Feature Extraction Based on EWT With Scale Space Threshold and Improved MCKD for Fault Diagnosis

LINGFENG LI¹, (Member, IEEE), AIBIN GUO², (Member, IEEE), AND HUAYUE CHEN³

1School of Electronics and Information Engineering, Dalian Jiaotong University, Dalian 116028, China
2Haifeng General Aviation Technology Company Ltd., Beijing 100070, China
3School of Computer Science, China West Normal University, Nanchong 637002, China

Corresponding author: Huayue Chen (sunnyxiaoyue@163.com)

This work was supported by the National Natural Science Foundation of China under Grant 51605068 and Grant 51475065.

ABSTRACT Aiming at the problem of feature extraction of non-stationary, non-linear and weak fault signals, a new feature extraction method based on empirical wavelet transform (EWT) with scale space threshold (STEWT) and improved maximum correlation kurtosis deconvolution (MCKD) with power spectral entropy and grid search (PGMCKD), namely STEWT-PGMCKD is proposed for rolling bearing faults in this paper. In the proposed STEWT-PGMCKD method, the scale space threshold method is designed to solve the problems of falling into local extremum and mode over decomposition caused by the local-max-min band decomposition method of EWT, which is used to decompose the frequency band of signal, and the correlation analysis is carried out between the decomposed modal components and the original signal to retain the modal components with high correlation. Then an adaptive MCKD based on power spectral entropy is proposed to solve the problem that the signal processing effect of MCKD is affected by filter size \(L\) and deconvolution period \(T\). Nextly, the parameters of the MCKD are optimized by grid search method. Finally, the power spectrum analysis of the enhanced signal is carried out to realize the feature extraction and fault diagnosis. The experiment results show that the proposed STEWT-PGMCKD method can effectively extract the weak fault information and accurately realize the fault diagnosis for rolling bearings.

INDEX TERMS Rolling bearing, feature extraction, empirical wavelet transform, maximum correlation kurtosis deconvolution, scale space threshold method, power spectral entropy, grid search.

I. INTRODUCTION

As the core component of typical rotating machinery, rolling bearing plays an important role in the effective operation of the whole mechanical system, and also becomes one of the components with frequent faults. It is necessary and significant to monitor the states of rolling bearings and find out the fault in time [1]–[5]. The fault signal of rolling bearings is non-stationary and non-linear noise signal, and the fault location is often in the position where the signal is not easy to collect, which leads to the problem of signal weakening and interference by the external environment in the transmission process of fault signal, which brings difficulties to the fault diagnosis of rolling bearings [6]–[11]. It is the key to find an effective method which can effectively extract fault features under weak signal and strong noise interference [12]–[16]. In recent year, a lot of methods are proposed by some researchers [17]–[21].

For the rolling bearing fault feature extraction, empirical mode decomposition (EMD), ensemble empirical mode decomposition (EEMD), wavelet transform (WT) and other time-frequency analysis methods are widely used. Lou and Loparo [22] proposed a bearing fault classification method combining wavelet transform with fuzzy inference. To increase the computation speed of biorthogonal WT, lifting wavelet transform was proposed by Sweldens [23], and it was successfully applied to detecting the fault of a helicopter’s planetary transmission system. Lin and Qu [24] proposed a feature extraction method based on Morlet wavelet, and this method was greatly applied to mechanical
fault diagnosis. Qin et al. [25] proposed M-band flexible wavelet transform for identifying the underlying fault features in measured signals. But these methods have certain limitations and are susceptible to interference. EMD is an adaptive signal time-frequency analysis method proposed by Huang et al. [26] in 1998. This method has achieved satisfactory results in dealing with nonlinear and non-stationary signals, and has been widely used. But EMD is limited to use due to mode mixing and end effects. Although EEMD alleviates the problem of mode mixing, but the number of iterations will increase significantly, resulting in a sharp increase in computation. Dragomiretskiy and Zosso [27] proposed Variational Mode Decomposition (VMD) in 2014. By constructing and solving variational problems, the centre frequency of each IMF and its bandwidth are determined. Finally, the signal is adaptively decomposed into K IMFs. This method not only retains the advantages of dealing with non-linear and non-stationary signals, but also overcomes the disadvantages of mode mixing and end effects. However, the processing effect of VMD is greatly affected by K, improper selection of K will lead to over decomposition or under decomposition. Empirical wavelet transform (EWT) is a new adaptive signal decomposition method proposed by Gilles [28]. This method is based on empirical mode decomposition and wavelet transform. The EWT overcomes the problems that wavelet bases cannot be selected adaptively according to signal characteristics, and EMD is lack of mathematical basis and prone to mode aliasing and endpoint effect. It can adaptively select a set of wavelet filter banks to extract different AM-FM components of the signal according to the maximum value of the Fourier spectrum. A lot of researchers have done relevant research based on EWT. Bhattacharyya et al. [29] proposed a signal decomposition method of the EEG signal using empirical wavelet transform technique. A sparsity guided empirical wavelet transform is proposed to automatically establish Fourier segments required in empirical wavelet transform for fault diagnosis of rolling element bearings [30]. Zheng et al. [31] proposed an adaptive parameterless empirical wavelet transform based time-frequency analysis method and its application to rotor rubbing fault diagnosis. Ou et al. [32] proposed a method of rolling bearing fault diagnosis using improved majorization-minimization-based total variation and empirical wavelet transform. Pan et al. [33] proposed a modified EWT (MEWT) method via data driven adaptive Fourier spectrum segment is proposed for mechanical fault identification. Yuan et al. [34] proposed a new technique that combines the second-order blind identification (SOBI) method with the EWT to delineate closely-spaced frequencies. However, due to the fact that the actual mechanical vibration signal is more complex and easily interfered by strong noise, the local-max-min method for interval segmentation is prone to modal redundancy or falling into local extremum. Therefore, it is necessary to study the optimization strategy of interval segmentation.

