

Article

Adaptive Multiclass Mahalanobis Taguchi System for Bearing Fault Diagnosis under Variable Conditions

Ning Wang^{1,2,3}, Zhipeng Wang^{1,2,3,*} , Limin Jia^{1,2,3,*}, Yong Qin^{1,2,3}, Xinan Chen^{1,2,3} and Yakun Zuo^{1,2,3}

¹ State Key Lab of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China; 17114235@bjtu.edu.cn (N.W.); yqin@bjtu.edu.cn (Y.Q.); 15114217@bjtu.edu.cn (X.C.); 18120772@bjtu.edu.cn (Y.Z.)

² National Engineering Laboratory for System Safety and Operation Assurance of Urban Rail Transit, Guangzhou 510000, China

³ Beijing Research Center of Urban Traffic Information Sensing and Service Technologies, Beijing Jiaotong University, Beijing 100044, China

* Correspondence: zpwang@bjtu.edu.cn (Z.W.); lmjia@bjtu.edu.cn (L.J.); Tel.: +86-010-5168-4281(Z.W.); +86-010-5168-4639 (L.J.)

Received: 16 November 2018; Accepted: 19 December 2018; Published: 21 December 2018



Abstract: Bearings are vital components in industrial machines. Diagnosing the fault of rolling element bearings and ensuring normal operation is essential. However, the faults of rolling element bearings under variable conditions and the adaptive feature selection has rarely been discussed until now. Thus, it is essential to develop a practicable method to put forward the disposal of the fault under variable conditions. Considering these issues, this paper uses the method based on the Mahalanobis Taguchi System (MTS), and overcomes two shortcomings of MTS: (1) MTS is an effective tool to classify faults and has strong robustness to operating conditions, but it can only handle binary classification problems, and this paper constructs the multiclass measurement scale to deal with multi-classification problems. (2) MTS can determine important features, but uses the hard threshold to select the features, and this paper selects the proper feature sequence instead of the threshold to overcome the lesser adaptivity of the threshold configuration for signal-to-noise gain. Hence, this method proposes a novel method named adaptive Multiclass Mahalanobis Taguchi system (aMMTS), in conjunction with variational mode decomposition (VMD) and singular value decomposition (SVD), and is employed to diagnose the faults under the variable conditions. Finally, this method is verified by using the signal data collected from Case Western Reserve University Bearing Data Center. The result shows that it is accurate for bearings fault diagnosis under variable conditions.

Keywords: fault diagnosis; bearing; SVD; VMD; adaptive Multiclass Mahalanobis Taguchi System

1. Introduction

Rolling element bearings have wide applications in industrial machines and are one of the most critical components. If faults occur in bearings, equipment could be damaged and disasters might happen consequently. Therefore, it is essential to monitor the health conditions of bearings. The analysis of vibration signals has been a hot research pot and used to detect faults of bearings. It is crucial to recognize faults occurring in bearings and avoid fatal breakdowns as early as possible. For decades, many researchers have conducted extensive research on fault diagnosis. At present, the fault diagnosis methods are divided into model-based methods and data-driven methods. The model-based methods generally build on the physics of the process, generating the residuals between the measure process variables and estimates [1], such as Hidden Markov Modeling (HMM) which

is successfully applied to bearing fault detection and diagnosis [2]. Autoregressive modelling [3] also has had excellent performance in bearing fault diagnosis. The accelerated degradation testing (ADT) [4] method is useful in fault diagnosis and lifespan prediction, and reference [5] presents a new approach using observer-based residual generation with no complicated design constraints to establish the relationship between the state estimation error and the fault signal. For the data-driven method, which is based on the historical data and does not need accurate mathematic and priori knowledge, it has wide applications in fault diagnosis. For example, Bayesian network is an excellent data-driven diagnosis method [6–8]. Machine learning algorithms, such as K-nearest Neighbor (KNN), Deep Convolutional Neural Networks (DCNN), and auto-encoders are effective in the fault diagnosis of various bearings [9–12]. There is also decision tree [13], Support Vector Machine (SVM) [14], and wavelet transform [15]. Artificial neural network (ANN) and Trace Ratio Linear Discriminant Analysis are other methods of diagnosing the bearings fault [16,17]. In addition, the vibration signal under the various operating conditions (especially in low rotational speed) is non-stationary and non-linear, the characteristic defect frequencies move continuously with the change of rotating speed, and is the same as the bearing that is going to break down. If the bearing progresses toward failure, the nonlinear features also start to be dominated by stochastic signal. With the vibration signal measured for diagnosing the bearing fault under variable condition, it will be difficult to diagnose the fault of bearing by using the traditional methods. Reference [18] proposed a signal selection scheme based upon two order tracking techniques from complicated non-stationary operational measured vibrations. Reference [19] reviewed features extraction methods and its application on bearing vibration signal and presents an empirical study of feature extraction methods in low rotational speed. Reference [20] proposes the estimation of instantaneous speed relative fluctuation (ISRF) in a vibration speed. Reference [21] proposes Stacked Convolutional Autoencoders (SCAE) together with DCNN in stationary and non-stationary speed operation. Graph-based rebalance semi-supervised learning (GRSSL) [22], weighted self-adaptive evolutionary extreme learning machine (WSaE-ELM) [23] and Singular Spectrum Analysis [24] are effective in diagnosing the fault under variable conditions.

However, although the aforementioned methods are effective for bearing faults diagnosis, the feature selection part is often non-adaptive or unexplainable. The conventional methods mainly extract and select features manually, which relies heavily on the experts' knowledge and experience. Since the signals acquired in the real world might be various in many different aspects, the features selected manually might be sensitive in the variations of operation conditions and import inevitable errors for fault diagnosis. The deep learning algorithms can extract features automatically and overcome this drawback. However, deep learnings are data-hungry and require plenty of training data which are hardly acquired in practice, especially the faulty data under different conditions. Besides, the features acquired by deep learnings are unexplainable. Therefore, the aforementioned methods can hardly diagnose the faults under various operation conditions in practice. In contrast, the Mahalanobis Taguchi System (MTS) is more robust than the other methods in various operating conditions [25].

MTS offers a tool to determine important features and optimize the system. It is a different form compared with the other classification methods, because this classification model of measurement scale is constructed by using the class samples. It is useful to diagnose bearing faults under various conditions because the different pattern could be identified by using the Mahalanobis distance (MD) and Taguchi method. In this paper, MD is used to calculate the distance of the correlations between the benchmark and others, and the distance could be measured without the volatility of data. The advantage of MD is that it takes into consideration the correlations between the features and this consideration is very important in pattern analysis, which is why MTS is suitable for bearing fault diagnosis under various conditions [26]. On the other hand, the Taguchi method is used to select features without manual intervention, which could improve the robustness of the algorithm. MTS also offers an effective tool for multivariate analysis [27], considering that the bearing faults can be classified according to locations, such as inner race, outer race and rolling element [28]. However, when the conventional MTS is used for bearing fault diagnosis, misclassifications might occur due to

less adaptivity of the threshold configuration for signal-to-noise gain. During the feature selection, the threshold is normally set as a constant, which might result in the overfitting problem. If the threshold value is too large, some critical features might be eliminated. On the contrary, if the threshold value is too low, some useless or harmful feature might be selected. Therefore, if the threshold value does not march the training data sufficiently, misclassifications emerge.

