Intelligent Fault Detection for Rotating Machinery Using Cyclic Morphological Modulation Spectrum and Hierarchical Teager Permutation Entropy

Junchao Guo, Qingbo He, *Senior Member*, *IEEE*, Dong Zhen, and Fengshou Gu

***Abstract*—Intelligent fault detection of rotating machines is essentially a pattern classification issue. At the same time, effectively obtaining fault features from the measured signals is a key step to timely diagnose the health status of rotating machinery and evaluate the results of fault classification. To accurately obtain effective fault information to enhance fault accuracy, this paper proposes a novel fault detection scheme based on cyclic morphological modulation spectrum (CMMS) and hierarchical Teager permutation entropy (HTPE). In this scheme, firstly, CMMS was developed to analyze the measured signal to obtain a series of CMMS slices with different frequency bands, which solved the deficiencies of manual empirical selection of frequency band bandwidth in the traditional cyclic modulation spectrum (CMS). Subsequently, by integrating Teager energy operator into hierarchical permutation entropy (HPE), an improved feature selection method named hierarchical Teager permutation entropy (HTPE) is presented to obtain fault information of different frequency band slices, which can improve the fault feature extraction capability of HPE. Finally, the acquired HTPE-based vectors are integrated into the extreme learning machine (ELM) classifier to achieve fault classification of rotating machinery under different working conditions. The proposed scheme is validated by experimental cases including cylindrical roller bearings and planetary gearboxes. The analysis results indicate that the proposed scheme not only can effectively obtain the fault features, but also accurately realize the classification and recognition of the fault mode. In addition, the proposed scheme can achieve higher detection accuracy than some existing algorithms.**

***Index Terms*—Cyclic morphological modulation spectrum; Hierarchical Teager permutation entropy; Rotating machinery; Fault detection.**

Manuscript received December 7, 2021; revised March 27, 2022; accepted June 15, 2022.This work was supported in part by the National Science and Technology Major Project under Grant J2019-IV-0018-0086, in part by the National Program for Support of Top-Notch Young Professionals, in part by the China Postdoctoral Science Foundation under Grant 2021M702122, and in part by the National Natural Science Foundation of China under Grant 12121002. Paper no. TII-22-1378. (*Corresponding author*: *Qingbo He*)

Junchao Guo and Qingbo He are with the State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China. (e-mail: jc_guo12@sjtu.edu.cn; qbhe@sjtu.edu.cn).

Dong Zhen is with the School of Mechanical Engineering, Hebei University of Technology, Tianjin 300401, China. (e-mail: d.zhen@hebut.edu.cn).

Fengshou Gu is with the Centre for Efficiency and Performance Engineering, University of Huddersfield, Huddersfield HD1 3DH, UK. (e-mail: f.gu@hud.ac.uk).

I. Introduction

FAULT diagnosis of rotating machinery is an essential topic in prognostics and health management systems. Unexpected failure of rotating machinery can cause damage to mechanical equipment, resulting in huge economic losses and even casualties [1], [2]. In addition, since the rotating machinery has been working in harsh environments, the measurement signals obtained by sensors are easily affected by the complex transmission path and strong background noise, which also greatly increases the difficulty of rotating machine's fault diagnosis [3], [4]. Therefore, timely and accurate diagnosis of rotating machines faults is of great significance to prevent serious equipment damage and downtime. Currently, a few signal processing techniques are applied in rotating machinery fault diagnosis. For instance, variational mode decomposition (VMD), maximum correlated kurtosis deconvolution (MCKD), singular value decomposition (SVD) and sparse decomposition, etc. Although these techniques have made certain achievements in improving the accuracy of rotating machinery fault detection, they are based on tracking the amplitude of the fault frequency and neglecting the inherent modulation characteristics in the measurement signal. If the modulation components and carrier components of the measurement signal cannot be effectively separated, it will be difficult to accurately obtain the fault features.

Spectral correlation (SC) analysis is a two-dimensional spectrogram that can simultaneously present modulation components and carrier components for fault detection in rotating machinery [5], [6]. However, the SC contains abundant discrete Fourier transforms in the analysis process of the measurement signals, which results in excessive calculational costs. In order to solve this problem, Antoni *et al*. [7] developed a cyclic modulation spectrum (CMS), which first performs a short-time Fourier transform (STFT) on the measured signal, and then implements a Fourier transform on the time axis of the spectrum. It not only greatly improves the performance of fault extraction, but also significantly outperforms SC in terms of computational efficiency. With its excellent performance, CMS provides an effective tool for rotating machinery fault diagnosis [8], [9]. However, the maximum cycle frequency α_{max} of CMS is chosen to be less than Fs/2 (Fs represents the sampling frequency), and its cycle frequency resolution is low because the window size applied used in STFT is fixed. Multi-scale combination morphological filter (MCMF) is a nonlinear filtering processing technique based on mathematical morphological transformation. It relies on structural elements (SE) with the same function as the filter window to match the geometric characteristics of the signal to be analyzed to achieve fault feature retention and interference noise removal. However,

because the SE of the same scale is applied, the output of MCMF will be biased [10]. Therefore, an improved MCMF operator with different scales is developed to construct an improved multi-scale combination morphological filter (IMCMF) and integrated into the CMS to replace the STFT window in this study, named cyclic morphological modulation spectrum (CMMS), which can effectively overcome the drawback of low cycle frequency resolution and expand the maximum cycle frequency range to improve the fault detection accuracy. Unfortunately, the choose of fault information bandwidth from CMMS relies on user experience choices. That is, improper selection of frequency bands will also affect the accuracy of the fault features. Sparsity measurement (SM) is a potential method to determine the optimal frequency, which involves the fault impulse component of the measured signal. Moreover, SM has been demonstrated to be an effective frequency band selection scheme [11], [12]. In view of this, SM is utilized to select the optimal frequency band of CMMS to reduce in-band noise.

Apparently, several frequency bands selected by CMMS using SM have different fault signatures. Therefore, it is necessary to adopt accurate measurement schemes to obtain the fault features of each frequency band. In recent years, entropy-based algorithms have been introduced into the fault feature extraction of rotating machinery, such as correlation dimension (CD), sample entropy (SE), approximate entropy (AE) and fuzzy entropy (FE). Although these algorithms have achieved certain benefits in fault feature extraction, they still have some limitations in their applications. The CD requires a longer data set when processing the measured data, so CD is not suitable for practical applications [13]. The SE is generated by the Heaviside step function, which is discontinuous and abrupt at the boundary [14]. The AE largely depends on record length and lower estimation value in fault feature extraction [15]. The FE is defined using the membership function, and the parameters are extremely hard to determine [16]. Recently, a new entropy-based algorithms named permutation entropy (PE) was put forward by Bandit [17], which can amplify weak signals and detect the operating mode of the mechanical devices. On the basis of PE, multi-scale permutation entropy (MPE) is put forward to improve the statistical significance of PE. However, the coarse-grained program applied in MPE indicates linear smoothing, which uses an averaging strategy to obtain low-frequencies while ignoring the feature information submerged in high-frequencies [18]. Therefore, Li *et al*. [19] developed a new indicator named hierarchical permutation entropy (HPE), which is specifically used to evaluate the dynamics of rotating machinery systems. In view of the excellent performance of HPE, a large number of researchers have made great contributions to promote the development of HPE-based weak fault feature extraction methods [20], [21]. Besides, the Teager energy operator (TEO) was selected as an effective demodulation fault feature method by Randall et al. [22] and Song *et al*. [23] to recognize fault types under different operating situations. Based on the advantages of HPE and TEO, a novel feature extraction algorithm is presented, called hierarchical Teager permutation entropy (HTPE), which can effectively overcome the modulation information that HPE cannot effectively capture the frequency band features. Hence, CMMS and HTPE are combined to obtain fault features of rotating machine in this study. Finally, the acquired feature vectors are integrated into the extreme learning machine (ELM) classifier to achieve the fault classification of rotating machine under different working situations. It should be pointed out that compared with these classification methods (e.g., back-propagation neural network (BPNN), k-nearest neighbor (KNN) and support vector machine (SVM)), ELM classifier has the advantages of fast learning speed and good generalization ability. Additionally, the weight matrix between the input layer and the hidden layer and the bias of the hidden layer neurons are randomly assigned, and no adjustment is required during the training process [24], [25].