Maximum correlated kurtosis deconvolution (MCKD), as an algorithm to enhance the periodic component in the signal, can highlight the periodic pulse signal in the original signal to the greatest extent. Therefore, it has been widely used in the field of fault diagnosis for bearings. Miao et al. [35] proposed an improved MCKD method for fault diagnosis of rolling element bearings. The parameterized time-frequency transform methods are proposed to solve extraction the IRF accurately due to the strong nonstationary properties of the signal [36]. Zhang et al. [37] proposed a bearing fault diagnosis and degradation analysis based on improved EMD and MCKD. McDonald et al. [38] proposed a new deconvolution method for the detection of gear and bearing faults from vibration data. Hong et al. [39] proposed a method combining customized balanced multiwavelets and adaptive MCKD for rotating mechanical compound faults diagnosis. Jia et al. [40] proposed an improved spectral kurtosis (SK) method based on maximum correlated kurtosis deconvolution (MCKD) to extract the weak fault characteristics of bearings from the signals. Tang et al. [41] proposed a new diagnosis method based on Adaptive maximum correlated kurtosis deconvolution (AMCKD) for accurate identification of compound faults of rolling bearings. Wang et al. [42] proposed a MCKD-CEEMD-ApEn method to denoise and classify the combined failure of slewing bearings. Lyu et al. [43] proposed an improved maximum correlated kurtosis deconvolution method based on quantum genetic algorithm, named QGA-MCKD for gear and bearing compound fault diagnosis. Cheng et al. [44] proposed the optimal minimum entropy deconvolution adjusted and multipoint optimal minimum entropy deconvolution adjusted (MOMEDA) for enhancing the impulse-like component in the fault signal. Zhang et al. [45] proposed a new detected method combining complementary advantages of variational mode decomposition (VMD) and adaptive maximum correlated kurtosis deconvolution, named VMD-AMCKD. Wang et al. [46] proposed multipoint optimal minimum entropy deconvolution adjusted to the fault diagnosis of gearbox. Shen et al. [47] proposed a novel signal processing method based on the particle swarm optimization, maximum correlated kurtosis deconvolution, variational, mode decomposition and fast spectral kurtosis (PSO-MCKD-VMD-FSK) to extract fault characteristics for the signal-to-noise ratio and uneven energy distribution problems. Zhang et al. [48] proposed an improved maximum correlation kurtosis deconvolution based on grasshopper optimization algorithm. Liang et al. [49] employed an optimized Morlet wavelet as the initial filter in the deconvolution process, which contributes to improving both the efficiency and performance of MAKD. However, the parameters of MCKD are strict. Only when the unwinding period matches the corresponding failure period, the MCKD can give full play to the optimal effect. Therefore, it is necessary to study the parameter selection of the MCKD. In recent year, some optimization algorithms are proposed to determine the parameters for it [50]–[54].
Aiming at the deficiency of local local-max-min method of EWT for interval segmentation, the EWT with scale space threshold is deeply studied, and the an adaptive MCKD based on power spectral entropy is proposed to solve the problem that the signal processing effect of MCKD is affected by filter size $L$ and deconvolution period $T$. The grid search method is used to optimize the parameters of the MCKD. Firstly, the vibration signal is decomposed by EWT with scale space threshold, and the AM-FM component is analyzed by envelope spectrum. Then the fault feature is enhanced and extracted by the MCKD with power spectral entropy and grid search. The rolling bearing fault signal of QPZZ-II is used to prove the effectiveness of the proposed STEWT-PRGMCCKD method.

The main contributions are summarized as follows:

- A novel feature extraction method based on integrating EWT with scale space threshold and improved MCKD with energy entropy and grid search was proposed.
- The scale space threshold method is designed to improve the EWT to avoid to fall into local extremum and mode over decomposition.
- The energy entropy and grid search method are combined to adaptively select and optimize the filter order $L$ and deconvolution period $T$ of the MCKD.

The rest of this paper is arranged as follows. Section 2 presents empirical wavelet transform with scale space threshold. The maximum correlation kurtosis deconvolution with power spectral entropy and grid search is described in Section 3. Section 4 introduced a novel fault diagnosis method. The experimental results on rolling bearing fault signal of QPZZ-II are shown and analyzed in Section 5. Finally, the conclusion is given in Section 6.

II. EMPIRICAL WAVELET TRANSFORM WITH SCALE SPACE THRESHOLD

EWT is a new signal processing method proposed by Gilles, which extracts different AM-FM components of signal by adaptively selecting a set of wavelet filter banks according to the Fourier spectrum characteristics of signal. In order to select a suitable wavelet filter, the Fourier spectrum is defined in $[0, \pi]$ in signal analysis according to Shannon criterion, and the Fourier spectrum is segmented adaptively. After the partition interval is determined, empirical wavelet is defined as band-pass filter in each interval. Gilles constructs empirical wavelet with empirical wavelet function and empirical scaling function according to Meyer wavelet construction method.

