Gearbox Fault Diagnosis Based on Hierarchical Instantaneous Energy Density Dispersion Entropy and Dynamic Time Warping

Guiji Tang, Bin Pang *, Yuling He and Tian Tian

Department of Mechanical Engineering, North China Electric Power University, Baoding 071000, China; tanggjlk@ncepu.edu.cn (G.T.); heyuling1@ncepu.edu.cn (Y.H.); nceputt316@163.com (T.T.)
* Correspondence: 1162102031@ncepu.edu.cn

check for updates

Abstract: The accurate fault diagnosis of gearboxes is of great significance for ensuring safe and efficient operation of rotating machinery. This paper develops a novel fault diagnosis method based on hierarchical instantaneous energy density dispersion entropy (HIEDDE) and dynamic time warping (DTW). Specifically, the instantaneous energy density (IED) analysis based on singular spectrum decomposition (SSD) and Hilbert transform (HT) is first applied to the vibration signal of gearbox to acquire the IED signal, which is designed to reinforce the fault feature of the signal. Then, the hierarchical dispersion entropy (HDE) algorithm developed in this paper is used to quantify the complexity of the IED signal to obtain the HIEDDE as fault features. Finally, the DTW algorithm is employed to recognize the fault types automatically. The validity of the two parts that make up the HIEDDE algorithm, i.e., the IED analysis for fault features enhancement and the HDE algorithm for quantifying the information of signals, is numerically verified. The proposed method recognizes the fault patterns of the experimental data of gearbox accurately and exhibits advantages over the existing methods such as multi-scale dispersion entropy (MDE) and refined composite MDE (RCMDE).

Keywords: hierarchical instantaneous energy density dispersion entropy; dynamic time warping; hierarchical dispersion entropy; gearbox; fault diagnosis

1. Introduction

As an important component for transmitting energy and adjusting speed, gearbox is used in a wide variety of industrial machinery, such as wind turbines, aircraft engines, and automobiles [1,2]. Its failure is one of the main factors that cause the machinery to stop [3]. Thus, the accurate fault diagnosis of gearbox is highly required.

Generally, the existing fault diagnosis approaches can be classified into two types: Model-based and data-driven [4–6]. The model-based approach depends on an expansive system and dynamic knowledge [6]. The data-driven approach does not require an exhaustive system or dynamic knowledge, which is more conducive to intelligent diagnosis. Thus, the data-driven approaches have been widely used in gearbox fault diagnosis. Fault feature extraction is a key step in the data driven fault diagnosis methods and has a decisive impact on the final result [7]. At present, the typical and mature fault feature extraction algorithms are mainly based on processing the measurements such as vibration signal [8–10], acoustic signal [11,12], and temperature [13]. Among them, the algorithms based on vibration signal processing are favored by researchers because of their obvious advantages, such as low sensor price, simple signal acquisition process, and rich information contained in the signal. The complicated transmission path and running environment contribute to the nonlinear–nonstationary property and noise interference of the gearbox vibration signal, which increases the difficulty of fault

feature extraction [14]. Hence, the conventional spectra analysis algorithms based on Fourier transform (FT) is not enough to accurately diagnose gearbox faults [15].

Based on the excellent characteristics of time-frequency analysis (TFA) in processing nonlinear–nonstationary signals, scholars have been committed to the development and introduction of various TFA algorithms to fault feature extraction of gearbox in the past two decades. Classical TFA algorithms include wavelet transform (WT) [16], windowed FT (WFT) [17], and Hilbert-Huang transform (HHT) [18]. Among them, WT and WFT are less adaptive. HHT is exploited by integrating an adaptive signal separation method named empirical mode decomposition (EMD) and Hilbert transform (HT). Compared with WT and WFT, HHT is more adaptive. Moreover, the TFA results of WT and WFT contain the information on the interference components, while HHT can exclude the unwanted information of the interference signals by EMD. The signal separation of EMD is based on the local extreme points fitting, which causes modal aliasing when dealing with signals with anomalous events such as noise, discontinuities, and shocks [19]. To improve the immunity of EMD to anomalous events, some noise-assisted forms of EMD, such as ensemble EMD (EEMD) [20] and complementary EEMD (CEEMD) [21], have been invented and present advantages over EMD in fault feature extraction of gearbox. However, how to adaptively select the added noise is an obstacle of EEMD and CEEMD. Some scholars tried to modify EMD by improving the extreme point fitting [22]. Moreover, some variants of EMD, such as intrinsic time-scale decomposition (ITD) [23], local mean decomposition (LMD) [24], and local oscillatory-characteristic decomposition (LOD) [25] were developed by modifying the envelope definition. However, they are also sensitive to noise. Some recently proposed signal decomposition algorithms such as variational mode decomposition (VMD) [26], morphological component analysis (MCA) [27], and empirical wavelet transform (EWT) [28] make great breakthroughs in overcoming the modal aliasing problem compared with those signal separation algorithms based on extreme point fitting. However, they are highly sensitive to parameter selection and are not conducive to online analysis. Considering the limitations of the previous signal separation algorithms, Bonizzi [29] put forward the singular spectrum decomposition (SSD) approach. SSD inherits the signal separation function of singular spectrum analysis (SSA) and does not need to manually select parameters. The effectiveness of SSD for processing vibration signals of rotating machinery and its superiority over the previous signal separation algorithms has been fully illustrated in previous studies [30–32]. Thus, it is combined HT for TFA in this work. In the introduction of the principle of HHT, Huang also gives the concept of instantaneous energy density (IED). The IED analysis has shown its advantages for enhancing impact fault features [33,34]. In this paper, a novel IED analysis method based on SSD and HT is proposed to inhibit the useless interferences.

With the concept of data fusion, we expect to use reliable fault feature extraction methods and intelligent classifiers to realize intelligent diagnosis [35]. Complexity and regularity reflect the characteristic information of the signal [36]. Dispersion entropy (DE) is a novel proposed entropy indicator, which is able to quantify the complexity and regularity of the signal [37]. It makes some improvements compared with the previous entropy indicators such as approximate entropy (AE) and permutation entropy (PE). For example, DE is more efficient than AE. The information relates to the amplitude of the signal, which is considered in DE, while it is discarded by PE [38]. To better measure the information of the complex vibration signal of rotating machinery, some multi-scale forms of DE, such as multi-scale DE (MDE) [39] and refined composite MDE (RCMDE) [40], are proposed by introducing the coarse-grained and refined composite coarse-grained procedures. However, the coarse-grained and refined composite coarse-grained procedures can only detect the low-frequency information of the signal [41]. Thus, the hierarchical decomposition [42,43], which can probe the low-frequency and high-frequency information of the signal is combined with DE to develop the hierarchical DE (HDE) algorithm. Based on the merits of the IED analysis and HDE, we put forward the hierarchical instantaneous energy density dispersion entropy (HIEDDE) to better extract the fault information of gearbox. After the process of fault feature extraction, dynamic time warping (DTW) [44] is adopted to recognize the fault patterns. Compared with the other classifiers, fewer samples are

required to form the template signal [45]. A novel gearbox fault diagnosis framework based on HIEDDE and DTW is formulated. Finally, the proposed fault diagnosis framework is tested by the experimental data of gearbox.

The outline of this study is as follows. In Section 2, the details of the proposed HIEDDE algorithm based on the IED analysis and the HDE algorithm are introduced. Section 3 briefly illustrates the principle of DTW. Section 4 introduces the specific implementation procedures of the proposed method. Section 5 presents the experimental test. Finally, a few conclusions are presented in Section 6.

2. Hierarchical Instantaneous Energy Density Dispersion Entropy

2.1. IED Analysis

The vibration signal of gearbox is usually accompanied by interferences. Thus, a novel IED analysis algorithm based on SSD and HT is introduced to reinforce the fault feature signal. The SSD algorithm is used to obtain the useful potential components. Then, HT is employed to detect the time-frequency information of these components. Finally, the IED signal is obtained based on the time-frequency information.

2.1.1. SSD

SSD is a parameter-less signal-processing algorithm that can reach the signal decomposition adaptively. It is regarded as a modified version of SSA as it achieves the adaptive selection of the embedding dimension and can complete the automatic principal components grouping. The SSD decomposition for a composite signal is mainly composed of four steps as follows [29].

Step 1: Trajectory matrix construction.

For a discrete signal $a(n)$ that has N samples, $a(n) = (a_1, a_2, \ldots, a_N)$, the form of its trajectory matrix A is closely related to the selection of embedding dimension. If the embedding dimension is selected as K, A will be a $K \times N$ matrix, whose k-th row is $a_k = (a_k, \ldots, a_N, a_1, \ldots, a_{k-1})$. That is, $A = \left[a_1^T, a_2^T, \cdots, a_K^T\right]^T$.

How to choose the embedding dimension reasonably is a difficult problem of the original SSA algorithm. The following procedures are adopted in SSD to adaptively choose the embedding dimension at iteration j:

First, perform the calculation of the power spectral density (PSD) of the residual component $v_j(n)$ for iteration j, $v_j(n) = a(n) - \sum_{k=1}^{j-1} v_k(n)$ ($v_0(n) = a(n)$). Then, the frequency $f_{\max}$ at the maximum peak of the calculated PSD is distinguished. When $j = 1$, the residual component is regarded as a trend item if $f_{\max}$ is smaller than a given threshold, which is set as $0.01F_s$ in this paper, (F_s represents the sampling frequency). In this situation, K is recommended as $N/3$. In other situations, $K = 1.2F_s/f_{\max}$.

Step 2: Singular valued decomposition.

The trajectory matrix A obtained in **Step 1** carries the mixed information of all component signals. Thus, it is subjected to singular valued decomposition (SVD):

$$
\begin{aligned}
A &= P\Sigma Q^{\mathrm{T}} \\
&= \left[p_1, p_2, \cdots, p_K\right]
\begin{bmatrix}
\alpha_1 & 0 & 0 & 0 & \mathbf{0} \\
0 & \alpha_2 & 0 & 0 & \mathbf{0} \\
\vdots & \vdots & \ddots & \vdots & \mathbf{0} \\
0 & 0 & 0 & \alpha_K & \mathbf{0}
\end{bmatrix}
\begin{bmatrix}
q_1^{\mathrm{T}} \\
q_2^{\mathrm{T}} \\
\vdots \\
q_N^{\mathrm{T}}
\end{bmatrix} \\
&= \alpha_1 p_1 q_1^{\mathrm{T}} + \alpha_2 p_2 q_2^{\mathrm{T}} + \cdots + \alpha_K p_K q_K^{\mathrm{T}} \\
&= A_1 + A_2 + \cdots + A_K
\end{aligned}
\tag{1}
$$

where $P \in R^{K \times K}$, $\sum \in R^{K \times N}$, $Q \in R^{N \times N}$. $\{\alpha_1, \alpha_2, \ldots, \alpha_K\}$ are the singular values in matrix$\sum$, $A_i = \alpha_i p_i q_i$ represents the i-th principal component of A.