To overcome this drawback, this paper presents a novel method named adaptive Multiclass Mahalanobis Taguchi system (aMMTS) for bearing fault diagnosis. This method employs the MTS for multi-classification by considering different conditions as different benchmarks respectively. The results are based on the minimum MDs between the data and each benchmark data, and the label of the data is determined to be consistent with the label of the benchmark data whose MD is minimum. The method can be described briefly as follows: Firstly, after the Mahalanobis space (MS) is constructed by the two-level orthogonal array of the Taguchi method, aMMTS calculates the MDs from the data to the benchmark data and obtains features' signal-to-noise ratios (SNRs) and gain values by using two-level orthogonal array. Secondly, the features are selected adaptively by recalculating the MDs via rearranging the order of features' gain values by ascending and descending. Therefore, the proposed method is able to select the best classification result according to the adaptive chosen sequence of features' SNRs instead of a hard threshold. Here, the sequence of SNR is determined by a function, which selects several maximum or minimum features to calculated MDs. Finally, a set of features with the best results is selected as the final feature vector. In this method, two different sets of training samples are employed to calculate the SNRs respectively and obtain the final feature vectors respectively. By the aforementioned improvement, the proposed aMMTS is capable to overcome the drawback of the conventional MTS and prevent the over-fitting problem. Therefore, the aMMTS is insensitive to the operation conditions and can be employed for bearing fault diagnosis.

Moreover, this method is combined with variational mode decomposition (VMD) [29] and singular value decomposition (SVD) to diagnose the faults. VMD is an entirely non-recursive algorithm, and is used to decompose the signal. It has been proven that due to the characteristics of nonlinear vibration in the bearings, VMD is more efficient than empirical mode decomposition (EMD) and Fourier transform (FT) under variable condition. SVD is used to extract the features. Therefore, VMD and SVD are employed in this paper.

The rest of this article is organized as follows: Section 2 introduces the algorithms involved in this paper. Section 3 illustrates the experiments to validate the proposed method. Section 4 is the conclusions.

2. Methodology

In this paper, the main steps of fault diagnosis are signal decomposition, feature extraction and fault detection. The detailed process and method are as follows:

Step1: The VMD in conjunction with wavelet denoising is employed to eliminate the noises and decompose the raw signals;

Step2: Extracting features from the decomposed signals by SVD;

Step3: The proposed aMMTS is employed for the fault diagnosis.

The steps of this method are shown in Figure 1.

2.1. VMD

After the wavelet denoising is used to estimate the noise of the raw signal, this paper employs VMD to decompose non-stationary signals. VMD can decompose a signal into different simple intrinsic mode functions, whose frequency center and bandwidth are band-limited and determined by iterative searching for the optimal solution of the variational model. The constrained formula is given as:

$$\min_{\mu_k, \omega_k} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * \mu_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \quad s.t. \sum \mu_k = f \quad (1)$$

where μ_k is the sub-signals, ω_k represents the center frequency of sub-modes. The optimal solution can be solved as the minimization problem, which could be addressed by introducing a quadratic penalty and Lagrangian multipliers [30]:

$$L(\{\mu_k\}\{\omega_k\}, \lambda) = \alpha \sum_k \|\partial_t[(\delta(t) + \frac{j}{\pi t}) * \mu_k(t)]|e^{-j\omega_k t}\|_2^2 + \|f(t) - \sum_k \mu_k(t)\|_2^2 + \langle \lambda(t), f(t) - \sum_k \mu_k(t) \rangle \quad (2)$$

α denotes the balancing parameter of the constraint.

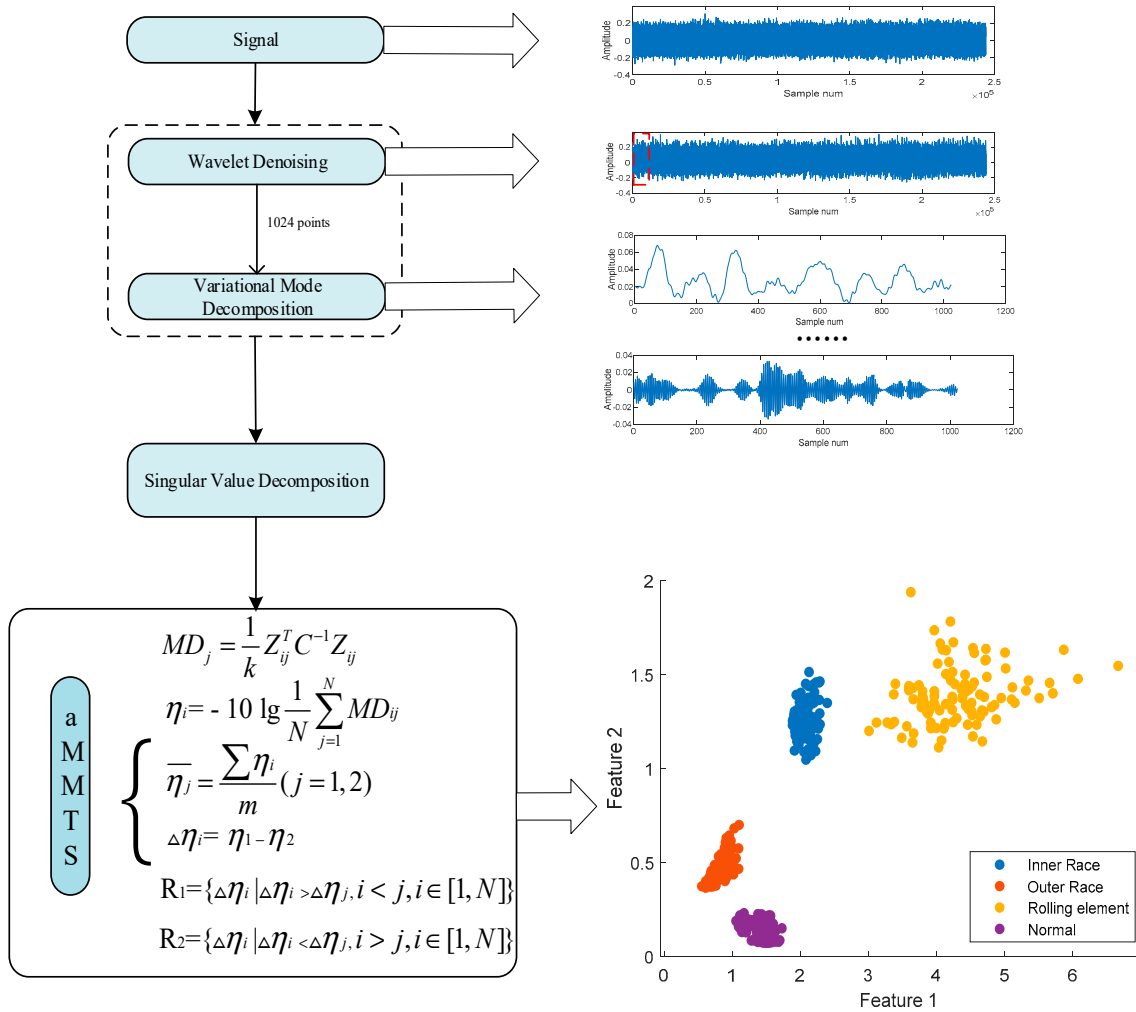


Figure 1. The scheme of the proposed fault diagnosis method.

All sub-signals μ_k are updated for all $\omega \geq 0$ as follow:

$$\hat{\mu}_k^{n+1} \leftarrow \frac{\hat{f} - \sum_{i < k} \hat{\mu}_i^{n+1} - \sum_{i > k} \hat{\mu}_i^n + \frac{\hat{\lambda}^n}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \quad (3)$$

$\hat{\mu}_k^1, \hat{\omega}_k^1$ and $\hat{\lambda}^1$ are initialized to all zeroes.