The dividing the central frequency band of the traditional EWT method is based on the local minimum between two adjacent maxima in the Fourier spectrum, which is easy to fall into the local minimum and cause modal redundancy. In this paper, the central frequency band is divided based on the scale space threshold method. The scale space is based on the Fourier transform of the signal, through further internal operation to obtain the spectrum segmentation information. Based on the scale space representation, the STEWT method is used to classify the spectrum of the original signal by using the fuzzy C-means method in order to obtain the spectrum division interval. In this paper, the fuzzy C-means method divides the minimum length curve set into 2 clusters (meaningful and non-meaningful). Thus, the original signal is decomposed into several intrinsic mode functions (IMFs). Finally, Pearson correlation coefficient is made between the IMFs and the original signal, and the threshold value is set to retain the IMF component which is larger than the threshold value, and then the power spectral density is analyzed to extract the fault characteristic frequency.

A. SCALE SPACE REPRESENTATION

Let $f(x)$ in $[0, x_{\text{max}}]$, define Gaussian kernel function

$$g(x; t) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/(2t)}$$  \hspace{1cm} (1)

then the scale space is expressed as

$$L(x, t) = g(x; t) * f(x)$$  \hspace{1cm} (2)

where, $*$ represents convolution, $t$ represents the scale parameter.

In practical application, the Fourier transform $f(\omega)$ of the original signal and the sampled Gaussian kernel function are used to describe the scale space representation, which is shown as follow.

$$L(\omega, t) = \sum_{n=-M}^{+M} g(n; t) * f(\omega - n)$$  \hspace{1cm} (3)

When $M$ is large enough, the error of Gaussian approximation is negligible. In this paper, $M = C \sqrt{T} + 1$, $3 < C \leq 6$. In order to satisfy the requirement that the error is small enough, $C = 6$.

Under the initial scale parameter ($\sqrt{T_0} = 0.5$), the minimum between the local maximum values can be obtained in the scale space representation $L(\omega, \omega_0)$. When $t = 0$, the number of local minimums is recorded as $N_0$. These local minimums constitute “the scale space curve” $C_t (i \in \{1, N_0\})$. In the scale space representation plane $L(\omega, t)$, $N_0$ calculated at $t = 0$ is recorded as the number of minimum length curves $L_t$, and as the scale $t$ increases, the number of local minimums is decreasing until the number of minimums is reduced to 0. At this time, the value of the minimum length curve $L_t$ is calculated, starting with the first “scale space curve” $C_{1t}$ of length 1 represented by the local minimum of $f(\omega)$ at $t = 0$, and sequentially accumulating “the scale space curve” at different scales $\sqrt{T_k}$ ($k = 1, 2, \cdots 2N_0$) and gets the first “minimum length curve” $L_{ti}$ ($i = 1, 2, N_0$). From this, a proportional threshold $T$ can be set from the minimum length curve $L_t$, that is to say all minimum length curves corresponding to the proportional threshold $T$ are found. The problem is transformed into a clustering problem. So that we can use fuzzy C-mean method to divide the set $\{L_{ti}\}_{i=1}^{N_0}$ clusters into two clusters (meaningful and non-meaningful minima).
B. THE IMF SELECTION
Due to the strong noise interference in the actual operation environment, IMFs have the following characteristics after EWT decomposition: 1) there are differences in the fault information contained in every IMF; 2) there are still interference components in the IMF containing fault information, which cause interference to the power spectrum analysis of the signal, so it is unable to extract the fault frequency from the IMF.

Therefore, this paper introduces Pearson correlation coefficient, by comparing the correlation between the IMF and the original signal, to retain the IMF which has greater correlation with the original signal in order to further process the signal. For the two set of data \( \{x_i\} \) and \( \{y_i\}, i = 1, 2, \cdots, n \), their correlation coefficient can be expressed as:

\[
r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}}
\]

where are \( \bar{x} \) and \( \bar{y} \) are the mean of data. The correlation coefficient \( r \) ranges between \(-1 \) and 1. If \( r = \pm 1 \), it indicates a completely positive or negative linear correlation between \( x \) and \( y \); If \( r = 0 \), it means that there is no linear correlation between them.

When the Pearson correlation coefficients of every IMF are obtained, it is necessary to set a threshold to retain IMFs with more fault information as much as possible. The value of the threshold affects the signal processing results. Based on the experience of the experiment, it is usually better to set the threshold to 0.2 for signal processing.

C. THE STEPS OF EWT WITH SCALE SPACE THRESHOLD
In this paper, the scale space threshold method is used to segment the frequency band, and the IMF components after reconstruction are analyzed by correlation analysis, which is more relevant to the original signal. The specific steps are described as follows.

**Step 1.** Get Fourier transform \( f(\omega) \) of the original signal \( f(x) \).

**Step 2.** Convolute \( f(\omega) \) according to the equation (3).

**Step 3.** When there is \( t = 0 \), the local minimum of \( f(\omega) \) is calculated. Take these positions as the starting points, the scale space curves under different scales \( t_k (k = 1, 2, \ldots, 2N_0) \) are accumulated in turn, and the minimum length curve set \( \{L_i\}_{i=1}^{N_0} \) is obtained.

**Step 4.** Fuzzy c-means is used to cluster the minimum length curve in set \( \{L_i\}_{i=1}^{N_0} \). The minimum length curves can be divided into two clusters (meaningful and non-meaningful).

