



Article Rolling-Bearing Fault-Diagnosis Method Based on Multimeasurement Hybrid-Feature Evaluation

Jianghua Ge¹, Guibin Yin¹, Yaping Wang¹, Di Xu^{1,*} and Fen Wei²

- Key Laboratory of Advanced Manufacturing and Intelligent Technology, Ministry of Education, Harbin University of Science and Technology, Harbin 150080, China; gejianghua@sina.com (J.G.); guibin-yin@foxmail.com (G.Y.); wypbl@163.com (Y.W.)
- ² Hunan Provincial Key Laboratory of Mechanical Equipment Maintenance, Hunan University of Science and Technology, Hunan 411201, China; weifen@hunst.edu.cn
- * Correspondence: 15114642186@163.com

Received: 2 November 2019; Accepted: 15 November 2019; Published: 19 November 2019



Abstract: To improve the accuracy of rolling-bearing fault diagnosis and solve the problem of incomplete information about the feature-evaluation method of the single-measurement model, this paper combines the advantages of various measurement models and proposes a fault-diagnosis method based on multi-measurement hybrid-feature evaluation. In this study, an original feature set was first obtained through analyzing a collected vibration signal. The feature set included time- and frequency-domain features, and also, based on the empirical-mode decomposition (EMD)-obtained time-frequency domain, energy and Lempel–Ziv complexity features. Second, a feature-evaluation framework of multiplicative hybrid models was constructed based on correlation, distance, information, and other measures. The framework was used to rank features and obtain rank weights. Then the weights were multiplied by the features to obtain a new feature set. Finally, the fault-feature set was used as the input of the category-divergence fault-diagnosis model based on kernel principal component analysis (KPCA), and the fault-diagnosis model was based on a support vector machine (SVM). The clustering effect of different fault categories was more obvious and classification accuracy was improved.

Keywords: rolling bearing; feature evaluation; fault diagnosis; hybrid measurements

1. Introduction

As one of the most widely used components in rotating machinery, the health status of rolling bearings has an important impact on the working conditions of the entire mechanical equipment. Once a failure occurs, the performance of the equipment is greatly reduced and even has catastrophic consequences [1]. Therefore, the status monitoring and fault diagnosis of rolling bearings have become an important part to ensure the normal operation and safety of machinery and equipment.

As a key part of the fault diagnosis of rolling bearings, feature extraction aims to extract various parameter indices that reflect the fault characteristics by analyzing the original vibration signals in the time, frequency, and time-frequency domains. In recent years, it has become a popular trend to use integrated multidomain and multicategory features to characterize the fault modes of rolling bearings [2–4]. Feature extraction plays an important role in subsequent data processing. This is not only reflected in the processing of vibration and sound signals, but also in machine learning and deep learning [5–7]. The importance of feature extraction is more obvious in the diagnosis of rotating machinery faults [8–10]. Now, many scholars have invested in the research of rolling bearing fault diagnosis based on feature evaluation and feature extraction [11–13]. Usually, fault-classification input is composed of a high-dimensional feature vector. However, the high-dimensional feature set

increases computational demands and reduces diagnostic efficiency [14,15]. On the other hand, the advantage of sensitive features to fault classification cannot be highlighted, as correlation between nonsensitive features of the classification effect is weakened, which makes it difficult to improve the accuracy of fault diagnosis. Therefore, before a feature-vector set is input into the classifier, each feature value needs to be evaluated. According to the feature score, each feature parameter is weighted. The existing evaluation criteria are mainly divided into the following categories: distance measures [16], correlation measures [17], information measures [18], and consistency measures [19]. The evaluation criteria of each category also include a number of feature-evaluation models, such as the average distance between classes, within-class-between-class integrated distance, and the Fisher score [20] belonging to the category of distance measurement. The Pearson correlation coefficient belongs to a class of correlation-measure criteria, and information gain, minimum description length, and mutual information are the evaluation criteria based on information measures. In addition, the Laplacian score (LS) proposed by He et al. [21,22] retentively scores features based on the local information and the variance of the features, which provides a new idea for feature evaluation and has successfully been applied to fault diagnosis. The above-mentioned various characteristic evaluation models evaluate characteristics from different aspects, respectively, which has a certain one-sidedness that then affects subsequent classification performance [23–28]. Therefore, based on the combination of multiple models under different measurement criteria, it is possible to obtain relatively comprehensive and objective evaluation results.

