

Article

Rapid Multi-Sensor Feature Fusion Based on Non-Stationary Kernel JADE for the Small-Amplitude Hunting Monitoring of High-Speed Trains

Jing Ning ^{1,2,*}, Mingkuan Fang ¹, Wei Ran ¹, Chunjun Chen ¹ and Yanping Li ¹

¹ School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, Sichuan, China; 13971658422@163.com (M.F.); ranwei@swjtu.edu.cn (W.R.); cjchen@swjtu.edu.cn (C.C.); signal2020_swj@163.com (Y.L.)

² Technology and Equipment of Rail Transit Operation and Maintenance Key Laboratory of Sichuan Province, Chengdu 610031, Sichuan, China

* Correspondence: ningjing@swjtu.cn; Tel./Fax: +86-28-87600690

Received: 29 March 2020; Accepted: 16 June 2020; Published: 18 June 2020



Abstract: Joint Approximate Diagonalization of Eigen-matrices (JADE) cannot deal with non-stationary data. Therefore, in this paper, a method called Non-stationary Kernel JADE (NKJADE) is proposed, which can extract non-stationary features and fuse multi-sensor features precisely and rapidly. In this method, the non-stationarity of the data is considered and the data from multi-sensor are used to fuse the features efficiently. The method is compared with EEMD-SVD-LTSA and EEMD-JADE using the bearing fault data of CWRU, and the validity of the method is verified. Considering that the vibration signals of high-speed trains are typically non-stationary, it is necessary to utilize a rapid feature fusion method to identify the evolutionary trends of hunting motions quickly before the phenomenon is fully manifested. In this paper, the proposed method is applied to identify the evolutionary trend of hunting motions quickly and accurately. Results verify that the accuracy of this method is much higher than that of the EEMD-JADE and EEMD-SVD-LTSA methods. This method can also be used to fuse multi-sensor features of non-stationary data rapidly.

Keywords: high-speed trains; hunting; non-stationary; feature fusion; multi-sensor fusion

1. Introduction

Hunting motion is a self-excited vibration that is a serious obstacle to the safety of high-speed trains [1]. Monitoring systems are designed to detect hunting only after it has developed to a specific degree. Besides, in most cases, the recognition result is only obtained using a single observation. Therefore, the accuracy and real-time performance of monitoring systems need to be further improved [2]. With the increasing performance requirements of high-speed trains, it is important to establish an accurate and rapid feature extraction method for hunting detection through multi-characterizations before the phenomenon has developed to any significant degree.

The structure of high-speed trains is very complicated and their working conditions are very poor, resulting in non-stationary vibration signals [3]. In the existing feature extraction research, a variety of data extraction algorithms are utilized. These algorithms can be roughly divided into two categories: those designed for stationary data and those for non-stationary data. Feature extraction methods for stationary data include Singular Value Decomposition (SVD), Linear Discriminant Analysis (LDA) [4,5], Principal Component Analysis (PCA) [6], Locality Preserving Projection (LPP) [7], and so on. However, it is hard to extract features of non-stationary data using these methods. For non-stationary data, manifold learning is a good feature extraction method [8–11]; however, it incurs a high computational cost and has extremely long calculation times, which means that the diagnostic information cannot be

fed back to the system in time. Therefore, a rapid yet precise method to extract the non-stationary signal in practical engineering applications is needed. Cardoso [12,13] proposed a method named Joint Approximate Diagonalization of Eigen-matrices (JADE) in the field of blind source separation, which is used to quickly separate multiple features. Because the method is simple and effective, it is also widely used in pattern recognition. Liu [14] improved the JADE method and achieved a good feature fusion performance by simplifying the calculations, and then developed a method based on kernel JADE to identify the rolling bearing fault [15]. The method based on JADE was adopted to diagnose the faulty parts of the rolling bearing in [16], and a similar method was applied to predict coaxial bearing performance degradation in [17]. However, in this application, the non-stationary and rapid character of the signal is not considered.

The structure of high-speed trains is complex. The vibration signals from high-speed trains are affected by many factors, such as the propagation path and the measuring point location, which lead the signal to couple with different characteristic information. Pattern recognition based on single sensor vibration signals has limitations, which cannot completely render the evolutionary characteristics of high-speed train vibrations [18]. Therefore, it is necessary to utilize a multi-sensor fusion method to extract the features of the train's operating state.

Furthermore, the evaluation parameters of lateral stability of railway passenger trains in different countries are different. Lateral force on the rail and on the wheel axis, lateral acceleration of the bogie frame, acceleration of the axle-box, and lateral acceleration of the vehicle body can be used as evaluation parameters, respectively [19–22]. However, only one parameter is used in each standard, and this parameter has different limits. For example, the influence of frequency is not considered when lateral acceleration of the bogie frame is used. In fact, the acceleration of the axle-box is a particularly important index in hunting monitoring. Therefore, in this paper, we try to fuse the signal from the bogie frame and axle box to extract the hunting motion features. To monitor the state of hunting of high-speed trains, four basic states are classified: normal, small convergence, small divergence, and hunting. Our aim is to identify the small divergence state before hunting occurs. Thus, in this paper, the monitoring of small amplitude hunting enables us to rapidly distinguish between these four basic states online [23].

Moreover, for real-time classification, it is extremely important to ensure that once the small amplitude appears to diverge, the entire calculation process can be completed immediately. Therefore, the time of the whole calculation process must be short enough for real-time classification.

Considering the above problems, a rapid multi-sensor feature fusion method based on the Non-stationary Kernel JADE method is proposed. The JADE method is a fast and accurate feature fusion algorithm, but it is generally used in stable environments. In order to use the algorithm in a non-stationary environment, the whole time series is divided into M time periods [24] and the kernel function is introduced. Then, M kernel matrices are obtained by the M time periods, which are jointly decomposed. After this, Jacobian rotation is used to obtain the unitary matrix by diagonalizing multiple kernel matrices simultaneously to extract the non-stationary fusion features. In addition, in order to visualize the data features, the extracted fusion features are expressed in three dimensions. The between-class indicator and within-class indicators [25] are also employed to describe the clustering performance of the features quantitatively.

In this paper, a multi-sensor data feature extraction framework is provided, in which a rapid feature fusion method using Non-stationary Kernel JADE (NKJADE) is proposed. This framework consists of the following series of steps. First, the Ensemble Empirical Mode Decomposition (EEMD) method is utilized to decompose the preprocessed signals to Intrinsic Mode Functions (IMFs). Then, the energy matrices are obtained using the IMFs and the fusion features are obtained through NKJADE, followed by inputting the extracted features to an LSSVM [26,27] for training and recognition.

Data from Case Western Reserve University (CWRU) are used to verify the performance of this method against that of the SVD and JADE methods. In this paper, the proposed method is also applied

to identify the evolutionary trend of the hunting motion quickly and accurately. The results verify that the accuracy of this method is much higher than that of the JADE method.

The remainder of this paper is organized as follows: In Section 2, the theoretical backgrounds of Ensemble Empirical Mode Decomposition (EEMD) [28] and a separability metric are introduced. The proposed method of non-stationary Kernel JADE is also described. The framework of the proposed method is introduced in Section 3. In Section 4, data from Case Western Reserve University (CWRU) are used to test the proposed method, and the operational data of multiple states of high-speed trains are used to verify the accuracy and rapidity of the proposed method. Finally, the conclusion is presented in Section 5.