All center frequency of sub-modes ω_k are updated as follow:

$$\omega_k \leftarrow \frac{\int_0^\infty \omega \left| \hat{\mu}_k^{n+1}(\omega) \right|^2 d\omega}{\int_0^\infty \left| \hat{\mu}_k^{n+1}(\omega) \right|^2 d\omega} \quad (4)$$

End for

$$\hat{\lambda}^{n+1} \leftarrow \hat{\lambda}^n + \tau(\hat{f} - \sum_k \hat{\mu}_k^{n+1}) \quad (5)$$

Until:

$$\frac{\sum_k \|\hat{\mu}_k^{n+1} - \hat{\mu}_k^n\|_2^2}{\|\hat{\mu}_k^n\|_2^2} < \varepsilon \quad (6)$$

2.2. SVD

After the signals are decomposed into several modes by VMD, the features are extracted from the modes by SVD, and can be constructed as a feature matrix. SVD is a powerful tool for feature extraction in linear algebra. According to SVD, the matrix could be decomposed as follow:

$$X = U\omega V^T \quad (7)$$

where X represents a $m \times n$ matrix. There are two orthogonal matrices: matrix $U(m \times m)$ and $V(n \times n)$, and a singular diagonal matrix $\omega(\omega_{ij} \neq 0, i = j \text{ and } \omega_{11} \geq \omega_{22} \geq \dots \geq 0)$, the diagonal element $\omega_{11}, \omega_{22}, \dots, \omega_{mm}$ is the singular value of X . U is called left singular vector, and the columns of U . V is called right singular vector. The columns of U and V especially are orthogonal to each other, and are base vector. To obtain more intrinsic information in the matrix, the singular vectors are selected. As a consequence, SVD is employed to decompose the eigenmatrix, and obtained the singular value vectors $(\omega_{11}, \omega_{22}, \dots, \omega_{mm})$.

2.3. Mahalanobis–Taguchi System

Mahalanobis–Taguchi System is a pattern recognition method integrated by the MD, orthogonal table and other tools such as SNR that are proposed by Taguchi [31], who introduces the experimental design of the field SNR to pattern recognition, which can reduce the dimensions of data, and use the orthogonal table to construct the MS. The MSs are used to calculate the MDs of the experimental data, and the valid features are distinguished by the SNR. Then, the MSs are recalculated by using the valid features. Finally, the results are obtained. The calculation of the MD is described as follows:

2.3.1. Mahalanobis Distance

The MD is a method of using normal data to normalize the fault data to compute the average distance between points and groups using normal data, the calculation formula of MD is as follows:

$$MD_j = \frac{1}{k} Z_{ij}^T C^{-1} Z_{ij} \quad (8)$$

$$Z_{ij} = \left\{ z_{1j}, z_{2j}, z_{3j}, \dots, z_{kj} \right\}, z_{ij} = \frac{x_{ij} - \bar{x}_i}{s_i} \quad (9)$$

MD_j represents the MD of the j th sample, k represents the number of the feature, x_{ij} represents the i th feature's value of the of the j th sample \bar{x}_i represents the mean of i th feature, s_i represents the Standard deviation of i th feature, C^{-1} represents the inverse matrix of the correlation coefficient matrix.

2.3.2. Taguchi Method

In the Mahalanobis–Taguchi System, the MD measures the deviation of the test value from the normal value. The Taguchi method is able to select the features which have a larger contribution to identifying bearing faults, and then use the selected valid features to calculate the MD; the method is as below:

Orthogonal Array

Selecting the appropriate two-level orthogonal array, and then the k -original features obtained by VMD and SVD, are arranged into each column of the orthogonal array. In the orthogonal array, “1” indicates that the feature is selected, “2” indicates that the feature is not selected, and a MS is generated according to each row of the orthogonal array.

SNR and Its Gain

The main function of signal noise ratio is to select a valid feature, the calculation formula of generating SNR η_i in the i line based on the orthogonal array.

$$\eta_i = -10 \lg \frac{1}{N} \sum_{j=1}^N MD_{ij} \quad (10)$$

$j \in [1, N]$ represents the number of training samples.

η_i represents the recognition effects of the characteristic feature, the valid feature is selected by comparing the mean of SNR of each characteristic feature at two levels. The formula is as follow:

$$\bar{\eta}_j = \frac{\sum \eta_i}{m} \quad j = 1, 2 \quad (11)$$

$$\overline{\Delta\eta}_j = \sum \left(\frac{\bar{\eta}_j}{m} \right) \quad (12)$$

$j = 1, 2$ represents two levels, i represents the number of rows in the MS, $\bar{\eta}_1$ represents that in the level of ‘1’, the recognition effects to identify abnormal conditions of using this feature. $\bar{\eta}_2$ represents that in the level of ‘2’, the recognition effects to identify abnormal conditions of not using this feature. If $\overline{\Delta\eta}_j$ represents the SNR gain, $R = \{ \overline{\Delta\eta}_j | \overline{\Delta\eta}_j > 0 \}$ that indicates that this feature is a valid feature, if $\overline{\Delta\eta}_j < 0$ that indicates that this feature is not a valid feature.

2.3.3. Adaptive Multiclass MTS

After SNR gain is calculated, to overcome the drawback of the conventional MTS during the threshold selection, this paper presents the adaptive multiclass MTS. There are several following improvements:

(1) Solving the multiple-classification problem. Selecting samples from each kind as the benchmark data. Then, the distances between other data and each benchmark data are calculated by MTS. Therefore, the label of the benchmark data with the minimum distance is selected as the label of the training data.

(2) Selecting the features adaptively. The feature sequence is selected several times, and the best fault recognition is the one which is minus MDs between it and benchmark. It solves the error problem caused by hard threshold selection.

(3) Avoiding the overfitting problem. Since the SNR gains are calculated by training samples, different training samples are employed to calculate the MDs and identify bearing faults, and the difference validation samples are set to validate the identified result and recalculate MDs.

This method is shown in Figure 2.

This adaptive multiclass MTS can be described as follow:

First, the data are labeled, the multi-class MSs are constructed, MDs between MSs are calculated and SNR gains are obtained. This step is shown as Figure 3; m is the number of samples, n is the number of features, and j is the number of the label. The training data are divided into three parts, and one of them is named as Benchmark. The data A_j and C_j represents one kind of data respectively (such as normal, fault of inner race, outer race, rolling element), and data B includes all kinds of data.

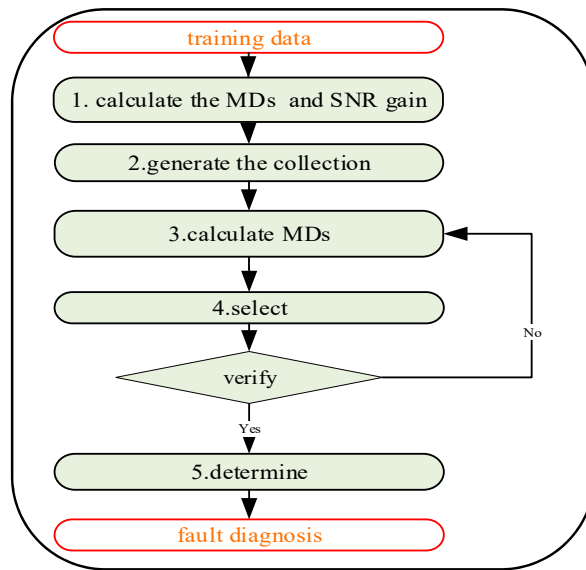


Figure 2. The step of Multiclass–Mahalanobis–Taguchi system (aMMTS).