In view of the above analysis, this paper proposes a fault-diagnosis method based on multi-measurement hybrid-feature evaluation. By using a four-feature evaluation model of comprehensive distance, correlation, and information, the original feature set composed of the time domain, frequency domain, and time-frequency domain features parameters that are used to obtain the feature score. The new weighted feature set is formed for each feature weight by each feature score, and then applied to the fault diagnosis of the rolling bearing. Finally, the proposed diagnostic method is applied to two different sets of experimental data of rolling-bearing failure. The comparative-analysis results verified the effectiveness and superiority of the proposed method. The paper is organized as follows: Feature extraction and the weighting scheme are introduced in Section 2. Section 3 describes the method and process of fault diagnosis results. Finally, conclusions are drawn in Section 5.

2. Feature Extraction and Multimeasurement Hybrid-Feature Weighting Scheme

2.1. Multiple-Type Feature Extraction from Multiple Domains

When rotating machinery falls into a faulty state, the amplitude and probability distribution of the collected vibration signal changes, as well as the frequency components and the position of spectral peak. Thus, statistical features describing distribution information of the time-domain waveform and frequency spectrum have been efficacious indicators of failure occurrence in rotating machinery. Due to the advantages of clear physical meanings of simple computation and strong practicability, various statistical features of the time and frequency domains have been successfully utilized in the fault diagnosis of rotating machinery in many studies. Among these, 10 time-domain features, namely, peak, mean, variance, root mean square (RMS), skewness, kurtosis, waveform index, crest index, pulse index, and peak indicator, and five frequency-domain features, namely, mean frequency, gravity frequency, RMS frequency, standard deviation, and kurtosis frequency, were adopted in this paper to construct the original time- and frequency-domain feature vector set; the detailed computation equations of the above features are provided in Table 1.

Notation	Equation
f_1^t	$\max(\mathbf{x}(n))$
f_2^t	$\frac{\sum_{n=1}^{N} x(n)}{N}$
f_3^t	$\frac{\sum_{n=1}^{N} (\hat{\mathbf{x}}(n) - f_2^t)}{N}$
f_4^t	$\sqrt{\sum_{n=1}^{N} (\mathbf{x}(n))^2 / N}$
f_5^t	$\frac{\sum_{n=1}^{N} (\boldsymbol{x}(n) - f_2^t)^3}{N_{12} \left(\sqrt{f_1^t} \right)^3}$
f_{c}^{t}	$\frac{\sum_{n=1}^{N} (\mathbf{x}(n) - f_2^t)^4}{\sum_{n=1}^{N} (\mathbf{x}(n) - f_2^t)^4}$
f_7^t	$\frac{N \times (f_3^t)^2}{f_4^t / f_2^t}$
f_8^t	$f_1^{\dagger} / f_4^{\dagger}$
J9 ct	$\underbrace{\begin{smallmatrix} J_1 \\ J_2 \\ f_1^t \end{smallmatrix}}_{f_1^t}$
J_{10}^{*}	$\frac{1}{N} \sqrt{\sum_{n=1}^{N} \mathbf{x}(n) ^2}$
f_1^p	$\frac{\sum_{k=1}^{\kappa} \boldsymbol{s}(k)}{K}$
f_2^p	$\frac{\sum_{k=1}^{K} f_k \times \boldsymbol{s}(k)}{\sum_{k=1}^{K} \boldsymbol{s}(k)}$
f_3^p	$\sqrt{rac{\sum_{k=1}^{K}(f_k)^2 imes oldsymbol{s}(k)}{\sum_{k=1}^{K}oldsymbol{s}(k)}}$
f_4^p	$\sqrt{rac{\sum_{k=1}^{K}\left(f_{k}-f_{2}^{p} ight)^{2} imes m{s}(k)}{K}}$
f_5^p	$\frac{\sum_{k=1}^{K}\left(f_k - f_2^p\right)^4 \! \times \! \boldsymbol{s}(k)}{K \! \times \! \left(f_4^p\right)^4}$

Table 1. Detailed description of statistical features.

x(n), time series of a signal; $n = 1, 2, \dots, N$; N, length of the series; s(k), frequency spectrum of x(n) wherein $k = 1, 2, \dots, K$; K, number of spectral lines; and f_k , frequency value of the k-th spectral line.

