrum entropy (IESE) is extracted from these signals at each scale to form an original feature set. Subsequently, joint approximate diagonalization eigen (JADE) is applied to fuse above feature set for effectively eliminating redundancy and generated a refined feature set. Finally, the newly generated feature set is input into support vector machines (SVM) to effectively diagnose bearing health status. Two cases studies are employed to demonstrate the reliability of the proposed method. The results illustrate the proposed method can improve the stability of extracted features and increase the computational efficiency.

Keywords: Effective feature extraction, multi-scale improved envelope spectrum entropy

(MIESE), feature fusion, fault diagnosis, rolling bearing

1. Introduction

Rolling machinery is vital to modern manufacturing industry. The health status of bearing is critical for the safe operation of rotating machinery due to its key role in

rotating machinery [1, 2]. However, bearings typically operate in environments with high speeds, heavy loads and high temperatures, which makes them prone to damage. Thus, accurate and timely monitoring of bearing faults is essential to ensure the safety and reliability of rotating machinery [3-5].

Effective feature extraction is the prerequisite to ensure the accuracy and real-time performance of bearing fault diagnosis. Over the past few decades, various methods have been proposed and utilized for extracting features in bearing fault diagnosis. These methods can be categorized into time domain analysis, frequency domain analysis, time-frequency domain analysis, etc. To reduce noise interference in bearing vibration signal, several signal decomposition approaches have been applied in feature extraction for bearing fault. These include empirical mode decomposition (EMD) [6], EMD-based improve decomposition approaches [7-9], singular spectrum decomposition (SSD)[10], and variational mode decomposition (VMD) [11], etc. Besides, with advancements in artificial intelligence (AI) and computer hardware, numerous deep learning-based models have been leveraged to extract features from bearings, and achieved satisfied results [12-15].

Entropy, a measure of disorder within a system, was proposed by Clausius. On this basis, some other entropies have been proposed, such as information entropy (IE)[16], approximate entropy (ApEn)[17], energy entropy (EE)[18], fuzzy entropy (FE)[19], permutation entropy (PE)[20], dispersion entropy (DE)[21], etc. Entropy-based methods has also been widely applied in mechanical fault diagnosis. To address the nonlinear characteristics of bearing fault signals, Zhu et.al [22] proposed an improved FE to extract degradation indexes, they then established a model for bearing degradation assessment. Considering that the features of bearing fault signals correlate closely with the fault types, Li et.al [23] combined DE and improved complete ensemble EMD with adaptive noise (ICEEMDAN) to extract the features. They subsequently established a fault identification model for bearings based on support vector machines (SVM). Zhou et.al [24] introduced wavelet packet energy entropy (WPEE) to capture essential bearing features and constructed a model for assessing bearing degradation based on radial basis function neural network (RBFNN).

It is worth stating that valuable features of a signal are typically distributed across multiple frequency bands, indicating that single-band analysis has its limitations due to the entropy of the signal being spread out[25]. To overcome this defect, several entropy-based hierarchical analysis approaches have been proposed and applied to extract bearing features. For instance, Xue et.al [25] introduced hierarchical DE (HDE) in bearing fault diagnosis and combined it with JADE to extract health status features for fault diagnosis. To reduce the interference in the transmission path of inter-shaft bearing signals, Tian et.al [26] combined hierarchical permutation entropy (HPE) with locally linear embedding (LLE) for fault feature extraction. Moreover, recognizing that signal features are usually embedded across multiple time scales, feature extraction based on multiscale entropy has been widely applied, including multiscale entropy (MSE)[27], multiscale FE (MFE)[28], multiscale DE (MDE)[29], and multiscale PE (MPE)[30]. However, tradition entropy methods based on multiscale or hierarchical analysis typically involve a cumbersome selection process for numerous input parameters. To reduce the number of input parameters, a new entropy named hierarchical improve envelope spectrum entropy (HIESE) was proposed and tested in our previous work[31], requiring only the selection of a single hierarchical node as an input parameter. However, although HIESE is more efficient

than hierarchical entropy (HE)[32], HDE, MPE and MFE, it still spends too much time in features extraction due to the hierarchical decomposition process is still complex. In addition, the data length of HIESE method is chosen as exponential power of 2, restricts its applicability. Therefore, it is necessary to develop a feature index that has higher computational efficiency, not dependence on sample length, and has little input parameters should be selected. To mitigate these challenges, multiscale analysis is introduced to replace hierarchical decomposition. Then, a novel entropy, i.e., multiscale improved envelope spectrum entropy (MIESE) is proposed in this work, which only has one input parameter, and no too much restriction about the sample length.

Notably, some time-scale signals lack effective information regarding the health status of bearings, leading to the generation of redundant original features during extraction[33]. It is crucial to reduce redundant within the original feature set. In our earlier research, joint approximate diagonalization eigen (JADE) is selected as feature fusion method due to its effectiveness in eliminating redundancy[25]. Consequently, JADE was utilized to fuse the original feature set and create a refined set that eliminates redundancy in this study. To explain the choice of JADE as fusion approach, several prominent feature fusion methods, including principal component analysis (PCA)[34], kernel PCA (KPCA)[35], linear discriminant analysis (LDA) [36] are introduced for comparison.

To evaluate the performance of the refined feature set, some evaluate approaches should be introduced. Feature clustering stands out as a primary method for evaluating the spatial distribution of feature sets. Therefore, clustering is selected as evaluation approach. In addition, to quantitatively measure the features distribution, between-class and within-class scatters (I_{SS}) is adopted [37]. Larger I_{SS} means better performance of feature distribution.

Lastly, in order to determine the health status of the bearing via the extracted refined feature set, a fault identification model should be developed. In our previous work, the performance of support vector machine (SVM) is tested, results illustrated that the SVM has strong generalization ability, suitability for small sample learning and no local minimum [38]. Therefore, SVM is also selected in this work. Meanwhile, commonly used classifiers such as probabilistic neural networks (PNN)[39], back probabilistic neural network (BPNN) [40], and radial basis function neural networks (RBFNN) [41]are employed to analyze the stability of extracted feature set.

The contributions of this study can be summarized as:

(1) Propose a novel entropy, namely MIESE, which requires the selection of only one input parameter: the scale factor.
(2) Based on MIESE, design a novel framework for bearing features extraction that boasts high computational efficiency and overcomes the limitations of traditional methods in real-time monitoring.
(3) A refined feature is gotten by fusing original features using JADE, thus eliminating the redundancy generated during feature extraction.

The rest of this paper is structured as follows. The proposed entropy MIESE is presented in Section Ⅱ followed by the feature fusion method JADE. Section III outlines the framework of the methodology proposed, which is then validated with two cases in Section IV. Lastly, conclusion is drawn in Section Ⅴ.

2. Methodology

2.1 Improved ESE

Envelope spectrum entropy (ESE) is a novel entropy proposed by introducing entropy into the envelope spectrum, aimed at measuring the complexity of a time series. ESE demonstrates good performance in calculation efficiency and requires no input parameters to be set.

Assuming that $x(t)$ is a time series and $e(t)$ is its envelope signal, described as:

$$e(t) = \sqrt{[x(t)]^2 + [H(t)]^2} \quad (1)$$

where $x(t)$ is the time series, $H(t)$ denotes its corresponding Hilbert transform.

The envelope spectrum is obtained using fast Fourier transform (FFT) to process the envelope signal $e(t)$, followed by performing modulation operation, expressed as:

$$E = |FFT(e(t))|. \quad (2)$$

Then, introduce entropy into E, the ESE can be calculated as follows:

$$P = E / \sum_{i=1}^{N} E_i \quad (3)$$

$$ESE = -\sum_{j=1}^{N} P_j \log_2 P_j \quad (4)$$

where P refers to the ratio of E, N means the data length, Ei and P_j is the i-th data of E and j-th data of P, respectively.