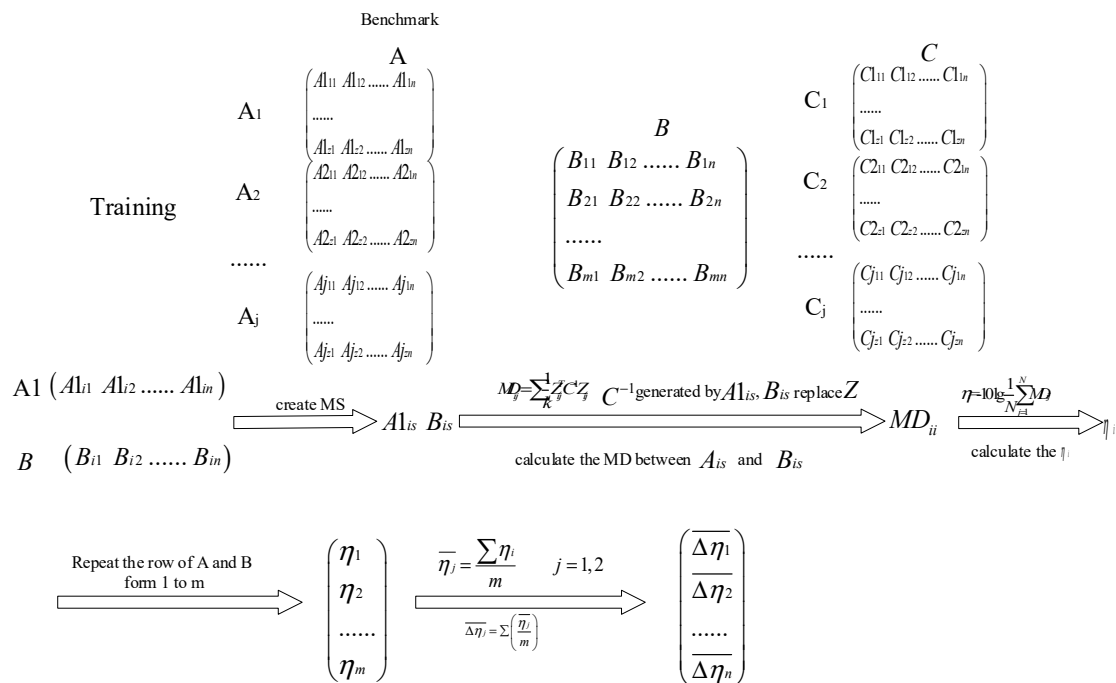


Figure 3. The first step of aMMTS.

Second, new sequences of feature parameters are generated by the sequence of features' SNR gains in ascending and descending order, then two collections are obtained from the above sequences based on the ascending or descending order, with the ascending and descending collection as follows:

$$r_k = (\overline{\Delta \eta_{ki}}, \dots, \overline{\Delta \eta_{kj}} \mid \overline{\Delta \eta_{ki}} > \overline{\Delta \eta_{kj}}, i < j, i \in [1, N]) \tag{13}$$

$$q_k = (\overline{\Delta \eta_{ki}}, \dots, \overline{\Delta \eta_{kj}} \mid \overline{\Delta \eta_{ki}} < \overline{\Delta \eta_{kj}}, i > j, i \in [1, N]) \tag{14}$$

$$R = \{r_k \mid k \in [1, N]\} \tag{15}$$

$$Q = \{q_k \mid k \in [1, N]\} \tag{16}$$

This step is shown in Figure 4:

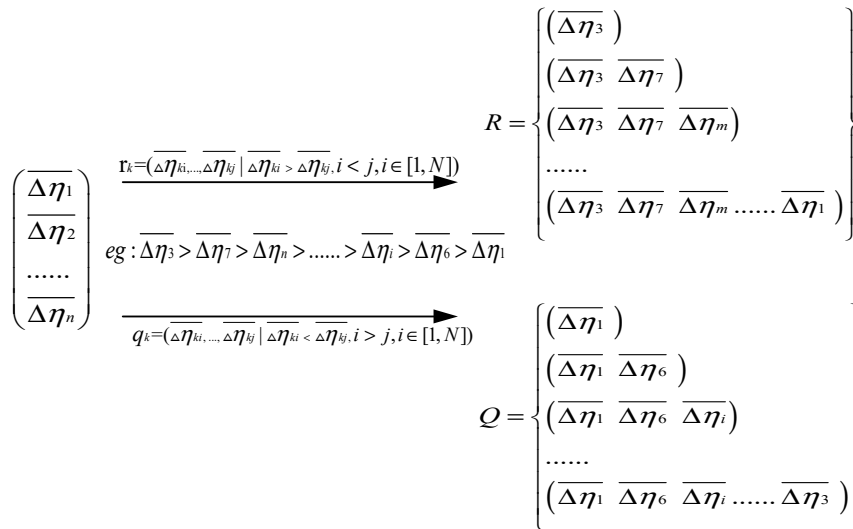


Figure 4. The second step of aMMTS.

Third, and the positions of feature’s SNR gain are the same as the positions of corresponding features in the sequences, the features that are used to recalculate the MDs are selected by the corresponding feature’s SNR gain. This step is shown in Figure 5. The MDs that are between each kind of A and C_i are calculated.

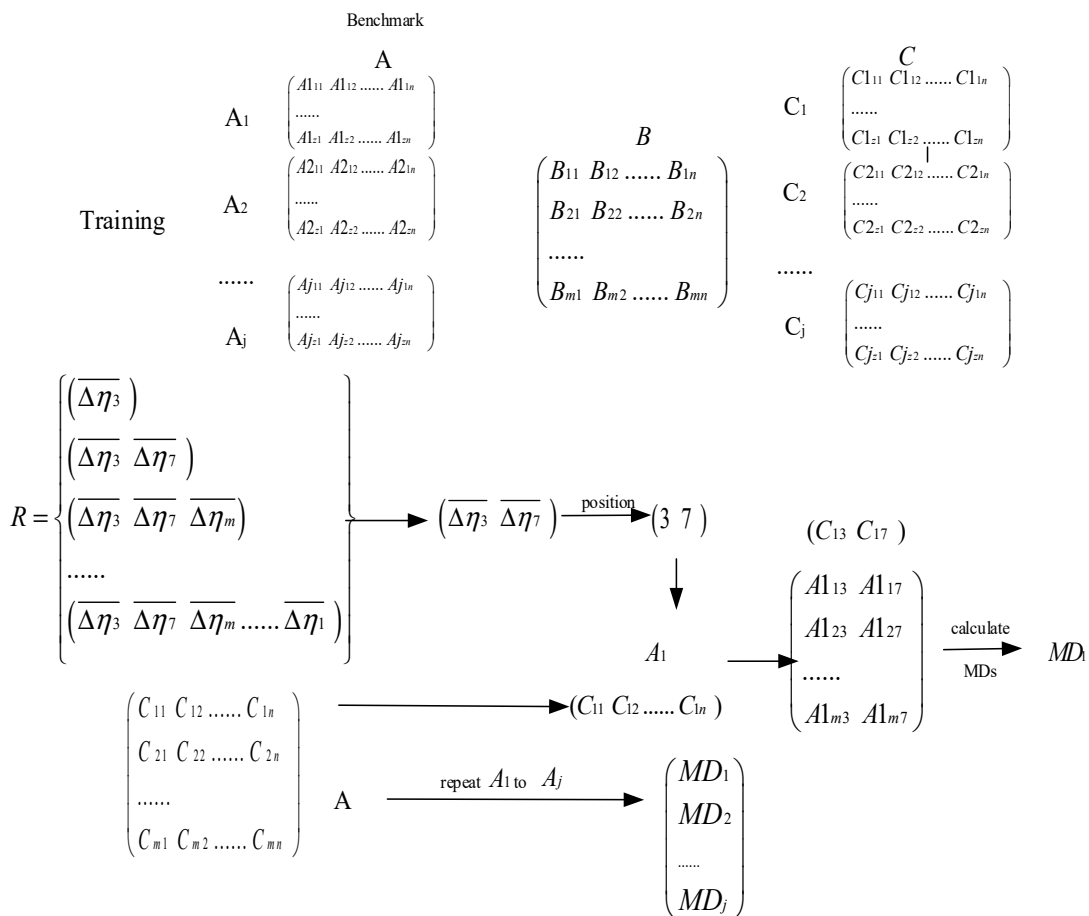


Figure 5. The third step of aMMTS.