To sum up, the main intention of this paper is to develop a new intelligent fault detection scheme that integrates CMMS, HTPE and ELM classifier to realize automatic fault feature recognition and classification. The main innovations and novelties of this paper are summarized as follows: (1) A new method named CMMS is proposed to decompose the measurement signal to obtain CMMS slices with different frequency bands, which overcomes the shortcomings of traditional CMS that require manual experience to select bandwidth and low cyclic frequency resolution to obtain more comprehensive and abundant fault feature information. (2) A novel fault feature indicator (i.e., HTPE) is developed to construct feature vectors, which can overcome the deficiency that HPE cannot effectively capture the modulation information of frequency band features and improve the fault identification ability of rotating machinery. The results on simulated signals and two experimental cases indicate the proposed scheme realizes the classification and recognition of the fault mode. In addition, compared with some existing methods (e.g., MSB-HTPE, Fast-SC-HTPE, CMMS-HPE, CMMS-HSE and CMMS-HPE), the CMMS-HTPE provides better detection performance when used for rotating machinery fault classification.

The remainder of the paper is organized as follows. In Section II, the theoretical aspects of CMMS are introduced. In Section III, the detailed diagnostic process of the proposed scheme is presented. Vibration signals collected from the cylindrical roller bearings and planetary gearboxes are provided to illustrate the effectiveness of the proposed scheme in Section IV. Finally, the conclusions of the research work are summarized in Section V.

II. Cyclic Morphological Modulation Spectrum

A. Cyclic Morphological Modulation Spectrum

Morphological filter (MF) is a non-linear filtering technology. Its core idea is to perform front-to-back translation matching or partial correction to the original signal by constructing specific structural elements, so as to suppress noise while retaining the morphological characteristics of the signal. According to the local morphological characteristics of the signal to be processed, the signal is separated from noise through MF. Let g is a unit SE, and $k\,(k=1,2,\ldots,K)$ is the scale, the SE used in the scale k can be expressed as:

$$kg=\underbrace{g\oplus g\oplus\cdots\oplus g}_{k-1\ times} \tag{1}$$

Assume that $f(n)$ is the input signal. Multi-scale basic morphological operators are defined as:

$$(f\oplus kg)(n)=f\oplus\underbrace{(g\oplus\cdots\oplus g)}_{k-1times} \tag{2}$$

$$(f\Theta kg)(n)=f\Theta\underbrace{(g\Theta\cdots\Theta g)}_{k-1times} \tag{3}$$

$$(f\circ kg)(n)=((f\Theta kg)\oplus kg)(n) \tag{4}$$

$$(f\bullet kg)(n)=((f\oplus kg)\Theta kg)(n) \tag{5}$$

where $\oplus$, Θ, $\circ$ and $\bullet$ are dilation, erosion, opening and closing operator, respectively. Based on the combination of opening and closing operator, two multi-scale morphological filters, namely, multi-scale opening-closing (F_{OC_k}) and multi-scale closing-opening (F_{CO_k}), which are defined as follows:

$$F_{OC_k}(n)=\left(\left(f\circ kg\right)\bullet kg\right)(n) \tag{6}$$

$$F_{CO_k}(n)=\left(\left(f\bullet kg\right)\circ kg\right)(n) \tag{7}$$

Both the $F_{OC_k}(n)$ and $F_{CO_k}(n)$ filters can obtain cyclic impulses. However, due to the statistical bias, satisfactory results cannot be obtained by the $F_{OC_k}(n)$ or $F_{CO_k}(n)$ filter alone. To solve the problem, the average combination of $F_{OC_k}(n)$ and $F_{CO_k}(n)$ (i.e., multi-scale combination morphological filter, MCMF) is proposed, which is expressed as:

$$y_k(n)=\frac{F_{CO_k}(n)+F_{OC_k}(n)}{2} \tag{8}$$

For the $F_{OC_k}(n)$ filter, the first opening operator can enhance negative impulses and filter out positive impulses. If the closing operator uses the same length of SE, it will not be able to effectively filter out all negative impulses. Similarly, the $F_{CO_k}(n)$ filter also has the same issue. To eliminate the above issue, the improved multi-scale opening-closing ($\hat{F}_{OC_k}(n)$) and multi-scale closing-opening ($\hat{F}_{CO_k}(n)$) filters are defined as:

$$\hat{F}_{OC_k}(n)=\left(\left(f\circ kg_1\right)\bullet kg_2\right)(n) \tag{9}$$

$$\hat{F}_{CO_k}(n)=\left(\left(f\bullet kg_1\right)\circ kg_2\right)(n) \tag{10}$$

where the lengths of g_1 and g_2 are L and $2L$-1, respectively. The improved multi-scale combination morphological filter (IMCMF), which is denoted as:

$$\hat{y}_k(n)=\frac{\hat{F}_{OC_k}(n)+\hat{F}_{CO_k}(n)}{2} \tag{11}$$

where $\hat{y}_k(n)(k=1,2,\ldots,K)$ is a two-dimensional matrix $\hat{y}(k,n)$, each row of which corresponds to a specified scale k, each column corresponds to a specific point n across entire scales. Due to the appearance of random noise and interference frequencies, it is difficult to accurately identify periodic impulse features in the time domain. Therefore, the discrete Fourier transform (DFT) is applied to transform $\hat{y}(k,n)$ into $\hat{Y}(k,f_x)$, which is defined as:

$$\hat{Y}(k,f_x)=\sum_{\varphi=-\infty}^{\infty}\hat{y}(k,n)e^{-j2\pi f_x\varphi} \tag{12}$$

where $\hat{Y}(k,f_x)$ represents spectrum of the IMCMF, f_x means spectral frequency. Consequently, cyclic morphological modulation spectrum (CMMS) is calculated from the DFT of IMCMF spectrum, which is defined in Eq. (13):

$$S_{CMMS}(a,f_x)=\frac{1}{KF_s}\sum_{k=1}^{K}\left|\hat{Y}(k,f_x)\right|^2 e^{-j2\pi k\frac{a}{F_s}} \tag{13}$$

where F_s means the sampling frequency, K represents the signal length, α indicates the cyclic frequency. The normalized form of CMMS can be given by Eq. (14), and the result is shown in Fig. 1.

$$\gamma_{CMMS}(a,f_x)=\frac{S_{CMMS}(a,f_x)}{\sqrt{S_{CMMS}(0,f_x)S_{CMMS}(0,f_x-a)}} \tag{14}$$

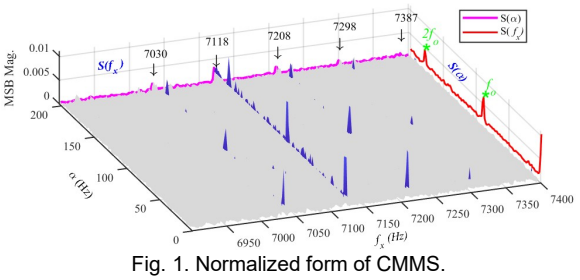

Fig. 1. Normalized form of CMMS.

To get the suboptimal f_x slices, the $\gamma_{CMMS}(a,f_x)$ is determined by CMMS slice $S(f_x)$, which is computed by averaging the main CMMS peaks:

$$S(f_x)=\frac{1}{M-1}\sum_{\vartheta=2}^{M}\gamma_{CMMS}(\vartheta\Delta g,f_x) \tag{15}$$

where Δg means the frequency resolution in the α direction. However, CMMS needs a better method to select the number of frequency bands. In other words, when the CMMS processes the measured signals to extract fault defect features, it needs to choose a reasonable bandwidth. Consequently, a new criterion called sparseness measurement (SM) is proposed to determine the number of CMMS slices to cover the useful resonance frequency band to reduce the loss of effective fault information. According to the analysis of CMMS slice $S(f_x)$, the SM can be defined as follows:

$$SM=\frac{\sqrt{\sum_{f_x=1}^{L}\left(S(f_x)\right)^2}}{\sum_{f_x=1}^{L}\left|S(f_x)\right|}=\frac{\left\|S(f_x)\right\|_2}{\left\|S(f_x)\right\|_1} \tag{16}$$

where L indicates the length of the CMMS slices. The $\left\|S(f_x)\right\|_1$ and $\left\|S(f_x)\right\|_2$ are L_1 norm and L_2 norm, respectively. If the measured signal has good sparsity, the SM will be larger. That is, the larger the SM, the better the CMMS fault feature extraction effect. Therefore, the largest SM is used to determine the number of CMMS slices.

To obtain more robust results, the CMMS detector is constructed by averaging the selected suboptimal slices marked with '↓' in Fig. 1, which is defined as:

$$S(\alpha)=\frac{1}{P}\sum_{p=1}^{P}\gamma_{CMMS}(\alpha,f_x^p),\qquad (f_x>0) \tag{17}$$

where P represents the total number of selected CMMS slices, which depends on the number of SM.