2. Theoretical Background

2.1. EEMD

The basic idea of the EEMD method is to use the sifting process to decompose the signal into several intrinsic mode functions, when Gaussian white noise is added. For the original signal, the specific steps to use EEMD to decompose it into IMFs are as follows:

- (1) Obtain the overall signal by adding Gaussian white noise to the original signal:

$$a'(t) = a(t) + \omega(t) \quad (1)$$

- (2) The overall signal is decomposed to obtain the IMF components of each order, where i represents the i -th component, r is the residual term, and n is the number of IMFs:

$$a'(t) = \sum_i^n c_i + r \quad (2)$$

- (3) Repeat step (1) and (2), each time adding different white noise sequences of the same amplitude:

$$a'_j(t) = \sum_{i=1}^n c_{ij} + r_j \quad (3)$$

In Equation (3), c_{ij} is the i -th IMF component of the decomposition to which white noise is added for the j -th time, while r_j is the residual value of the decomposition.

- (4) Using the zero-mean principle of the Gaussian white noise frequency, the effect of white noise can be eliminated, and the IMF component corresponding to the original signal can be expressed as:

$$c_i(t) = \frac{1}{N} \sum_{j=1}^N c_{ij}(t) \quad (4)$$

where n represents the number of times the white noise is added and c_i represents the i -th IMF component obtained by the EEMD decomposition of the original signal.

2.2. IMF Energy Matrix

For the IMF component $c'_i = [c'_i(1), \dots, c'_i(M)]$, the energy feature can be expressed as:

$$e_i = \sum_{j=1}^M [c'_i(j)]^2 \quad (5)$$

All IMF components of a vibration signal form a feature energy vector $e = [e_1, e_2, \dots, e_n]$, where n is the dimension of vector e and represents the number of IMF components.

2.3. Proposed Non-Stationary Kernel JADE

Since the original JADE algorithm is based on stationary signal analysis, considering the non-stationary nature of high-speed train signal, the kernel JADE method [15] is applied to a non-stationary environment.

The main idea of the kernel is to map the input matrix into the nonlinear space Φ . Suppose $X = \{x_1, x_2, \dots, x_m\}$; then, the processing can be defined as follows:

$$\{x_1, x_2, \dots, x_m\} \rightarrow \{\Phi(x_1), \Phi(x_2), \dots, \Phi(x_m)\} \quad (6)$$

During implementation, we need to calculate the inner product of two eigenvectors which have been mapped into a nonlinear space using a kernel function, and a kernel matrix will be calculated using Equation (7):

$$K_{ij} = k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle \quad (7)$$

where x_i, x_j are vectors. The commonly used kernel function includes the following [29]:

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\delta^2}\right) \quad (8)$$

$$k(x_i, x_j) = (\alpha x_i^T x_j + c)^d \quad (9)$$

$$k(x_i, x_j) = \tanh(\alpha x_i^T x_j + c) \quad (10)$$

Similar to KPCA [30], the centered kernel matrix K can be calculated through Equation (11):

$$K_{ij} = K_{ij} - \frac{1}{M} \sum_{r=1}^M K_{ir} - \frac{1}{M} \sum_{r=1}^M K_{rj} + \frac{1}{M^2} \sum_{r,s=1}^M K_{rs} \quad (11)$$

The whole time series is divided into M segment intervals T_1, \dots, T_M , and consequently M covariance matrices S_{T_1}, \dots, S_{T_M} can be generated. These M covariance matrices are jointly diagonalized to find a unitary matrix U which can diagonalize M covariance matrices simultaneously; then, the energy features can be extracted.

Step 1: For M segment intervals T_1, \dots, T_M , the covariance matrices of signal $x(t)$ can be expressed as:

$$Cov_{T_k} = S_{T_k} = \frac{1}{|T_m|} \sum_{t \in T_m} [k(s_t, s_t)] \quad (12)$$

where $s_t = x_t - E(x_t)$, $E(x_t)$ denotes the mean of x_t .

Step 2: The most common method to diagonalize matrices S_{T_1}, \dots, S_{T_M} is to diagonalize the first matrix, and then transform the remaining $M-1$ matrices into diagonalization. W is the diagonalized matrix of the covariance matrix S_{T_1} :

$$W = S_{T_1}^{-1/2} = V^H \Lambda^{-1/2} V \quad (13)$$

In Equation (13), V is the eigen-matrix of S_{T_1} , while Λ is the eigenvector of S_{T_1} .

For the remaining $M-1$ matrices, the diagonalization matrix can be defined respectively:

$$S_{T_m}^* = S_{T_1}^{-1/2} S_{T_m} (S_{T_1}^{-1/2})^H, m = 2, \dots, M \quad (14)$$

Step 3: The approximate joint diagonalization problem is equivalent to finding an orthogonal matrix \mathbf{U} that minimizes:

$$\sum_{m=2}^M \|\text{off}(\mathbf{US}_{T_m}^* \mathbf{U}^H)\|^2 = \sum_{m=2}^M \sum_{b \neq d} (\mathbf{US}_{T_m}^* \mathbf{U}^H)_{bd}^2 \quad (15)$$

where $\text{off}(\mathbf{US}_{T_m}^* \mathbf{U}^H)$ has the same off-diagonal elements as $\mathbf{US}_{T_m}^* \mathbf{U}^H$, and the diagonal element is zero, while b and d represent the b -th row and d -th column of the matrix, respectively.

Since the sum of squares remains the same when multiplied by an orthogonal matrix, the problem is equivalent to maximizing the sum of squares of the diagonal elements:

$$\sum_{m=2}^M \|\text{diag}(\mathbf{US}_{T_m}^* \mathbf{U}^H)\|^2 = \sum_{m=2}^M \sum_{b=1}^p (\mathbf{US}_{T_m}^* \mathbf{U}^H)_{bb}^2 \quad (16)$$

where p represents the dimension to which the feature is extracted.

Step 4: Givens rotation is used to transform the set of matrices to a more diagonal form, two rows and two columns at a time. The Givens rotation matrix is given by:

$$G(i, j, \theta) = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & \cos(\theta) & \cdots & -\sin(\theta) & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & \sin(\theta) & \cdots & \cos(\theta) & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} \quad (17)$$

where

$$\theta = \frac{1}{2} \text{arccot} \left[\frac{(\mathbf{S}_{T_m}^*)_{22} - (\mathbf{S}_{T_m}^*)_{11}}{2(\mathbf{S}_{T_m}^*)_{12}} \right] \quad (18)$$

The initial value for the orthogonal matrix \mathbf{U} is the identity matrix \mathbf{I} . Matrices $\mathbf{S}_{T_2}, \dots, \mathbf{S}_{T_m}$ are then updated using:

$$\begin{aligned} \mathbf{S}_{T_m}^* &\leftarrow \mathbf{G}(1, 2, \theta) \mathbf{S}_{T_m}^* \mathbf{G}(1, 2, \theta), m = 2, \dots, M \\ \mathbf{U} &\leftarrow \mathbf{U} \mathbf{G}(1, 2, \theta) \end{aligned} \quad (19)$$

When the values of all non-diagonal elements are less than a given threshold ε , the iteration is completed, and the joint approximation diagonalization is achieved. The unitary matrix $\hat{\mathbf{U}}$ is obtained so that multiple matrices are diagonalized jointly.