It should be noted that ESE consists of envelope spectrum and entropy, and shares the same problem as entropy, i.e., it exhibits poor stability when measuring the complexity of multiple signals [42]. In bearing fault diagnosis, this instability affects the recognition rate as the variety of samples increases. To address this challenge, an enhanced version called Improved ESE (IESE) was introduced in our prior research [31], formulated as:

$$IESE = -\ln(ESE / N) \quad (5)$$

where N denotes the data length.

2.2 MIESE

The effective health status information in bearing vibration signals typically spans multiple scales, implying that IESE is unable to extract full information from the signal. Therefore, to fully extract the effective information, multiscale analysis is usually introduced to process the signal for extracting the information across each scale. Inspired by traditional multiscale approaches, MIESE is proposed with the following computational steps:

Step 1: For a time series signal $S=[s_1, s_2, \ldots, s_n]$, its multiple scale p_m^n can be calculated as:

$$p_m^n = \frac{1}{n} \sum_{i=m}^{m+n-1} s_i \,|\, 1 < m < N - n + 1 \quad (6)$$

where m denotes scale number in multiscale analysis, n represents the dynamical feature measurements of the time series at each scale. If n is 1, p_1^1 is the original time series. The time series process for n=2 and n=3 is shown in Fig.1.

Step 2: The *MIESE* is derived by calculating the *IESE* for each multiscale time series as:

$$MIESE = [IESE_m^1, IESE_m^2, \ldots, IESE_m^n] \quad (7)$$

2.3 JADE

Joint approximate diagonalization of eigenmatrices (JADE) is a blind source separation approach that can separate mutually statistically independent signals,

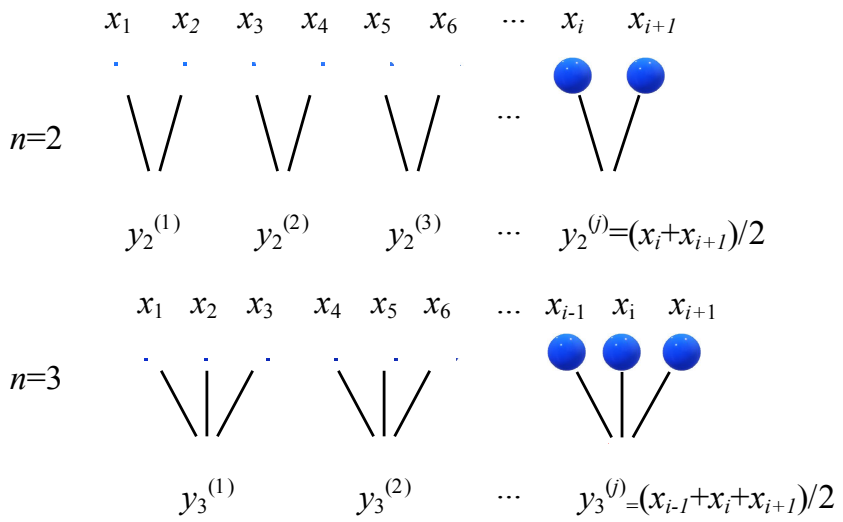

Fig. 1. Coarse-grained process of time series.

thereby enhancing signal observation in a low-dimensional space. Assuming that $x(t)$ is a standard linear signal mode, it can be expressed as:

$$m(t)=s(t)+n(t)=Ax(t)+n(t) \qquad (8)$$

where the signal $m(t)$ contains noise and serves as the source signal, while $s(t)$ denotes the resulting output signal, $n(t)$ and $x(t)$ refer to the noise and source signal respectively. A is used to describe the transformation between $s(t)$ and $x(t)$.

The detail separation of JADE can be divided into four steps as follows:

Step 1: Calculate whitening matrix W for sample covariance Rx with the following formula:

$$W=\left[(\upsilon_1-\theta)^{-\frac{1}{2}}h_1,\ldots,(\upsilon_n-\theta)^{-\frac{1}{2}}h_n\right]^H \qquad (9)$$

where v_n is the n-th largest eigenvalues, θ is the noise deviation, h_n means the n-th eigenvectors of R_x, and the signal after whitening is described as $v(t)=Wx(t)$.

Step 2: Calculate the fourth-order cumulants of $v(t)$ using the following formula:

$$Q_z=\{C(v_i,v_j,v_k,v_l),1\le i,j,k,l\le d\} \qquad (10)$$

where Q_z is the fourth-order cumulants of $v(t)$, while d stands for the dimension of v, C represents the calculation of the cumulants expressed as:

$$C(v_i,v_j,v_k,v_l)=E\overline{v_iv_jv_kv_l}-E\overline{v_iv_l}E\overline{v_jv_k}-E\overline{v_jv_l}E\overline{v_iv_k}-E\overline{v_iv_j}E\overline{v_kv_l} \qquad (11)$$

where $\overline{v_i},\overline{v_j},\overline{v_k},\overline{v_l}$ is the mean value of v_i,v_j,v_k,v_l, respectively.

Then, the cumulant with maximal set can be calculated as follows:

$$Q_z(M_i)=\lambda_iM_i,1\le i\le n \qquad (12)$$

where $Q_z(M_i)$ represents the cumulant with maximal set, M_i denotes the eigen matrix, and λ_i is the eigenvalue.

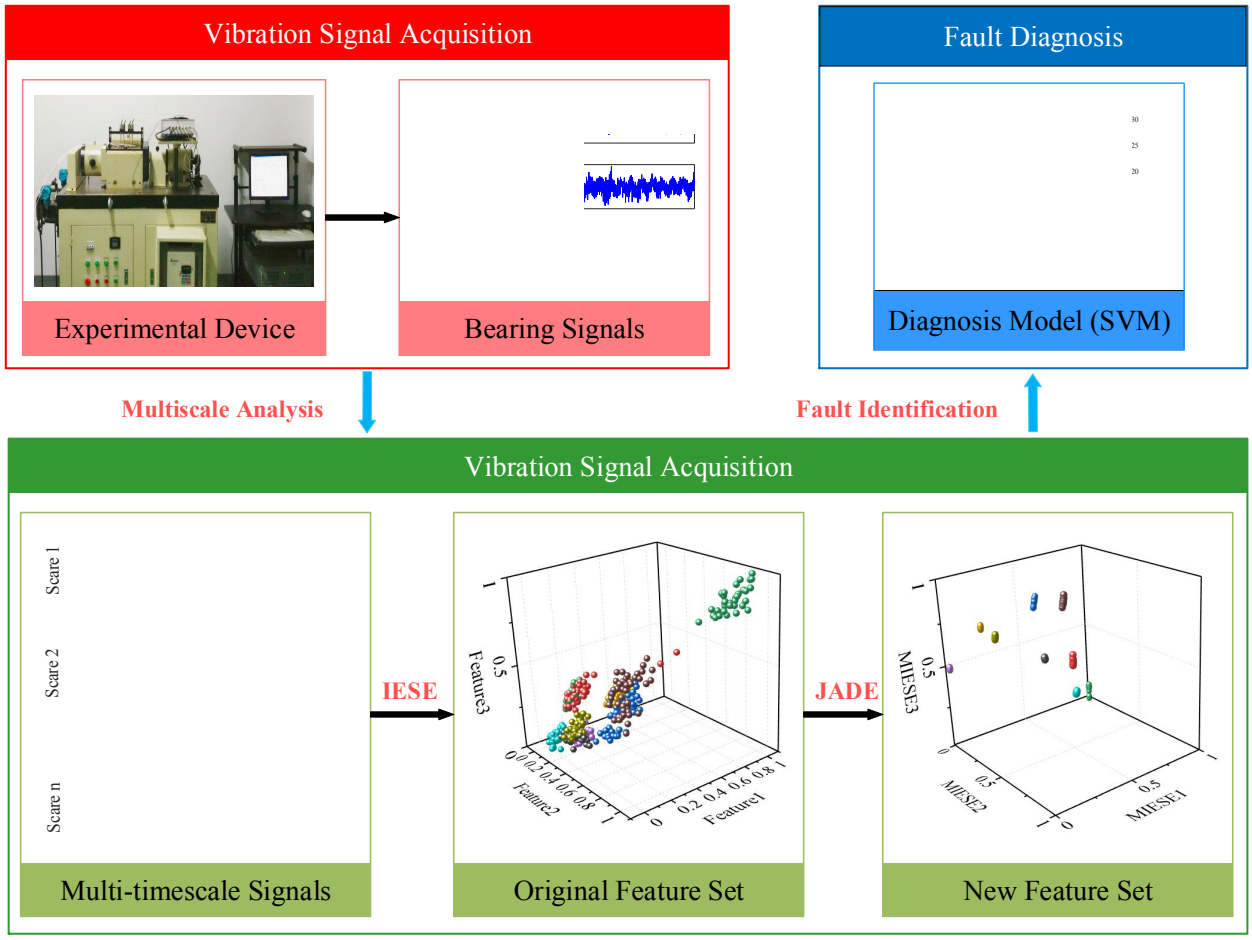

Fig.2. Description of the method.