B. Comparison of CMMS, MSB and Fast-SC by simulation signal

To evaluate the performance of the CMMS, the results are compared with modulation signal bispectrum (MSB) [26] and Fast spectrum correlation (Fast-SC) [27]. The rolling element bearing fault model $x(t)$ with a local fault is defined as follows [28]:

$$\begin{aligned}x(t)=&\sum_l A_l\cos(2\pi lf_o t+\theta_l)+\sum_i C_i G(t-iT-\tau_i)\\&+\sum_j R_j G(t-Tr_j)+n(t)\end{aligned} \tag{18}$$

where the first part represents the periodic interference components from the rotor or shaft, f_o indicates the rotational frequencies, set to 10 Hz, A_l and θ_l represent the measured signal amplitude and initial phase of the lth harmonic, respectively. The second part indicates the periodic impulse signal produced by a bearing defect, C_i indicates the ith impulse amplitude. $G(t)$ means the impulse response. The τ_i indicates a uniformly distributed random number for the random slippage effect of the bearing, which is often represented as 1%~2% of the T. The periodic impulse can be denoted as follow:

$$G(t)=e^{-\partial t}\sin(2\pi f_r t) \tag{19}$$

where f_r stands the resonant frequency, ∂ represents the decay parameter. The third part indicates random impulse aroused by external interference from the rolling bearing housing. Among them, R_j stands random impulse amplitude, Tr_j indicates occurrence time. The last part $n(t)$ represents Gaussian white noise with a signal-to-noise ratio (SNR) of -15 dB. The parameters of the

bearing fault model are illustrated in Table I. Meanwhile, the sampling frequency f_s= 96 kHz and sampling number N=960000. Fig. 2 displays the time waveform of $x(t)$ and its frequency spectrum.

TABLE I
PARAMETERS OF THE BEARING FAULT MODEL

Rotor or shaft		Periodic impulses			Random shocks		
A_1	θ_1	C_i	T	f_{r1}	β_1	f_{r2}	β_2
0.01	0°	1	1/60	4500	300	2500	500

The CMMS is utilized to process the simulated signal illustrated in Fig. 2(a) to extract the fault defect frequencies. First of all, the lengths of g_1 and g_2 are 3 and 5 [29], respectively. Subsequently, the CMMS slices $S(f_x)$ are computed using Eq. (15), and the suboptimal CMMS slices are selected according to the SM method. As illustrated in Fig 3, when the CMMS slice is 3, SM reaches its maximum value. Finally, the CMMS is obtained through averaging 3 suboptimal CMMS slices with '*' makers in Fig. 4(a), and the detection result is displayed in Fig. 4(b). It is clearly found that the CMMS can accurately distinguish the fault frequency 60 Hz and its first 4 harmonics.

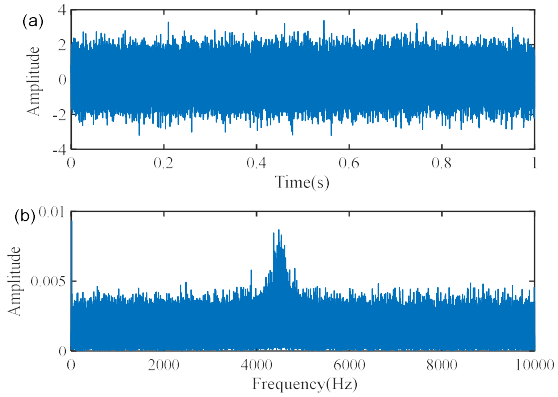

Fig. 2. Simulation case: (a) time waveform (b) frequency spectrum.

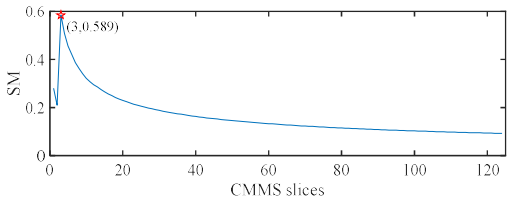

Fig. 3. The SM value of CMMS slices.

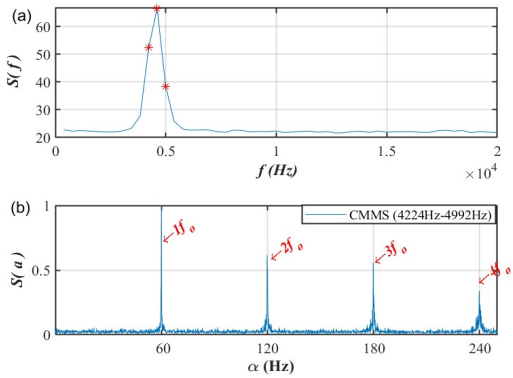

Fig. 4. Results of the CMMS: (a) CMMS slice (b) CMMS detector.

For comparison, the MSB and Fast-SC are utilized to process the simulated signal of defective bearing. The MSB is a novel modulation signal demodulation algorithm put forward in Ref. [26]. For MSB method, the MSB slice and window length N_w are set to 3 and 2^{14}. The algorithm uses the above parameters to suppress the interference of random noise and non-periodic components to clearly reflect the modulation components in the measurement signal. As depicted in Fig. 5, the detection result obtained by the MSB is difficult to distinguish the fault frequencies. For the Fast-SC algorithm [27], the window length N_w and the maximum cyclic frequency α_{max} are equal to 2^{14} and 250 Hz, the Fast-SC using the above parameters to process the fault model. Hereafter, the enhanced envelope spectrum (EES) is utilized to obtain the fault defect frequencies from Fast-SC filtered signals. Although the fault defect frequency 60 Hz and its first 4 harmonics can be extracted, the redundant noise and interference frequencies in higher harmonic frequencies still exist as illustrated in Fig. 6.

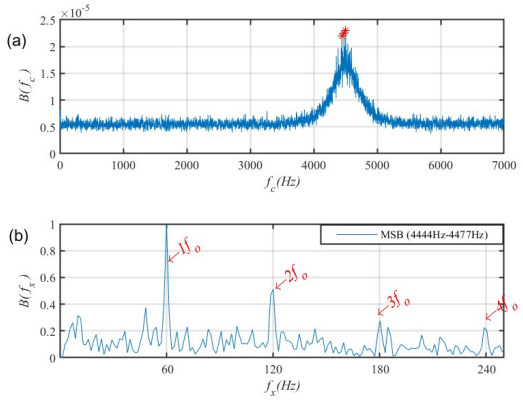

Fig. 5. Results of the MSB: (a) MSB slice (b) MSB detector.

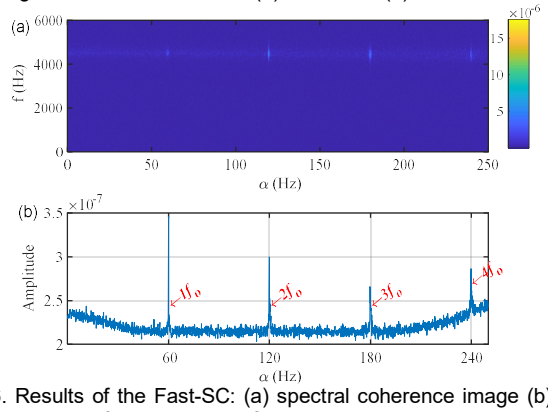

Fig. 6. Results of the Fast-SC: (a) spectral coherence image (b) EES generated in the frequency band from 4000 Hz to 5000 Hz.

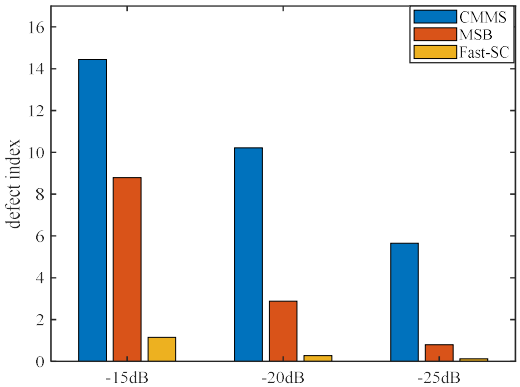

Fig. 7. The defect index of three algorithms under different SNRs.

In order to further highlight the effectiveness of the CMMS algorithm, the defect index u is presented to evaluate the performance of three algorithms. The defect index u is defined as follows:

$$u = 10\log_{10}\left(\frac{\|A_p\|_2^2}{\|A_n\|_2^2}\right) \tag{20}$$

where A_p and A_n indicate the magnitude of the fault frequency and white noise (or random impulse) interferences. Fig. 7 depicts the u of three algorithms under different SNRs. As depicted in Fig. 7, with the SNR decreases, the u of three algorithms also gradually decreases. Moreover, CMMS has stronger fault recognition capabilities than MSB and Fast-SC. Through the above comparison, the CMMS illustrates higher performance in terms of enhanced fault features and signal modulation.