Step 3: Principal components grouping and reconstruction of the j-th SSC.

As represented in Equation (1), the trajectory matrix is divided into K principal components by SVD. The key to the successful separation of a potential component is to sort these principal components reasonably. The following criterion is used in SSD for selecting principal components to reconstruct the SSC at the j-th iteration:

For $j = 1$, if a trend item is identified, only the first principal component A_1 is chosen as the useful principal component, and the first SSC is retrieved through the diagonal averaging of A_1.

Otherwise, and for $j > 1$, a subset principal components, whose left eigenvectors show a primary frequency in $[f_{max} - f_1, f_{max} + f_1]$ (where f_1 is estimated through the Gaussian interpolation of the PSD of $v_j(n)$), are combined to form a matrix A'. The j-th SSC is retrieved through the diagonal averaging of A'.

Step 4: Setting stopping criterion.

Once a component $SSC_j(n)$ is retrieved, a corresponding residual signal can be obtained as:

$$v_{j+1}(n) = v_j(n) - SSC_j(n). \tag{2}$$

$v_{j+1}(n)$ is then employed as the input to retrieve the $(j + 1)$-th SSC. If the result of the equation below is smaller than a given threshold, the iteration will be stopped.

$$NMSE_j = \frac{\sum_{i=1}^{N} (v_{j+1}(i))^2}{\sum_{i=1}^{N} (x(i))^2}. \tag{3}$$

2.1.2. IED Calculation Based on SSD and HT

After the SSD process, the composite signal $a(t)$ can be expressed as the sum of k SSCs and a residual signal:

$$a(t) = \sum_{j=1}^{k} SSC_j(t) + r(t), \tag{4}$$

where $SSC_j(t)$ represents the j-th SSC and $r(t)$ is the residual signal.

Previous studies have indicated that SSD has stronger signal decomposition capability than the traditional signal decomposition approaches such as EMD and EEMD [30–32]. Thus, it is integrated with HT to make the time-frequency analysis. HT is used to estimate the instantaneous amplitude (IA) and instantaneous frequency (IF) of the decomposed components by referring to Reference [18]. The SSD-HT spectrogram can be constructed by lumping the IAs and IFs of all SSCs:

$$SSD - HT(\omega, t) = \sum_{j=1}^{k} a_j(t) \exp(i \int \omega_j(t)). \tag{5}$$

where k represents the number of the decomposed components, $a_j(t)$ and $\omega_j(t)$ are the IA and IF of the j-th SSC, respectively.

We can further obtain the IED signal with the following form:

$$IED(t) = \int_{\omega} |SSD - HT(\omega, t)|^2 d\omega. \tag{6}$$

The IED reflects the energy fluctuation of the signal over time.

However, some of the decomposed SSCs may be false components because the vibration signal of gearbox contains interferences. In this paper, the correlation coefficient between the original signal $a(t)$ and the decomposed components is used to discard the false components [46]:

$$\rho_j = \frac{COV\big(SSC_j(t), a(t)\big)}{\sigma_j \sigma_a},$$ (7)

where ρ_j is the correlation coefficient between $a(t)$ and the j-th SSC, $COV(SSC_j(t), a(t))$ represents the covariance of $SSC_j(t)$ and $a(t)$, and σ_j and σ_a, respectively, correspond to the variance of $SSC_j(t)$ and $a(t)$.

If ρ_j is larger than a threshold (which is set to 0.2 max $(\rho_1, \rho_1, \ldots, \rho_k)$ in this paper), $SSC_j(t)$ is seen as a real component. All the real SSCs are selected out to calculate the IED to enhance the fault signatures of gearbox.

2.1.3. Verification of the Fault Feature Enhancement Capability of IED

A mixed signal containing the distributed fault signal of gear and noise is tested to examine the fault feature enhancement capability of the IED analysis. The fault signal is generated using the model proposed by McFadden [47]:

$$s_{\text{faulty}}(t) = \sum_{k=0}^{L} S_k(1 + a_k(t)) \cos(2\pi k z f_r t + \varphi_k + b_k(t)),$$ (8)

where L is the harmonic orders, S_k and φ_k are the amplitude and phase of the k-th meshing harmonic, z and f_r represent the number of teeth and rotating frequency of the faulty gear, $a_k(t)$ is the amplitude modulation (AM) item and $b_k(t)$ represents the phase modulation (PM) item. The AM and PM items are, respectively, defined as:

$$a_k(t) = \sum_{m=0}^{M} A_{km} \cos(2\pi m f_r t + a_{km}),$$ (9)

$$b_k(t) = \sum_{m=0}^{M} B_{km} \cos(2\pi m f_r t + b_{km}),$$ (10)

where M is the total orders of the k-th meshing harmonic, A_{km} and a_{km} denote the amplitude and phase of AM, B_{km} and b_{km} represent the amplitude and phase of PM.

The other parameters are adopted as: $L = 3$, $S_k = 1$, $z = 30$, $f_r = 12$ Hz, $\varphi_k = 0$, $A_{km} = 0.9$, $a_{km} = 0$, $B_{km} = 0.1$, $b_{km} = 0$. The meshing frequency can be calculated as: $f_m = 30 \times 12 = 360$ Hz. In order to meet the requirement of frequency spectrum analysis and sampling theorem, the sampling frequency should be at least 3–5 times the meshing frequency; we select the sampling frequency as 10,000 Hz. Moreover, the signal length is set as 8192 to ensure a higher accuracy of Fourier transform.

Equation (11) is adopted to produce the noise:

$$n(t) = 0.2 randn(1, N),$$ (11)

where N is the signal length.

By adding the noise to the fault signal, the tested signal is yielded as shown in Figure 1a. From its amplitude spectrum as represented in Figure 1b, there are three harmonics of meshing frequency. However, the rotating frequency of the faulty gear cannot be located.
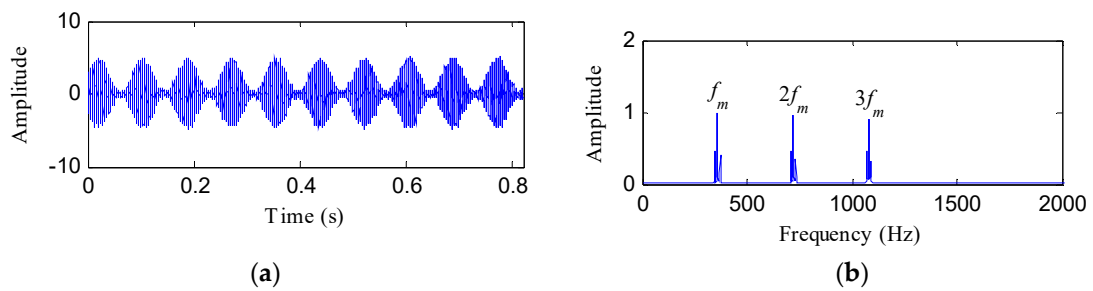
Figure 1. (**a**) The tested signal and (**b**) its amplitude spectrum.

Then, SSD is used to decompose the tested signal, and four SSCs are obtained as shown in Figure 2a. Table 1 shows the correlation coefficients between the tested signal and the SSCs. Based on the selection criterion as introduced in Section 2.2, SSC_2 is determined as a false signal. The other three SSCs are demodulated by HT to obtain the SSD-HT spectrogram, as shown in Figure 2b. Obviously, the SSD-HT spectrogram reflects three frequency components, i.e., f_m, $2f_m$ and $3f_m$. Figure 2c represents the IED signal. From its amplitude spectrum as shown in Figure 2d, the first and second harmonics of the rotating frequency of the faulty gear and the meshing frequency are clearly located. The fault feature enhancement capability of the IED analysis is verified.
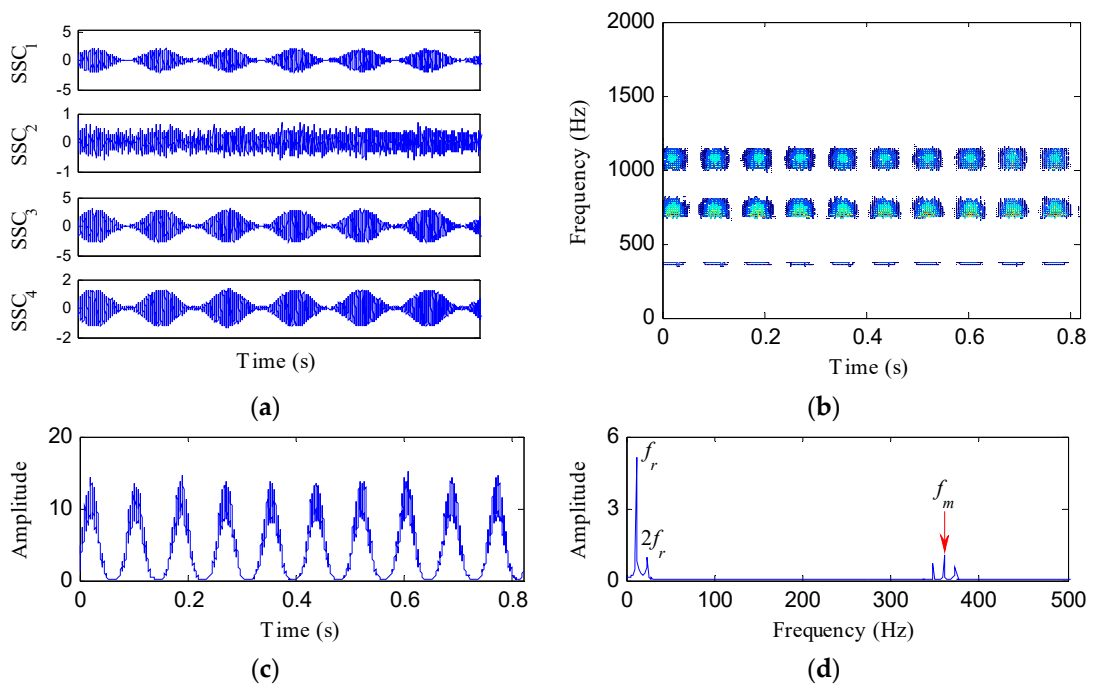
Figure 2. The IED analysis of the tested signal based on SSD and HT: (**a**) the decomposition results of SSD; (**b**) the SSD-HT spectrogram; (**c**) the IED signal; and (**d**) its amplitude spectrum.