Step 3: To make the cumulant diagonalized, unitary matrix $\widehat{U}$ is employed as follows:

$$\widehat{U} = \arg_{\min} \sum_{i} off(U^{\#} Q_z(M_i) U) \quad ,1 \le i \le n \tag{13}$$

where $\arg_{\min}$ denotes the plural argument, *off* pertains to the square of non-diagonal elements, U represents the rotation matrix, and # is the pseudo-inverse.

Step 4: Then, the separate matrix $\widehat{A}$ is as obtained:

$$\widehat{A} = \hat{U} W^{\#} \tag{14}$$

Finally, the separation result of $x(t)$ is given as:

$$x(t) = \widehat{A} m(t) \tag{15}$$

3. Description of the method

The vibration signal collected by accelerometers is usually one-dimensional. However, bearing vibration signal often contains noise due to the complex operating condition such as high speed, heavy load and high temperature. Therefore, it is difficult to identify the health status of the bearing by directly analyzing the vibration signal. Furthermore, fault information of the bearing does not manifest in a single time scale, complicating feature extraction. To overcome these problems, a novel bearing health status

feature extraction method using MIESE is proposed. The specific process of the proposed method can be described as Fig. 2, and which consists of five steps:

Step 1: Data collection.

Bearing vibration signals containing health status are collected from the designed experimental device. To balance computational efficiency and feature validity, the sample length of 2048 is selected in this work.

Step 2: Multiscale analysis.

To effectively extract features about the health status of the bearings, multiscale analysis is introduced to process the raw signal. Considering the scale factor is crucial to the performance of the proposed method, that are: a large-scale factor results in too much time cost in feature extraction, whereas a small one may hinder effective feature extraction. Therefore, the scale factor is set as 8 in this study.

Step 3: Original feature extraction.

The initial feature set can be derived by evaluating the IESE value of all scale signals, the results can be described as $MIESE = [IESE_m^1, IESE_m^2, ..., IESE_m^n]$.

Step 4: Feature fusion.

It should be noted that not all scale signals contain effective features, which means redundancy in the original feature set. To focus on effective features, the above set is fused by JADE to generate a refined feature set, denoted as:

$$F^3 = \widehat{A} F_{MIESE}^n \tag{16}$$

where F^3 represents new feature set, $\widehat{A}$ is separate matrix, F_{MIESE}^n refers to original feature set, 3 is the dimension of initial feature set, and n denotes the dimension of new feature set and is 8 in this study.

Step 5: Establish health status identification model of the bearings.

To identify the health status of the tested bearings via the refined feature set extracted in Step 4, a classifier should be selected to construct a status identification model. In this study, SVM is selected for its robust performance in training with limited data.

4. Experimental verification

Two datasets are analysed to validate the proposed method described above in this section. Furthermore, to better illustrate the superiority of the proposed method, other methods are introduced for comparison. It is important to note that only one condition in each comparative experiment is inconsistent with the proposed method, categorized as follows: (1) To test the performance of MIESE in extracting effective features, some prevalent entropies are introduced as comparisons including MFE, MPE, HDE, and HIESE. (2) Other methods for feature fusion, such as PCA, KPCA, and LDA, are presented for a comparison with the feature fusion method used (3) Lastly, for analysing the stability of the new feature set extracted, BPNN, PNN, and RBFNN are employed to further assess he effectiveness of the proposed method in feature extraction. Additionally, the I_{SS} index serves as a tool for assessing spatial distribution of the set for quantitative analysis. The performance of each method is evaluated by comparing recognition rates and processing times. As a note, the datasets are analysed using Matlab-R2018a. The

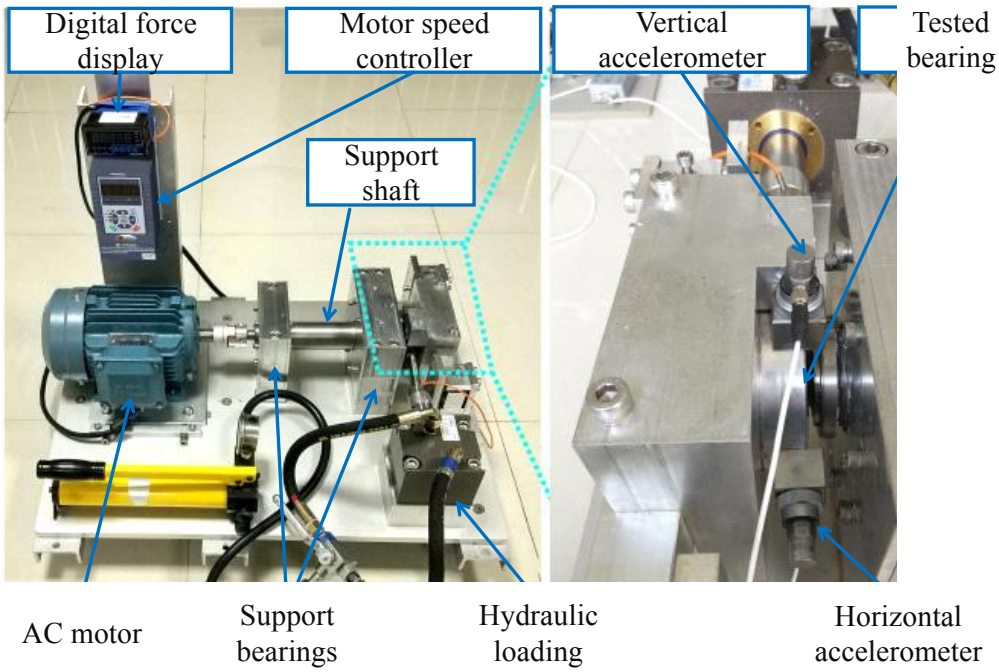

Fig.3. Test platform of XJTU.

algorithm runs on an Inter(R) Core(TM) i5-10210 CPU. The reported accuracy and processing time represent the average of five analysis runs.

4.1. Case 1

Dataset 1 was collected from bearing experiment diverse at Xian Jiaotong University (XJTU) [43, 44]. Fig. 3 shows the corresponding experimental setup, which consists of a digital force display, an AC motor, a hydraulic loading, a motor speed controller, a support shaft and tested bearings. To facilitate comparison, nine types of fault signals were measured, including outer ring defect, cage ring defect, inner element defect and mixed detects were measured. The corresponding operational conditions of the bearings are detailed in Table 1, with specifics summarized in Table 2. Each sample has a length of 2048, and a total of 450 samples are selected from exhibiting weak faults.

First, multiscale analysis is performed on all samples. The original feature set is obtained by calculating the IESE value across all time scale signals, which has a dimension of 450×8. Then, these features are fused by JADE to eliminate redundant information from their extraction process, producing a refined feature set sized at 450×3. For comparison, MFE, MPE, HDE and HIESE are also used to derive their respective refined feature sets. Fig. 4 illustrates the clustering of these new feature sets. It can be seen that the clustering of HIESE and MIESE outperforms the others, which can be explained by the fact that, for most features, those from the same samples are gathered to one point, while those from different samples are clustered. However, as seen in Fig. 4(d), B1-4, B2-1, B3-1 and B3-3 of HIESE are too close together, whereas this phenomenon does not exist for the clustering of MIESE. Therefore, the refined features generated by MIESE show better space distribution.