III. The Proposed Fault Detection Scheme

A. Hierarchical Teager permutation entropy for feature extraction

Teager energy operator (TEO) is a novel model of nonlinear operator that forecasts the entire energy required through the signal source to generate a dynamic signal by the nonlinear combination of the instantaneous value of the signal. For a given CMMS slice $S(f_x)(f_x=1,2,\cdots,P)$, the TEO can be defined as follows:

$$\Psi(S(f_x)) = S^2(f_x) - S(f_x-1)S(f_x+1) \tag{21}$$

where $\Psi(S(f_x))$ means the use of 3 samples to calculate the signal energy at frequency f_x. Then, use phase space reconstruction to obtain the matrix:

$$Y = \begin{bmatrix} \Psi(S(1)) & \Psi(S(1+\tau)) & \cdots & \Psi(S(1+(m-1)\tau)) \\ \Psi(S(2)) & \Psi(S(2+\tau)) & \cdots & \Psi(S(2+(m-1)\tau)) \\ \Psi(S(j)) & \Psi(S(j+\tau)) & \cdots & \Psi(S(j+(m-1)\tau)) \\ \vdots & \vdots & \cdots & \vdots \\ \Psi(S(n)) & \Psi(S(n+\tau)) & \cdots & \Psi(S(n+(m-1)\tau)) \end{bmatrix} j=1,2,\cdots,n \tag{22}$$

where m and τ are embedding dimension and delay time, respectively. Each row $Y(j)$ in the matrix Y is treated as a reconstructed component, with a total of n reconstructed components, and $n = P-(m-1)\tau$. Rearrange $Y(j)=\{\Psi(S(j)), \Psi(S(j+\tau)), \cdots, \Psi(S(j+(m-1)\tau))\}$ in ascending order, $j_1, j_2, \cdots, j_m$ indicates the index of the column where each element in the reconstructed component is located, that is:

$$\Psi\left(S(f_x+(j_1-1)\tau)\right) \le \Psi\left(S(f_x+(j_2-1)\tau)\right) \le \cdots \le \Psi\left(S(f_x+(j_m-1)\tau)\right) \tag{23}$$

If there are equal values exist in the reconstructed component, which is $\Psi\left(S(f_x+(j_p-1)\tau)\right) = \Psi\left(S(f_x+(j_q-1)\tau)\right)$. At this time, the j_p and j_q values are sorted. When $j_p < j_q$, the arrangement is expressed as:

$$\Psi\left(S(f_x+(j_p-1)\tau)\right) \le \Psi\left(S(f_x+(j_q-1)\tau)\right) \tag{24}$$

Therefore, for each line $Y(j)$ of any time series reconstructed, a set of symbol sequences can be obtained as follow:

$$B(l) = (j_1, j_2, \ldots, j_m) \tag{25}$$

where $l = 1,2,\cdots,k$ and $k \le m!$. The m dimensional phase space maps different $m!$ symbol sequences, and symbol sequence $B(l)$ is one of them.

The occurrence probability of each symbol sequence is denoted as $p_1, p_2, \cdots, p_k$ and the PE of $\Psi(S(f_x))$ can be denoted as:

$$H_p(m) = -\sum_{j=1}^{m!} p_j \ln p_j \tag{26}$$

where $0 \le H_p(m) \le \ln(m!)$, when $p_j = 1/m!$, $H_p(m)$ reaches the maximum value $\ln(m!)$. The PE value is normalized as:

$$H_p = H_p(m) / \ln(m!) \tag{27}$$

The value of H_p represents the randomness degree of $\Psi(S(f_x))$. The smaller value of H_p, the more orderly the $\Psi(S(f_x))$; otherwise, the more random it is. Therefore, H_p reflects a small change in the $\Psi(S(f_x))$. In order to describe the complexity of $\Psi(S(f_x))$ more accurately, the hierarchical permutation entropy (HPE) method is proposed in Ref. [20]. The averaging operator Q_0 and high operator Q_1 are expressed as:

$$Q_0(\Psi(S)) = \frac{\Psi(S(f_x)) + \Psi(S(f_x+1))}{2} \tag{28}$$

$$Q_1(\Psi(S)) = \frac{\Psi(S(f_x)) - \Psi(S(f_x+1))}{2} \tag{29}$$

Subsequently, the operator Q_δ^σ (δ=0 or 1) at the hierarchical layer σ can be formulated as:

$$Q_\delta^\sigma = \begin{bmatrix} \frac{1}{2} & \underbrace{0\cdots0}_{2^{\sigma-1}-1} & \frac{(-1)^\delta}{2} & 0 & \cdots & 0 & 0 & 0 \\ & \frac{1}{2} & \underbrace{0\cdots0}_{2^{\sigma-1}-1} & \frac{(-1)^\delta}{2} & \cdots & 0 & 0 & 0 \\ \cdots & & & & & & & \\ 0 & 0 & 0 & 0 & \cdots & \frac{1}{2} & \underbrace{0\cdots0}_{2^{\sigma-1}-1} & \frac{(-1)^\delta}{2} \end{bmatrix} \tag{30}$$

For a given vector $[\lambda_1, \lambda_2, \cdots, \lambda_\sigma]$, the integer γ is written as:

$$\gamma = \sum_{\zeta=1}^{\sigma} 2^{\sigma-\zeta}\lambda_\zeta \tag{31}$$

where $\{\lambda_\zeta, \zeta=1,\cdots,\sigma\} \in \{0,1\}$ means the Q_0 or Q_1 at the ζ^{th} layer. According to the $[\lambda_1, \lambda_2, \cdots, \lambda_\sigma]$, the hierarchical component of $\Psi(S(f_x))$ is written as:

$$X_{\sigma,\gamma} = Q_{\lambda_\sigma}^{\sigma} \square Q_{\lambda_{\sigma-1}}^{\sigma-1} \square \cdots \square Q_{\lambda_1}^{1} \square \Psi(S(f_x)) \tag{32}$$

The $X_{\sigma,\gamma}$ is used as the input of PE to calculate HTPE, as depicted in Eq. (33):

$$\begin{cases} H_{HPE}(m) = -\sum_{j=1}^{m!} p_j(X_{\sigma,\gamma}) \ln p_j(X_{\sigma,\gamma}) \\ H_{HPE} = H_{HPE}(m) / \ln(m!) \end{cases} \tag{33}$$

where H_{HPE} can more fully enhance the fault component and eliminate random noise of the measurement signal by combining the TEO and HPE, which is helpful for rotating machinery fault diagnosis.

B. Extreme learning machine for fault recognition

Extreme learning machine (ELM) is a classification algorithm based on single hidden layer feedforward neural network (SLFN), which has better generalization performance than traditional learning algorithms. Moreover, ELM is extremely insensitive to dynamic parameters and can complete the classification of data sets faster and more conveniently. Fig. 8 depicts the schematic diagram of ELM.

For M training samples $x_j (j = 1,\ldots,M)$, where $x_j = \left[x_{j1}, x_{j2}, \cdots, x_{jc}\right]^T \in R^c$ indicates the input data, $z_j = \left[z_{j1}, z_{j2}, \cdots, z_{js}\right]^T \in R^s$ indicates the desired output label. The output function is defined as:

$$\sum_{i=1}^{N} \beta_i g(w_i \square x_j + b_i) = o_j (j = 1,2,\ldots,M \text{ and } i = 1,2,\ldots,N) \tag{34}$$

where $g(\cdot)$ denotes the sigmoid function of ELM, $w_i=[w_{i1},w_{i2},\cdots,w_{ic}]^T$ and $\beta_i=[\beta_{i1},\beta_{i2},\cdots,\beta_{is}]^T$ are the weight vector between the ith hidden node and the input node, b_i means the bias of the ith hidden node and o_j means the output of ELM for jth sample. Next, the hidden nodes H is calculated as:

$$H=\begin{bmatrix} g(w_1\cdot x_1+b_1) & \cdots & g(w_N\cdot x_1+b_N) \\ \vdots & \ddots & \vdots \\ g(w_1\cdot x_M+b_1) & \cdots & g(w_N\cdot x_M+b_N) \end{bmatrix}_{M\times N} \tag{35}$$

$$\beta=\begin{bmatrix}\beta_1\\ \vdots \\ \beta_N\end{bmatrix} \text{ and } Z=\begin{bmatrix}z_1\\ \vdots \\ z_M\end{bmatrix} \tag{36}$$

where β and Z denote the output weight vector and target matrix. In view of ELM theories, the hidden nodes (w_i,b_i) is randomly distributed rather than adjusted. The β of the Eq. (37) is calculated as:

$$\beta=H^{+}Z \tag{37}$$

where H^{+} means Moore-Penrose generalized inverse of the hidden nodes H.