**Step 5.** The meaningful cluster is regarded as the segmentation point of spectrum segmentation interval.

**Step 6.** A band-pass filter based on wavelet transform is established in the spectrum division interval.

**Step 7.** The IMF component is obtained by decomposing the vibration signal.

**Step 8.** Based on Pearson correlation coefficient method, the correlation coefficient between each component and the original signal is calculated, and the threshold value is set to retain larger IMF components than the threshold value.

**Step 9.** The IMFs with larger Pearson correlation coefficient will be selected to reconstruct the signal.

III. MAXIMUM CORRELATION KURTOSIS DECONVOLUTION WITH POWER SPECTRAL ENTROPY AND GRID SEARCH
A. THE MCKD
For periodic impulse signal \( x(n) \), the attenuation response of the transmission path is \( h(n) \) and the noise component of doping is \( e(n) \). Then the measured signal \( y(n) \) collected by the sensor can be expressed as

\[
y(n) = h(n) \cdot x(n) + e(n)
\]

The expression for maximizing of the correlation kurtosis is described as follow.

\[
CK_M(T) = \max_f \frac{\sum_{n=1}^{N}(\prod_{m=0}^{M}x_{n-mT})^2}{\left(\sum_{n=1}^{N}x_{n}^2\right)^{M+1}}
\]

The formula (4) is derivatived as follow.

\[
\frac{d}{df_k}CK_M(T) = 0, \quad k = 1, 2, \cdots, L
\]

The filter coefficients are calculated.

\[
f = \frac{\|x^2\|}{2\|B\|^2}(Y_0y^T)\cdot\sum_{m=0}^{M}Y_mA_m
\]

where, \( T \) and \( M \) represent the period of impulse signal and the number of shifts. \( f \) represents the filter vector of finite impulse response. \( L \) represents the filter length. The coefficients in equation (7) are obtained as follow.

\[
Y_r = \begin{bmatrix}
y_{1-r} & y_{2-r} & \cdots & y_{N-r} \\
y_{r} & 0 & \cdots & y_{N-1-r} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & y_{N-L-r+1}
\end{bmatrix}_{L \times N}
\]

\[
r = \begin{bmatrix}
0 & 2T & \cdots & mT
\end{bmatrix}_{1 \times M}
\]

\[
A_m = \begin{bmatrix}
X_{1-mT}^{-1} & X_{1-2mT}^{-1} & \cdots & X_{1-mTN}^{-1} \\
X_{1-mTN}^{-1} & \cdots & X_{N-mTN}^{-1}
\end{bmatrix}_{N \times 1}
\]

\[
B = \begin{bmatrix}
x_{1}y_{1-T} & \cdots & x_{1-mT} \\
x_{N}y_{N-T} & \cdots & x_{N-mT}
\end{bmatrix}_{N \times 1}
\]

B. THE PARAMETER OPTIMIZATION OF THE MCKD
Filter size \( L \) and deconvolution period \( T \) are two important parameters of the MCKD, especially the deconvolution period \( T \). The selection of \( T \) determines whether MCKD can enhance the fault frequency. And the reasonable selection of
L. Li et al.: Feature Extraction Based on EWT With Scale Space Threshold and Improved MCKD for Fault Diagnosis

$L$ will affect whether the fault frequency is obvious in the power spectrum. When the MCKD is used to process the signal, it is necessary to determine the appropriate parameters according to the characteristics of the signal because of the processed signal with different characteristics. The performance of the MCKD depends on the selection of the parameters.

Theoretically, the deconvolution period $T$ can be calculated, and the formula is shown as follow.

$$T = \frac{f_s}{f_0}$$  \hspace{1cm} (10)

where, $f_s$ is the sampling frequency, $f_0$ is the fault frequency.

However, the value of $T$ is slightly different from the theoretical value in practical application, so it is necessary to optimize the parameter $T$ near the theoretical value. Generally, the selection range of filter size $L$ is 100-500. Because the filter size $L$ and deconvolution period $T$ affect each other, it is of great significance to deeply study the adaptive selection of two parameters in order to improve the effect of the MCKD.

The search range of parameter $L$ is $[a, b]$, step is $n_1$. The search range of parameter $T$ is $[c, d]$, step is $n_2$. The possible values of parameters $L$ and $T$ are arranged and combined, and all possible combination results generate grid, which is described as $(L, T) \in [a, b] \times [c, d]$.

Power spectral entropy reflects the distribution of signal energy, so the power spectral entropy is used to optimize the parameters. As a parameter optimization method, grid search method is simple and accurate, although the calculation efficiency of grid search method is lower than other methods. However, in the case of low precision requirement, the calculation speed can be improved by increasing the step size.

The proposed parameter optimization evaluation criteria based on power spectral entropy are defined as follows

$$\max_{(L,T)\in[a,b] \times [c,d]} - \sum_i E_i (L, T) * \log_2 (E_i (L, T))$$  \hspace{1cm} (11)

where, there is $i = 1, 2, \cdots, n$.

IV. A NOVEL FAULT DIAGNOSIS METHOD

In order to solve the problem that the fault signal of rolling bearings is weak and it is difficult to extract fault feature under strong noise and complex transmission path, a new feature extraction method based on EWT with scale space threshold and MCKD with power spectral entropy and grid search is proposed to realize the fault diagnosis of rolling bearings in this paper. Firstly, the AM-FM modal decomposition of vibration signal is carried out by using STEWT, and the correlation analysis of the decomposed modal components is carried out to remove the modal components whose correlation is less than the threshold. The remaining IMF will be used to reconstruct the signal. Then, the filter size $L$ and impulse signal period $T$ of the PGMCKD are adaptively selected, and then the reconstructed signal is enhanced in order to achieve feature extraction and fault diagnosis.