Table 1. The correlation coefficients between the tested signal and its singular spectrum components.

Decomposition Components	SSC_1	SSC_2	SSC_3	SSC_4
Correlation coefficients	0.4913	0.1084	0.7381	0.5731

For comparison, the simulated gear fault signal is also analyzed by the IED analysis based on EMD and HT. Twelve components are generated after the EMD process, as shown in Figure 3a. From the waveforms of these components, it can be seen that the fifth to twelfth components are false components. Thus, the first four components retrieved by EMD are analyzed by HT to acquire the EMD-HT spectrogram, as depicted in Figure 3b. Figure 3c,d represent the corresponding IED signal

and its amplitude spectrum, respectively. Only the rotating frequency of the faulty gear appears in Figure 3d, while its harmonics and the meshing frequency are invisible. The superiority of the IED analysis based on SSD is demonstrated.
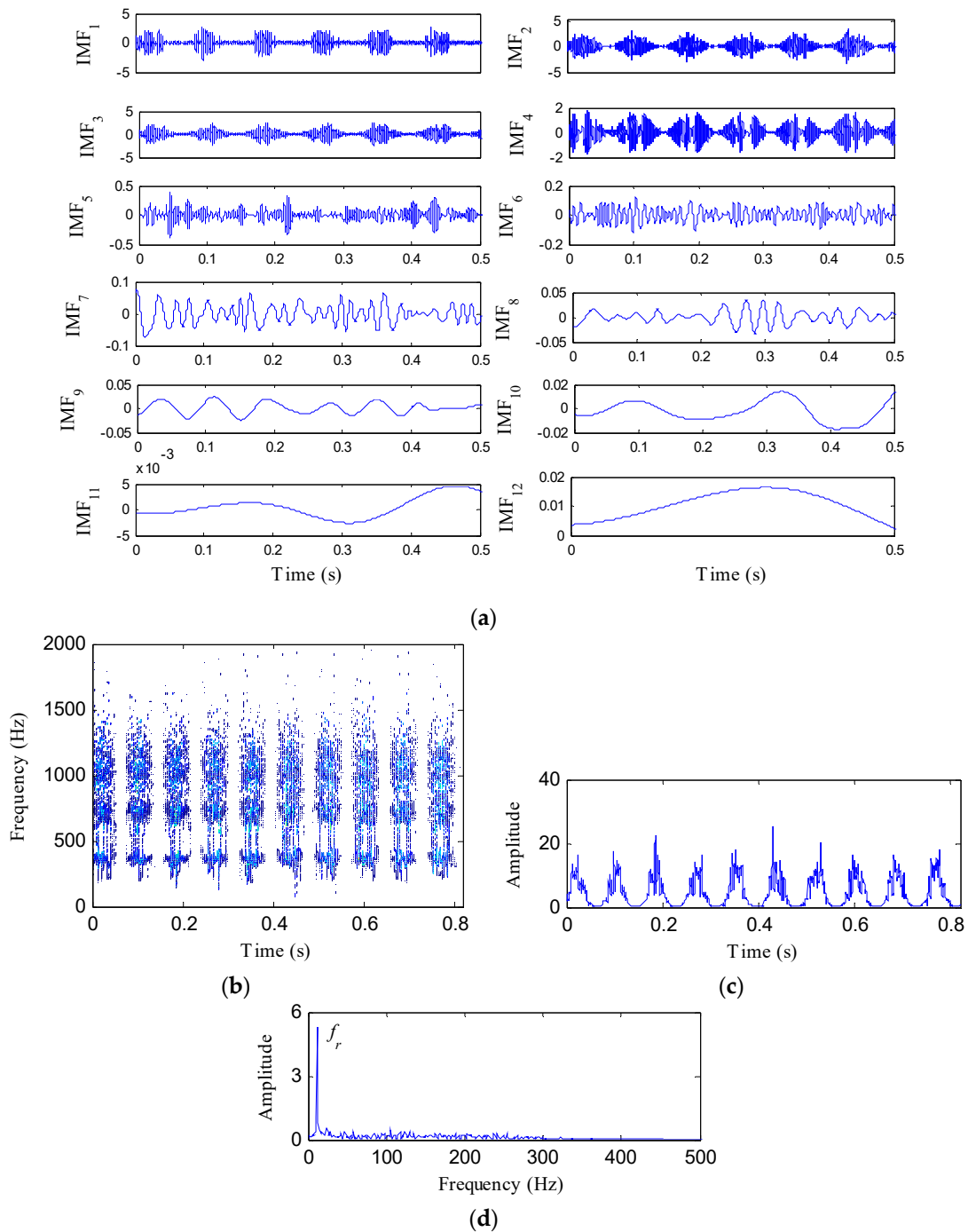

Figure 3. The IED analysis of the tested signal based on EMD and HT: (**a**) The decomposition results of EMD; (**b**) the EMD-HT spectrogram; (**c**) the IED signal; and (**d**) its amplitude spectrum.

2.2. Hierarchical Dispersion Entropy

2.2.1. Dispersion Entropy

DE is a novel entropy index for quantifying the certainty of signals. Its advantages over the traditional entropy methods have been investigated by many scholars [38–40]. For a signal that has N sampling points, $a(n) = (a_1, a_2, \ldots, a_N)$, its DE can be calculated through the following procedures [37]:

(1) A new signal $b(n) = (b_1, b_2, \ldots, b_N)$, with $b_i \in (0,1)$, is obtained by applying normal cumulative distribution functions (NCDF) to the original signal $a(n)$:

$$b_i = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{a_i} e^{\frac{-(t-u)^2}{2\sigma^2}} dt, \tag{12}$$

where σ^2 and u indicate the variance and mean of $a(n)$, respectively.

(2) $b(n)$ is mapped to the integer range of $[1, 2, \ldots, c]$ with the following formula:

$$Z_i^c = round(c \cdot b_i + 0.5). \tag{13}$$

(3) An embedded vector $Z_i^{m,c}$ is formed with embedding dimension m and time delay τ:

$$Z_i^{m,c} = \left\{ Z_i^c, Z_{i+\tau}^c, \ldots, Z_{i+(m-1)\tau}^c \right\}, i = 1, 2, \ldots, (m-1)\tau. \tag{14}$$

(4) The corresponding dispersion pattern $\pi_{v_0, v_1, \ldots, v_{m-1}}$ is obtained with the following definition:

$$Z_i^c = v_0, \; Z_{i+\tau}^c = v_1, \ldots, \; Z_{i+(m-1)\tau}^c = v_{m-1}. \tag{15}$$

(5) The relative frequency of each dispersion pattern is computed as:

$$P(\pi_{v_0, v_1, \ldots, v_{m-1}}) = \frac{N_a(\pi_{v_0, v_1, \ldots, v_{m-1}})}{N - (m-1)\tau}. \tag{16}$$

(6) With referring to the definition of Shannon entropy, the DE of $a(n)$ is computed as:

$$DE(a, m, c, \tau) = -\sum_{\pi=1}^{c^m} P(\pi_{v_0, v_1, \ldots, v_{m-1}}) \ln(P(\pi_{v_0, v_1, \ldots, v_{m-1}})). \tag{17}$$

Previous studies [36–40] recommend that the parameters selection for the DE algorithm should be: $m = 2$, $c = 6$, $\tau = 1$. Thus, we take this suggestion in all the calculation of DE in this paper.

2.2.2. Hierarchical Dispersion Entropy

In order to better detect the regularity of the complex signals, some multi-scale forms of DE, such as MDE and RCMDE, have been developed. However, previous studies reveal that the multi-scale analysis algorithms only consider the low-frequency components of the signal and cannot completely detect the hidden fault features [41–43]. Thus, a hierarchical DE (HDE) algorithm is proposed by integrating the hierarchical decomposition and DE to quantify the information of the signal more comprehensively.

For a 1 d signal $A = (a_1, a_2, \ldots, a_N)$ with the length of N ($N = 2^n$, n represents a positive integer), its calculation procedures of HDE include:

(1) An average operator q_0 and a difference operator q_1 are respectively constructed as follows:

$$q_0(A) = \frac{a_{2j} + a_{2j+1}}{2}, j = 1, 2, \ldots, 2^{n-1}, \tag{18}$$

$$q_1(A) = \frac{a_{2j} - a_{2j+1}}{2}, j = 1, 2, \ldots, 2^{n-1},$$

(19)

where $q_0(A)$ and $q_1(A)$ carry the low-frequency and high-frequency features of A at scale 2, respectively.

(2) The matrix expression of the operators q_j ($j = 0$ or 1) is obtained as the following form:

$$q_j = \begin{bmatrix} \frac{1}{2} & \frac{(-1)^j}{2} & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{(-1)^j}{2} & \ldots & 0 & 0 \\ 0 & 0 & 0 & 0 & \ldots & \frac{1}{2} & \frac{(-1)^j}{2} \end{bmatrix}.$$

(20)

(3) In order to perform the hierarchical analysis on the signal A, the above operators have to be employed iteratively. Let $p \in N$, a vector $[\beta_1, \beta_2, \ldots, \beta_p] \in \{0, 1\}$ can be designed to describe the integer d:

$$d = \sum_{j=1}^{p} \beta_j 2^{p-j}.$$

(21)

It can be inferred that for a given integer E, there is a unique vector $[\beta_1, \beta_2, \ldots, \beta_p]$ corresponding to it.