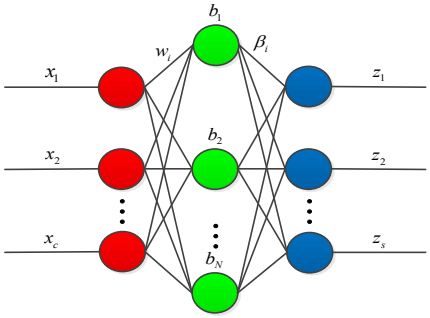

Fig. 8. Schematic diagram of ELM.

C. The proposed fault detection scheme

As mentioned above, this section develops a novel fault detection scheme based on CMMS, HTPE and ELM classifier. The framework of the fault detection scheme is illustrated in Fig. 9, and its diagnostic process is elaborated as follows:

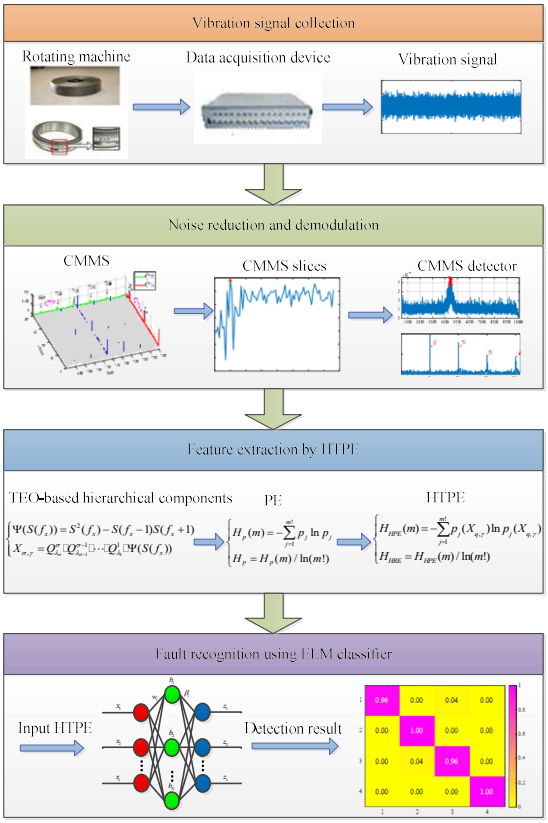

Fig. 9. Diagnostic process of the intelligent fault detection scheme.

(1) Noise reduction and demodulation. The collected vibration signals are pre-processed and demodulated using the CMMS to obtain a series of CMMS slices with different frequency bands.

(2) Frequency band selection. A novel indicator SM is utilized to determine the CMMS slices to optimize the bandwidth of CMMS to improve feature extraction performance.

(3) Feature extraction. The TEO of the CMMS slices is calculated and used as the import of the HPE to obtain the HTPE.

(4) Fault recognition. The obtained HTPE is integrated into the ELM classifier to achieve fault classification of rotating machinery under different working conditions.

IV. Experimental Validation

A. Case Ⅰ: Bearing Fault Diagnosis

To verify the availability of the proposed scheme for rolling element bearing, the faulty cylindrical roller bearing was performed to collect vibration signals. As depicted in Fig. 10, the test bench is composed of motor, bearing house, flexible type coupling and DC generator. During the experiment, three different fault defects (i.e., outer race, inner race and ball race) on cylindrical roller bearings were produced using the electrical discharge machining, and their depth was 0.5 mm and width was 0.2 mm. The sampling frequency of the faulty cylindrical roller bearing data was 71428 Hz and signal length were 300,000 data points, respectively. Geometric parameters of the faulty cylindrical roller bearing are described in Table II. The faulty cylindrical roller bearing data consists of four working conditions. There are 40 samples for each working condition and a total of 160 samples. The 80 samples are choosing randomly as the training data, and other 80 samples as the testing samples. These data messages are described in Table III.

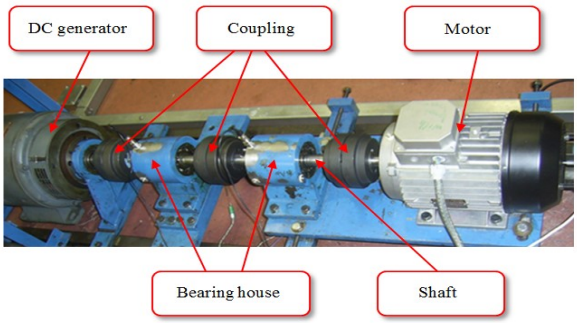

Fig. 10. The test bench of cylindrical roller bearings.

TABLE II
GEOMETRIC PARAMETERS OF CYLINDRICAL ROLLER BEARING

Bearing type	Ball numbers d (mm)	Pitch Diameter D_m (mm)	Ball Number z	Contact Angle β
N406	14	59	9	0°

Table III
DATA MESSAGES ABOUT CYLINDRICAL ROLLER BEARING

Working conditions	Class label	Number of training data	Number of testing data
Normal	1	20	20
Outer race	2	20	20
Inner race	3	20	20
Ball race	4	20	20

The waveform of the cylindrical roller bearing under four working conditions are depicted in Fig. 11. Obviously, the fault feature information is completely overwhelmed by random noise and interference frequencies. To obtain the fault feature information, the measured signal is pre-processed and demodulated using the CMMS. Subsequently, the TEO of the CMMS filtered signal is calculated and used as the import of the HPE to obtain the HTPE vector. Finally, the acquired HTPE vector is integrated into the ELM classifier. Before the implementation of the ELM classifier, the number of hidden neurons is set to 20, and the sigmoidal function is applied [30]. Fig.12 illustrates the confusion matrix of the proposed scheme in the first trial. As depicted in Fig 12, the minimum detection accuracy rate is 95% and the average detection accuracy rate is 98.75%. This reveals that the CMMS-HTPE and ELM classifier can distinguish the different working conditions of cylindrical roller bearings, and the diagnosis accuracy is also desirable.

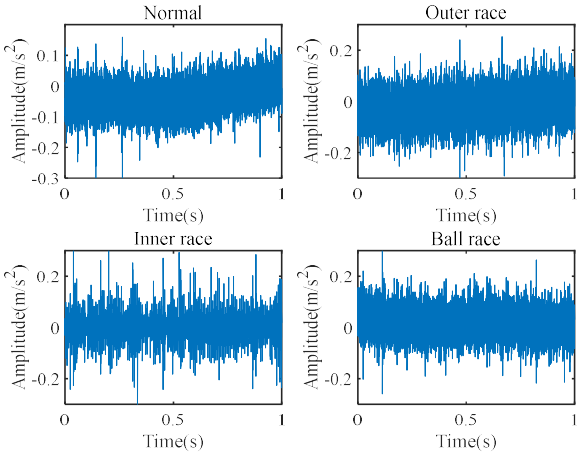

Fig. 11. Waveform of cylindrical roller bearing fault.

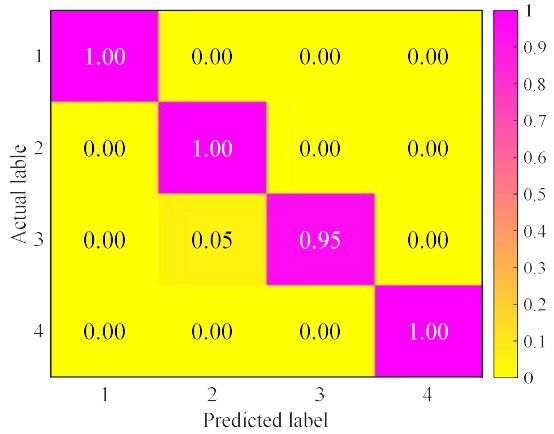

Fig. 12. Confusion matrix of the proposed scheme in the first trial.

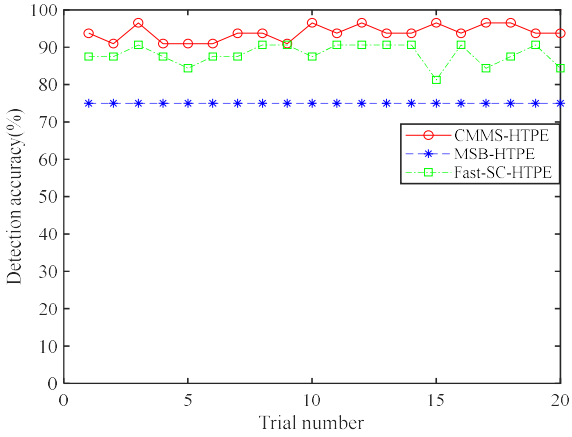

Fig. 13. Detection accuracy of three methods in 20 trials.