It is necessary to introduce evaluation indices to quantitatively compare the

Table 1. Working conditions of tested bearing

Condition	1	2	3
Rotation speed	2100	2250	2400
Radial force	12	11	10

Table 2. Descriptions of tested dataset

Condition	Bearing dataset	Fault location
1	B1_2	Outer
	B1_4	Cage
	B1_5	Inner and outer
2	B2_1	Inner
	B2_2	Outer
	B2_3	Cage
3	B3_1	Outer
	B3_2	Inner, outer, roller and cage
	B3_3	Inner

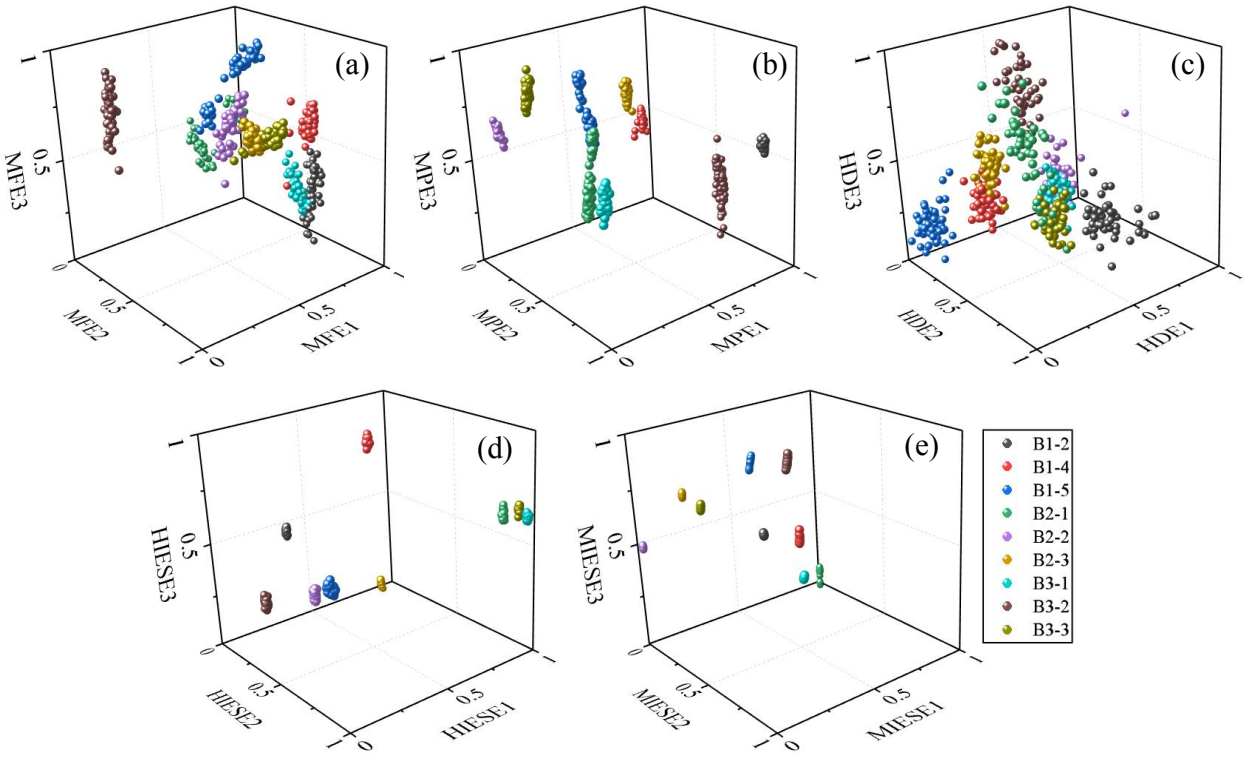

Fig.4. Comparison of clustering effects: (a) MFE, (b) MPE, (c) HDE, (d) HIESE, (e) MIESE.

clustering performance of the different methods. Effective feature clustering involves aggregating similar features while dispersing dissimilar ones. Therefore, I_{SS}, which combines inter-class and intra-class, serves as a quantitative evaluation index. Fig. 5 is the I_{SS} value of each entropy-based methods, from which it can be seen that most I_{SS} values of MIESE are larger than those of other methods. Consequently, the proposed method

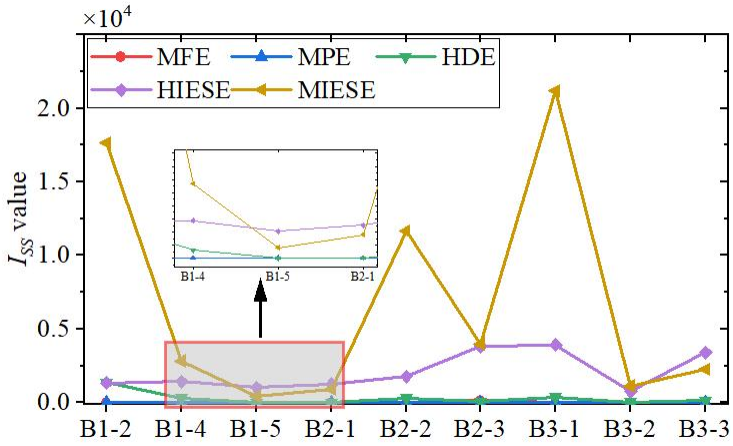

Fig.5. I_{SS} values of compared methods.

Table 3. Performance comparison of methods for feature extraction

Method	MFE	MPE	HDE	HIESE	MIESE
Accuracy (%)	71.75	96.83	84.76	**100**	**100**
Cost-time (s)	200.17	335.53	133.92	32.92	**6.50**

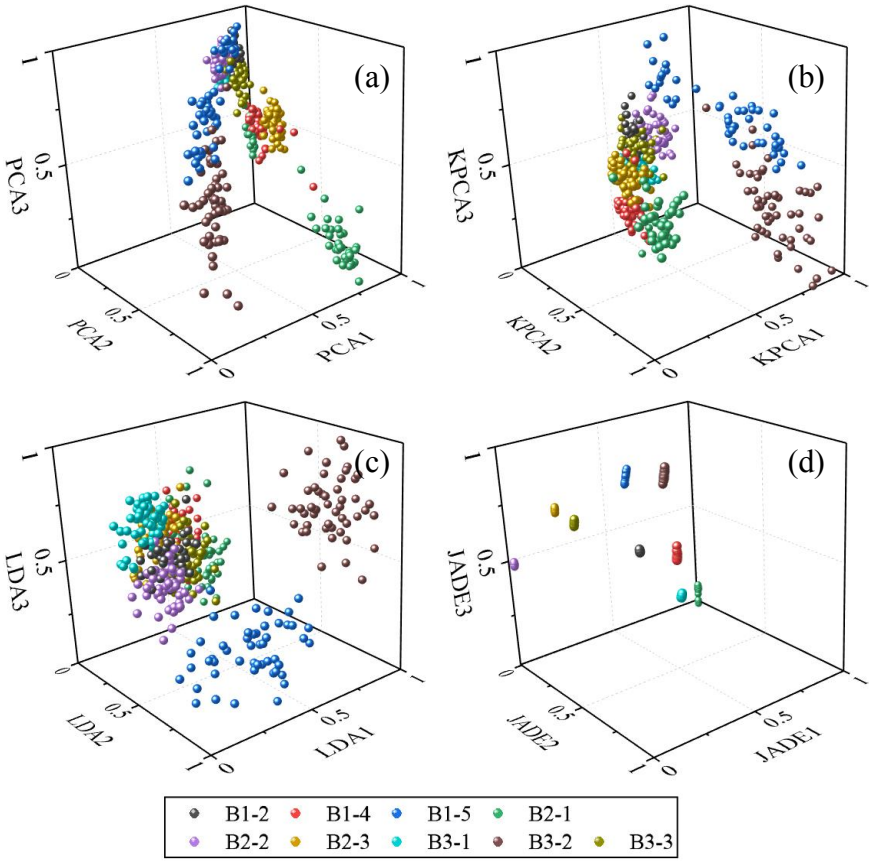

Fig.6. Comparison of clustering effects: (a)PCA, (b) KPCA, (c) LDA, (d) JADE.