The specific process of the proposed STEWT-PGMCKD method is shown in Figure 1.

![Figure 1. The flow of the STEWT-PGMCKD method.](image-url)
value of the search range needs to be calculated, and the search in a small range near the theoretical value is obtained.

**Step 4.** The grid search method is used to optimize the two parameters of the MCKD, and the power spectral entropy evaluation criterion is used to determine the optimal parameters.

**Step 5.** The optimized MCKD is used to enhance the reconstructed signal.

**Step 6.** The enhanced signal is transformed by Hilbert transform, and analyzed by power spectrum in order to identify the fault feature frequency.

**V. THE EXPERIMENT RESULTS AND ANALYSIS**

**A. THE EXPERIMENT PLATFORM**

In this paper, the rolling bearing fault signal of QPZZ-II test-bed is used to obtain experiment data. The bearing model is N205, the pitch diameter is 39 mm, there are 13 rolling elements with a diameter of 7.5mm, the rotating speed is 1500r / min, and the sampling frequency is 12000 Hz. According to the above information, the fault characteristic frequency of bearing rolling element is about 125.2 Hz, the inner race fault characteristic frequency is about 194.2 Hz, and the rotating frequency $f_r$ is about 25Hz. The QPZZ-II bearing fault test-bed is shown in Figure 2.

**FIGURE 2.** QPZZ-II fault test-bed.

**B. THE EXPERIMENT RESULTS**

Because the fault feature of outer race is obvious, the inner race and rolling element fault feature extraction are only considered to prove the effectiveness of the proposed fault diagnosis method in this paper.

1) **INNER RACE FAULT**

The time domain and power spectrum of inner race fault signal of rolling bearings are shown in Figure 3 and Figure 4.

As shown in Figure 3, the power spectrum contains strong noise interference, and there is no obvious periodic component of fault information. However, for the power spectrum in Figure 4, it cannot find any useful information for fault diagnosis of inner race.

The proposed STEWT-PGMCKD method is used to further analyze in this paper. Firstly, the EWT with scale space threshold is used to decompose the signal in order to obtain AM-FM components. The result of spectrum segregation is shown in Figure 5.

**FIGURE 3.** Time domain diagram of inner race fault signal.

**FIGURE 4.** Power spectrum of inner race fault signal.

**FIGURE 5.** Spectrum segmentation of inner race fault signal.

According to the spectrum segmentation, the signal is decomposed into several IMFs. The frequency components of each IMF are obtained by analyzing the power spectrum of IMFs, the result and the Pearson coefficient of each IMF are shown in Table 1.

**TABLE 1.** Pearson coefficient and frequency component of each IMF of inner race fault signal.

| IMF | IMF1 | IMF2 | IMF3 | IMF4 | IMF5 | IMF6 | IMF7 |
|-----|------|------|------|------|------|------|------|
| IMF(Hz) | 23.44 | 99.61 | 23.44 | 76.17 | 76.17 | 76.17 | 29.3 |
| P | 0.19 | 0.89 | 0.13 | 0.11 | 0.11 | 0.26 | 0.23 |

As can be seen from Table 1, EWT cannot directly get the fault frequency of the signal, so it is necessary to select the IMF which contains more fault information for signal reconstruction and further process the reconstructed signal.
In this paper, the scale space threshold is set to 0.2 according to experience. Therefore, IMF2, IMF6 and IMF7 are selected for signal reconstruction.

Then the reconstructed signal is processed by using the MCKD with power spectral entropy and grid search. The theoretical value of $T$ is 62, so the search range of $T$ is [60, 64] and the step size is 1. The value range of $L$ is [100, 500], and the step size is 5. According to the principle of minimum entropy of energy spectrum, a set of optimization parameters [62, 400] is obtained. The final processing results by power spectrum analysis are shown in Figure 6.

2) ROLLING ELEMENT FAULT

The time domain and power spectrum of rolling element fault signal of rolling bearings are shown in Figure 7 and Figure 8.

It can be seen from Figure 7 and Figure 8 that the power spectrum contains strong noise interference, and there is no obvious periodic component of fault information. However, for the power spectrum in Figure 7, it can find the useless information for fault diagnosis of rolling element.

The proposed STEWT-PGMCKD method is used to further analyze in this paper. Firstly, the EWT with scale space threshold is used to decompose the signal in order to obtain AM-FM components. The result of spectrum segregation is shown in Figure 9.

The frequency components of each IMF are obtained by analyzing the power spectrum of IMFs, the result and the Pearson coefficient of each IMF are shown in Table 2.

As can be seen from Table 2, EWT also cannot directly get the fault frequency of the signal, so IMF2, IMF3, IMF7 and IMF8 are selected for signal reconstruction. Then the reconstructed signal is processed by using the MCKD with power spectral entropy and grid search. The theoretical value of $T$ is 96, so the search range of $T$ is [94, 98] and the step size is 1. The value range of $L$ is [100, 500], and the step size is 5. According to the principle of minimum entropy of energy spectrum, a set of optimization parameters [98, 460] is obtained for MCKD. The final processing results by power spectrum analysis are shown in Figure 10.