To prove the availability of the proposed scheme, the measured signals of cylindrical roller bearing are applied to contrast the performance between CMMS-HTPE, MSB-HTPE and Fast-SC-HTPE. In addition, considering the randomness of the fault diagnosis results in one trial, the cylindrical roller bearing data is carried out twenty times in this study. Fig. 13 presents the detection accuracy of three algorithms for twenty trials. It can be found that compared with MSB-HTPE and Fast-SC-HTPE, CMMS-HTPE has the most outstanding detection accuracy. Thus, it is deduced that CMMS-HTPE is more robust than MSB-HTPE and Fast-SC-HTPE.

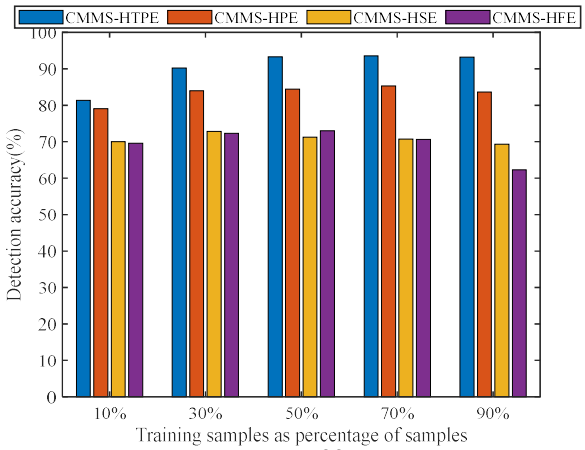

Fig. 14. Average detection accuracy of four methods in 20 trials under different percentages of training samples.

Table IV
DETECTION ACCURACY OF THE FOUR METHODS UNDER DIFFERENT PERCENTAGES OF TRAINING SAMPLES

Methods		Accuracy rate acquired under different percentages of training samples (%)					Average CPU time (s)
		10%	30%	50%	70%	90%	
CMMS-HTPE	Max	89.58	95.54	95.00	94.94	96.53	4.07
	Min	73.96	83.63	90.63	91.96	85.42	
	Mean	81.39	90.20	93.28	93.57	93.19	
	SD	4.54	2.26	1.14	0.75	3.07	
CMMS-HPE [20]	Max	87.84	88.69	87.50	87.50	91.32	3.82
	Min	69.10	77.23	78.13	82.29	71.88	
	Mean	79.08	83.95	84.38	85.30	83.63	
	SD	5.66	2.93	2.48	1.81	4.29	
CMMS-HSE [14]	Max	80.56	78.13	75.00	79.46	78.82	4.21
	Min	62.15	68.30	62.50	64.58	52.08	
	Mean	70.05	72.80	71.25	70.76	69.34	
	SD	4.88	2.86	3.45	3.51	5.67	
CMMS-HFE [16]	Max	79.17	76.93	81.25	77.83	76.74	4.14
	Min	56.94	61.01	62.50	61.76	56.60	
	Mean	69.57	72.29	72.97	70.60	62.28	
	SD	6.26	3.65	4.56	3.88	5.81	

For comparison, the advanced fault detection method based on permutation entropy (e.g., Hierarchical permutation entropy (HPE), Hierarchical sample entropy (HSE) and Hierarchical fuzzy entropy (HFE)) are utilized to analyze the same data samples. Moreover, we set the parameters of all algorithms to the same value to process the experimental data, where m means embedding dimension, equal to 6, τ indicates time delay, set to 1, σ represents standard deviation (SD) of signal, q means hierarchical layer, set to 3, and r indicates the tolerance and is equal to 0.15 times of c. Concretely, for HTPE and HPE, we set m=6, τ=1 and q=3. For HSE and HFE, we set m=6, q=3 and r=0.15×σ, according to [19]. Here, we will set the percentage of training samples as: 10%, 30%, 50%, 70% and 90%. Hereafter, the detection accuracy of the above four methods is calculated in 20 trials, as illustrated in Fig. 14. As depicted in Fig. 14, the CMMS-HTPE has the highest detection accuracy compared with the other three algorithms (CMMS-HPE, CMMS-HSE and CMMS-HFE) regardless of the percentage of training samples. In addition, Table IV lists the comparison indicators of the four algorithms in 20 trials, including the maximum (Max), minimum (Min), mean and SD of classification accuracy, as well as the average CPU time. As displayed in Table IV, the minimum and maximum recognition accuracy of the detection result of the proposed scheme is higher than that of the three algorithms (e.g., CMMS-HPE, CMMS-HSE and CMMS-HFE) under different percentages of training samples. At the same time, the average CPU time of the proposed scheme is 4.07 s, which is lower than the CPU time of CMMS-HSE and CMMS-HFE, but higher than that of CMMS-HPE. The main reason that the calculation efficiency of the proposed scheme is lower than that of CMMS-HPE is that TEO is added to HPE. Moreover, the proposed scheme has the smallest SD, which reveals that the proposed scheme has the best detection ability.

B. Case Ⅱ: Planetary Gearbox Fault Diagnosis

In this section, the effectiveness of the proposed scheme will be further investigated on the planetary gearbox vibration signal obtained from the test platform presented in Fig. 15. Test platform consists of motor, planetary gearboxes, helical gearbox, load generator and vibration sensor are horizontally fixed on top of the planetary gearbox. In this experiment, the failure modes mainly include sun gear chipping and sun gear misalignment. The sampling frequency of the experiment data was 96 kHz and signal length were 300,000 data points, respectively. Geometric parameters of the planetary gearbox are described in Table V. There are 40 samples under each working condition for a total of 160 samples. The 80 samples are selected randomly as training samples and the other 80 samples as testing samples. The data messages are given in Table VI.

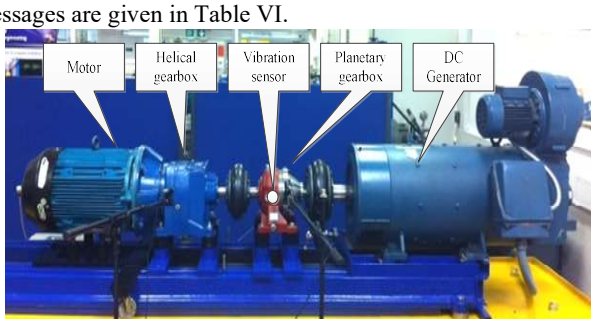

Fig. 15. The test platform for the planetary gearbox facility.

Table V
GEOMETRIC PARAMETERS OF PLANETARY GEARBOX

Parameter	Number of teeth
Sun gear	10
Planet gear (number)	26(3)
Ring gear	62
Carrier	-

Table VI
DATA MESSAGES ABOUT PLANETARY GEARBOX

Working conditions	Class label	Number of training data	Number of testing data
Normal	1	20	20
Sun gear chipping	2	20	20
Sun gear with 0.4 mm misalignment	3	20	20
Sun gear with 1.0 mm misalignment	4	20	20

Fig. 16 presents the waveform of the sun gear under four working conditions. From the waveform, the transient impulse information cannot be identified. The proposed scheme is utilized to process the faulty signals of the sun gear. Firstly, CMMS is used to perform noise reduction and signal demodulation on the measurement signal. Subsequently, the HTPE of the CMMS filtered signal is calculated to establish the CMMS-HTPE vector. Finally, the CMMS-HTPE vector is input into the ELM classifier to identify the type of sun gear failure. The confusion matrix of the proposed scheme in the first trial is presented in Fig. 17. It can be observed that several samples are misclassified, and the overall detection accuracy rate is 95%. This indicates that the CMMS-HTPE and ELM classifier can effectively recognize sun gear's fault features under different operating conditions.

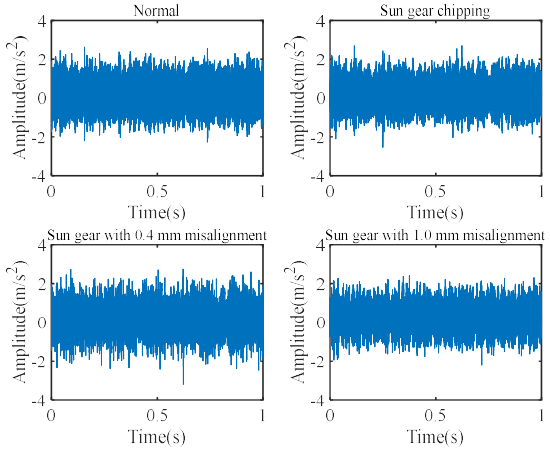

Fig. 16. Waveform of sun gear.