(4) The hierarchical component $A_{k,d}$ (where k and d are the layer number and node number, respectively) is expressed as:

$$A_{k,d} = q_{\beta_1} \bullet q_{\beta_2} \bullet \ldots \bullet q_{\beta_p} \bullet A.$$

(22)

(5) The DE of the hierarchical component at node d and layer k is calculated to obtain the HDE:

$$HDE(A, k, d, m, c, \tau) = DE(A_{k,d}, m, c, \tau).$$

(23)

The hierarchical decomposition of signal A with three layers is displayed in Figure 4 to better illustrate its principle. It can be found that there are 2^k hierarchical components at layer k. Figure 5 represents the calculation flowchart of HDE. In this paper, the HDE in layer 3 (which contains 8 nodes), which is used to characterize the information of the signal.
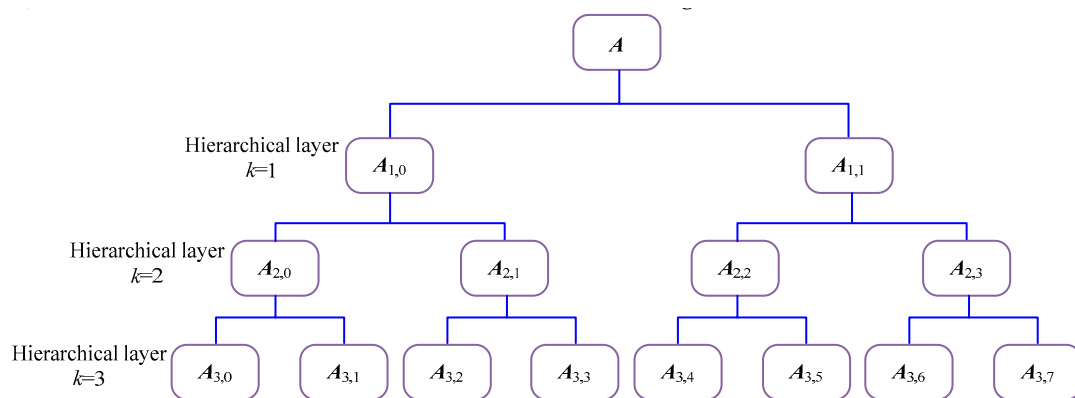

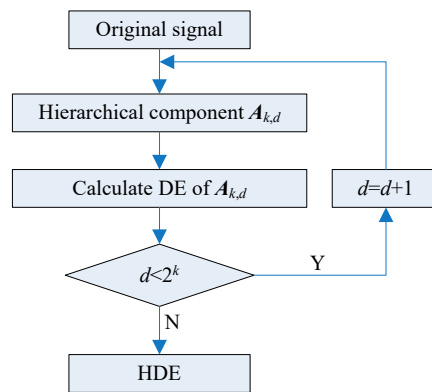**Figure 4.** Hierarchical decomposition with three layers.

Figure 5. Flowchart of the HDE algorithm.

2.2.3. Effectiveness Evalution of HDE

A Gaussian white noise (GWN) with 4096 sampling points, as depicted in Figure 6a is studied to compare the effectiveness of HDE and MDE. From its magnitude spectrum, as shown in Figure 6b, we can deduce that the complexity of GWN in different frequency bands is almost the same. The HDE (with three layers) and MDE (with eight scales) of the tested GWN are computed, as displayed in Figure 6c. The HDE values of eight hierarchical components are almost equal. Whereas, the MDE value gradually decreases as the scale increases. In theory, the DE values that correspond to different frequency bands should not change much. The HDE more realistically reflects the complexity of the studied GWN signal compared to the MDE.
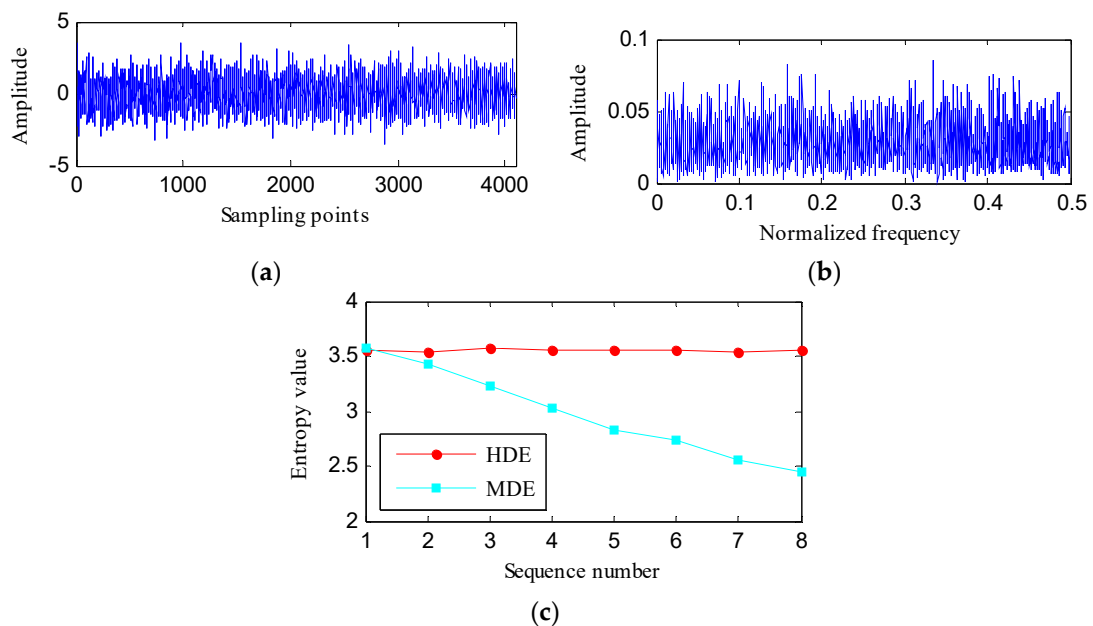
Figure 6. (**a**) The studied GWN signal; (**b**) its amplitude spectrum; and (**c**) its HDE and MDE.

In addition, 100 independent sets of GWN (each set comprises 4096 samples) are tested to reveal the advantage of HDE over MDE. The median and standard deviation (SD) of HDE and MDE are calculated to obtain the error bars as shown in Figure 7. The SD of HDE in different nodes is very tiny, which illustrates the HDE algorithm has strong stability. However, the SD of MDE increases monotonically with scale. Starting from sequence number 2, the SD value of MDE is significantly larger than the SD value of HDE, which implies that the HDE algorithm is more stable.
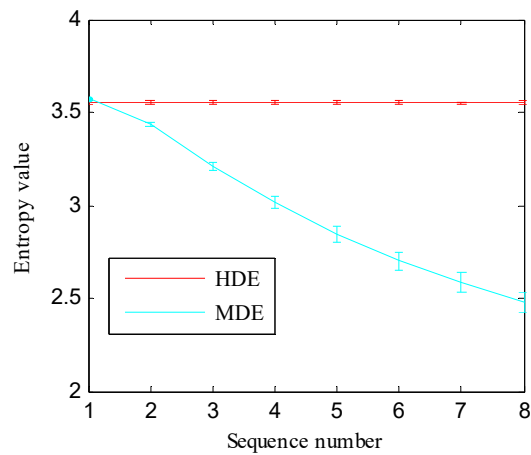
Figure 7. The error bars of HDE and MDE.

2.3. Hierarchical Instantaneous Energy Density Dispersion Entropy

Integrating the merits of IED and HDE, a novel fault signatures algorithm named hierarchical instantaneous energy density dispersion entropy (HIEDDE) is put forward. In this algorithm, the IED analysis is first used for fault features enhancement. The HDE process is then employed to quantify the information of the IED signal to get the HIEDDE. Figure 8 shows the flowchart of the HIEDDE algorithm. Total five procedures are needed to implement the HIEDDE algorithm:

(1) Apply the SSD algorithm to the vibration signal to generate k SSCs.

(2) Calculate the correlation coefficient between the original signal and the j-th SSC, ρ_j (j = 1, 2, …, k). If ρ_j is larger than a threshold, as given in Section 2.2, $SSC_j(t)$ is seen as a real component. Otherwise, $SSC_j(t)$ is excluded as a false component.

(3) Apply HT to all real SSCs identified in step (2) to obtain the SSD-HT spectrogram.

(4) Calculate the IED signal of the SSD-HT spectrogram.

(5) Calculate the HDE of the IED signal to obtain the proposed HIEDDE.
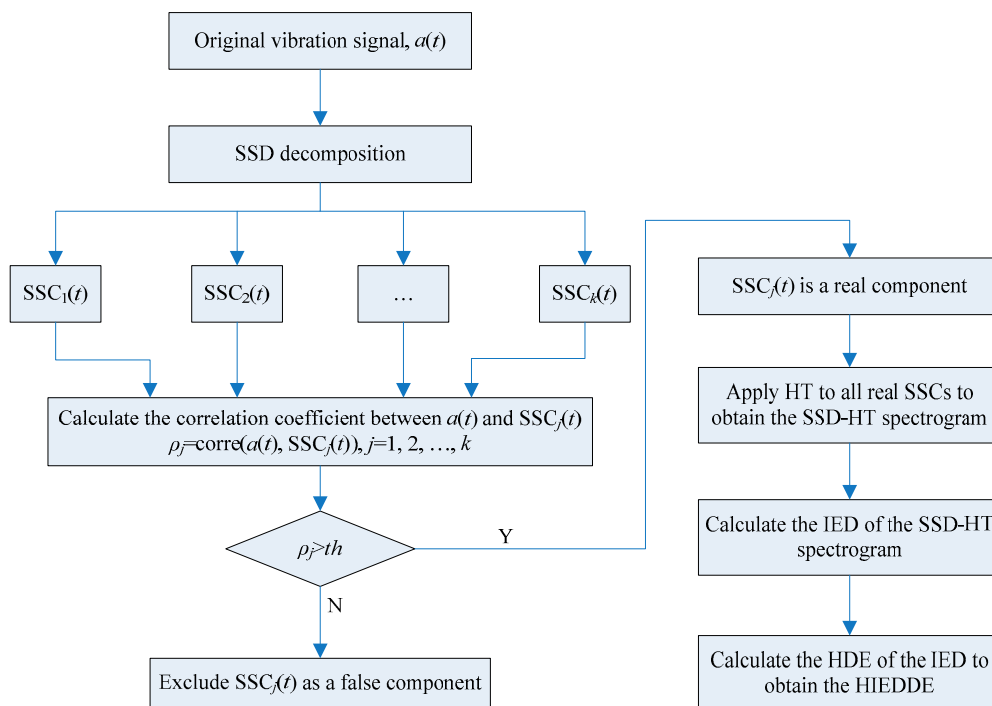
Figure 8. Flowchart of the HIEDDE algorithm.