Step 5: The transform matrix \mathbf{A} can be calculated as $\mathbf{A} = \hat{\mathbf{U}} \mathbf{W}^\#$, where superscript # denotes the pseudo-inverse.

As such, the original signal $s(t)$ can be expressed as:

$$s(t) = \mathbf{A} \cdot \mathbf{K} \quad (20)$$

Since the NKJADE method can extract the nonlinear relationships hidden in the high-dimensional feature space, the fusion features can be estimated through the joint feature decomposition using multiple inputs. The fusion features can express the nonlinear and non-stationary relationships hidden in the inputs well, so they can represent the characteristic relationship of different states, and then distinguish different feature states quickly and accurately.

2.4. Separability Evaluation

To illustrate the merit of our proposed algorithm, the separability J is utilized to demonstrate the algorithm's ability to form distinct classes. The capability of the feature extraction in pattern classification can be described quantitatively using between-class scatter S_b , within-class scatter S_w [31], and separability [32]. Assuming that the data have a total of C classifications, the vector of the i -th classification is:

$$x^i = (x_1^i, x_2^i, \dots, x_{n_i}^i) \quad (21)$$

where n_i is the number of i -th classification.

Between-class scatter S_b and within-class scatter S_w can be calculated as follows:

$$S_b = \sum_{i=1}^C p_i (m_i - m)(m_i - m)^T \quad (22)$$

$$S_w = \sum_{i=1}^C \left[\frac{p_i}{n_i} \sum_{k=1}^{n_i} (x_k^i - m_i)(x_k^i - m_i)^T \right] \quad (23)$$

where $p_i = n_i / \sum_{j=1}^C n_j$, $m_i = \text{mean}(x_{n_i}^i)$, $m = \sum_{i=1}^C p_i m_i$.

The between-class scatter S_b describes how far different classes are separated, and the within-class scatter S_w indicates how compactly each class of samples is distributed. Based on between-class scatter and within-class scatter, separable evaluation J is introduced to describe the clustering ability of different methods. Separable evaluation J could be calculated as follows:

$$J = \text{trace}(S_b / S_w) \quad (24)$$

where function *trace* refers to the sum of diagonal elements.

3. Methodology

For the different characteristics of vibration signals, a multi-sensor data feature extraction framework is provided, in which a rapid feature fusion method using NKJADE is proposed. The framework of the feature extraction method is shown in Figure 1. The method can quickly and accurately extract different features of information contained in non-stationary vibration signals. The main steps of this framework are as follows:

- (1) The m position sensor data and l classes are selected, and a total of m sample matrices are obtained.
- (2) The EEMD algorithm is used to separately decompose the sample matrices of m positions, then n IMF components of each sample signal can be obtained.
- (3) By extracting the energy features of the n components of IMF obtained using decomposition, an IMF energy matrix is obtained from each position. In this paper, a total of m energy matrices of IMF were obtained.
- (4) For the obtained energy matrices of IMF of m positions, a rapid feature fusion method using NKJADE is proposed, so the dimensionality is reduced to three for a better-observed performance.
- (5) The LSSVM is trained on and used to test the fusion features to verify the accuracy of the method.

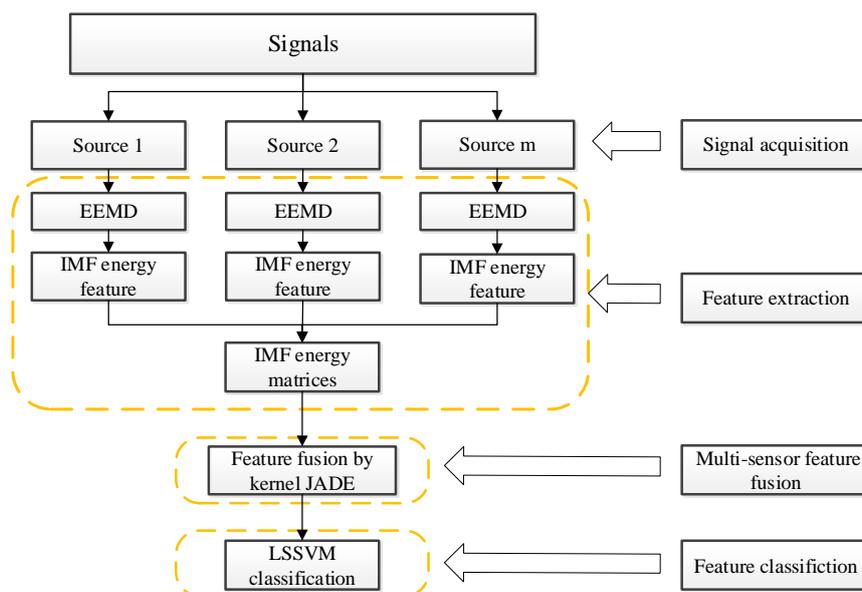


Figure 1. Method framework.

4. Experiment Results and Analysis

In order to verify the valid of the method, we applied it to bearing fault identification (Case I) and small amplitude hunting monitoring of high-speed trains (Case II). Some traditional approaches, such as Ensemble Empirical Mode Decomposition Joint Approximate Diagonalization of Eigen-matrices (EEMD-JADE) and Ensemble Empirical Mode Decomposition Singular Valuable Decomposition Learning Technology Systems Architecture (EEMD-SVD-LTSA), were compared with the proposed method.

4.1. Case I—Case Western Reserve University Data

4.1.1. Data Description

The bearing test data for normal and faulty bearings were from the Case Western Reserve University (CWRU). The test signal contained normal state, ball fault, and inner and outer race faults; for the latter three, the fault diameters were 0.07, 0.14 and 0.21 inches, respectively.

The data used in this paper were collected from the drive end. The sampling frequency was 12 KHZ and the speed of the shaft was 1725 r/min, corresponding to 400 points collected per revolution. In order to reduce the influence of equipment fluctuations, each sample contained 800 points.

As shown in Table 1, two datasets were selected. Dataset A had the same fault location (inner race fault), but the fault size was different. Dataset B had different fault locations, but the fault size was the same, and the outer race fault was at the location of 3 o'clock. Each of the datasets was divided into three categories.

Table 1. Description of dataset.

Datasets	Class	Number of Training Samples	Number of Testing Samples	Label
Dataset A	Inner race, 0.07"	70	30	1
	Inner race, 0.14"	70	30	2
	Inner race, 0.21"	70	30	3
Dataset B	Inner race, 0.07"	70	30	1
	Ball, 0.07"	70	30	4
	Outer Race, 0.07"	70	30	5

4.1.2. Signal Processing Results

The scatter plots obtained by applying the EEMD-SVD-LTSA, EEMD-JADE method and EEMD-NKJADE methods on dataset A are shown in Figure 2a–c, while the corresponding scatter plots for dataset B are shown in Figure 2d–f. The results of selecting parameters for kernel are shown in Figure 3. The results of the application of the three algorithms in dataset A are shown in Table 2, while the results for dataset B are shown in Table 3.