It can be found from Figure 10 that 1/4 rotating frequency is 5.859Hz, fault feature frequency is 123Hz, and its double frequency, triple frequency, quadruple frequency and quintuple frequency can be used to accurately diagnose the rolling element fault of rolling bearings in the signal power spectrum processed by using the MCKD with power spectral entropy and grid search.

The frequency components of each IMF are obtained by analyzing the power spectrum of IMFs, the result and the Pearson coefficient of each IMF are shown in Table 2.

As can be seen from Table 2, EWT also cannot directly get the fault frequency of the signal, so IMF2, IMF3, IMF7 and IMF8 are selected for signal reconstruction. Then the reconstructed signal is processed by using the MCKD with power spectral entropy and grid search. The theoretical value of $T$ is 96, so the search range of $T$ is [94, 98] and the step size is 1. The value range of $L$ is [100, 500], and the step size is 5. According to the principle of minimum entropy of energy spectrum, a set of optimization parameters [98, 460] is obtained for MCKD. The final processing results by power spectrum analysis are shown in Figure 10.

It can be found from Figure 10 that 1/4 rotating frequency is 5.859Hz, fault feature frequency is 123Hz, and its double frequency, triple frequency, quadruple frequency and quintuple frequency can be used to accurately diagnose the rolling element fault of rolling bearings in the signal power spectrum processed by using the MCKD with power spectral entropy and grid search.
spectral entropy and grid search. Therefore, the experiment results show that the proposed STEWT-PGMCKD method can effectively diagnose the fault of rolling element.  

C. THE COMPARISON AND ANALYSIS

1) THE COMPARISON OF SIGNAL DECOMPOSITION METHODS

In order to illustrate the effectiveness of the EWT in signal decomposition, the EWT is compared with the EMD, EEMD and VMD. All the signal enhancement methods adopt the MCKD method with power spectral entropy and grid search.

For the decomposed signal by using EMD, some IMFs are selected for signal reconstruction according to the Pearson coefficient. And then the reconstructed signal is selected to enhance by the MCKD with power spectral entropy and grid search. The search range of $L$ is $[100, 500]$, the theoretical value of $T$ is 96, the search range is $[94, 98]$, and the optimized parameter is $[94, 110]$. The power spectrum is shown in Figure 11.

For the decomposed signal by using EEMD, some IMFs are selected for signal reconstruction according to the Pearson coefficient. And then the reconstructed signal is selected to enhance by the MCKD with power spectral entropy and grid search. The search range of $L$ is $[100, 500]$, the theoretical value of $T$ is 96, the search range is $[94, 98]$, and the optimized parameter is $(94, 100)$. The power spectrum is shown in Figure 12.

From Figure 11 and Figure 12, it can be seen that $1/3$ harmonic of fault frequency, but it is not obvious. Compared with the power spectrum by using the STEWT in Figure 10, the fault feature of rolling element cannot be extracted obviously. Therefore, the comparison results show that the SWEWT is more effective than the EMD and EEMD.

For the decomposed signal by using VMD, some IMFs are selected for signal reconstruction according to the Pearson coefficient. Considering that VMD needs to preset the number of mode decomposition $K$, this paper sets $K$ as the number of IMFs after EMD. And then the reconstructed signal is selected to enhance by the MCKD with power spectral entropy and grid search. The search range of $L$ is $[100, 500]$, the theoretical value of $T$ is 96, the search range is $[94, 98]$, and the optimized parameter is $[97, 145]$. The power spectrum is shown in Figure 13.

From Figure 13, it can be seen that the $1/4f_r$ and the fault frequency $123$Hz, but the fault frequency is not the most obvious frequency component in the figure, and there is no harmonic of fault frequency in the figure. Although VMD has made great progress compared with EMD and EEMD, but compared with the power spectrum by using the STEWT in Figure 10, the fault feature of rolling element still cannot be extracted obviously. The comparison results show that STEWT is effective for signal feature extraction.

2) THE COMPARISON OF OPTIMIZATION METHODS FOR MCKD

Firstly, in order to illustrate the necessity of MCKD parameter optimization and the influence of the $L$ and $T$, the reconstructed signal is reprocessed by MCKD with random parameters selection (the parameter is $[66, 250]$), the result is shown in Figure 14.
It can be seen from the figure that there is no fault information except rotating frequency, and the distribution of frequency components is disordered. This phenomenon shows that the reasonable selection of parameters has a great influence on the signal processing effect of MCKD.

In order to further verify the effectiveness of the improved MCKD with power spectral entropy and grid search, it is compared with the MCKD and the MCKD with Shannon entropy. The $T$ value of the MCKD is 96, and $L$ is 460, which is the same as the improved MCKD with power spectral entropy and grid search in this paper.

The MCKD with power spectral entropy is set as the same search range and step size with the improved MCKD with power spectral entropy and grid search. The final parameter optimization result is $[95, 315]$. The power spectrums of reconstructed signal by using two methods are shown in Figure 15 and Figure 16.

It can be seen from Figure 15 that MCKD can extract the related information with the rotating frequency and double harmonic of the fault frequency. Figure 16 shows that the fault frequency (123 Hz) can be extracted by using the improved MCKD with Shannon entropy. However, compared with Figure 10, it is found that the extraction effect of the improved MCKD with Shannon entropy is not obvious. The improved MCKD with power spectral entropy and grid search is more effective for fault diagnosis.