Forth, the proper sequence of SNR gain is chosen by a function, which is according to the minimum MD. If two labels of data corresponding to minimum MD are the same, the recognition result is right, and accumulate the number of right recognition results. S is the recognition result. This step is shown in Figure 6.

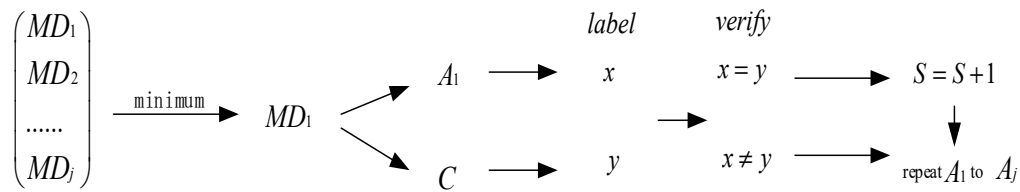


Figure 6. The forth step of aMMTS.

Finally, the recognition result is verified if there is a unique optimal recognition result, and the SNR gain's position in the sequence is the feature's order. If there are several optimal recognition results, repeat step 3 to recalculate the result. Afterwards, a set of sequences with the best recognition effect is determined as the feature sequence, which solves the self-adaptation problem of thresholds.

3. Results

In this paper, the experimental data are from Case Western Reserve University Bearing Data Center. This experiment involved three different faults that occurred on three components: inner race, outer race and rolling element. The vibration signals were acquired under four different speeds: 1797 r/min, 1772 r/min, 1750 r/min, and 1730 r/min, and the sampling frequency was set to 12 kHz. To demonstrate the aMMTS, this study randomly selected the data in the dataset under the defect of 0.07 inches. The number of samples are shown in Table 1.

Table 1. The number of samples.

	Label	Motor Load (hp)	Speed (r/min)	Training Samples			Validation Samples	Test Samples
				BenchMark	Group A	Group B		
Inner Race	1	0 (0W)	1797	27	27	29	27	30
	1	1 (735W)	1772	27	27	29	27	30
	1	2 (1470W)	1750	27	27	29	27	30
	1	3 (2205W)	1730	27	27	29	27	30
Outer Race	2	0 (0W)	1797	27	27	29	27	30
	2	1 (735W)	1772	27	27	29	27	30
	2	2 (1470W)	1750	27	27	29	27	30
	2	3 (2205W)	1730	27	27	29	27	30
Rolling Element	3	0 (0W)	1797	27	27	29	27	30
	3	1 (735W)	1772	27	27	29	27	30
	3	2 (1470W)	1750	27	27	29	27	30
	3	3 (2205W)	1730	27	27	29	27	30
Normal	0	0 (0W)	1797	27	27	29	27	30
	0	1 (735W)	1772	27	27	29	27	30
	0	2 (1470W)	1750	27	27	29	27	30
	0	3 (2205W)	1730	27	27	29	27	30

There were 2192 samples; 548 for inner race, 548 for outer race, 548 for rolling element and 548 for normal. The data are divided into three parts: training data, validation data and test data. In order to avoid the overfitting caused by the training data, training samples were used to construct MS, generate the SNR gain and calculate MDs by using the sequences of SNR gains, and were divided into three parts, with one of the parts set as benchmark group. To avoid the over-fitting problem, group A was used to construct MS and generate the SNR gain, and group B were used to calculate the MDs with the sequences of SNR gain and identify faults.

Validation samples were used to verify the recognition result if there exists the same minimum MDs, and the sequence was selected according to the best result.

Test samples were used to validate the proposed method.

3.1. Signal Decomposition by Using VMD and Wavelet Denoising

Above all, this study employed wavelet denoising to remove the noise from the raw signals. First, the Daubechies 5 (db5) was used to decompose the signal, and obtained the wavelet decomposition vector and the bookkeeping vector. Second, thresholds wavelet coefficient was calculated by setting the detail vector which would be compressed as [1–3] and the vector which is the corresponding percentages of lower coefficients as [100,90,80], and using the wavelet decomposition vector and the bookkeeping vector. Lastly, the thresholds, Daubechies 5 (db5) and decomposed signals were used to reconstruct the denoising signal. Then the VMD was used to decompose the signal, and was needed to give the preset IMF component number K and penalty parameter α which constrained the moderate bandwidth. The value of α took the default value 1024, the value of K was 8. An example is shown in Figure 7.

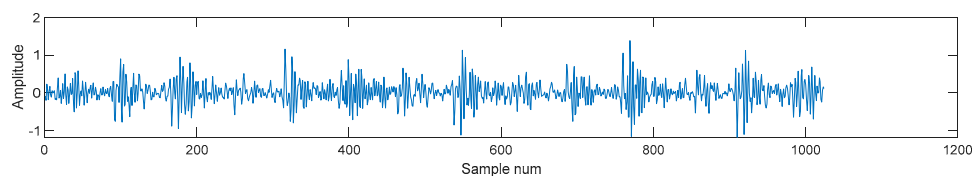


Figure 7. The intercepted signal.

IMFs are as shown in Figure 8.