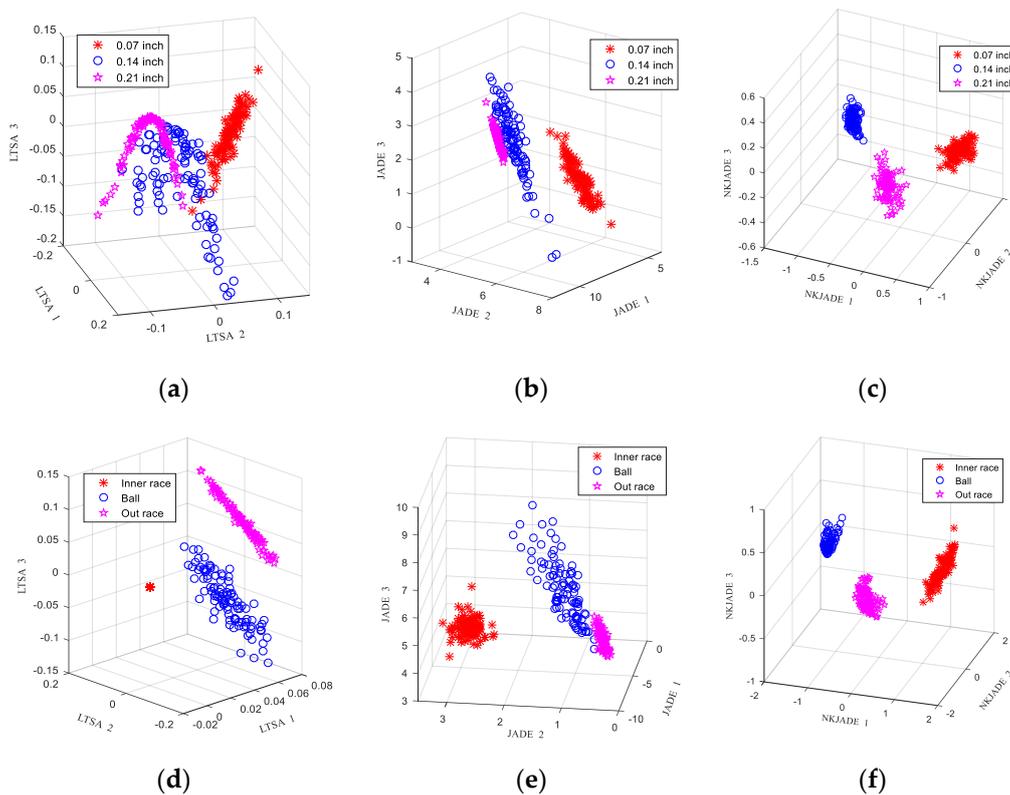


Figure 2. Scatter plots of faults. (a) EEMD-SVD-LTSA on dataset A; (b) EEMD-JADE on dataset A; (c) EEMD-NKJADE on dataset A; (d) EEMD-SVD-LTSA on dataset B; (e) EEMD-JADE on dataset B; (f) EEMD-NKJADE on dataset B.

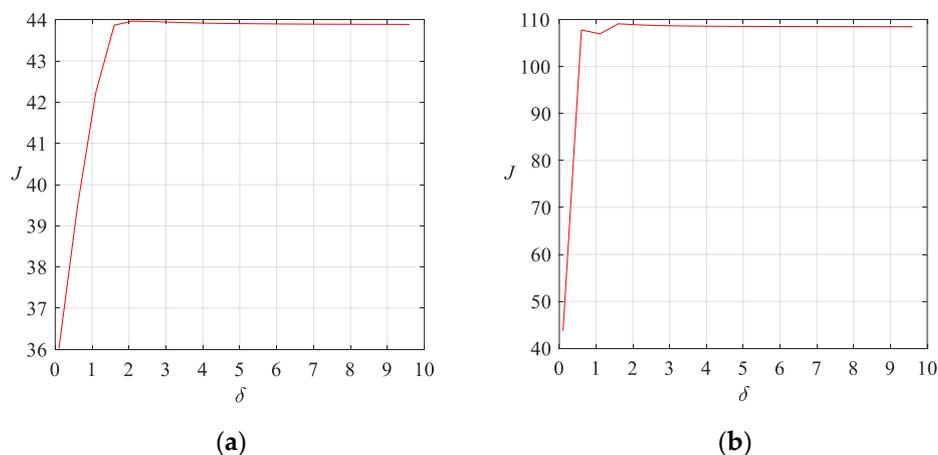


Figure 3. The results of selecting parameters for the Gaussian kernel. (a) Parameter selection results of Gaussian kernel of EEMD-NKJADE method on dataset A; (b) Parameter selection results of Gaussian kernel of EEMD-NKJADE method on dataset B.

Table 2. Results from dataset A.

Method	EEMD-SVD-LTSA	EEMD-JADE	EEMD-NKJADE
Accuracy (%)	98.33	98.89	100
J	36.26	39.24	43.91
Time (s)	1.7593	0.7820	1.0070

Table 3. Results from dataset B.

Method	EEMD-SVD-LTSA	EEMD-JADE	EEMD-NKJADE
Accuracy (%)	100	98.89	100
J	1.195×10^{24}	47.96	87.96
Time (s)	1.5513	0.7190	0.9650

4.1.3. Discussion

We found that the generalization ability of the Gaussian kernel is better than that of the other kernels. Therefore, we focused on the parameters of the Gaussian kernel.

According to the results of selecting parameters in [15], we tried to set the range of σ to (1, 10). The parameter was incremented step by step (parameter step is 0.5, shown in X-axis), and the separability J was calculated (shown in Y-axis). The larger J is, the better the classification result will be. Therefore, for the parameter selection of EEMD method, the optimal parameter of the Gaussian kernel on dataset A and dataset B is 2.

From Figure 2 and Table 2, we see that all three methods could extract the features with satisfactory results. The accuracy of the three methods was nearly 100% in all cases. The result may be attributed to the CWRU bearing fault data having great differences, which are quite easy to classify.

However, the separability J in Table 2 is different, which is consistent with (a)–(c) in Figure 2. In this respect, the classification ability of EEMD- NKJADE algorithm is obviously better than other algorithms. In addition, the results of the running time in Table 2 show that the time required for EEMD-JADE calculation was relatively short (the PC configuration was as follows: Intel Core i5-4460, 12GB of memory, NVIDIA GeForce GT720).

In order to further verify the robustness of the algorithm, we also tested the results of the algorithm on dataset B.

In Table 3, the accuracy of the three methods in dataset B was almost the same as that in dataset A. However, the separability of dataset B was the highest after being processed using the EEMD-SVD-LTSA algorithm, while in dataset A, the separability achieved using EEMD-SVD-LTSA was the lowest, which indicates that the selection of dataset had a greater impact on EEMD-SVD-LTSA. Compared with EEMD-SVD-LTSA, the results from the EEMD-NKJADE and EEMD-JADE algorithms were less affected by the different dataset, and therefore seem to be more robust. Therefore, the method of EEMD-NKJADE offers superior performance with respect to the classification effect and the robustness, but its calculation time is longer than EEMD-JADE.

4.2. Case II—Small Amplitude Hunting Monitoring of High-Speed Trains

4.2.1. Problem Description

The stability of hunting has always been a key problem in the study of vehicle lateral motion stability [33]. Small amplitude hunting is a sign of hunting instability. In China, hunting phenomena are considered to occur when the amplitude of lateral acceleration signals from the bogie frame reaches or exceeds $8\text{--}10 \text{ m/s}^2$ more than 5 times continuously after a 10 Hz low-pass filter [34]. In Figure 4, the lateral acceleration of the bogie frame signals will sometimes go through a normal operation/small amplitude hunting/normal operation periodic cycle, which is a gradually convergent

process. The hunting amplitude in this situation is small and convergent. Sometimes, the signals will go through a normal operation/small amplitude hunting/critical hunting process, which is a gradually divergent process. The hunting amplitude in this situation starts small and then diverges.