3. Dynamic Time Warping

DTW was originally developed for speech classification [44]. In recent years, it has been introduced to the field of mechanical fault diagnosis [45]. The advantage of DTW over existing classifiers is that it requires only a small number of samples to create the template signal while ensuring high classification accuracy. It is used to qualify the similarity between two time sequences based on the DTW distance, which is calculated by finding the optimal alignment. The smaller DTW distance between the two signals represents the higher similarity. The following gives the specific description of DTW:

Given two time series $P = \{p_i\}$, $I = 1, \ldots, m$ and $Q = \{q_i\}$, $I = 1, \ldots, n$, where m and n represent the length of P and Q, respectively. A $m \times n$ distance matrix D is firstly constructed with the elements $d(p_i, q_j)$, which are the Euclidean distance between the points p_i and q_j. Then, a path through the matrix D, which minimizes the cumulative distance between the two sequences can be found. DTW distance corresponds to the path with minimal warping cost:

$$DTW(P, Q) = \min \sqrt{\sum_{k=1}^{K} w_k}, \qquad (24)$$

where w_k is the matrix element and also belongs to the k-th element of a warping path W. The warping path needs to reach the following three constraints:

(a) Boundary condition: the starting point is $w_1 = (1,1)$ and the ending point is $w_K = (m, n)$.

(b) Continuity: if $w_k = (a, b)$ and $w_{k-1} = (a', b')$, then $(a - a') \leq 1$ and $(b - b') \leq 1$.

(c) Monotonicity: if $w_k = (a, b)$ and $w_{k-1} = (a', b')$, then $(a - a') \geq 0$ and $(b - b') \geq 0$.

The Warping path can be determined by using dynamic programming as follows:

$$\gamma(i, j) = d(p_i, q_j) + \min \begin{cases} \gamma(i-1, j-1) \\ \gamma(i-1, j) \\ \gamma(i, j-1) \end{cases}, \qquad (25)$$

where $\gamma(i, j)$ is the sum of $d(p_i, q_j)$ and the minimum cumulative distance of three adjacent elements.

4. The Proposed Fault Diagnosis Method

A novel fault diagnosis method of gearbox based on the HIEDDE and DTW is proposed in this study. Figure 9 represents its implementation process, which is mainly composed of three steps:

(1) Collect the vibration data of gearbox under different health conditions. The vibration data of each condition is divided into some non-overlapping data groups. For each state, one data group is used for template sample and the other groups are used as testing samples. The number of the template samples is equal to the number of fault types.

(2) Calculate the HIEDDE of the vibration data as the fault features. The hierarchical layer is set to three so that the HIEDDE of each data sample contains the information of eight hierarchical components.

(3) Employ DTW for fault pattern recognition. Based on the HIEDDE fault features, the DTW distance between the testing samples and template samples are calculated, and the fault type of the testing sample can be recognized based on the label of the template sample with the smallest distance.
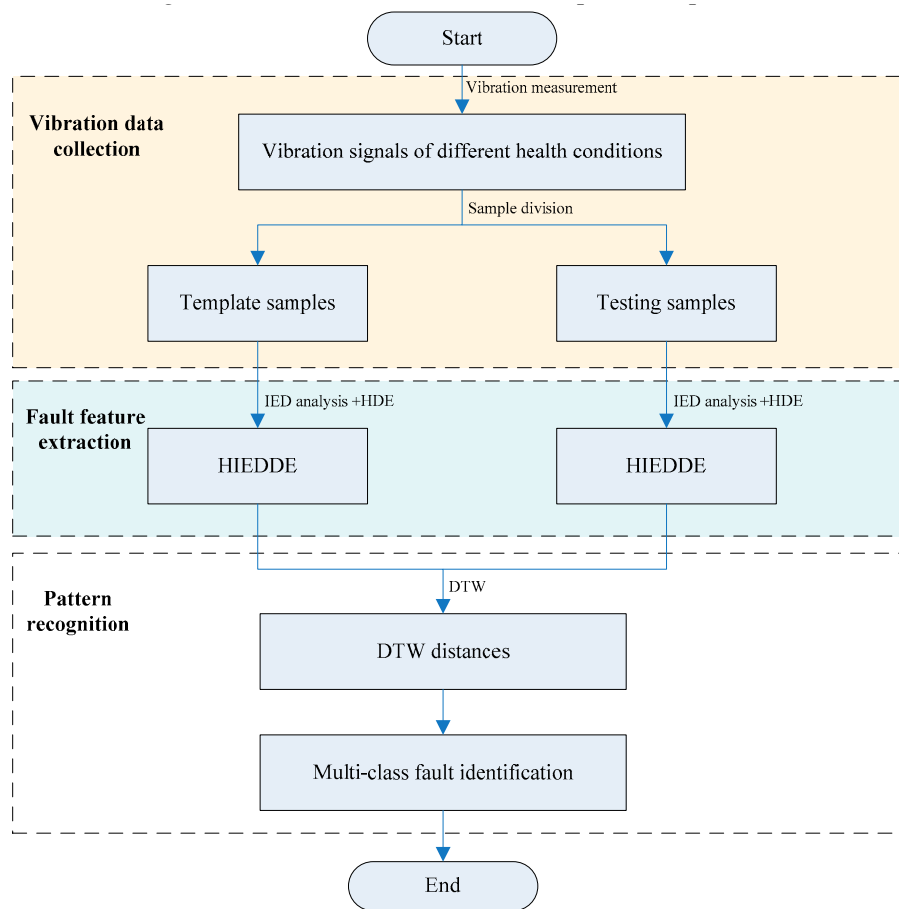
Figure 9. Implementation process of the proposed method.

5. Experimental Analysis

In this section, the presented fault diagnosis approach is employed to analyze the experimental data of gearbox under different states to test its validity. The data is obtained through the QPZZ-II rotating machinery fault simulation experiment platform. From the schematic diagram of the experiment platform, as depicted in Figure 10a, it can be found that the platform for gearbox faults simulation mainly consist of electric motor, belt drive, shaft support, gearbox, and power twister. The gearbox contains a small gear with 55 teeth installed on the input shaft and a big gear with 75 teeth mounted on the output gear. Table 2 represents the parameters illustration of the two gears. Four accelerometers were mounted on the housing of gearbox to measure the vertical vibration and two eddy current sensors are responsible for monitoring the radial vibration of the input bearing. The installation positions of the sensors are marked in Figure 10a. Figure 10b shows the sensor distribution from two perspectives. As displayed, the four accelerometers are numbered ①~④, respectively. The rotating speed of the input shaft in the experiment is 880 r/min. The meshing frequency of the small gear and big gear is 807 Hz. The sampling frequency is 5120 Hz.
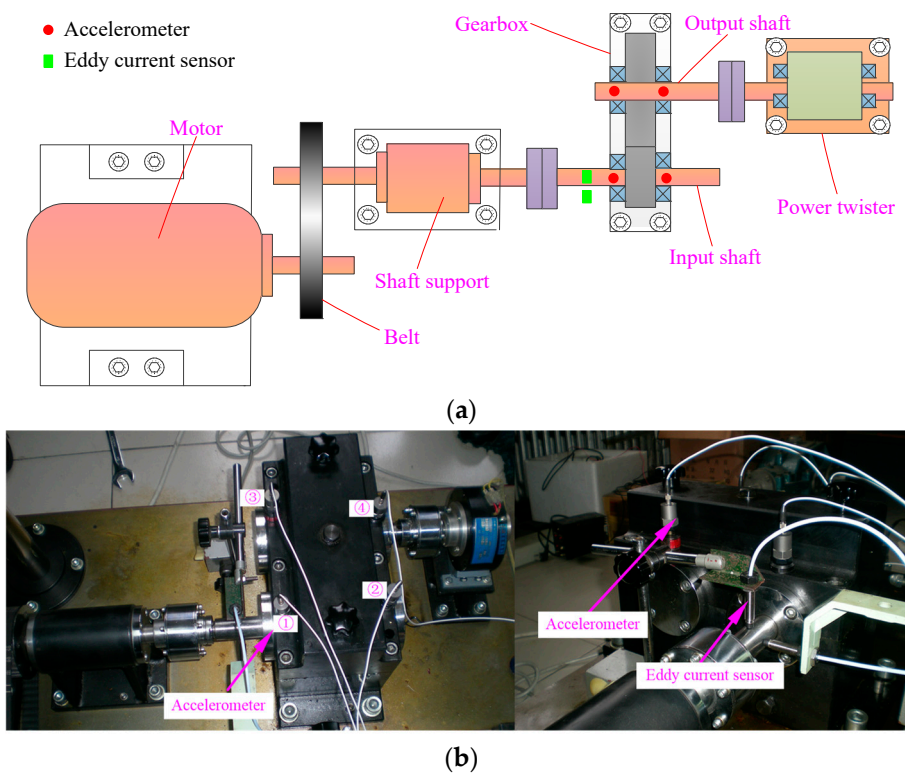
(a)

(b)

Figure 10. (**a**) The structure sketch of the test stand; (**b**) the sensors placement.

Table 2. Parameters illustration of the two gears.