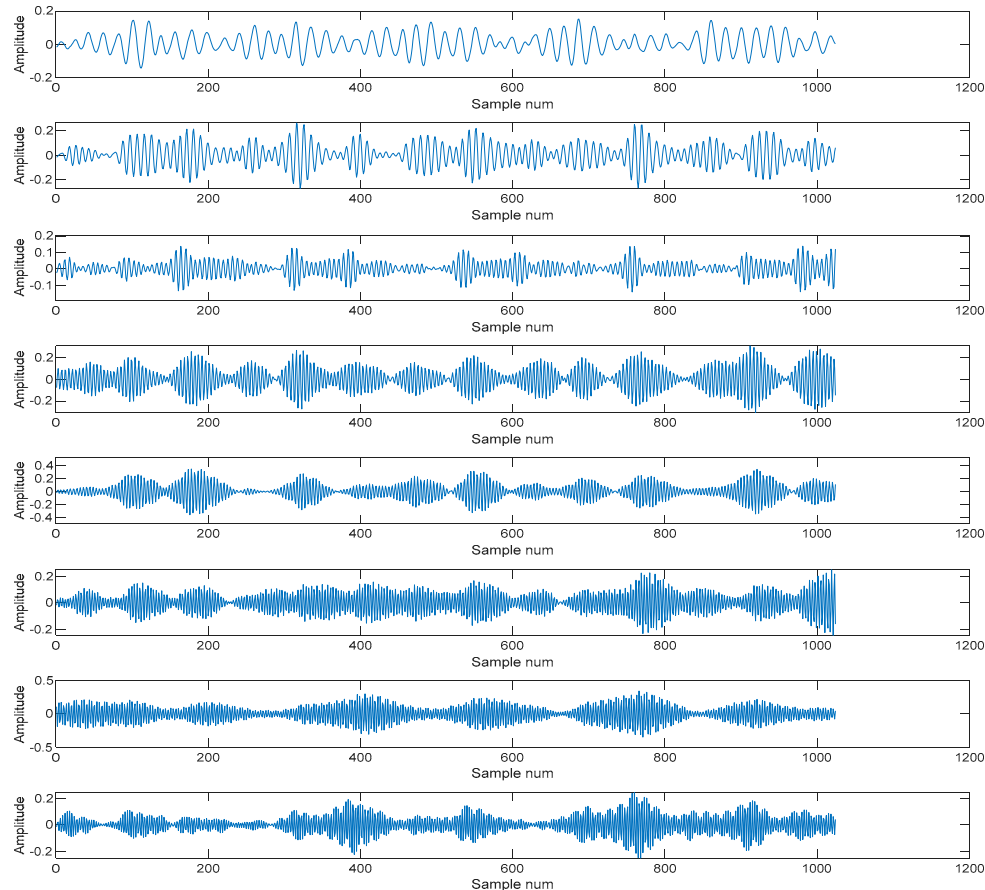


Figure 8. The IMFs.

3.2. Feature Extraction by Using SVD

SVD was used to analyze the IMFs. After the signal decomposition, the IMF matrix was decomposed by SVD, and obtained singular value vectors. The singular value vectors were considered as features and formed the feature matrix. Then, the feature matrix was used to diagnose the fault by aMMS. To avoid the over-fitting problem, the features were divided into training samples, validation samples and test samples. The features of the above IMFs of those were shown in Table 2.

Table 2. The feature of IMFs.

	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8
Features	2.019	1.968	1.836	1.582	1.491	1.459	1.264	1.168

The features obtained by SVD are shown in Table 3.

Table 3. The features of decomposed signals.

	Features							
	1	2	3	4	5	6	7	8
Normal	1.554	1.116	0.848	0.479	0.412	0.324	0.214	0.095
	1.472	1.230	0.473	0.425	0.292	0.258	0.197	0.083
	1.094	1.034	0.897	0.428	0.321	0.270	0.185	0.082
	1.173	0.939	0.931	0.683	0.387	0.287	0.229	0.080
Inner Race	2.020	1.968	1.836	1.582	1.491	1.459	1.264	1.168
	2.034	1.993	1.916	1.603	1.583	1.306	1.195	0.913
	2.115	2.023	1.918	1.763	1.570	1.410	1.376	1.143
	1.941	1.841	1.798	1.592	1.482	1.430	1.393	1.151
Rolling Element	0.811	0.793	0.667	0.577	0.562	0.542	0.451	0.223
	0.932	0.702	0.585	0.583	0.520	0.505	0.464	0.408
	0.740	0.657	0.556	0.515	0.501	0.487	0.435	0.356
	0.968	0.773	0.686	0.623	0.591	0.528	0.476	0.467
Outer Race	5.349	4.209	3.833	3.772	3.146	1.943	1.469	1.163
	5.312	4.468	4.003	3.505	2.988	2.751	1.416	1.145
	4.141	3.274	3.024	2.945	2.440	1.743	1.631	1.019
	3.334	2.901	2.560	2.382	2.036	1.697	1.236	0.941

3.3. Fault Diagnosis Using aMMS

After the feature extraction, the aMMS was used to identify and diagnose fault modes. The steps of aMMS are as follow:

Firstly, the MS of training and benchmark were constructed, the eight-factor and two-level orthogonal array is shown in Table 4, and the MS based on Table 2 is shown in Table 5;

Table 4. The eight-factor and two-level orthogonal array.

	A	B	C	D	E	F	G	H
1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	2	2	2
3	1	1	2	2	2	1	1	1
4	1	2	1	2	2	1	2	2
5	1	2	2	1	2	2	1	2
6	1	2	2	2	1	2	2	1
7	2	1	2	2	1	1	2	2
8	2	1	2	1	2	2	2	1
9	2	1	1	2	2	2	1	2

Table 5. The Mahalanobis space (MS) based on the inner race.

	A	B	C	D	E	F	G	H
1	2.020	1.968	1.836	1.582	1.491	1.459	1.264	1.168
2	2.020	1.968	1.836	1.582	1.491			
3	2.020	1.968				1.459	1.264	1.168
4	2.020		1.836			1.459		
5	2.020			1.582			1.264	
6	2.020				1.491			1.168
7		1.968			1.491	1.459		
8		1.968		1.582				1.168
9		1.968	1.836				1.264	

Secondly, the MD was calculated, and SNR gain was also obtained by benchmark samples and training samples. The SNR gain is shown in Table 6;

Table 6. The signal-to-noise ratio (SNR) gain of features.

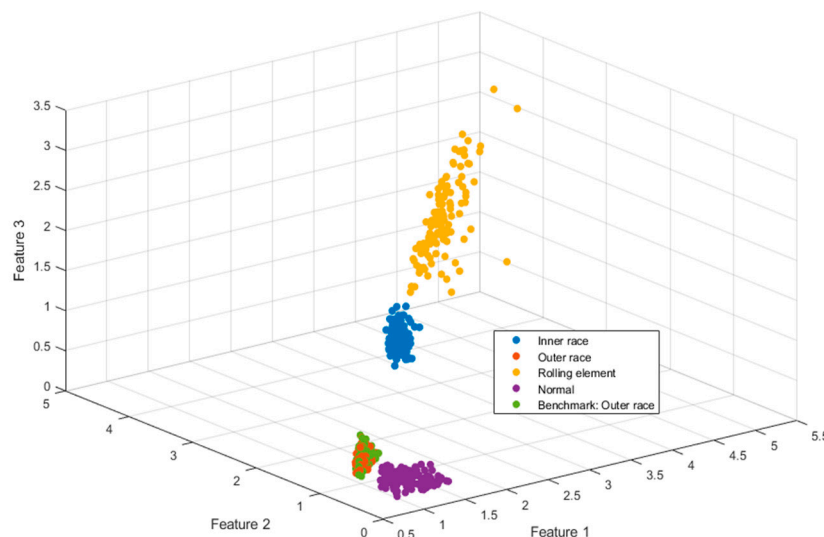
	Features							
	1	2	3	4	5	6	7	8
Normal	2.512	−0.471	0.390	0.726	0.789	−0.049	1.519	1.469
Inner Race	2.512	−0.471	0.390	0.726	0.789	−0.049	1.519	1.469
Rolling element	2.397	0.137	1.007	0.432	1.078	0.482	1.283	0.770
Outer Race	2.439	0.970	3.801	−0.104	0.676	2.216	1.501	−1.809

Thirdly, the MDs between the benchmark samples and validation samples were calculated by using the ascending and descending order of SNR;

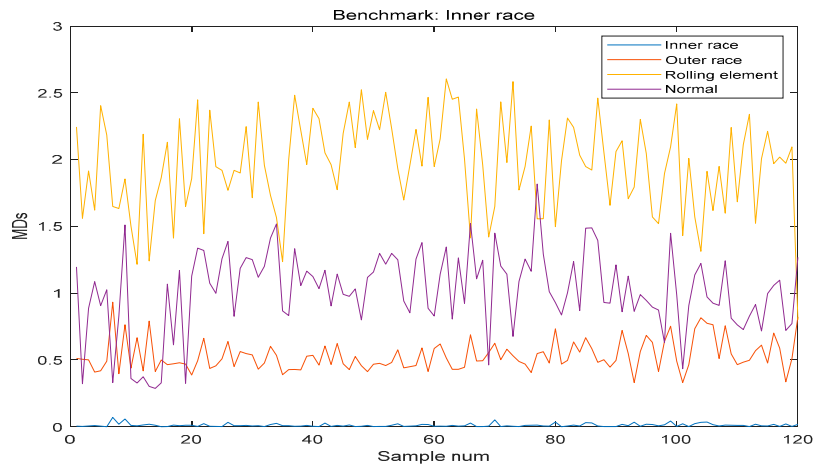
Fourthly, the validation samples were used to verify the correctness of feature selection which existed more than one smallest MD;

Fifthly, the best sequence was chosen and set as the sequence of features.