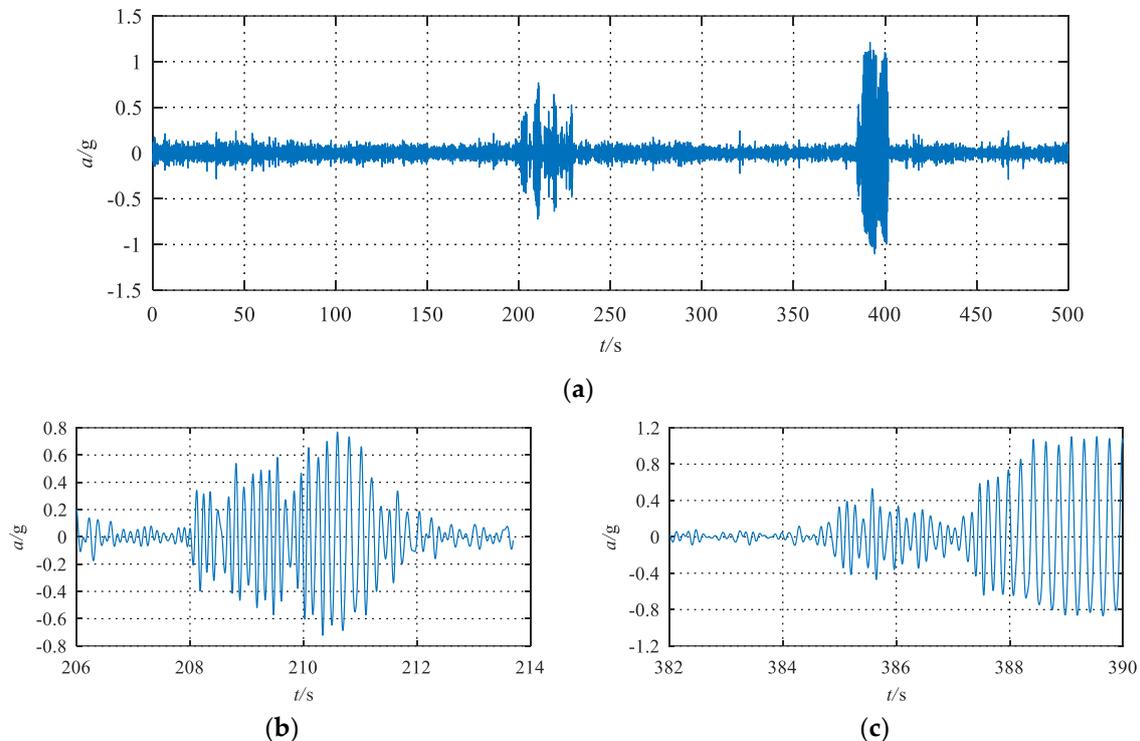


Figure 4. Lateral acceleration signal of the bogie frame when small-amplitude hunting motion occurs during an online test. (a) Bogie frame acceleration; (b) Small amplitude convergent hunting; (c) Small amplitude divergent hunting.

Therefore, it is necessary to extract the features of different states rapidly and accurately, especially the small amplitude divergent hunting states, to guarantee the system is alerted in time to ensure the safe operation of the train.

4.2.2. Data Acquisition

The data used in this paper were lateral acceleration signals of the bogie frame and axle box from an online tracking experiment. The CRTS II ballastless track and seamless rail were used on the whole line. The speed of the train was 320–350 km/h. The sampling frequency was 2500 Hz. All the data were acquired in accordance with China’s Railway Passenger Traffic Safety Monitoring Standard [35]. Although in China the amplitude of lateral acceleration signals from the bogie frame is used as the testing parameter to monitor hunting motion, research has proven that many other testing parameters are also important for hunting monitoring. As such, in this paper, acceleration signals of the bogie frame and the axle box were used.

The installation locations of the accelerometers are shown in Figure 5, where two accelerometers are installed in the diagonal direction on the H-shaped bogie frame. The lateral accelerometers from the bogie are respectively denoted as S1 and S2. Also, considering that the vibration of the axle box is important for hunting [36], a sensor located on the axle box was used. In Figure 6, the lateral accelerometer on the axle box is denoted as S3. Figure 5 shows a photograph of the installation site.

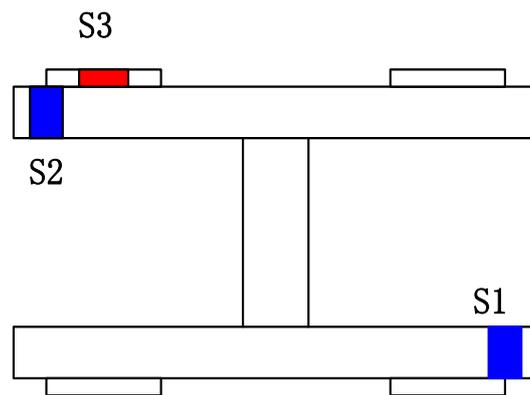


Figure 5. Installation locations of the accelerometers. S1, S2: Accelerometer on the bogie frame; S3: Accelerometer on the axle box.



Figure 6. Site installation drawing.

Considering the applications of other researchers, the sampling frequency of the original signal was set to 2500 Hz. However, for hunting, the characteristic frequency of the lateral acceleration of the frame is only 3–7 Hz. Therefore, in the preprocessing stage, a 250 Hz resampling method was applied [37], and a low-pass filter of 0–10 Hz was applied. Then, a zero-mean smoothing process was used for preprocessing to eliminate trend terms. In accordance with commonly followed practices, parameters such as the amplitude of the lateral acceleration signals from the bogie frame and axle box were used to obtain a synthetic assessment of the lateral stability of the high-speed train tested. In this paper, the filtered lateral acceleration signals were divided into four states: normal, small amplitude convergent hunting, small amplitude divergent hunting, and hunting. Ten groups of sample data were used for each of the four states and each sensor, yielding a total of 120 sample groups. The length of each sample was 500 points, corresponding to a sample time of 2 s.

Nonlinear factors [38,39] have been proven to affect the bifurcation evolution of small amplitude hunting and, according to [23], all the values of the Lyapunov exponent of the lateral acceleration are greater than 1, which means that the lateral acceleration signals from the bogie frame have non-stationary characteristics.

4.2.3. Feature Fusion

First, EEMD was applied on each signal. Then, 7 IMFs and 1 residue were obtained using the EEMD method. Figure 7 shows an EEMD decomposition view of the three signals at different points during the same time period.