Gears	Modulus	Number of Teeth	Rotating Frequency	Materials
Small gear	2	55	14.67 Hz	S45C
Big gear	2	75	10.76 Hz	S45C

Experimental signals were collected under four health conditions including: (1) Normal condition; (2) a pitting tooth (PT) on the output gear; (3) a broken tooth (BT) on the output gear; (4) some wear teeth (WT) on the input gear; and (5) a compound fault condition with a BT on the output gear and a WT on the input gear (BT-WT). Figure 11 represents the damaged gears with different fault types employed in the test. For each state, the experimental signal is segmented into 41 data groups with the length of 4096. In order to ensure the independence of each group and test the effectiveness of the proposed method more reasonably and strictly, the data groups are non-overlapping. One data group of each state is selected as the template sample and the other 40 groups are regarded as the testing samples. Table 3 represents the details of the experimental signals.

| **(a)** | **(b)** | **(c)** |
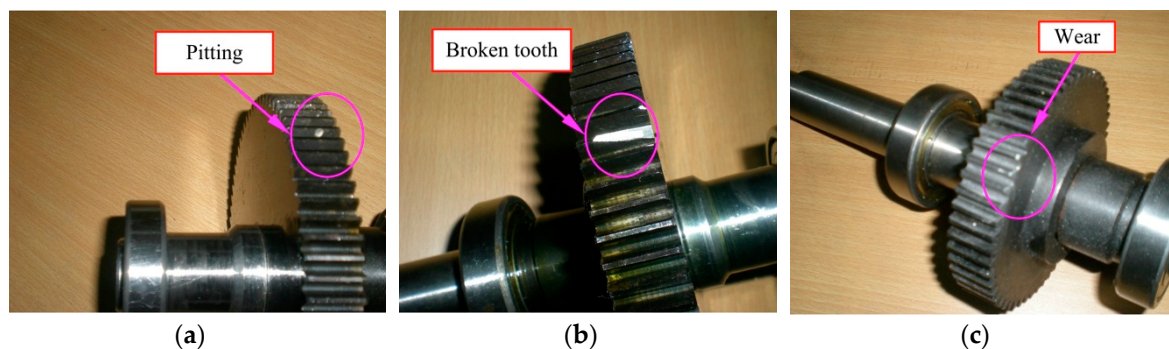
Figure 11. Gear damage faults: (**a**) pitting fault; (**b**) broken tooth fault; and (**c**) wear fault.

Table 3. Description of the experimental data.

Fault Type	Label	Number of Template Signal	Number of Testing Samples
Normal	1	1	40
PT	2	1	40
BT	3	1	40
WT	4	1	40
BT-WT	5	1	40

One testing sample of each health condition is used to illustrate the process of the proposed fault diagnosis approach. Figure 12 shows these samples under different health conditions. Despite the impacts of the data samples under the four fault states (PT, BT, WT, and BT-WT) being more prominent than the normal state, the fault states cannot be recognized because their impact features are somewhat similar.
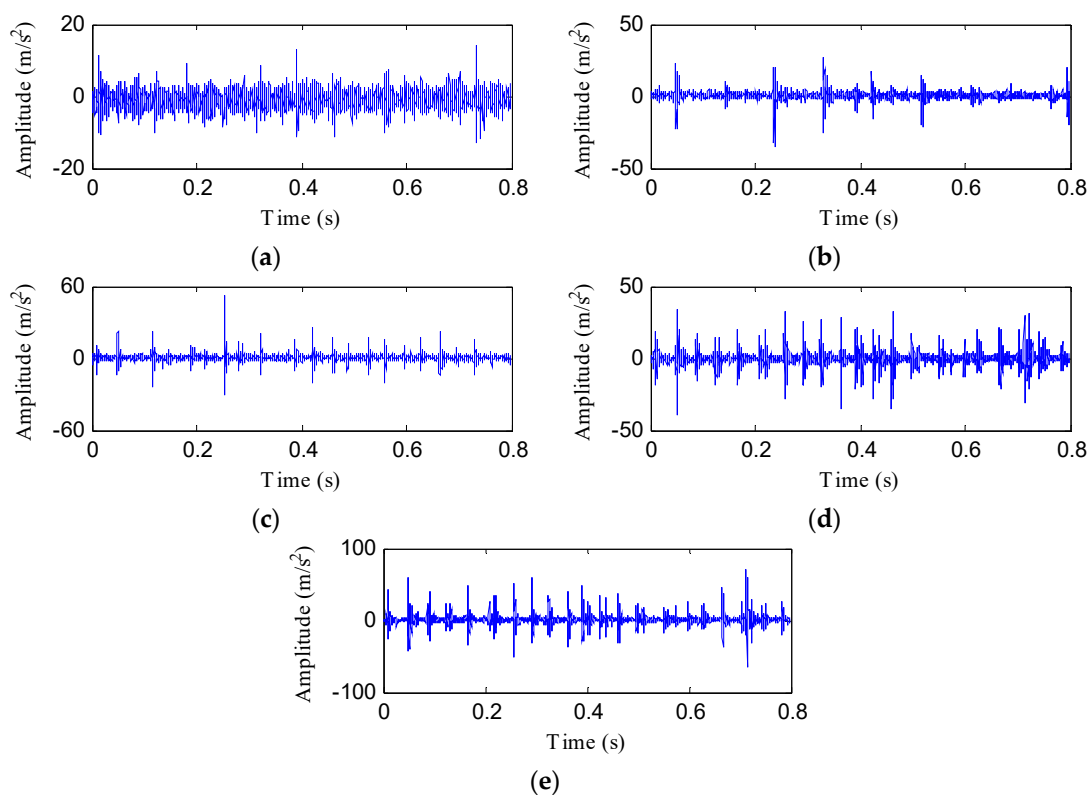
Figure 12. Data samples of the four health conditions: (**a**) Normal; (**b**) PT fault; (**c**) BT fault; (**d**) WT fault; and (**e**) BT-WT.

Then, the IED analysis is applied to the five selected data samples and the acquired corresponding IED signals are shown in Figure 13. Compared to the original signals, the impacts of the IED signal are more obvious as the noise is inhibited by the IED analysis. The HDE of the IED signals are calculated to obtain the HIEDDE fault features, as illustrated in Figure 14. The HIEDDE fault features under the four health conditions can be distinguished easily. The HIEDDE values of the eight nodes under health state are bigger than the three fault states. This implies that the signals under fault states are more regular. Based on the HIEDDE fault features, the DTW distances between the testing samples and the template samples are calculated and the results are represented in Figure 15. The labels of the five template samples correspond to the normal condition, PT fault, BT fault, WT and BT-WT faults are 1, 2, 3, 4, and 5, respectively. Figure 15a shows that the DTW distances between the testing sample under the normal condition and the five template samples have the minimum value at template label

1, which indicates that the health condition type of the normal state is accurately recognized. At the same time, we can find that the fault types of the testing samples under the four fault states are all identified based on the results shown in Figure 15b–e. Figure 16 shows the classification rate of the total 200 groups of testing data samples under the four health conditions by using the proposed method (HIEDDE-DTW). The classification rate is 100%.
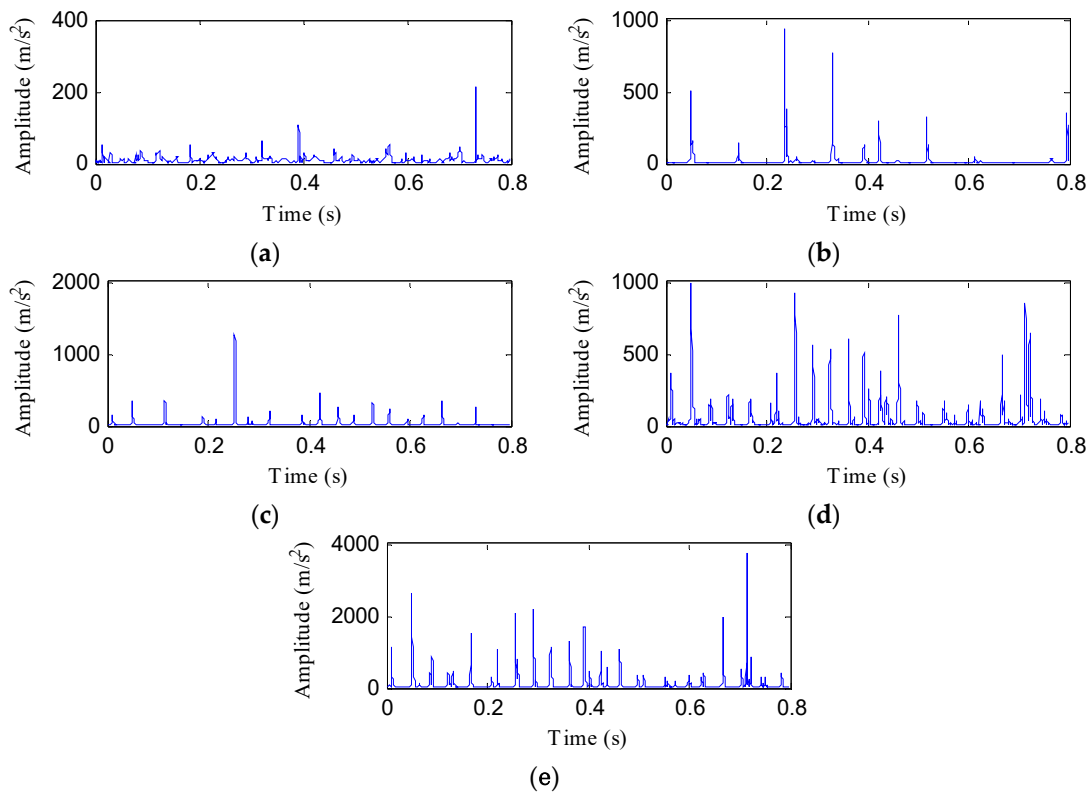