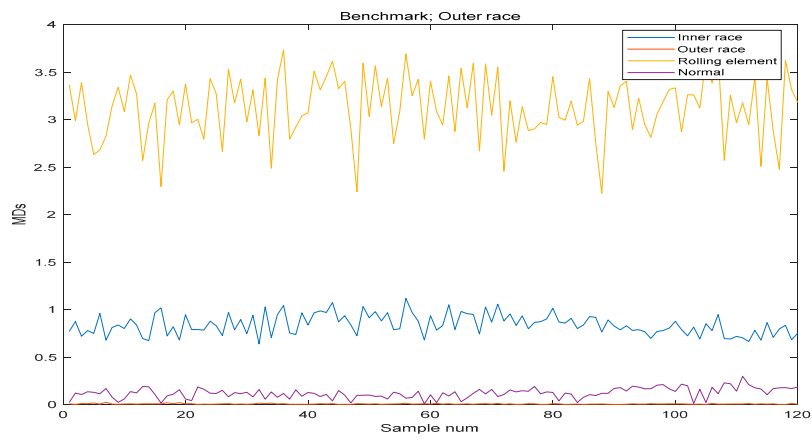
Lastly, the best sequence was used to identify the test samples. Took the benchmark is outer race as the example shown in Figure 9.

**Figure 9.** The classification result of outer race.

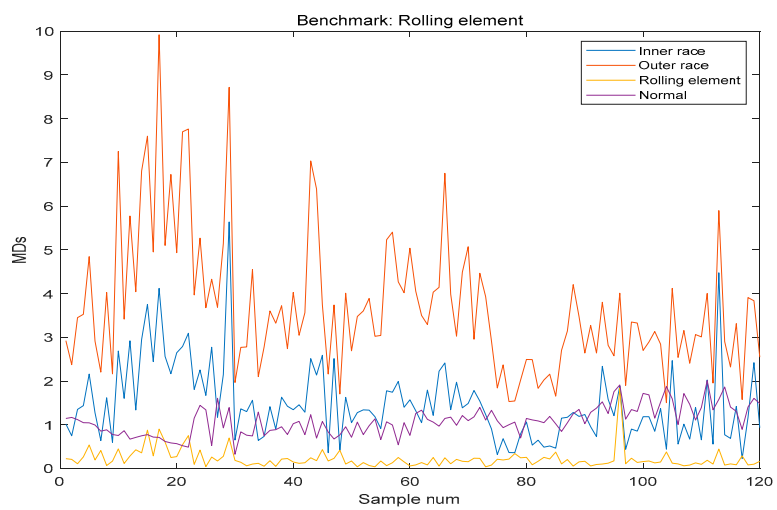
Finally, the test sample was used to test the result of the method, and the benchmarks were inner race, rolling element, outer race and normal. The results are shown in Table 7 and the MDs between benchmark and test sample are shown in Figure 10.



(a)

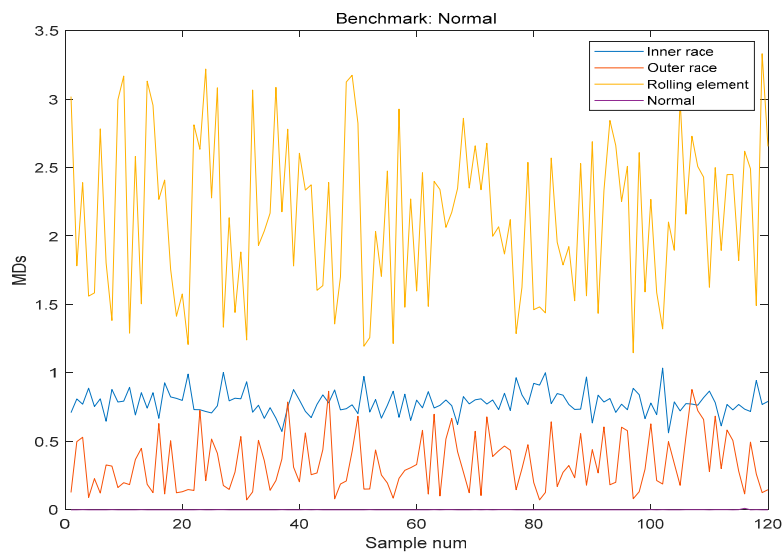


(b)



(c)

Figure 10. Cont.



(d)

Figure 10. The Mahalanobis distances (MDs) between benchmark and testing data: (a) The MDs between benchmark (Inner race) and testing data; (b) The MDs between benchmark (Outer race) and testing data; (c) The MDs between benchmark (Rolling element) and testing data; (d) The MDs between benchmark (Normal) and testing data.

Table 7. The recognition result.

	Inner Race	Outer Race	Rolling Element	Normal	Total
Result	100%	99.16%	95%	100%	98.54%

As shown in Table 7 and Figure 10, this method accurately classified and diagnosed the fault of the bearing by using the different benchmarks. The recognition results of normal and outer race reached 100%. However, it is not accurate enough to diagnose the fault of inner race and rolling element. However, in the normal and the fault of inner race, it is effective in industrial application.

4. Discussion

Rolling element bearings are one of the most frequently used components in rotating machineries. This paper presents the method based on the wavelet denoising VMD-SVD-aMMS to diagnose the fault of bearings under the variable conditions. Firstly, VMD is used to decompose the signal. Secondly, SVD is used to extract the feature. The adaptive aMMS uses the feature sequences and multi-benchmarks to overcome the drawback of MTS for adaptive feature selection, multi-classification and over-fitting. The experimental result shows that the method could accurately diagnose faults effectively.

However, in the actual situation, there is an imbalance between fault data and normal data. In this method, aMMS lacks research on the imbalanced study. The absence of faulty data may create a new problem, such over-fitting. Therefore, additional experiments under imbalanced data should be done to improve the method.

Author Contributions: N.W. collected and analyzed the data, made charts and diagrams, conceived and performed the experiments and wrote the paper; Z.W. and L.J. conceived the structure, provided guidance and modified the manuscript; X.C. analyzed the data and contributed analysis tools. Y.Q. provided guidance. Y.Z. revised the reviews.