Table 4. Performance comparison of methods for feature fusion

Method	PCA	KPCA	LDA	JADE
Accuracy (%)	53.65	59.05	86.35	**100**
Cost-time (s)	**6.43**	7.15	7.04	6.50

Table 5. Performance comparison of classifiers

Method	SVM	PNN	RBFNN	BPNN
Accuracy (%)	**100**	97.46	100	97.30
Cost-time (s)	**0.004**	0.043	0.186	0.187

demonstrates strong performance in feature extraction.

To determine bearing health via the extracted refined set, a bearing health status identification model is established. SVM is selected to establish the model due to its strong performance in small sample classification and stability. 30% of the refined set is randomly selected as training samples, with the remaining 70% designated as testing samples. The accuracy and cost time of each entropy-based methods are listed in Table 3. It is easy to see that the recognition rate of HIESE-based and MIESE-based methods are highest. However, the MIESE-based requires approximately one-ninth of the time compared to the HIESE-based method. This suggests that the proposed method outperforms comparative methods by being less time-consuming while achieving a higher fault recognition rate, making it suitable in real-time bearing health status monitoring.

Further, PCA, KPCA and LDA are selected to highlight the superiority of JADE in feature fusion. The clustering results of each fusion methods are shown in Fig. 6, revealing that PCA, KPCA, and LDA struggle to effectively separate features from different categories. In stark contrast, JADE demonstrates outstanding clustering performance. Then, SVM is used to test the refined sets extracted by each method, and the results are presented in Table 4. The findings indicate that the proposed method achieves a test accuracy of 100%, surpassing that of the comparison methods. This suggests that JADE is able to enhance the features of MIESE than that of other methods.

Finally, PNN, RBFNN and BPNN are introduced to build fault diagnosis model of the bearing, aimed to verify the feature stability of the proposed method. The hidden layer of BPNN is 10 neurons, and its spread is 1. Other input parameters are maintained at default values [25]. The results are described in Table 5. It can be noticed that the test accuracy of SVM and RBFNN are 100%, and both PNN and BPNN also have good test accuracy. These results indicate that the feature set extracted using proposed method has good stability and its performance is minimally influenced by classifier selection. In this work, SVM is selected as the classifier due to it has good theoretical foundation, robustness and more suitable for small sample training.

4.2 Case 2

For further testing the proposed method, the signals with different fault types and operation conditions collected from bearing fatigue experiment device designed by ourselves are used for analyzing. The device is shown in Fig. 7, which consists of the tested bearings, load, sensors, drive system, test piece, simulation module, lubrication system and spindle system. The sampling frequency is 10240 Hz, with a radial load of 0.9 kN applied to the tested SKF-6011

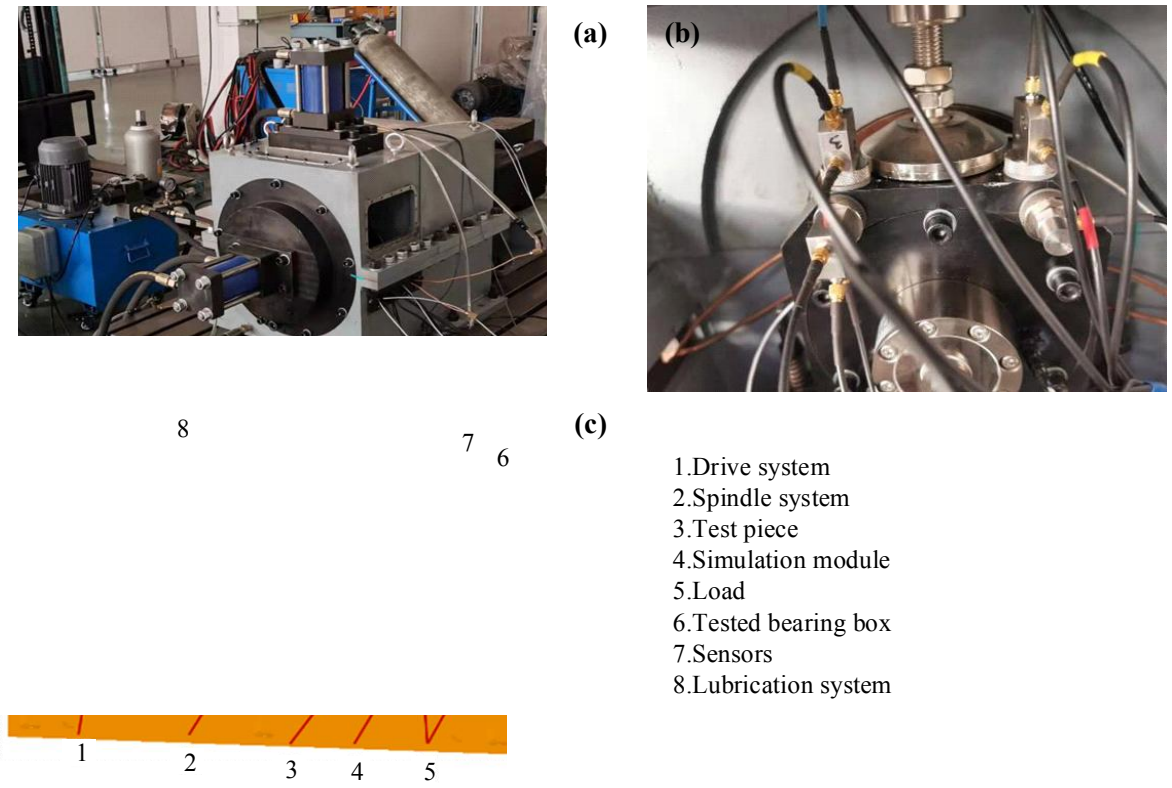

Fig.7. Experimental details: (a) Test platform; (b) Sensors positions; (c) Description of device structure.

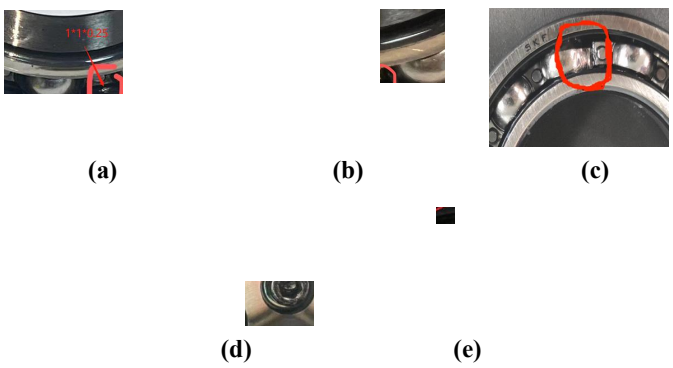

Fig.8. Photos of fault bearings: (a) Outer minor fault, (b) Outer severe fault, (c) Cage fault and contact with inner, (d) Cage fault and contact with outer, (e) Only cage fault.

bearings. Fig. 8 illustrates five bearings tested in the fatigue experiment, exhibiting minor outer defects, severe outer defects, cage defects with inner contact, cage defects with outer contact, and cage defects. Each bearing operates at two speeds: 3000 r/min and 8000 r/min, resulting in a total of 10 fault signals used to validate the proposed method, summarized in Table 6.