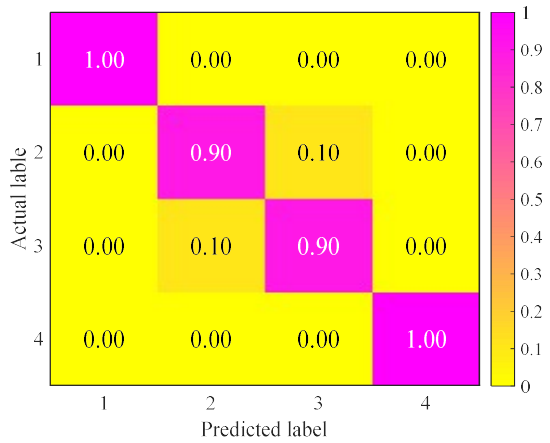

Fig. 17. Confusion matrix of the proposed scheme in the first trial.

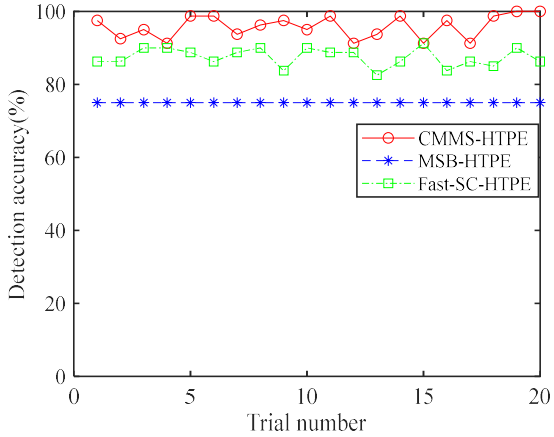

Fig. 18. Detection accuracy of three methods in 20 trials.

Likewise, considering the randomness of the fault detection results in one trial, the sun gear data was carried out twenty times. In addition, other two algorithms (MSB-HTPE and Fast-SC-HTPE) are utilized to process the measured signals of sun gear under different working conditions. Fig. 18 illustrates the detection accuracy of three algorithms for twenty trials. It can be found that compared with (MSB-HTPE and Fast-SC-HTPE), CMMS-HTPE has the most outstanding detection accuracy. Similarly, Section IV. A, for HTPE and HPE, we set m=6, τ=1 and q=3. For HSE and HFE, we set m=6, q=3 and r=0.15×σ. Besides, we also will set the percentage of training samples as: 10%, 30%, 50%, 70% and 90%. Fig. 19 depicts the average detection accuracy of the above four algorithms in 20 trials. As illustrated in Fig. 19, compared with CMMS-HPE, CMMS-HSE and CMMS-HFE, the CMMS-HTPE has the highest detection accuracy under different percentages of training samples. Moreover, Table VII shows the statistical indicators of the detection results, which are Min, Max, SD and average CPU time. As depicted in Table VII, the minimum and maximum recognition accuracy of the detection result of the proposed scheme is higher than that of the other three algorithms (e.g., CMMS-HPE, CMMS-HSE and CMMS-HFE) under different percentages of training samples. Meanwhile, the average CPU time of the proposed scheme is 1.48 s, which is lower than the CPU time of CMMS-HSE and CMMS-HFE, but higher than that of CMMS-HPE. In addition, the proposed scheme has the smallest SD, which demonstrates that the proposed scheme has the best detection performance.

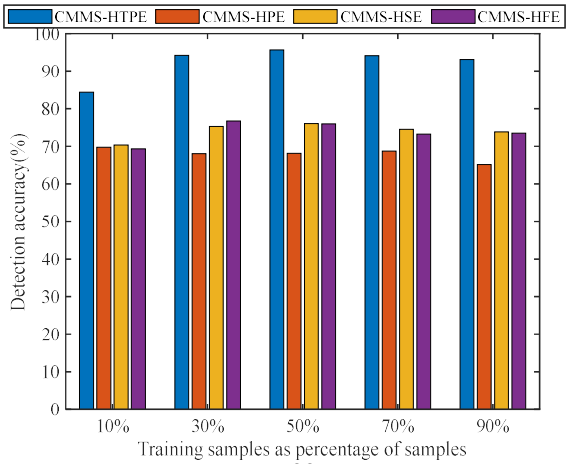

Fig. 19. Average detection accuracy of four methods in 20 trials under different percentages of training samples.

Table VII
DETECTION ACCURACY OF THE FOUR METHODS UNDER DIFFERENT PERCENTAGES OF TRAINING SAMPLES

Methods		Accuracy rate acquired under different percentages of training samples (%)					Average CPU time (s)
		10%	30%	50%	70%	90%	
CMMS-HTPE	Max	89.58	100	100	100	100	1.48
	Min	73.96	90.48	90.63	89.73	88.19	
	Mean	84.38	94.21	95.63	94.15	93.09	
	SD	3.98	2.75	2.36	2.93	3.23	
CMMS-HPE [20]	Max	78.13	72.77	71.88	73.36	74.31	1.26
	Min	62.85	62.05	59.38	62.35	55.90	
	Mean	69.74	68.09	68.13	68.71	65.19	
	SD	5.36	3.14	3.30	3.49	5.80	
CMMS-HSE [14]	Max	79.17	82.00	84.38	82.44	84.03	1.63
	Min	62.15	68.90	71.88	64.58	68.40	
	Mean	70.38	75.26	76.09	74.55	73.87	
	SD	4.54	3.32	2.92	4.24	4.03	
CMMS-HFE [16]	Max	77.43	84.23	81.25	82.44	82.29	1.55
	Min	53.13	71.73	68.75	65.18	53.82	
	Mean	69.34	76.71	75.94	73.26	73.47	
	SD	5.05	3.65	3.67	4.69	7.83	

V. CONCLUSION

In this paper, a novel intelligent detection algorithm based on CMMS, HTPE and ELM is put forward, which can automatically realize fault classification of rotating machinery. Case studies of rolling bearings and planetary gearboxes verify the effectiveness of the proposed scheme in fault classification and identification. The main conclusions of this paper are summarized as follows:

(1) A new cyclic morphological modulation spectrum (CMMS) is put forward for fault feature extraction. Simulation and experimental results reveal that CMMS has stronger denoising and demodulation performance than MSB and Fast-SC algorithms, indicating the superiority of using CMMS in fault feature extraction.

(2) The hierarchical Teager permutation entropy (HTPE) is proposed to select CMMS features. Experimental analysis shows that CMMS-HTPE has better detection performance than some existing algorithms (e.g., CMMS-HPE, CMMS-HSE and CMMS-HPE).

(3) The case study of rotating machinery shows that the proposed scheme realizes the classification and recognition of failure modes under different working conditions, and the detection accuracy is higher than that of the methods (e.g., MSB-HTPE, Fast-SC-HTPE, CMMS-HPE, CMMS-HSE and CMMS-HPE) under the ELM classifier. Thence, the proposed scheme is proved to be effective for identifying different operating conditions of rotating machinery.

REFERENCES

[1] Q. Li, X. Ding, Q. He, W. Huang and Y. Shao, "Manifold sensing-based convolution sparse self-learning for defective bearing morphological feature extraction," *IEEE Trans. Ind. Informat*., vol. 17, no. 5, pp. 3069-3078, May. 2021.

[2] S. Xing, Y. Lei, S. Wang and F. Jia, "Distribution-invariant deep belief network for intelligent fault diagnosis of machines under new working conditions," *IEEE Trans. Ind. Electron*., vol. 68, no. 3, pp. 2617-2625, Mar. 2021.

[3] K. Zheng, T. Li, Z. Su, and B. Zhang, "Sparse elitist group lasso denoising in frequency domain for bearing fault diagnosis," *IEEE Trans. Ind. Informat*., vol. 17, no. 7, pp. 4681-4691, Jul, 2021.

[4] B. Yang, R. Liu, X. Chen, "Fault diagnosis for a wind turbine generator bearing via sparse representation and shift-invariant K-SVD," *IEEE Trans. Ind. Informat*., vol. 13, no. 3, pp. 1321-1331, Jun. 2017.

[5] H. Sun, S. Yuan and Y. Luo, "Cyclic spectral analysis of vibration signals for centrifugal pump fault characterization," *IEEE. Sens. J*., vol. 18, no. 7, pp. 2925-2933, Apr. 2018.

[6] Q. Li and X. Jing, "Fault diagnosis of bolt loosening in structures with a novel second-order output spectrum–based method," *Struct. Health Monit*., vol. 19, no. 1, pp. 123-141, Mar. 2019.