VI. CONCLUSION

Aiming at the problem of feature extraction of rolling bearing fault signal, a feature extraction method based on EWT with scale space threshold and improved MCKD with power spectral entropy and grid search is proposed to realize the fault diagnosis of rolling bearings in this paper. The scale space threshold is designed to improve the EWT, which is used to decompose the vibration signal. The correlation coefficient method is used to analyze the correlation of decomposed components in order to further remove redundant components. It can avoid the phenomenon of mode over decomposition or local extremum in the EWT. Then the power spectral entropy and grid search are used to improve the MCKD, which is used to enhance the reserved component in order to improve the enhancement effect of weak fault signal. The effectiveness of the proposed STEWT-PGMCKD method is proved by the rolling bearing fault signal of QPZZ-II. The comparison results with EMD, EEMD, VMD, MCKD and the MCKD with information entropy show that the proposed STEWT-PGMCKD method is an effective method for fault diagnosis.

ACKNOWLEDGMENT

The authors would like to thank all the reviewers for their constructive comments. The program for the initialization, study, training, and simulation of the proposed algorithm in this article was written with the tool-box of MATLAB 2018b produced by the Math-Works, Inc.

REFERENCES

[1] X. Gu, S. Yang, Y. Liu, and R. Hao, “Rolling element bearing faults diagnosis based on kurtogram and frequency domain correlated kurtosis,” *Meas. Sci. Technol.*, vol. 27, no. 1, Dec. 2016, Art. no. 135019.

[2] N. Sawalhi, R. B. Randall, and H. Endo, “The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis,” *Mech. Syst. Signal Process.*, vol. 21, no. 6, pp. 2616–2633, Aug. 2007.

[3] H. Zhao, H. Liu, Y. Jin, X. Dang, and W. Deng, “Feature extraction for data-driven remaining useful life prediction of rolling bearings,” *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, Feb. 2021, doi: 10.1109/TIM.2021.3059500.

[4] W. Deng, R. Yao, H. Zhao, X. Yang, and G. Li, “A novel intelligent diagnosis method using optimal LS-SVM with improved PSO algorithm,” *Soft Comput.*, vol. 23, no. 7, pp. 2445–2462, Apr. 2019.

[5] T. Han, C. Liu, L. Wu, S. Sarkar, and D. Jiang, “An adaptive spatiotemporal feature learning approach for fault diagnosis in complex systems,” *Mech. Syst. Signal Process.*, vol. 117, pp. 170–187, Feb. 2019.

[6] X. Cai, H. Zhao, S. Shang, Y. Zhou, W. Deng, H. Chen, and W. Deng, “An improved quantum-inspired cooperative co-evolution algorithm with multi-strategy and its application,” *Expert Syst. Appl.*, vol. 171, Jun. 2021, Art. no. 114629.

[7] C. Liu, K. G. Lore, Z. Jiang, and S. Sarkar, “Root-cause analysis for time-series anomalies via spatiotemporal graphical modeling in distributed complex systems,” *Knowl.-Based Syst.*, vol. 211, Jan. 2021, Art. no. 106527.
[8] T. Han, C. Liu, W. Yang, and D. Jiang, “Deep transfer network with joint distribution adaptation: A new intelligent fault diagnosis framework for industry application,” ISA Trans., vol. 97, pp. 269–281, Feb. 2020.

[9] Y. Hu and C. Zhao, “Fault diagnosis with dual cointegration analysis of common and specific nonstationary fault variations,” IEEE Trans. Autom. Sci. Eng., vol. 17, no. 1, pp. 237–247, Jan. 2020.

[10] T. Han, C. Liu, W. Yang, and D. Jiang, “Learning transferable features in deep convolutional neural networks for diagnosing unseen machine conditions,” ISA Trans., vol. 93, pp. 342–353, Oct. 2019.

[11] W. Qian, S. Li, and J. Wang, “A new transfer learning method and its application on rotating machine fault diagnosis under variant working conditions,” IEEE Access, vol. 6, pp. 69907–69917, 2018.

[12] H. Zhao, H. Liu, J. Xu, and W. Deng, “Performance prediction using high-order differential mathematical morphology gradient spectrum entropy and extreme learning machine,” IEEE Trans. Instrum. Meas., vol. 69, no. 7, pp. 4165–4172, Jul. 2020.

[13] D. Wang, H. Liu, J. Xu, H. Zhao, and Y. Song, “An improved quantum-inspired differential evolution algorithm for deep belief network,” IEEE Trans. Instrum. Meas., vol. 69, no. 10, pp. 7319–7327, Oct. 2020.

[14] H. Zhao, J. Zheng, J. Xu, and W. Deng, “Fault diagnosis method based on principal component analysis and broad learning system,” IEEE Access, vol. 7, pp. 99263–99272, 2019.

[15] S. Haidong, C. Junsheng, J. Hongkai, Y. Yu, and W. Zhantao, “Enhanced deep gaited recurrent unit and complex wavelet packet energy moment entropy for early fault prognosis of bearing,” Knowl.-Based Syst., vol. 188, Jan. 2020, Art. no. 105022.

[16] T. Han, D. Jiang, Y. Sun, N. Wang, and Y. Yang, “Intelligent fault diagnosis method for rotating machinery via dictionary learning and sparse representation-based classification,” Measurement, vol. 118, pp. 181–193, Mar. 2018.

[17] X. Li, W. Zhang, N.-X. Xu, and Q. Ding, “Deep learning-based machinery fault diagnostics with domain adaptation across sensors at different places,” IEEE Trans. Ind. Electron., vol. 67, no. 8, pp. 6785–6794, Aug. 2020.