As previously mentioned, the signals are first processed using MIESE to generate an original feature set (500×8). Subsequently, JADE is employed to fuse this initial feature set into a refined feature set (500×3), aiming

Table 6 Descriptions of tested dataset

Fault description	Rotating speed	Bearing dataset
Outer minor fault	3000 r/min	OM3
	8000 r/min	OM8
Outer severe fault	3000 r/min	OS3
	8000 r/min	OS8
Cage fault and contact with inner	3000 r/min	CI3
	8000 r/min	CI8
Cage fault and contact with outer	3000 r/min	CO3
	8000 r/min	CO8
Only cage fault	3000 r/min	C3
	8000 r/min	C8

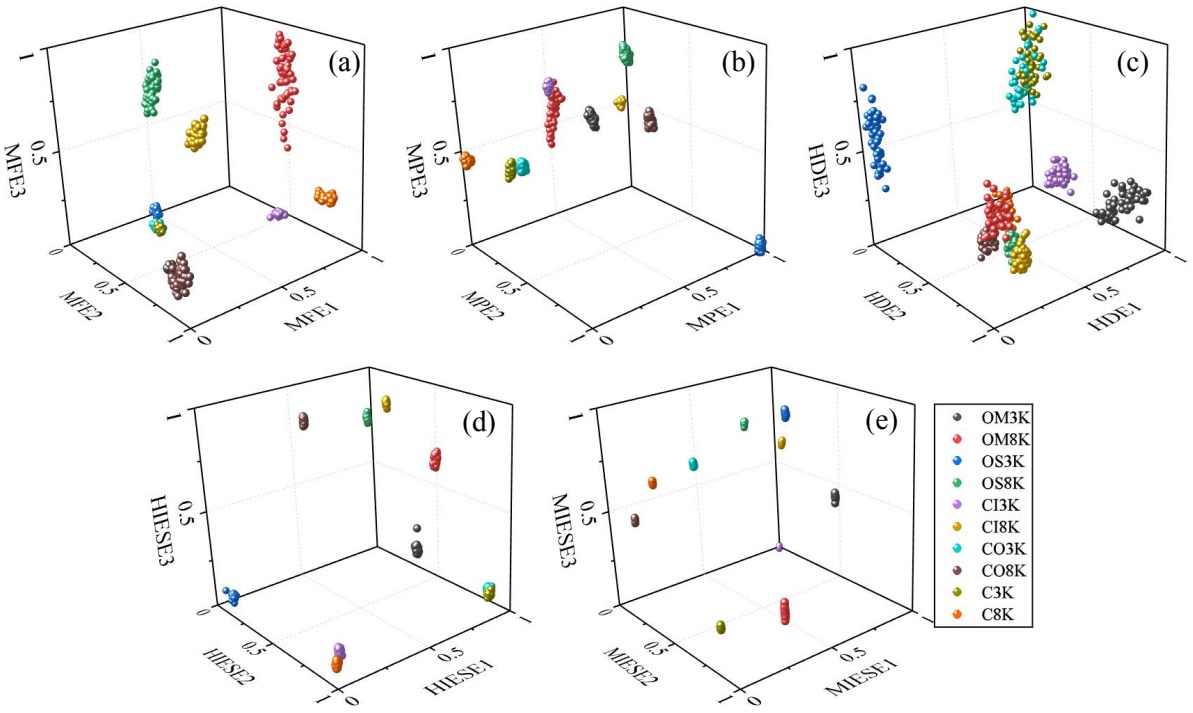

Fig.9. Comparison of clustering effects: (a) MFE, (b) MPE, (c) HDE, (d) HIESE, (e) MIESE.

to eliminate redundancy and reduce feature dimensions. Similar to Case 1, MFE, MPE, HDE, and HIESE are utilized for comparison to demonstrate the superior capability of MIESE in extracting bearing fault features. Fig. 9 shows the feature clustering results, highlighting that both HIESE and MIESE excel in clustering similar features together, indicated by the convergence of identical features into single clusters. However, upon comparing HIESE and MIESE results, it can be found that the overlapping of HIESE is more obvious, such as CO3K and C3K. In addition, some of the features extracted by HIESE are close, such as CI3K and C8K. Whereas from the clustering results of MIESE, it can be seen that only CS8K and CI8K are close, while the other features can be separated.

I_{SS} is still used as a contrast index to evaluate the spatial distribution performance

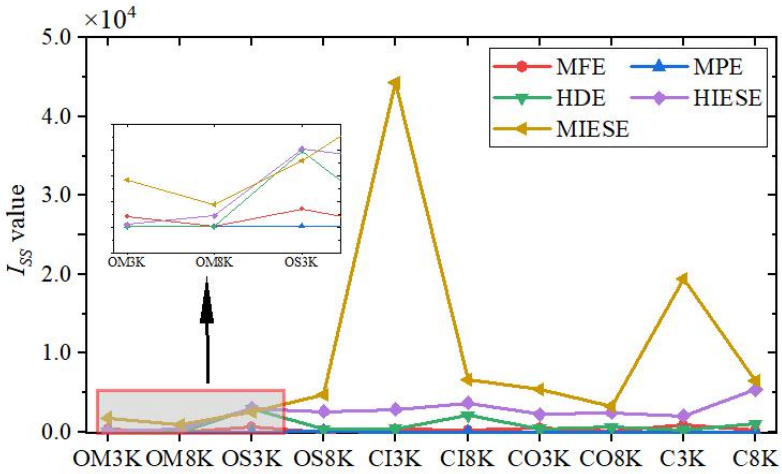

Fig.10. I_{SS} values of compared methods.

Table 7. Performance comparison of methods for feature fusion.

Method	MFE	MPE	HDE	HIESE	MIESE
Accuracy (%)	92.29	**100**	99.14	**100**	**100**
Cost-time (s)	224.19	378.68	149.43	37.24	**7.43**

of each method. In Fig. 10, the I_{SS} values of each method demonstrate that both HIESE and MIESE notably exceed those of MFE, MPE, and HDE. Furthermore, the I_{SS} values of MIESE are bigger than that of HIESE, except for OM8K and OS8K. Hence, the features extracted by proposed method have better spatial distribution.

The SVM is chosen as the classifier for detecting the health status of the bearing. Table 7 displays the accuracy and processing time of the four methods. It is evident from the findings that the accuracy of MPE, HIESE, and MIESE is notably superior to that of MFE and HDE. Nonetheless, the processing time for MIESE, MPE, and HIESE is 10.66s, 523.28s, and 65.91s, respectively. Consequently, the suggested approach exhibits superior performance in extracting features for determining the bearing health.

Similarly, SVM is chosen as the classifier to detect the status of the bearing. The diagnostic results of four methods are listed in Table 7. The results clearly show that the accuracy of MPE, HIESE and MIESE is significantly higher than that of MFE and HDE. Nonetheless, the cost time of MIESE, MPE and HIESE is 10.66s, 523.28s and 65.91s, respectively. Consequently, the method proposed has better performance in extracting fault features for bearing health status.

Same as Case 1, PCA, KPCA and LDA are introduced comparison with JADE. The refined feature set is generated by processing original feature set extracted by MIESE, as shown in Fig. 11. It can be seen that only the refined features processed by JADE can distinguish the bearing faults. For further comparison, Table 8 lists the accuracy and cost time of each method. The processing times are similar across all methods. However, the test accuracy of JADE is 100%, which is significantly higher than that of PCA, KPCA and LDA.