Funding: This research was funded by National Key R&D Program of China grant number (No.2016YFB1200402).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tidiriri, K.; Chatti, N.; Verron, S.; Tiplica, T. Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring: A review of researches and future challenges. *Annu. Rev. Control* **2016**, *42*, 63–81. [[CrossRef](#)]
2. Xin, G.; Hamzaoui, N.; Antoni, J. Semi-automated diagnosis of bearing faults based on a hidden Markov model of the vibration signals. *Measurement* **2018**, *127*, 141–166. [[CrossRef](#)]
3. Al-Bugharbee, H.; Trendafilova, I. A fault diagnosis methodology for rolling element bearings based on advanced signal pretreatment and autoregressive modelling. *J. Sound Vib.* **2016**, *369*, 246–265. [[CrossRef](#)]
4. Park, B.; Jeong, H.; Huh, H.; Kim, M.; Lee, S. Experimental study on the life prediction of servo motors through model-based system degradation assessment and accelerated degradation testing. *J. Mech. Sci. Technol.* **2018**, *32*, 5105–5110. [[CrossRef](#)]
5. Jeong, H.; Park, B.; Park, S.; Min, H.; Lee, S. Fault detection and identification method using observer-based residuals. *Reliab. Eng. Syst. Saf.* **2018**. [[CrossRef](#)]
6. Cai, B.; Huang, L.; Xie, M. A data-driven fault diagnosis methodology in three-phase inverters for PMSM drive systems. *IEEE Trans. Power Electron.* **2017**, *32*, 5590–5600. [[CrossRef](#)]
7. Cai, B.; Huang, L.; Xie, M. Bayesian Networks in Fault Diagnosis. *IEEE Trans. Ind. Inform.* **2017**, *13*, 2227–2240. [[CrossRef](#)]
8. Cai, B.; Kong, X.; Liu, Y.; Lin, J.; Yuan, X.; Xu, H.; Ji, R. Application of Bayesian networks in reliability evaluation. *IEEE Trans. Ind. Inform.* **2018**. [[CrossRef](#)]
9. Chen, Z.; Deng, S.; Chen, X.; Li, C.; Sanchez, R.-V.; Qin, H. Deep neural networks-based rolling bearing fault diagnosis. *Microelectron. Reliab.* **2017**, *75*, 327–333. [[CrossRef](#)]
10. Pandya, D.H.; Upadhyay, S.H.; Harsha, S.P. Fault diagnosis of rolling element bearing with intrinsic mode function of acoustic emission data using APF-KNN. *Expert Syst. Appl.* **2013**, *40*, 4137–4145. [[CrossRef](#)]
11. Shao, H.; Jiang, H.; Lin, Y.; Li, X. A novel method for intelligent fault diagnosis of rolling bearings using ensemble deep auto-encoders. *Mech. Syst. Signal Process.* **2018**, *102*, 278–297. [[CrossRef](#)]
12. Jia, F.; Lei, Y.; Guo, L.; Lin, J.; Xing, S. A neural network constructed by deep learning technique and its application to intelligent fault diagnosis of machines. *Neurocomputing* **2018**, *272*, 619–628. [[CrossRef](#)]
13. Jagath Sri Lal Senanayaka, H.V.K.; Kjell, G. Robbersmyr Towards online bearing fault detection using envelope analysis of vibration signal and decision tree classification algorithm. In Proceedings of the 2017 20th International Conference on Electrical Machines and Systems (ICEMS), Sydney, Australia, 11–14 August 2017. [[CrossRef](#)]
14. Rama Krishna, K.; Ramachandran, K.I. Machinery Bearing Fault Diagnosis Using Variational Mode Decomposition and Support Vector Machine as a Classifier. *IOP Conf. Ser. Mater. Sci. Eng.* **2018**, *310*, 012076. [[CrossRef](#)]
15. Yan, R.; Gao, R.X.; Chen, X. Wavelets for fault diagnosis of rotary machines: A review with applications. *Signal Process.* **2014**, *96*, 1–15. [[CrossRef](#)]
16. Ben Ali, J.; Fnaiech, N.; Saidi, L.; Chebel-Morello, B.; Fnaiech, F. Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. *Appl. Acoust.* **2015**, *89*, 16–27. [[CrossRef](#)]
17. Muruganatham, B.; Sanjith, M.A.; Krishnakumar, B.; Satya Murty, S.A.V. Roller element bearing fault diagnosis using singular spectrum analysis. *Mech. Syst. Signal Process.* **2013**, *35*, 150–166. [[CrossRef](#)]
18. Ke, F.; KeSheng, W.; Mian, Z.; Qing, N.; Ming, J.Z. A diagnostic signal selection scheme for planetary gearbox vibration monitoring under non-stationary operational conditions. *Meas. Sci. Technol.* **2017**, *28*, 035003.
19. Caesarendra, W.; Tjahjowidodo, T. A Review of Feature Extraction Methods in Vibration-Based Condition Monitoring and Its Application for Degradation Trend Estimation of Low-Speed Slew Bearing. *Machines* **2017**, *5*, 21. [[CrossRef](#)]
20. Francois Combet, R.Z. A new method for the estimation of the instantaneous speed relative fluctuation in a vibration signal based on the short time scale transform. *Mech. Syst. Signal Process.* **2009**, *23*, 16. [[CrossRef](#)]
21. Cabrera, D.; Sancho, F.; Li, C.; Cerrada, M.; Sánchez, R.V.; Pacheco, F.; de Oliveira, J.V. Automatic feature extraction of time-series applied to fault severity assessment of helical gearbox in stationary and non-stationary speed operation. *Appl. Soft Comput.* **2017**, *58*, 53–64. [[CrossRef](#)]

22. Chen, X.; Wang, Z.; Zhang, Z.; Jia, L.; Qin, Y. A Semi-Supervised Approach to Bearing Fault Diagnosis under Variable Conditions towards Imbalanced Unlabeled Data. *Sensors* **2018**, *18*, 2097. [[CrossRef](#)]
23. Wang, Z.; Jia, L.; Kou, L.; Qin, Y. Spectral Kurtosis Entropy and Weighted SaE-ELM for Bogie Fault Diagnosis under Variable Conditions. *Sensors* **2018**, *18*, 1705. [[CrossRef](#)]
24. Ciabattoni, L.; Ferracuti, F.; Freddi, A.; Monteriù, A. Statistical Spectral Analysis for Fault Diagnosis of Rotating Machines. *IEEE Trans. Ind. Electron.* **2018**, *65*, 4301–4310. [[CrossRef](#)]
25. Chao-Ton Su, Y.-H.H. An Evaluation of the Robustness of MTS for Imbalanced Data. *IEEE Trans. Knowl. Data Eng.* **2007**, *19*, 12. [[CrossRef](#)]
26. Cudney, E.A.; Hong, J.; Jugulum, R.; Paryani, K.; Ragsdell, K.M.; Taguchi, G. An Evaluation of Mahalanobis-Taguchi System and Neural Network for Multivariate Pattern Recognition. *J. Ind. Syst. Eng.* **2007**, *1*, 12.
27. Soylemezoglu, A.; Jagannathan, S.; Saygin, C. Mahalanobis Taguchi System (MTS) as a Prognostics Tool for Rolling Element Bearing Failures. *J. Manuf. Sci. Eng.* **2010**, *132*. [[CrossRef](#)]
28. Idriss El-Thalji, E.J. A summary of fault modelling and predictive health monitoring of rolling element bearings. *Mech. Syst. Signal Process.* **2015**, *60*, 252–272. [[CrossRef](#)]
29. Dragomiretskiy, K.; Zosso, D. Variational Mode Decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [[CrossRef](#)]
30. Wang, Y.; Markert, R.; Xiang, J.; Zheng, W. Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system. *Mech. Syst. Signal Process.* **2015**, *60–61*, 243–251. [[CrossRef](#)]
31. Taguchi, G.; Konishi, S.; Konishi, S. *Taguchi Methods: Orthogonal Arrays and Linear Graphs. Tools for Quality Engineering*; American Supplier Institute: Dearborn, MI, USA, 1987.



© 2018 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/>).