[18] J. Li, X. Li, D. He, and Y. Qu, “A domain adaptation model for early gear pitting fault diagnosis based on deep transfer learning network,” Proc. Inst. Mech. Eng. O. J. Risk Rel., vol. 234, no. 1, pp. 168–182, Feb. 2020.

[19] B. Cai, Y. Liu, and M. Xie, “A dynamic-Bayesian-network-based fault diagnosis methodology considering transient and intermittent faults,” IEEE Trans. Autom. Sci. Eng., vol. 14, no. 1, pp. 276–285, Jan. 2017.

[20] B. Cai, H. Liu, and M. Xie, “A real-time fault diagnosis methodology of complex systems using object-oriented Bayesian networks,” Mech. Syst. Signal Process., vol. 80, pp. 51–64, Dec. 2016.

[21] B. Cai, X. Shao, Y. Liu, X. Kong, H. Wang, H. Xu, and W. Ge, “Remaining useful life estimation of structure systems under the influence of multiple causes: Subsea pipelines as a case study,” IEEE Trans. Ind. Electron., vol. 67, no. 7, pp. 5737–5747, Jul. 2020.

[22] X. Lou and K. A. Loparo, “Bearing fault diagnosis based on wavelet transform and fuzzy inference,” Mech. Syst. Signal Process., vol. 18, no. 5, pp. 1087–1095, May 2004.

[23] W. Sweldens, “The lifting scheme: A custom-design construction of biorthogonal wavelets,” Appl. Comput. Harmon. Anal., vol. 3, no. 2, pp. 186–200, Apr. 1996.

[24] J. Lin and L. Qu, “Feature extraction based on Morlet wavelet and its application for mechanical fault diagnosis,” J. Sound Vib., vol. 234, no. 1, pp. 135–148, Jun. 2000.

[25] Y. Qin, Y. Yao, B. Tang, Y. Wang, and H. Chen, “M-band flexible wavelet transform and its application to the fault diagnosis of planetary gear transmission systems,” Mech. Syst. Signal Process., vol. 134, Dec. 2019, Art. no. 106298.

[26] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” Proc. Roy. Soc. London A, Math., Phys. Eng. Sci., vol. 454, no. 1971, pp. 903–905, Mar. 1998.

[27] K. Dragomiretskiy and D. Zosso, “Variational mode decomposition,” IEEE Trans. Signal Process., vol. 62, no. 3, pp. 531–544, Feb. 2014.

[28] J. Gilles, “Empirical wavelet transform,” IEEE Trans. Signal Process., vol. 61, no. 16, pp. 3999–4010, Aug. 2013.

[29] A. Bhattacharyya, M. Sharma, R. B. Pachori, P. Sircar, and U. R. Acharya, “A novel approach for automated detection of focal EEG signals using empirical wavelet transform,” Neural Comput. Appl., vol. 29, no. 8, p. 47, Apr. 2018.
L. Li et al.: Feature Extraction Based on EWT With Scale Space Threshold and Improved MCKD for Fault Diagnosis

[50] Y. Song, D. Wu, W. Deng, X.-Z. Gao, T. Li, B. Zhang, and Y. Li, “MPPCEDE: Multi-population parallel co-evolutionary differential evolution for parameter optimization,” Energy Convers. Manage., vol. 228, Jan. 2021, Art. no. 113661.

[51] M. Wang and H. Chen, “Chaotic multi-swarm whale optimizer boosted support vector machine for medical diagnosis,” Appl. Soft Comput., vol. 88, Mar. 2020, Art. no. 105946.

[52] W. Deng, J. Xu, Y. Song, and H. Zhao, “Differential evolution algorithm with wavelet basis function and optimal mutation strategy for complex optimization problem,” Appl. Soft Comput., vol. 100, Mar. 2021, Art. no. 106724.

[53] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, “Harris hawks optimization: Algorithm and applications,” Future Gener. Comput. Syst., vol. 97, pp. 849–872, Aug. 2019.

[54] Y. Song, D. Wu, A. W. Mohamed, X. Zhou, B. Zhang, and W. Deng, “Enhanced success history adaptive DE for parameter optimization of photovoltaic models,” Complexity, vol. 2021, Jan. 2021, Art. no. 6660115.

LINGFENG LI (Member, IEEE) received the B.S. degree in electrical engineering and automation from the Dalian University of Technology, Dalian, China, in 2018. He is currently pursuing the master’s degree with Dalian Jiaotong University, Dalian. His research interests include deep learning and fault diagnosis.

AIBIN GUO (Member, IEEE) received the B.S. degree in mechatronic engineering and the M.S. degree in mechanical engineering and automation from Dalian Jiaotong University, Dalian, China, in 2000 and 2003, respectively, and the Ph.D. degree in mechanical engineering and automation from Beihang University, Beijing, China, in 2008. Since 2008, he has been a Senior Engineer with Haileng General Aviation Technology Company Ltd., Beijing. He is engaged in aviation emergency rescue and logistics support informatization.

HUAYUE CHEN received the M.S. degree in computer software and theory from Chongqing University, Chongqing, China, in 2005. She is currently pursuing the Ph.D. degree in geodetection and information technology with the Chengdu University of Technology, Chengdu, China. Since 2012, she has been an Associate Professor with China West Normal University, Nanchong, China. Her research interests include artificial intelligence, optimization methods, and hyperspectral image processing.

* * *