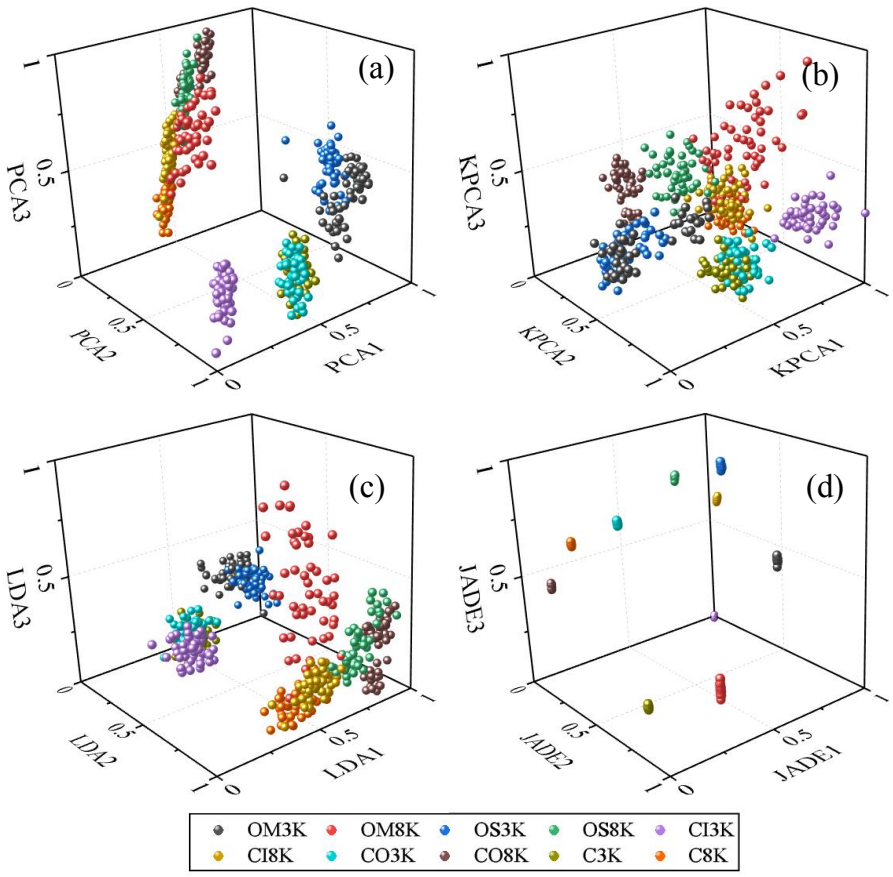

Fig.11. Comparison of clustering effects: (a)PCA, (b) KPCA, (c) LDA, (d) JADE.

Table 8. Performance comparison of methods for feature fusion

Method	PCA	KPCA	LDA	JADE
Accuracy (%)	78.29	72.57	81.14	**100**
Cost-time (s)	7.64	8.14	8.03	**7.43**

Lastly, PNN, RBFNN and BPNN are used to train and test the refined feature set extracted by the proposed method. Table 9 displays the test accuracy and cost time for each classifier. It can be found that the test accuracy of all classifier is 100%. The cost times of four classifiers are also similar due to the small sample of new feature set. Therefore, the features extracted by proposed method has good stability.

4.3 Further discussion

To illustrate the impact of parameter selection on the proposed method, further discussion is conducted based on XJTU dataset. It is worth noting that IESE not require input parameter settings. Therefore, only the time scale needs to be considered in MIESE. Besides, the dimension of the refined features, i.e., the output dimension of JADE, also needs to be taken into account.

First, the time scale of MIESE is analysed. The results of the proposed method with different time scale of MIESE are listed in Table 10. From the table, it is observed that as the time scale increases, both the recognition rate and processing time also increase.

Table 9. Performance comparison of classifiers

Method	SVM	PNN	RBFNN	BPNN
Accuracy (%)	100	100	100	100
Cost-time (s)	0.009	0.083	0.371	0.205

Table10. Results of the proposed method with different time scale

Time scale	3	4	5	6	8	10	15	20
Accuracy (%)	89.21	88.89	94.92	99.68	100	100	100	77.78
Cost-time (s)	4.88	5.41	5.58	5.93	6.50	7.34	8.15	8.74

Table 11 Results of the proposed method with different output dimension of JADE

Fused dimension	1	2	3	4	5	6	7
Accuracy (%)	30.48	98.41	100	100	100	100	100
Cost-time (s)	6.49	6.47	6.50	6.56	6.57	6.60	6.72

However, if the time scale is set too large, the recognition rate significantly decreases. The reason is that the signal with too little time scale contains limited information about the bearing fault status, which hinders model training. Additionally, despite the time consumption of MIESE increases with the increase in time scale, the overall time consumption does not significantly increase due to the efficient feature extraction capability of MIESE. Taking both accuracy and time consumption into consideration, this study selects a time scale of 8.

Meanwhile, the dimension of refined feature set is also analysed, and the corresponding results are shown in Table 11. From the table, it is evident that when the dimensionality of the refined features is greater than 1, the proposed methods achieve satisfactory recognition rates. Observation indicates a significant increase in recognition rate as the fusion dimension increases from 1 to 2. The reason is that although feature fusion can eliminate redundant information to some extent, excessive feature compression can also result in the loss of useful information. In addition, there is no significant difference in time consumption for different choices of refined feature dimensions, indicating that JADE does not significantly increase the time consumption of the proposed methods. Furthermore, it is beneficial for spatial distribution representation and observation while the fused dimension is 2 or 3. Therefore, considering comprehensively, the output dimension of JADE is set to 3 in this study.

5. Conclusion

In this study, to effectively analysis the complexity of bearing signals generated under different operating conditions and to improve the computational efficiency of HIESE, a novel feature index named MIESE was proposed. On the basis of MIESE, a novel effective bearing feature extraction method is devised, which analysis bearing signals from multiscale and calculate IESE values of each time scale to generate original feature set. Then, JADE is introduced to refine the original feature set into a new one to eliminate redundancy generated during feature extraction. The effectiveness of the proposed

method is tested by two cases, and the results can be summarized as:

(1) The clustering and I_{SS} reveal that the proposed MIESE exhibits strong feature extraction ability. Specifically, features belonging to the same category cluster together, while those from different classes are effectively separated.
(2) Tested accuracy and cost time of different methods are compared, results illustrate the designed method not only has good effective performance in features extraction ability about health status of bearings, but also cost least time.
(3) The input parameters, i.e, time scale of MIESE and fused dimension of JADE, are analysed to validate the applicability of the proposed method.

In future, MIESE will be considered to be combined with deep learning to solve the bearing fault diagnosis problem.

ACKNOWLEDGEMENTS

This work was supported in part by the Key Basic Research Project MKF20210008

References

[1] Singh J, Darpe A, Singh S P, "Use of Proximity-Based Shaft Displacement Sensors for Rolling Element Bearing Fault Diagnosis," Journal of Vibration Engineering & Technologies, vol. 12, no. 3, pp. 4571-4592, 2023.

[2] He C, Cao Y, Yang Y, Liu Y, Liu X and Cao Z, "Fault Diagnosis of Rotating Machinery Based on the Improved Multidimensional Normalization ResNet," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1-11, 2023.

[3] Hou B, Wang D, Peng Z and Tsui K, "Adaptive Fault Components Extraction by Using an Optimized Weights Spectrum Based Index for Machinery Fault Diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 71, pp. 985-995, 2024.

[4] Xu F, Fang Y, Kong Z, "A Fault Diagnosis Method Based on MBSE and PSO-SVM for Roller Bearings," *Journal of Vibration Engineering & Technologies*, vol. 4, no. 4, pp. 383-394, 2016.

[5] Liu F, Liu Y, Chen F and He B, "Residual life prediction for ball bearings based on joint approximate diagonalization of eigen matrices and extreme learning machine," *Proceedings of the Institution of Mechanical Engineers Part C: Journal of Mechanical Engineering Science*, vol. 231, pp. 1699-1711, 2015.

[6] Huang N, Shen Z, Long S, Wu M, Shih H, Zheng Q, Yen N, Tung C and Liu H, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London Series A: Mathematical," *Physical and Engineering Sciences*, vol. 454, pp. 903-995, 1998.

[7] Wu Z and Huang N, Ensemble empirical mode decomposition_a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, pp. 1-41, 2009.

[8] Fan X and Zuo M, "Machine fault feature extraction based on intrinsic mode functions," *Measurement Science and Technology*, vol. 19, 2008.