Figure 13. The IED signal of the four health conditions: (**a**) Normal; (**b**) PT fault; (**c**) BT fault; (**d**) WT fault; and (**e**) BT-WT.
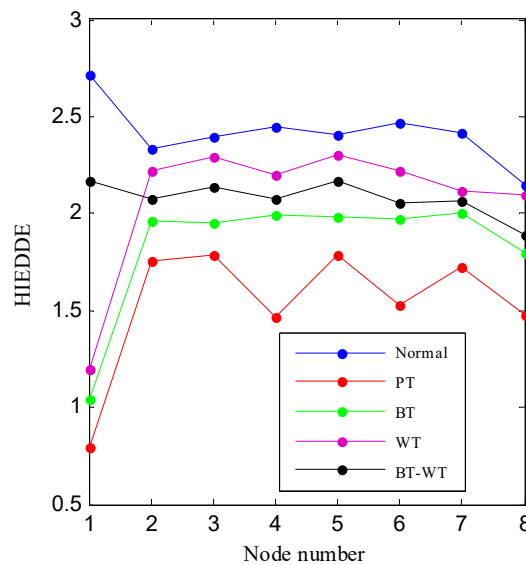
Figure 14. The HIEDDE fault features of the four data samples shown in Figure 11.
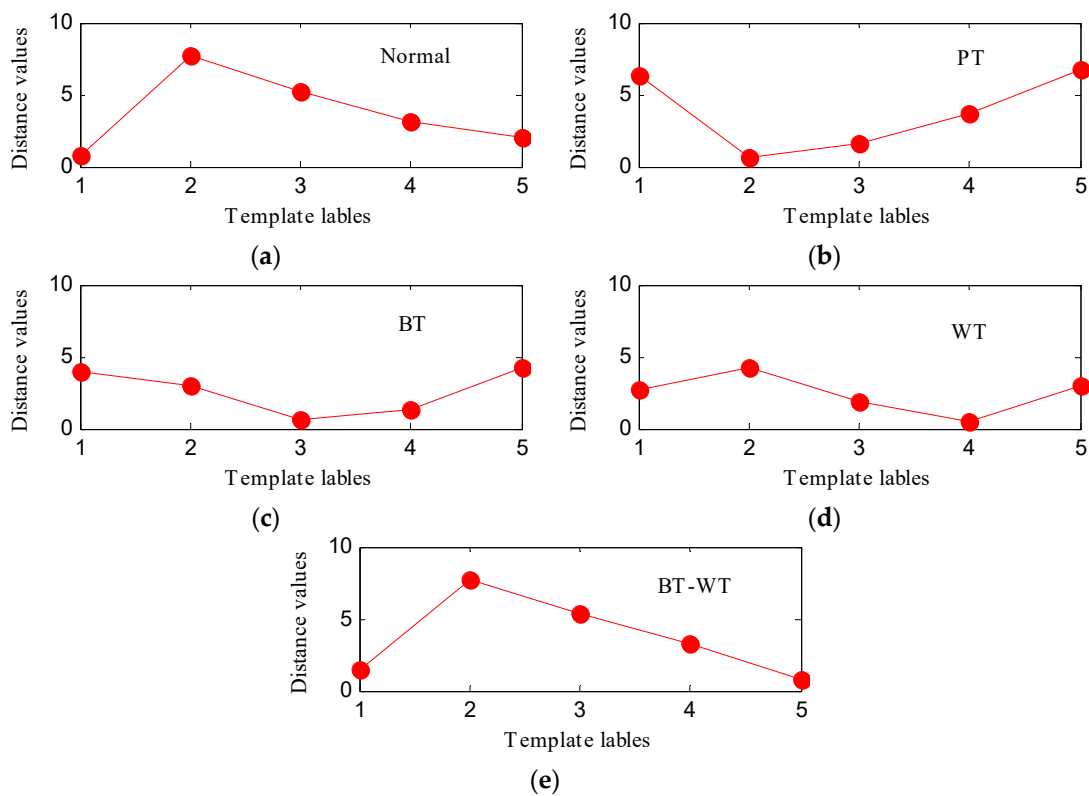
Figure 15. The similarity comparing results based on DTW: (**a**) Normal state; (**b**) PT; (**c**) BT; (**d**) WT; and (**e**) BT-WT.
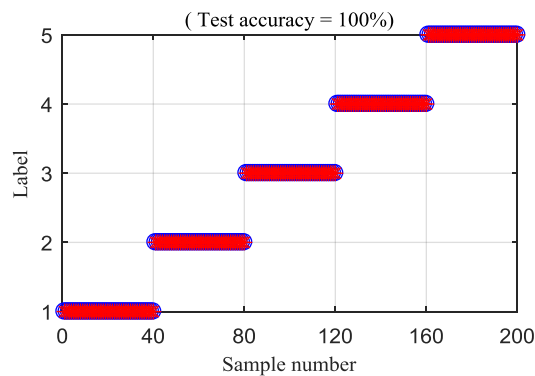
Figure 16. Classification rate of the proposed HIEDDE (IED-HDE) algorithm.

HIEDDE is the combination of these two algorithms, i.e., the IED analysis and the proposed HDE algorithm. In order to illustrate the significance of the combination of the two algorithms, the HDE algorithm is directly applied to the five original data samples, as shown in Figure 12, and the corresponding HDE fault features are displayed in Figure 17. The HDE values of some nodes under different health conditions are almost coincidental. The difference of the HDE fault feature under the four health conditions shown in Figure 17 is not as high as that of the HIEDDE fault features under the five health conditions displayed in Figure 14, which illustrates that the IED analysis can help to enhance the fault features. Figure 18 shows the classification rate of all testing samples by using the HDE fault features. Five testing samples of the PT fault are judged as the BT fault, and five testing samples of the BT-WT fault are misclassified, which enables the classification rate to be 95%.
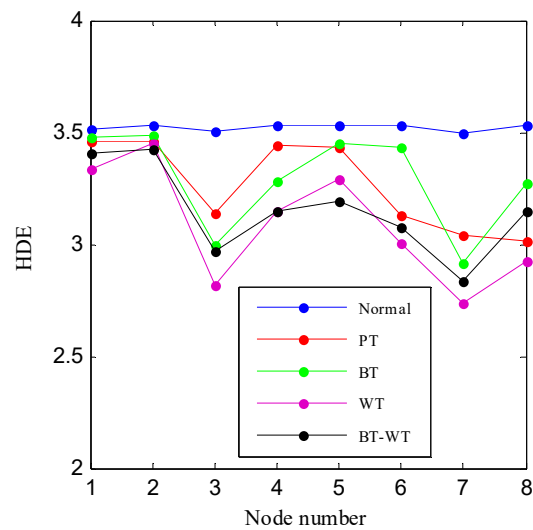
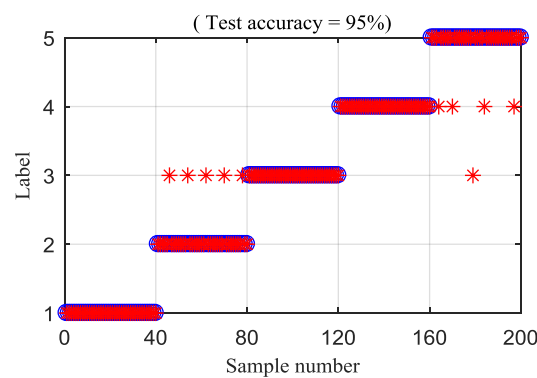**Figure 17.** The HDE fault features of the four data samples shown in Figure 11.

Figure 18. Classification rate of direct application of HDE.

To explain the superiority of the proposed method, the MDE and RCMDE algorithms are integrated with the IED analysis to analyze the experimental signals. Figures 19 and 20 illustrate the DTW classification results of the IED-MDE and IED-RCMDE approaches, respectively. The classification rate of the IED-MDE and IED-RCMDE approaches are 90.5% and 93%, respectively. The superiority of the proposed method is emphasized by the comparing results.
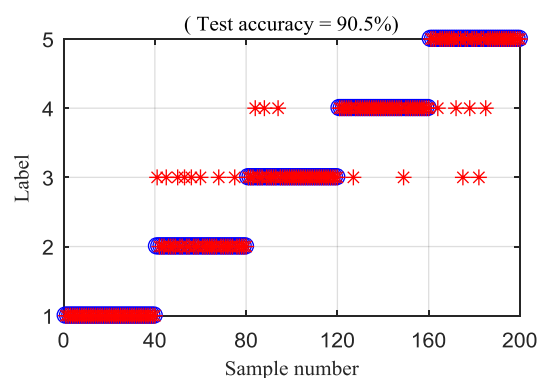
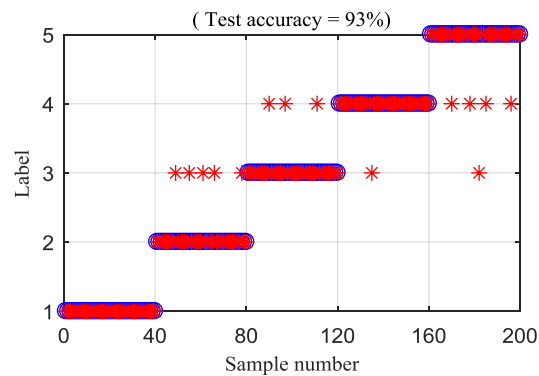**Figure 19.** Classification rate of the IED-MDE method.

Figure 20. Classification rate of the IED-RCMDE method.

6. Conclusions

To make an accurate fault diagnosis of gearbox, a novel fault diagnosis framework is put forward by combing HIEDDE and DTW. The major innovations can be summarized as follows: (1) The IED analysis based on SSD and HT is introduced for fault feature enhancement of gearbox signals; (2) a novel entropy algorithm named HDE, which is able to quantify the information of the low-frequency and high-frequency components of signals, is presented; and (3) the advantages of IED and HDE are integrated to develop the novel fault feature extraction algorithm, i.e., HIEDDE.

The analysis results of a simulated gear fault signal illustrate the fault feature enhancement ability of the IED analysis. The analysis results of the GWN signal demonstrate that the HDE algorithm is more stable and has higher accuracy compared with the MDE algorithm. The proposed method is used to identify the experimental fault patterns of gearbox. The results indicate the proposed method is able to recognize the fault patterns of gearbox accurately and has advantages over the existing methods.

It should be noticed that the validity of the proposed method is only studied by the experimental data obtained on the laboratory test bench. While satisfactory results have been obtained, the validity of the proposed method for practical engineering data needs to be further studied. The degree of fault difference between fault samples and the complexity of the fault may affect the accuracy of the proposed method. Moreover, the number of fault types that the proposed method can diagnose depends on the number of fault templates in the sample library. Only by establishing a rich sample library can the proposed method better solve the problem of gearbox fault diagnosis in engineering equipment.