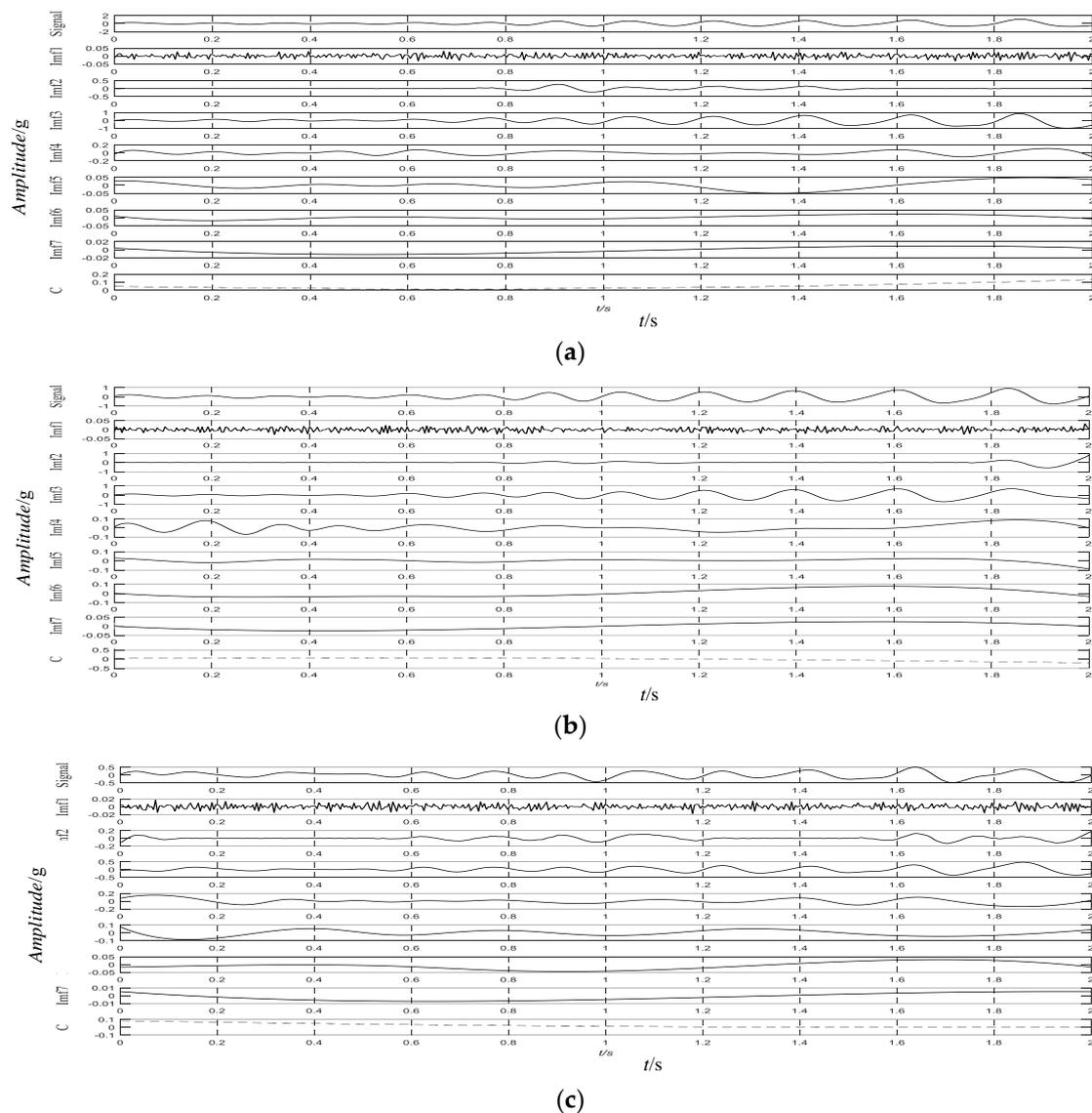


Figure 7. Ensemble Empirical Mode Decomposition (EEMD) diagram of signals at different positions. (a) S1 position (bogie frame); (b) S2 position (bogie frame); (c) S3 position (axle box).

The samples of the three sensors were processed, and the three corresponding energy matrices E_1 , E_2 , and E_3 , with a size of 40×8 were obtained. The NKJADE method was used to fuse the three high-dimensional energy matrices and the data dimensionality was reduced.

4.3. Result and Discussion

4.3.1. Single Sensor Classification Using NKJADE

A scatter plot of the features extracted using a single sensor and NKJADE is shown in Figure 8. The separability J and accuracy of the NKJADE method are shown in Table 4.

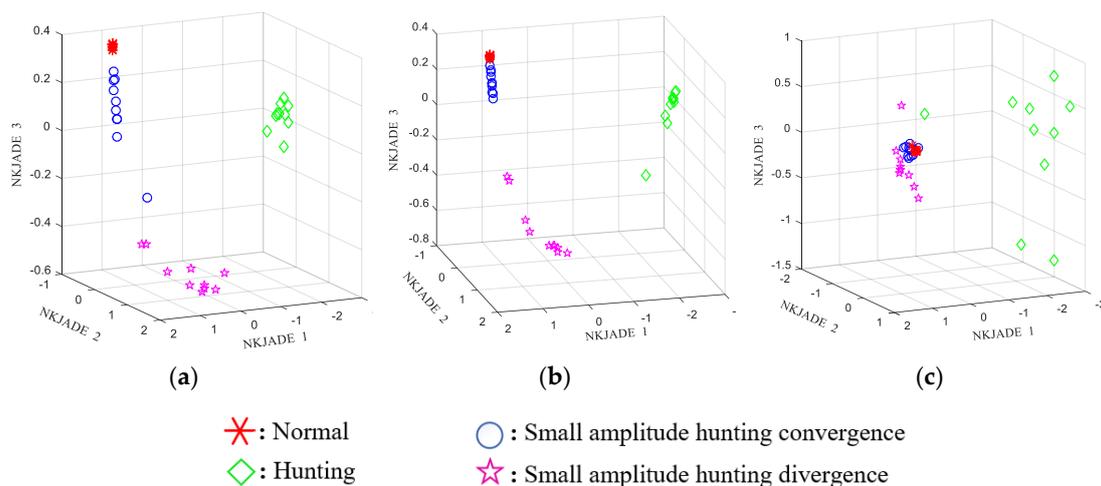


Figure 8. Scatter plot of feature extraction from single sensor using Non-stationary Kernel JADE (NKJADE). (a) S1 (bogie frame); (b) S2 (bogie frame); (c) S3 (axle box).

Table 4. Separability J running time and accuracy of NKJADE method.

Sensors	Only S1	Only S2	Only S3	S1, S2 and S3
J	65.6	61.1	21.1	155.7
Accuracy (%)	97.23	96.54	29.85	100
Run Time (s)	0.0392	0.0415	0.0387	0.1293

From Table 4, the accuracy achieved using S1 data only was 97.23%, which was greater than that attained at the other sensor locations. However, when the three sensors were used together, the accuracy rate became 100%, and J reached a much greater value.

4.3.2. Multi-Sensor Fusion Using NKJADE

The EEMD-SVD-LTSA and EEMD-JADE methods were used to compare the identification accuracy and calculation speed of the feature fusion method in multi-sensor conditions. The results of the methods are shown in Figure 9. The separability J and accuracy obtained using the different sensor fusion methods with data from all three sensors are shown in Table 5.

The JADE and SVD-LTSA methods were compared with the proposed method in the multi-sensor feature fusion conditions. From Table 5, the accuracy rate of the SVD-LTSA method was 93.75%, which used the non-stationary method LTSA. The accuracy rate of the JADE method was only 30.12%, which used the stationary method. The accuracy rate of NKJADE (Gaussian kernel, $\delta = 0.6$) was 100%, while the separability J of NKJADE was greater than those of the other methods. Considering that data from high-speed trains have more typical non-stationary characteristics than the bearing fault data from CWRU, this result shows that the proposed method may more be suitable for non-stationary data. The separability J of the JADE method was only 26.67, which is even lower than that achieved using single sensor S2 or S3 data. The reason for this is probably that in the JADE method, the non-stationary condition is not considered.