[7] J. Antoni and D. Hanson, "Detection of surface ships from interception of cyclostationary signature with the cyclic modulation coherence," *IEEE. J. Ocean. Eng*., vol. 37, no. 3, pp. 478-493, Jul. 2012.

[8] Z. Feng, M. Zuo, R. Hao, F. Chu and M.E. Badaoui, "Gear damage assessment based on cyclic spectral analysis," *IEEE Trans. Rel*., vol. 60, no. 1, pp. 21-32, Mar. 2011.

[9] Z. Chen, A. Mauricio, W. Li and K. Gryllias, "A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks," *Mech. Syst. Signal Process*., vol. 140, pp. 106683, Jun. 2020.

[10] J. Guo, D. Zhen, H. Li, Z Shi, F. Gu and A.D. Ball, "Fault detection for planetary gearbox based on an enhanced average filter and modulation signal bispectrum analysis," *ISA. Trans*., vol. 101, pp. 408-420, Jun, 2020.

[11] P. W. Tse and D. Wang, "The design of a new sparsogram for fast bearing fault diagnosis: Part 1 of the two related manuscripts that have a joint title as "Two automatic vibration-based fault diagnostic methods using the novel sparsity measurement - Parts 1 and 2," *Mech. Syst. Signal Process*., vol. 40, no. 2, pp. 499-519, Nov. 2013.

[12] D. Wang, X. Zhao, L. Kou, Y. Qin, Y. Zhao and K. Tsui, "A simple and fast guideline for generating enhanced/squared envelope spectra from spectral coherence for bearing fault diagnosis," *Mech. Syst. Signal Process*., vol. 122, pp. 754-768, Jan. 2019.

[13] D. Wang, J. Zhong, C. Shen, E. Pan, Z. Peng and C. Li, "Correlation dimension and approximate entropy for machine condition monitoring: Revisite," *Mech. Syst. Signal Process*., vol. 152, pp. 107497, May. 2021.

[14] Y. Li, K. Feng, X. Liang and M. J. Zuo, "A fault diagnosis method for planetary gearboxes under non-stationary working conditions using improved Vold-Kalman filter and multi-scale sample entropy," *J. Sound. Vib*., vol. 439, pp. 271-286, Jan. 2019.

[15] H. Xie, Z Gao and H Liu, "Classification of ventricular tachycardia and fibrillation using fuzzy similarity-based approximate entropy," *Expert. Syst. Appl*., vol. 38, no. 4, pp. 3973-3981, Apr. 2011.

[16] Y. Li, M. Xu, R. Wang and W. Huang, "A fault diagnosis scheme for rolling bearing based on local mean decomposition and improved multiscale fuzzy entropy," *J. Sound. Vib*., vol. 360, pp. 277-299, Jan. 2016.

[17] C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Phys. Rev. Lett*., vol. 88, pp. 174102. Apr. 2002.

[18] J. Zheng, H. Pan, S. Yang and J. Cheng, "Generalized composite multiscale permutation entropy and Laplacian score based rolling bearing fault diagnosis," *Mech. Syst. Signal Process*., vol. 99, pp. 229-243, Jan. 2018.

[19] Y. Li, G. Li, Y. Yang, X. Liang and M. Xu, "A fault diagnosis scheme for planetary gearboxes using adaptive multi-scale morphology filter and modified hierarchical permutation entropy," *Mech. Syst. Signal Process*., vol. 115, pp. 319-337, May. 2018.

[20] C. Yang and M. Jia, "Hierarchical multiscale permutation entropy-based feature extraction and fuzzy support tensor machine with pinball loss for bearing fault identification," *Mech. Syst. Signal Process*., vol. 149, pp. 107182, Feb. 2021.

[21] Y. Li, Y. Yang, X. Wang, B. Liu and X. Liang, "Early fault diagnosis of rolling bearings based on hierarchical symbol dynamic entropy and binary tree support vector machine," *J. Sound. Vib*., vol. 428, pp. 72-86, Aug. 2018.

[22] R. B. Randall and W. A. Smith, "Uses and mis-uses of energy operators for machine diagnostics," *Mech. Syst. Signal Process*., vol. 133, pp. 106199, Nov. 2019.

[23] J. Song, J. Zhao, X. Zhang, F. Dong, J. Zhao, L. Xu and Z. Yao, "Accurate demagnetization faults detection of dual-sided permanent magnet linear motor using enveloping and time-domain energy analysis," *IEEE Trans. Ind. Informat*., vol. 16, no. 10, pp. 6334-6346, Oct. 2020.

[24] C. Zhao, L. Kai, Y. Li, L. Wang, Y. Luo, X. Xu, X. Ding and Q. Meng, "Novel method based on variational mode decomposition and a random discriminative projection extreme learning machine for multiple power quality disturbance recognition," *IEEE. T. Ind. Informat.*, vol. 15, no. 5, pp. 2915-2926, May. 2019.

[25] X. Lu, F. Yin, C. Liu and M. Huang, "Online spatiotemporal extreme learning Machine for complex time-varying distributed parameter systems," *IEEE. T. Ind. Informat*., vol. 13, no. 4, pp. 1753-1762, Aug. 2017.

[26] X. Tian, J. Gu, I. Rehab, G. Abdalla, F. Gu and A.D. Ball, "A robust detector for rolling element bearing condition monitoring based on the modulation signal bispectrum and its performance evaluation against the Kurtogram," *Mech. Syst. Signal Process*., vol. 100, pp. 167-187, Feb. 2018.

[27] J. Antoni, G. Xin and N. Hamzaoui, "Fast computation of the spectral correlation," *Mech. Syst. Signal Process*., vol. 92, pp. 248-277, Feb. 2017.

[28] J. Guo, Z Shi, D. Zhen, Z. Meng, F. Gu and A.D. Ball, "Modulation signal bispectrum with optimized wavelet packet denoising for rolling bearing fault diagnosis," *Struct. Health Monit*., vol. 21, pp. 984-1011, Feb. 2022.

[29] L. Meng, J. Xiang, Y. Wang, Y. Jang and H. Gao, "A hybrid fault diagnosis method using morphological filter-translation invariant wavelet and improved ensemble empirical mode decomposition," *Mech. Syst. Signal Process*., vol. 50-51, pp. 101-115, Jan. 2015.

[30] X. Yan and M. Jia, "Intelligent fault diagnosis of rotating machinery using improved multiscale dispersion entropy and mRMR feature selection," *Knowl-based. Syst*., vol. 163, pp. 450-471, Jan. 2019.

Junchao Guo received the B.S. degree from in mechanical engineering from the North China University of Science and Technology, Tangshan, China, in 2016, received the Ph.D. degree in mechanical engineering from the Hebei University of Technology, Tianjin, China, in 2021. He is currently an assistant researcher, Shanghai Jiao Tong University, Shanghai, China.

His current research interests include fault diagnosis, signal processing, intelligent monitoring and pattern recognition.

Qingbo He (Senior Member, IEEE) received the B.S. and Ph.D. degrees in mechanical engineering from the University of Science and Technology of China, Hefei, China, in 2002 and 2007, respectively.

He was a Research Associate with The Chinese University of Hong Kong, Shatin, Hong Kong, from 2007 to 2008. He was a Post-Doctoral Researcher with the University of Massachusetts, Amherst, MA, USA, and the University of Connecticut, Storrs, CT, USA, from 2008 to 2009. He was a faculty with the University of Science and Technology of China, Hefei, China, for eight years. He is currently a Professor with the State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai, China. His current research interests include a combination of vibration analysis, signal processing, and metamaterials design for intelligent monitoring, diagnosis, and control in complex machines.

Dong Zhen received the B.S. degree from Shandong University of Science and

Technology, China, in 2006, received the M.S. degree in signal and information processing from Shandong University of Science and Technology, China, in 2009, received the Ph.D. degree in mechanical engineering from University of Huddersfield, UK, in 2012. He is currently a professor and associate dean of the School of Mechanical Engineering, Hebei University of Technology, Tianjin, China.

His main research interests include machinery condition monitoring and fault diagnostics, signal processing and intelligent algorithms development.

Fengshou Gu received the B.S. degree from Taiyuan University of Technology, China, in 1979, received the M.S. degree form Taiyuan University of Technology, China, in 1985, received the Ph.D. degree in mechanical engineering from Manchester University, UK, in 2008. He is currently a professor and the Deputy Dean of the Centre for Efficiency and Performance Engineering (CEPE), University of Huddersfield, UK.

His research interests include condition monitoring and fault diagnosis of rotating machines, signal processing, sensing technology and the applications of intelligent diagnosis technology.