References

1. Zhang, M.; Wang, K.; Wei, D.; Zuo, M.J. Amplitudes of characteristic frequencies for fault diagnosis of planetary gearbox. *J. Sound Vib.* **2018**, *432*, 119–132. [CrossRef]
2. Figlus, T. A method for diagnosing gearboxes of means of transport using multi-stage filtering and entropy. *Entropy* **2019**, *21*, 441. [CrossRef]
3. Sun, R.; Yang, Z.; Chen, X.; Tian, S.; Xie, Y. Gear fault diagnosis based on the structured sparsity time-frequency analysis. *Mech. Syst. Signal Process.* **2018**, *102*, 346–363. [CrossRef]

4.   Benmoussa, S.; Djeziri, M.A. Remaining useful life estimation without needing for prior knowledge of the degradation features. *Iet. Sci. Meas. Technol.* **2017**, *11*, 1071–1078. [CrossRef]

5.   Tang, J.; Shi, B.; Li, Z. Asymmetric delay feedback stochastic resonance method based on prior knowledge particle swarm optimization. *Chin. J. Phys.* **2018**, *56*, 2104–2118. [CrossRef]

6.   Guo, J.; Shi, Z.; Li, H.; Zhen, D.; Gu, F.; Ball, A.D. Early Fault Diagnosis for Planetary Gearbox Based Wavelet Packet Energy and Modulation Signal Bispectrum Analysis. *Sensors* **2018**, *18*, 2908. [CrossRef] [PubMed]

7.   Merainani, B.; Rahmoune, C.; Benazzouz, D.; Ould-Bouamama, B. A novel gearbox fault feature extraction and classification using Hilbert empirical wavelet transform, singular value decomposition, and SOM neural network. *J. Vib. Control* **2018**, *24*, 2512–2531. [CrossRef]

8.   Adamczak, S.; Stepien, K.; Wrzochal, M. Comparative study of measurement systems used to evaluate vibrations of rolling bearings. *Procedia Eng.* **2017**, *192*, 971–975. [CrossRef]

9.   Guo, Y.; Zhao, L.; Wu, X.; Na, J. Vibration separation technique based localized tooth fault detection of planetary gear sets: A tutorial. *Mech. Syst. Signal Process.* **2019**, *129*, 130–147. [CrossRef]

10. Li, Y.; Ding, K.; He, G.; Jiao, X. Non-stationary vibration feature extraction method based on sparse decomposition and order tracking for gearbox fault diagnosis. *Measurement* **2018**, *124*, 453–469. [CrossRef]

11. Glowacz, A.; Glowacz, W.; Glowacz, Z.; Kozik, J. Early fault diagnosis of bearing and stator faults of the single-phase induction motor using acoustic signals. *Measurement* **2018**, *113*, 1–9. [CrossRef]

12. Glowacz, A. Fault diagnosis of single-phase induction motor based on acoustic signals. *Mech. Syst. Signal Process.* **2019**, *117*, 65–80. [CrossRef]

13. Younus, A.M.; Yang, B.-S. Intelligent fault diagnosis of rotating machinery using infrared thermal image. *Expert Syst. Appl.* **2012**, *39*, 2082–2091. [CrossRef]

14. Jiang, G.Q.; He, H.B.; Yan, J.; Xie, P. Multiscale Convolutional Neural Networks for Fault Diagnosis of Wind Turbine Gearbox. *IEEE Trans. Ind. Electron.* **2018**. [CrossRef]

15. Pang, B.; Tang, G.; Tian, T.; Zhou, C. Rolling bearing fault diagnosis based on an improved HTT transform. *Sensors* **2018**, *18*, 1203. [CrossRef] [PubMed]

16. Parey, A.; Singh, A. Gearbox fault diagnosis using acoustic signals, continuous wavelet transform and adaptive neuro-fuzzy inference system. *Appl. Acoust.* **2019**, *147*, 133–140. [CrossRef]

17. Feng, Z.; Zhu, W.; Zhang, D. Time-frequency demodulation analysis via vold-kalman filter for wind turbine planetary gearbox fault diagnosis under nonstationary speeds. *Mech. Syst. Signal Process.* **2019**, *128*, 93–109. [CrossRef]

18. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lon. Ser. A* **1998**, *454*, 903–995. [CrossRef]

19. Moore, K.J.; Kurt, M.; Eriten, M.; McFarland, D.M.; Bergman, L.A.; Vakakis, A.F. Wavelet-bounded empirical mode decomposition for measured time series analysis. *Mech. Syst. Signal Process.* **2018**, *99*, 14–29. [CrossRef]

20. Park, S.; Kim, S.; Choi, J. Gear fault diagnosis using transmission error and ensemble empirical mode decomposition. *Mech. Syst. Signal Process.* **2018**, *108*, 262–275. [CrossRef]

21. Chen, J.Y.; Zhou, D.; Lyu, C.; Lu, C. An integrated method based on CEEMD-SampEn and the correlation analysis algorithm for the fault diagnosis of a gearbox under different working conditions. *Mech. Syst. Signal Process.* **2018**, *113*, 102–111. [CrossRef]

22. Li, H.; Qin, X.; Zhao, D.; Chen, J.; Wang, P. An improved empirical mode decomposition method based on the cubic trigonometric B-spline interpolation algorithm. *Appl. Math. Comput.* **2018**, *332*, 406–419. [CrossRef]

23. Feng, Z.P.; Lin, X.F.; Zuo, M.J. Joint amplitude and frequency demodulation analysis based on intrinsic time-scale decomposition for planetary gearbox fault diagnosis. *Mech. Syst. Signal Process.* **2016**, *72–73*, 223–240. [CrossRef]

24. Gao, Y.; Villecco, F.; Li, M.; Song, W. Multi-Scale Permutation Entropy Based on Improved LMD and HMM for Rolling Bearing Diagnosis. *Entropy* **2017**, *19*, 176. [CrossRef]

25. Zhang, K.; Chen, X.; Liao, L.; Tang, M.; Wu, J. A new rotating machinery fault diagnosis method based on local oscillatory-characteristic decomposition. *Digital Signal Processing* **2018**, *78*, 98–107. [CrossRef]

26. Li, Y.; Cheng, G.; Liu, C.; Chen, X. Study on planetary gear fault diagnosis based on variational mode decomposition and deep neural networks. *Measurement* **2018**, *130*, 94–104. [CrossRef]

27. Yu, D.; Wang, M.; Cheng, X. A method for the compound fault diagnosis of gearboxes based on morphological component analysis. *Measurement* **2016**, *91*, 519–531. [CrossRef]

28. Kong, Y.; Wang, T.Y.; Chu, F.L. Meshing frequency modulation assisted empirical wavelet transform for fault diagnosis of wind turbine planetary ring gear. *Renew. Energy* **2018**, *132*, 1373–1388. [CrossRef]

29. Bonizzi, P.; Karel, J.M.H.; Meste, O.; Peeters, R.M. Singular spectrum decomposition: A new method for time series decomposition. *Adv. Adapt. Data Anal.* **2014**, *6*, 1–34. [CrossRef]

30. Chen, X.; Peng, L.; Cheng, G.; Luo, C. Research on degradation state recognition of planetary gear based on multiscale information dimension of SSD and CNN. *Complexity* **2019**, *2019*. [CrossRef]

31. Pang, B.; Tang, G.; He, Y. Research on singular spectrum decomposition and its application to rotor failure detection. *J. Vibroengineering* **2018**, *20*, 2336–2351.

32. Du, W.; Zhou, J.; Wang, Z.; Li, R.; Wang, J. Application of Improved Singular Spectrum Decomposition Method for Composite Fault Diagnosis of Gear Boxes. *Sensors* **2018**, *18*, 3804. [CrossRef] [PubMed]

33. Shakya, P.; Darpe, A.K.; Kulkarni, M. Bearing damage classification using instantaneous energy density. *J. Vib. Control* **2015**, *23*, 2578–2618. [CrossRef]

34. Loutridis, S. Instantaneous energy density as a feature for gear fault detection. *Mech. Syst. Signal Process.* **2006**, *20*, 1239–1253. [CrossRef]

35. Luwei, K.C.; Yunusa-Kaltungo, A.; Sha'aban, Y.A. Integrated Fault Detection Framework for Classifying Rotating Machine Faults Using Frequency Domain Data Fusion and Artificial Neural Networks. *Machines* **2018**, *6*, 59. [CrossRef]

36. Villecco, F.; Pellegrino, A. Entropic Measure of Epistemic Uncertainties in Multibody System Models by Axiomatic Design. *Entropy* **2017**, *19*, 291. [CrossRef]

37. Rostaghi, M.; Azami, H. Dispersion Entropy: A Measure for Time-Series Analysis. *IEEE Signal Process. Lett.* **2016**, *23*, 610–614. [CrossRef]

38. Rostaghi, M.; Ashory, M.R.; Azami, H. Application of dispersion entropy to status characterization of rotary machines. *J. Sound Vib.* **2019**, *438*, 291–308. [CrossRef]

39. Zhang, Y.; Tong, S.; Cong, F.; Xu, J. Research of Feature Extraction Method Based on Sparse Reconstruction and Multiscale Dispersion Entropy. *Appl. Sci.* **2018**, *8*, 888. [CrossRef]

40. Azami, H.; Rostaghi, M.; Abasolo, D.; Escudero, J. Refined Composite Multiscale Dispersion Entropy and its Application to Biomedical Signals. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 2872–2879.

41. Li, Y.; Yang, Y.; Wang, X.; Liu, B.; Liang, X. Early fault diagnosis of rolling bearings based on hierarchical symbol dynamic entropy and binary tree support vector machine. *J. Sound Vib.* **2018**, *428*, 72–86. [CrossRef]

42. Jiang, Y.; Peng, C.K.; Xu, Y.S. Hierarchical entropy analysis for biological signals. *J. Comput. Appl. Math.* **2011**, *236*, 728–742. [CrossRef]

43. Li, Y.; Xu, M.; Zhao, H.; Huang, W. Hierarchical fuzzy entropy and improved support vector machine based binary tree approach for rolling bearing fault diagnosis. *Mech. Mach. Theory* **2016**, *98*, 114–132. [CrossRef]

44. Sakoe, H.; Chiba, S.; Waibel, A.; Lee, K. Dynamic programming algorithm optimization for spoken word recognition. *Read. Speech Recognit.* **1990**, *159*, 224.

45. Tian, Y.; Wang, Z.; Lu, C. Self-adaptive bearing fault diagnosis based on permutation entropy and manifold-based dynamic time warping. *Mech. Syst. Signal Process.* **2016**. [CrossRef]

46. Li, L.; Liu, G.; Zhang, L.; Li, Q. Sensor fault detection with generalized likelihood ratio and correlation coefficient for bridge shm. *J. Sound Vib.* **2019**, *442*, 445–458. [CrossRef]

47. McFadden, P. Detecting fatigue cracks in gears by amplitude and phase demodulation of the meshing vibration. *J. Vib. Acoust. Stress Reliab. Des.* **1986**, *108*, 165–170. [CrossRef]