Besides, because real-time processing is a significant factor for the small amplitude hunting monitoring of high-speed trains, the calculation time is a very important factor. If the calculation time is too long, the diagnostic information cannot be fed back to the system in time. Compared with the SVD-LTSA method, fusion features can be extracted very quickly using the proposed NKJADE method. The run time of the NKJADE method was nearly the same as that of the JADE method, and the separability of J of NKJADE was greater than JADE. This shows that NKJADE is a rapid multi-sensor feature fusion method based on non-stationary condition, which outperforms the SVD-LTSA and

JADE methods. It can be applied to the small amplitude hunting bifurcation evolution monitoring in high-speed trains.

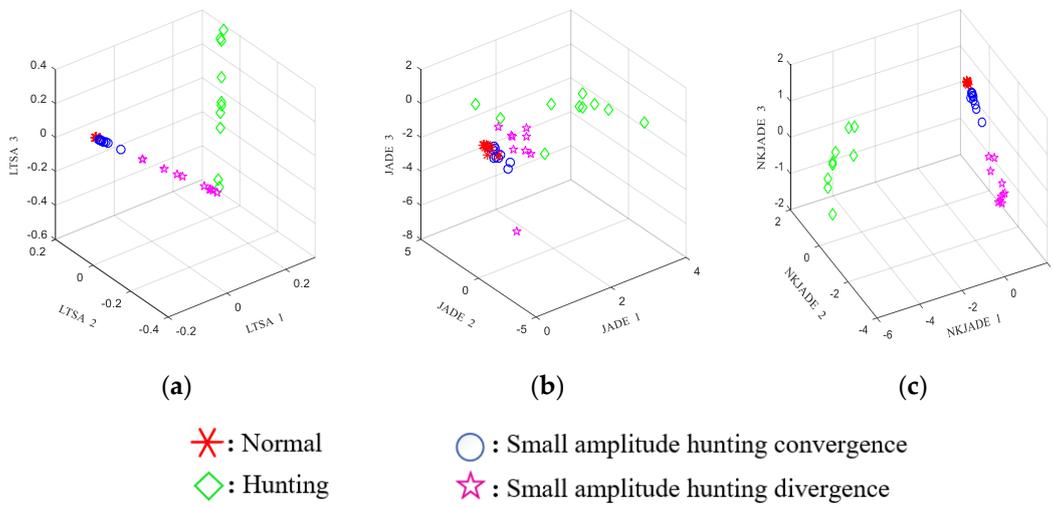


Figure 9. Scatter plot of feature extraction from multi-sensor data. (a) EEMD-SVD-LTSA; (b) EEMD-JADE; (c) EEMD-NKJADE.

Table 5. The separability J , running time, and accuracy using different sensor fusion methods.

Method	EEMD-SVD-LTSA	EEMD-JADE	EEMD-NKJADE
J	52	26.67	155.7
Accuracy (%)	93.75	30.12	100
Run time (s)	0.9972	0.1216	0.1298

As shown in Figure 10, the parameter is incremented step by step (parameter of step is 0.01, shown in X-axis), and the separability index J is calculated (shown in Y-axis). The larger J is, the better the classification result will be.

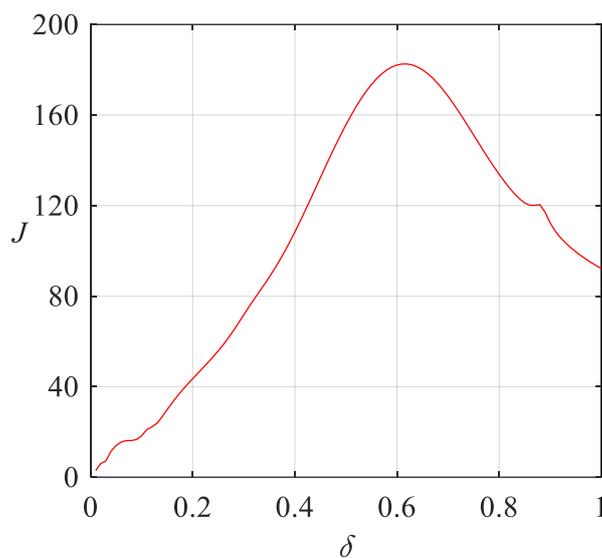


Figure 10. Parameter selection results of Gaussian kernel of EEMD-NKJADE method.

Because the raw data collected are mainly distributed in the range of $-1-1$ g (gravity). In Equation (8), σ is the width parameter of the kernel function, so we tried to set the range of σ to (0, 1). As shown in Figure 10, with the increase of δ , J increases at first and then decreases. It indicates that the range of parameter selection is reasonable. When $\delta = 0.6$, the classification result is the best.

5. Conclusions

In this paper, a rapid multi-sensor feature fusion method based on NKJADE is proposed, with which the features of multiple sensors can be extracted quickly and accurately. In order to use the algorithm in a non-stationary environment, the whole time series is divided into M time periods, and the kernel function is introduced. Then, Jacobian rotation is used to obtain the unitary matrix by diagonalizing multiple kernel matrices simultaneously to extract the non-stationary fusion features.

Our main findings are that:

1. The fusion features can be extracted quickly and efficiently using the proposed method NKJADE compared to the SVD-LTSA and the JADEs methods with bearing fault data from Case Western Reserve University.
2. The NKJADE method can extract the fusion features from non-stationary data effectively compared to the JADE method. In case I, the accuracy rate of the three methods (SVD-LTSA, JADE, and NKJADE) was nearly the same (100%), but in case II, the accuracy rate of the three methods was very different.
3. The NKJADE method can extract the multi-sensor fusion features effectively. The data from hunting monitoring of high-speed trains were used to verify the validity of the method in multi-sensor conditions.

Author Contributions: Data analysis and funding acquisition, J.N.; Writing—original draft preparation, M.F.; data collection: W.R.; experiment design, C.C.; detection equipment, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundations of China, grant number 51975486 and 51975487.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. De Pater, A.D. The approximate determination of the hunting movement of a railway vehicle by aid of the method of krylov and bogoljubow. *Appl. Sci. Res.* **1961**, *10*, 205–228. [[CrossRef](#)]
2. Yao, J.; Sun, L.; Hou, F. Study on evaluation methods for lateral stability of high-speed trains. *J. China Railway Sci.* **2012**, *33*, 132–139.
3. Chu, F.; Peng, Z.; Feng, Z. *Modern Signal Processing Methods in Machinery Fault Diagnosis*; Science Press: Beijing, China, 2009.
4. Guo, K.; Zhu, Y.; San, Y. Analog circuit fault diagnosis using LDA and OAOSVM approach. *Adv. Mater. Res.* **2012**, *490*, 1130–1134. [[CrossRef](#)]
5. Park, W.; Lee, S.; Joo, W.; Song, J. A mixed algorithm of PCA and LDA for fault diagnosis of induction motor. In *Advanced Intelligent Computing Theories and Applications*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 934–942.
6. Moura, E.; Souto, C.; Silva, A.; Irmao, M. Evaluation of principal component analysis and neural network performance for bearing fault diagnosis from vibration signal processed by RS and DF analyses. *Mech. Syst. Signal Process.* **2011**, *5*, 1765–1772. [[CrossRef](#)]
7. He, Q.; Ding, X.; Pan, Y. Machine fault classification based on local discriminant bases and locality preserving projections. *Math. Probl. Eng.* **2014**, *5*, 1–12. [[CrossRef](#)]