Aiming at the characteristics of low signal to noise ratio (SNR), that is, the nonlinearity and nonstationarity of the faulty signal, various types of time-frequency domain features were employed to mine more fault information than could be found with statistical parameters in the time and frequency domains. In this paper, two types of time-frequency domain features, i.e., energy and Lempel–Ziv complexity based on empirical-mode decomposition (EMD), were employed to build the original time-frequency-domain feature set. As one of the common nonlinear time-frequency analysis technologies, EMD adaptively decomposes a complicated signal x(t) into a set of complete and almost orthogonal components, namely, intrinsic mode functions (IMFs) $c_i(t)$ and residue $r_n(t)$. The decomposition procedure is defined as:

$$\mathbf{x}(t) = \sum_{i=1}^{n} \mathbf{c}_i(t) + \mathbf{r}_n(t)$$
(1)

The energies of IMFs including different frequency bands have the ability to more fully and effectively reveal the original vibration signal in view of time-frequency amplitude and distribution; thus, they are considered as efficacious indicators of the faulty state of rotating machinery. In this paper, after correlation analysis, the first five IMFs containing almost all valuable information were selected. The energy of the *i*-th IMF ($i = 1, 2 \cdots, 5$) was calculated as:

$$E_i = \int_{-\infty}^{+\infty} |\boldsymbol{c}_i(t)| dt \tag{2}$$

Furthermore, in order to eliminate the impact of the physical dimension, E_i was normalized by E_i/E , wherein $E = \sum_{i=1}^{5} E_i$.

Lempel–Ziv complexity measurement is a tool to weigh the complexity of finite sequences, which reflects the occurrence rate of a new mode in a time sequence along with the increase of length. Aiming at this topic, the Lempel–Ziv complexity indicator is expected to respond to changes in the condition of rotating machinery. Considering the dispersed sequence of the $S_{iN} = [S_{i1}, S_{i2}, \dots, S_{iN}]$, then the Lempel–Ziv value of S_{iN} can be calculated by $C_{iN}(r)$ ($r \le N$) through N cycles; $C_{iN}(r)$ is the Lempel–Ziv complexity of the *i*-th IMF, and N is the length of the binary data that is occupied by IMF when solving the complexity. The value of N needs to be defined based on the length of the IMF. The size of Lempel–Ziv complexity is affected by the length of the sequence N. To be robust, normalized

$$L_{i} = C_{i,N}(r) = (C_{i,N}(N)\log_{2}N)/N$$
(3)

Summarily, an original 25-dimensional fault-feature set, composed of a 15-dimensional time-domain and frequency-domain statistical-feature vector, and a 10-dimensional time-frequency-domain feature vector $\mathbf{F}_2 = [E_1/E, E_2/E, \dots, E_5/E, L_1, L_2, \dots, L_5]$, was constructed for subsequent fault diagnosis.

2.2. Hybrid Feature-Weighting Scheme

Lempel–Ziv complexity is defined as:

To strengthen the performance of the subsequent classification module, the obtained feature set was weighted with a novel hybrid model based on the feature-weighting scheme proposed in this section before they were input into the subsequent intelligent classier. The proposed feature-weighting scheme synthetically evaluated the original features from the perspectives of distance, correlation, and information; the corresponding flowchart is shown in Figure 1.



Figure 1. Flowchart of proposed feature-weighting scheme.

As seen in Figure 1, there are three steps for the capture of the weighted-feature set in the proposed scheme: the computation of feature weights based on the single evaluation model, the hybrid product of feature weights originating from the four available feature-evaluation models, and the feature-weighting phase through combining the integrated score vector of the feature set with the original feature set. From these steps, the second, namely, acquiring the integrated sensitivity vector of the feature set, is the key to the proposed scheme.

The hybrid feature-evaluation model embraces four mature feature-evaluation models, w_j^a , w_j^c , w_j^i , w_j^n , which are the distance, correlation, traditional-information, and special-information operators of the *j*-th feature, and calculated by inner- and interclass integrated distance, an integrated Pearson correlation coefficient, regularization information gain and Laplacian score, respectively, as follows.

Given a sample matrix $\mathbf{X} = \{x_{ij}\} \in \mathbb{R}^{N \times M}$ and the corresponding classification-label vector $\mathbf{Y} = \{y_i\} \in \mathbb{R}^N$, where *N* and *M* respectively denote the number of samples and dimensions of features, x_{ij} is the *j*-th feature value of the *i*-th sample and $y_i \in [1, 2, \dots, C]$ is the corresponding classification label of the *i*-th sample, where *C* denotes the class number in the sample set.