[9] Yeh J, Shieh J and Huang N, "Complementary Ensemble Empirical Mode Decomposition: A Novel Noise Enhanced Data Analysis Method,"

Advances in Adaptive Data Analysis, vol. 2, pp.135-156, 2011.

[10] Bonizzi P, Karel J, Meste O and Peeters R, "Singular Spectrum Decomposition: A New Method for Time Series Decomposition," *Advances in Adaptive Data Analysis*, vol. 6, 2015.

[11] Dragomiretskiy K and Zosso D, "Variational Mode Decomposition," *IEEE Transactions on Signal Processing*, vol. 62, pp. 531-544, 2014.

[12] Feng-Ping A and Zhi-Wen L, "Medical image segmentation algorithm based on feedback mechanism convolutional neural network," *Biomedical Signal Processing and Control*, vol. 53, 2019.

[13] Huang Z, Yang Y, Hu Y, Ding X, Li X and Liu Y, "Attention-augmented recalibrated and compensatory network for machine remaining useful life prediction," *Reliability Engineering & System Safety*, vol. 235, 2023.

[14] Li C, Zhang W, Peng G and Liu S, "Bearing Fault Diagnosis Using Fully-Connected Winner-Take-All Autoencoder," *IEEE Access*, vol. 6, pp. 6103-6115, 2018.

[15] Li R, Zhuang L, Li Y and Shen C, "Intelligent Bearing Fault Diagnosis Based on Scaled Ramanujan Filter Banks in Noisy Environments," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-13, 2021.

[16] Ye G, Pan C, Huang X, Zhao Z, He J, A Chaotic Image Encryption Algorithm Based on Information Entropy International," J*ournal of Bifurcation and Chaos*, 28, 2018.

[17] Delgado-Bonal A and Marshak A, "Approximate Entropy and Sample Entropy: A Comprehensive Tutorial," *Entropy*, vol. 21, 2019.

[18] Liu C, Zhu L and Ni C, "Chatter detection in milling process based on VMD and energy entropy," *Mechanical Systems and Signal Processing*, vol. 105, pp. 169-182, 2018.

[19] Jin Z, Sun Y, "Bearing Fault Diagnosis Based on VMD Fuzzy Entropy and Improved Deep Belief Networks," *Journal of Vibration Engineering & Technologies*, vol. 11, no. 2, pp. 577-587, 2023.

[20] Xue W, Dai X, Zhu J, Luo Y and Yang Y, "A Noise Suppression Method of Ground Penetrating Radar Based on EEMD and Permutation Entropy," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, pp. 1625-1629, 2019.

[21] Rostaghi M and Azami H, "Dispersion Entropy: A Measure for Time-Series Analysis" *IEEE Signal Processing Letters*, vol. 23, pp. 610-614, 2016.

[22] Zhu K, Jiang X, Chen L and Li H, "Performance Degradation Assessment of Rolling Element Bearings using Improved Fuzzy Entropy," *Measurement Science Review*, vol. 17, pp. 219-225, 2017.

[23] Li R, Ran C, Luo J, Feng S and Zhang B, "Rolling bearing fault diagnosis method based on dispersion entropy and SVM," *2019 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC)*, pp. 596-600, 2019.

[24] Zhou J, Wang F, Zhang C, Zhang L and Li P, "Evaluation of Rolling Bearing Performance Degradation Using Wavelet Packet Energy Entropy and RBF Neural Network," *Symmetry*, vol. 11, 2019.

[25] Xue Q, Xu B, He C, Liu F, Ju B, Lu S and Liu Y, "Feature Extraction Using Hierarchical Dispersion Entropy for Rolling Bearing Fault Diagnosis," *IEEE*

Transactions on Instrumentation and Measurement, vol. 70, pp. 1-11, 2021.

[26] Tian J, Zhang Y, Zhang F, Ai X and Wang Z, "A novel intelligent method for inter-shaft bearing-fault diagnosis based on hierarchical permutation entropy and LLE-RF," *Journal of Vibration and Control*, vol. 29, pp. 5357-5372, 2022.

[27] Wu S, Wu C, Lin S, Wang C and Lee K, "Time Series Analysis Using Composite Multiscale Entropy," *Entropy*, vol. 15, pp. 1069-1084, 2013.

[28] Zhao H, Sun M, Deng W and Yang X, "A New Feature Extraction Method Based on EEMD and Multi-Scale Fuzzy Entropy for Motor Bearing," *Entropy*, vol. 19, 2016.

[29] Shao K, Fu W, Tan J and Wang K, "Coordinated approach fusing time-shift multiscale dispersion entropy and vibrational Harris hawks optimization-based SVM for fault diagnosis of rolling bearing," *Measurement*, vol. 173, 2021.

[30] Li Y, Xu M, Wei Y and Huang W, "A new rolling bearing fault diagnosis method based on multiscale permutation entropy and improved support vector machine based binary tree," *Measurement*, vol. 77, pp. 80-94, 2016.

[31] Chen Z, Yang Y, He C, Liu Y, Liu X and Cao Z, "Feature Extraction Based on Hierarchical Improved Envelope Spectrum Entropy for Rolling Bearing Fault Diagnosis," IEEE Transactions on Instrumentation and Measurement, vol. 72, pp. 1-12, 2023.

[32] Jiang Y, Peng C and Xu Y, "Hierarchical entropy analysis for biological signals," *Journal of Computational and Applied Mathematics*, vol. 236, pp. 728-742, 2011.

[33] Wang Y, Ren X, Nan G, Yang Y and Deng W, "Rotating machine fault diagnosis based on denoising source separation," *2012 IEEE Fifth International Conference on Advanced Computational Intelligence (ICACI)*, pp. 1124-1127, 2013.

[34] Lever J, Krzywinski M and Altman N, "Principal component analysis," *Nature Methods*, vol. 14, pp. 641-642, 2017.

[35] Wu W, Massart D and Jong S, "The kernel PCA algorithms for wide data," *Part I: Theory and algorithms Chemometrics and Intelligent Laboratory Systems*, vol. 36, pp. 165-172, 2018.

[36] Fan Z, Xu Y and Zhang D, "Local Linear Discriminant Analysis Framework Using Sample Neighbors," *IEEE Transactions on Neural Networks*, vol. 22, pp. 1119-1132, 2011.

[37] He Q, "Vibration signal classification by wavelet packet energy flow manifold learning," *Journal of Sound and Vibration*, vol. 332, pp. 1881-1894, 2013.

[38] Hadroug N, Iratni A, Hafaifa A, Alili B, Colak I, "Implementation of Vibrations Faults Monitoring and Detection on Gas Turbine System Based on the Support Vector Machine Approach," *Journal of Vibration Engineering & Technologies*, vol. 12, no. 3, pp. 2877-2902, 2024.

[39] Saad E, Prokhorov D and Wunsch D, "Comparative study of stock trend prediction using time delay recurrent and probabilistic neural networks," *IEEE Transactions on Neural Networks*, vol. 9, no. 6, pp. 1456-1470, 1998.

[40] Li J, Yao X, Wang X, Yu Q and Zhang Y, "Multiscale local features learning based on BP neural network for rolling bearing intelligent fault diagnosis," *Measurement*, vol. 153, 2020.

[41] Karamichailidou D, Kaloutsa V and Alexandridis A, "Wind turbine power curve modeling using radial basis function neural networks and tabu search

Renewable," *Energy*, vol. 163, pp. 2137-2152, 2021.

[42] Ju B, Zhang H, Liu Y, Liu F, Lu S and Dai Z, "A Feature Extraction Method Using Improved Multi-Scale Entropy for Rolling Bearing Fault Diagnosis," *Entropy*, vol. 20, 2018.

[43] Wang B, Lei Y, Li N and Li N, "A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings," *IEEE Transactions on Reliability*, vol,69, pp. 401-412, 2020.

[44] Wang B, Lei Y, Li N and Yan T, "Deep separable convolutional network for remaining useful life prediction of machinery," *Mechanical Systems and Signal Processing*, vol. 134, 2019.