8. He, Q. Vibration signal classification by wavelet packet energy flow manifold learning. *J. Sound Vib.* **2013**, *7*, 1881–1894. [[CrossRef](#)]
9. Hettiarachchi, R.; Peters, J.F. Multi-manifold LLE learning in pattern recognition. *Pattern Recognit.* **2015**, *48*, 2947–2960. [[CrossRef](#)]
10. Bu, Y.; Chen, F.; Pan, J. Stellar spectral subclasses classification based on Isomap and SVM. *New Astron.* **2014**, *28*, 35–43. [[CrossRef](#)]
11. Sun, W.; Halevy, A.; Benedetto, J.J.; Czaja, W.; Li, W.; Liu, C.; Shi, B.; Wang, R. Nonlinear dimensionality reduction via the ENH-LTSA method for hyperspectral image classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 375–388. [[CrossRef](#)]
12. Cardoso, J. High-order contrasts for independent component analysis. *Neural Comput.* **1999**, *1*, 158–191. [[CrossRef](#)]
13. Cao, S.; Ouyang, H. Multi-damage identification based on joint approximate diagonalization and robust distance measure. *J. Phys. Conf. Ser.* **2017**, *842*, 012022. [[CrossRef](#)]
14. Liu, F.; Liu, Y.; Chen, F.; He, B. Residual life prediction for ball bearings based on joint approximate diagonalization of eigen-matrices and extreme learning machine. *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2017**, *9*, 1699–1711. [[CrossRef](#)]
15. Liu, Y.; He, B.; Liu, F. Feature fusion using kernel joint approximate diagonalization of eigen-matrices for rolling bearing fault identification. *J. Sound Vib.* **2016**, *385*, 389–401. [[CrossRef](#)]
16. Wu, T.; Liu, C.C.; He, C. “Fault Diagnosis of Bearings Based on KJADE and VNWOA-LSSVM Algorithm”. *Math. Probl. Eng.* **2019**, *14*, 1–19. [[CrossRef](#)]
17. Liu, F.; Li, L.; Liu, Y.; Cao, Z.; Yang, H.; Lu, S. HKF-SVR Optimized by Krill Herd Algorithm for Coaxial Bearings Performance Degradation Prediction. *Sensors* **2020**, *20*, 660. [[CrossRef](#)] [[PubMed](#)]
18. Ning, J.; Liu, Q.; Ouyang, H. A multi-sensor fusion framework for detecting small amplitude hunting of high-speed trains. *J. Vib. Control* **2018**, *17*, 3797–3808. [[CrossRef](#)]
19. European, British Standard Norme. *Testing and Approval of Railway Vehicles from the Point of View of Their Dynamic Behavior-Safety-Track Fatigue-Ride Quality*; UIC Code 518; International Union of Railways: Paris, France, 2005.
20. Czech Institute for Normalisation. *EN14363, B.S. Railway Applications-Testing for the Acceptance of Running Characteristics of Railway Vehicles-Testing of Running Behaviour and Stationary Tests*; Czech Institute for Normalisation: London, UK, 2016.
21. TSI; HSRST. Technical Specification for Interoperability Relating to the ‘Rolling Stock’ Sub-System of the Trans-European High-Speed Rail System. *Off. J. Eur. Union L* **2008**, *25*, 22–282.
22. Federal Railroad Administration. *Vehicle/Track Interaction Safety Standards, High-Speed and High Cant Deficiency Operations*; National Archives and Records Administration: College Park, MD, USA, 2013.
23. Ning, J. Feature recognition of small amplitude hunting signals based on the MPE-LTSA in high-speed trains. *Measurement* **2019**, *131*, 452–460. [[CrossRef](#)]
24. Miettinen, J.; Nordhausen, K.; Taskinen, S. Blind source separation based on joint diagonalization in R: The packages JADE and BSSasyp. *J. Stat. Softw.* **2017**, *76*, 1–31. [[CrossRef](#)]
25. Ding, X.; He, Q.; Luo, N. A fusion feature and its improvement based on locality preserving projections for rolling element bearing fault classification. *J. Sound Vib.* **2015**, *335*, 367–383. [[CrossRef](#)]
26. Ye, Y.; Ning, J. Small Amplitude Hunting Instability of High-speed Train Diagnosis Method Based on Modified Ensemble Empirical Mode Decomposition, Shannon Entropy and Least Square Support Vector Machine. In Proceedings of the 2016 4th International Conference on Mechanical Materials and Manufacturing Engineering, Wuhan, China, 15–16 October 2016; Atlantis Press: Paris, France, 2016.
27. Luo, B.; Wang, H.; Liu, H.; Li, B.; Peng, F. Early fault detection of machine tools based on deep learning and dynamic identification. *IEEE Trans. Ind. Electron.* **2018**, *66*, 509–518. [[CrossRef](#)]
28. Wang, J.; Gao, R. Integration of EEMD and ICA for wind turbine gearbox diagnosis. *Wind Energy* **2014**, *5*, 757–773. [[CrossRef](#)]
29. Cheng, C.; Peng, Z.; Dong, X.; Zhang, W.; Meng, G. A novel damage detection approach by using Volterra kernel functions based analysis. *J. Frankl. Inst.* **2015**, *8*, 3098–3112. [[CrossRef](#)]
30. Hoffmann, H. Kernel PCA for novelty detection. *Pattern Recognit.* **2007**, *3*, 863–874. [[CrossRef](#)]
31. He, Q. Time-frequency manifold for nonlinear feature extraction in machinery fault diagnosis. *Mech. Syst. Signal Process.* **2013**, *35*, 200–218. [[CrossRef](#)]

32. Zhang, S. *Research on Methods of Machinery Condition Recognition Based on Manifold Learning*; Southwest Jiaotong University: Chengdu, China, 2014.
33. Dong, H. *Study on Stability and Bifurcation Types of Railway Vehicles*; Southwest Jiaotong University: Chengdu, China, 2014.
34. TB/T3188-2007, *Technical Specification for Railway Car Safety Monitor and Diagnosis*; China Academy of Railway Sciences Locomotive and Car and Research Institute: Beijing, China, 2007.
35. TB 10761, *Ministry of Railways of the People's Republic of China, Technical Regulations for Dynamic Acceptance for High-Speed Railways Construction*; China Academy of Railway Sciences Locomotive and Car and Research Institute: Beijing, China, 2013.
36. Ning, J.; Lin, J.; Zhang, B. Time-frequency processing of track irregularities in high-speed train. *Mech. Syst. Signal Process.* **2016**, *66*, 339–348. [[CrossRef](#)]
37. Madiena, C. Color and vector flow imaging in Parallel Ultrasound with Sub-Nyquist sampling. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2018**, *5*, 795–802. [[CrossRef](#)]
38. Bruni, S.; Vinolas, J.; Berg, M. Modelling of suspension components in a railvehicle dynamics context. *Veh. Syst. Dyn.* **2011**, *49*, 1021–1072. [[CrossRef](#)]
39. Wen, Z.F.; Jin, X.S. Effects of lateral deformations of wheelset/track on creepforces of wheel/rail. *J. Mech. Strength* **2002**, *24*, 383–387.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).