2.2.1. Four Basic Measure Schemes

1. Category distance is an important criterion for the separability of the samples of different categories. The commonly used category distance mainly includes inner- and interclass average distance, inner--inter-class comprehensive distance, etc. Therefore, the calculation expression based on the category-distance model is as follows:

$$w_{j}^{d} = \frac{d_{j}^{b}}{d_{j}^{w}} = \frac{\frac{1}{C \times (C-1)} \sum_{c,e=1}^{C} |\mu_{e,j} - \mu_{c,j}|}{\frac{1}{C} \sum_{c=1}^{C} \frac{1}{N_{c}(N_{c}-1)} \sum_{l,m=1}^{N_{c}} |x_{m,c,j} - x_{l,c,j}|}$$

$$(4)$$

$$l \neq m$$

In Equation (4), $x_{m,c,j}$ is the *j*-th feature value of the *m*-th sample affiliated with the *c*-th class, N_c is the number of samples affiliated with the *c*-th class in the sample set, $N = \sum_{c=1}^{C} N_c$, d_j^b and d_j^w are, separately, the average distance between different class samples and the one within C classes, and the *j*-th feature mean value of the *c*-th class samples is defined as $\mu_{c,j} = \frac{1}{N_c} \sum_{m=1}^{N_c} x_{m,c,j}$. According to the principle that class distance is an important criterion to the divisibility of different class samples, in this paper, the class distance of samples under a certain feature dimension, especially inner- and interclass integrated distance, is regarded as an effective indicator of feature-score degree to distinguish different faulty classes. It is noteworthy that, the bigger the value of w_i^d , the stronger the score of the *j*-th feature.

Inner-and interclass integrated distance of samples in some features can be used as measurement criteria for the score of samples to different faults. This has the advantage of comprehensively considering similarity in the same category and difference between different categories, but this method does not reflect the corresponding relationship between features and classification labels.

2. The Pearson correlation coefficient measures the contribution of individual features to classification by using the correlation between features, or features and categories. The formula is as follows: $\Sigma^{N} = \Sigma^{N}$

$$w_{j}^{c} = \frac{p_{j}^{f-c}}{p_{j}^{f-f}} = \frac{(M-1)\frac{\sum_{i=1}^{N} x_{ij}y_{i}\frac{\sum_{i=1}^{N} x_{ij}y_{i}\frac{\sum_{i=1}^{N} y_{ij}}{N}}{\sqrt{\left(\sum_{i=1}^{N} x_{ij}^{2}\frac{\left(\sum_{i=1}^{N} x_{ij}\right)^{2}}{N}\right)\left(\sum_{i=1}^{N} y_{i}^{2}\frac{\left(\sum_{i=1}^{N} y_{i}\right)^{2}}{N}\right)}}{\sum_{\substack{s \neq j}}^{M} \left[\frac{\sum_{i=1}^{N} x_{ij}x_{is}\frac{\sum_{i=1}^{N} x_{ij}\sum_{i=1}^{N} x_{ij}}{N}}{\sqrt{\left(\sum_{i=1}^{N} x_{ij}^{2}\frac{\left(\sum_{i=1}^{N} x_{ij}\right)^{2}}{N}\right)\left(\sum_{i=1}^{N} x_{is}^{2}\frac{\left(\sum_{i=1}^{N} x_{ij}\right)^{2}}{N}\right)}}\right]}$$
(5)

In Equation (5), p_j^{f-c} and p_j^{f-f} , respectively, denote the Pearson correlation coefficient between features and classes, and the one between features. The above-defined integrated Pearson correlation coefficient synchronously considers the correlation of the *j*-th feature with classes and the correlation of the *j*-th feature with other features to measure the score of the *j*-th feature. In other words, if the correlation of the *j*-th feature with a class is the highest and the one of the *j*-th feature with other features is the lowest, then this feature is preferred to other features.

An integrated Pearson correlation coefficient uses the correlation between features or between features and category labels to measure the contribution of individual features to classification. This method measures linear correlation and is only sensitive to linear relations. If there is a nonlinear relationship between features and category labels, the integrated Pearson correlation coefficient is not able to correctly screen for it.

3. Regularization-information gain is a typical characteristic-evaluation model based on information measurement, whose basic idea is to measure the importance of characteristics for

classification by calculating the value of useful information brought by a certain characteristic. Therefore, the calculation expression based on regularization-information gain is as follows:

$$w_j^i = 2 \times \frac{H(\mathbf{Y}) - H(\mathbf{Y}|\mathbf{x}_j)}{\left(H(\mathbf{x}_j) + H(\mathbf{Y})\right)H(\mathbf{Y})}$$
(6)

where $H(x_j)$ is the information entropy of the *j*-th feature in the sample set, H(Y) is the information entropy of the class-label vector, and $H(Y|x_j)$ is the conditional entropy. The central theme of the regularization-information gain defined in Equation (6) is to measure the importance degree of the *j*-th feature by calculating the size of useful information brought by the feature for classification. The bigger the value of w_i^i is, the more useful the information of the *j*-th feature.

Regularization-information gain calculates the useful information brought by a certain characteristic for classification to measure the importance of this characteristic for classification. Regularization-information gain can only deal with attribute values of continuity, but not the characteristics of continuous values. The natural bias of the algorithm is to choose properties with more branches, which easily leads to overfitting.

4. The Laplacian score is based on Laplacian feature-value mapping and local retention projection. Its basic idea is to measure features through local-information-retention ability and variance information. Therefore, the calculation expression based on the Laplacian score model is as follows:

$$w_{j}^{ni} = \frac{\sum_{i,n=1}^{N} (x_{ij} - x_{nj})^{2} s_{in}}{\operatorname{var}(x_{j})}$$
(7)

where $var(x_j)$ is the estimated variance of the *j*-th feature, s_{in} is the element of the *i*-th row and the *n*-th column in similarity matrix S, which denotes the similarity between the *i*-th and *n*-th samples. If node i and n are connected to each other in the neighborhood graph, $s_{in} = exp(-||x_i - x_n||^2/t)$; otherwise, $s_{in} = 0$. Equation (7), namely, the Laplacian score calculation equation, is considered a new information-measurement-based feature-evaluation model due to its characteristic of reflecting locality-preserving power and variance information. By contrast to the aforementioned three operators, value w_i^{ni} is inversely proportional to the classification contribution of the feature.

The Laplacian score measures features through local-information retention and variance information, and describes the inherent local geometric structure of the data space. It is convenient and quick to operate. However, the Laplacian score ignores the relationship between features. It analyzes each feature independently and discards the relation between features, which makes it difficult for the Laplacian score to seek global optimization.

2.2.2. Weight Calculation of Hybrid-Feature Evaluation

It is worth noting that the above-mentioned classical feature-evaluation models are based on a single measure criterion to obtain the feature score, of which the results have some limitations, as they are not good at reflecting the relationship between features and labels and the linear correlation between features, do not have single-category feature-selection capabilities, and pose difficulties in seeking global optimality. Feature evaluation and the weighting framework of hybrid models based on distance, correlation, and information measurement can make up for the shortcomings of single-measurement feature evaluation. Among them, 25 characteristics were evaluated by four single-measurement models, and four feature evaluation results were obtained. The results were multiplied to obtain comprehensive feature-evaluation results, and comprehensive feature-evaluation weights were further calculated. The integrated feature weights and original feature values were calculated. This combination constituted a new fault-feature set that could combine the advantages of various measurement methods. Feature weighting has the characteristics of strengthening the contribution to the classification and weakening

the contribution of the classification while retaining the integrity of the original feature set. When the scores of the single-measurement methods are consistent, the multi-measurement hybrid-feature evaluation model increases as the score of the single-measurement model increases, and lower as the score of the single-measurement model decreases. This can screen out features with a large classification contribution rate, and features with a small classification contribution rate. When the scores of various single-measurement methods are inconsistent, the multiplication product comprehensively balances the scores of various measurement scores, the characteristics of the medium contribution rate can be further subdivided, and the dominant features can be screened out. The new fault-feature set obtained by the established multi-measurement hybrid-feature evaluation model is input into the subsequent diagnosis model to obtain the diagnosis result. The multi-measurement hybrid-feature evaluation model can more accurately evaluate features and further improve diagnostic accuracy. The integrated weight $w^H = \left\{ w^H_j \right\}_{j=1}^M$ of the new fault feature set is calculated by:

$$w_j^H = w_j^d \times w_j^c \times w_j^i / w_j^{ni} \tag{8}$$

Finally, the new fault-feature set is obtained by computing integrated score vector w^H of the feature set to multiply with the original feature set. From the definitions in Equations (4)–(7), it is remarkable that the value ranges of correlation measurement index $w_{j'}^c$ traditional information evaluation index $w_{j'}^i$ and special information index w_{j}^{ni} are 0 to 1. The value range of distance measurement index w_{j}^d is also larger than 1. To achieve more reliable multi-measurement hybrid feature weights, the four indices were respectively normalized before constructing hybrid weighting index w_{j}^H . Then, the obtained integrated feature weights were further normalized to eliminate the value-range fluctuation that is caused by the product admixture formula shown as Equation (8). The weight value of the *j*-th feature was normalized by the following formula:

$$w_j' = \frac{w_j - \min(w)}{\max(w) - \min(w)}$$
(9)

where w_j can represent measurement indices w_j^c , w_j^i , w_j^{ni} , w_j^d or w_j^H , respectively, and ω is the corresponding feature-weight vector w^c , w^i , w^{ni} , w^d or w^H . According to Equation (8), the four score vectors after normalization are used to calculate the score vector of the hybrid feature. After normalization, the hybrid-feature score vector is weighted to the original matrix to obtain the hybrid-feature weighted-feature matrix.

The hybrid model proposed in this paper scores the original feature set and weights the original feature set according to the given score. The new fault feature set highlights superiority over features with a high classification contribution rate at fault classification and weakens interference in features with a small classification contribution rate set at fault classification. Compared with the original feature set, the new fault-feature set is weighted for each feature. The feature of large classification contribution rates and the feature of small classification contribution rates and the feature of interference obtains smaller or even zero weight values. The new fault-feature set gives play to the advantage of features and improves classification accuracy.

3. Method and Process of Fault Diagnosis Based on Multimeasurement Hybrid-Feature Evaluation

3.1. Fault-Diagnosis Method Based on Multimeasurement Hybrid-Feature Evaluation

In order to verify the effectiveness of the new fault-feature set based on the multi-measurement hybrid-feature evaluation model in this paper, a category-divergence fault-diagnosis model based on kernel principal component analysis (KPCA) [29] and a fault-diagnosis model based on a support vector machine (SVM) [30], are proposed. Category-divergence fault diagnosis based on KPCA quantifies the

clustering ability of different feature sets using category divergence (including within-, between-, and within-class-between-class scatter) as the evaluation index. It can be considered a direct diagnostic model. The second diagnosis scheme is used to support the vector to take the output results of the SVM model with different feature sets under the same structural parameters. That is to say, classification accuracy is a measure index of different feature evaluation and weighted schemes. It can be considered as an indirect diagnostic model.

3.1.1. Fault-Diagnosis Model Based on KPCA

In order to further evaluate the influence of the proposed feature-weighting scheme on follow-up classification from a quantitative sense, three commonly used clustering-performance measurement indicators, namely, within-class scatter S_w , between-class scatter S_b , and synthesized within-class-between-class scatter *SS* were employed into the first three-dimensional principal components, namely, three KPCA-based reduction features of the five aforementioned different feature sets. Their detailed definitions are shown as follows:

$$S_w = \frac{1}{C} \sum_{i=1}^{C} \left(\frac{1}{N_i} \sum_{j=1}^{N_i} \|f_j^i - \mu_f^i\|^2 \right)$$
(10)

$$S_b = \frac{1}{C} \sum_{i=1}^{C} \left\| \mu_f^i - \overline{\mu}_f \right\|^2$$
(11)

$$SS = \frac{S_b}{S_w} \tag{12}$$

where *C* is the number of fault status categories, $N = \sum_{i=1}^{C} N_i$ is the number of total samples in which N_i is the number of samples belonging to the *i*-th class, $\overline{\mu}_f$ is the feature value of the *j*-th sample in the *i*-th class, μ_f^i is the mean feature value of the *i*-th class, $\overline{\mu}_f = 1/C\sum_{i=1}^{C} \mu_f^i$ is the total mean feature value of all classes. As known from the above definitions, within-class scatter S_w characterizes the average concentrated level of samples belonging to the same class, between-class scatter S_b indicates the average dispensability among different classes, and synthesized within-class–between-class scatter S_b and S_b . It is noteworthy that the clustering performance of a certain feature is proportional to the values of between-class scatter S_b and synthesized within-class–between-class scatter SS, but inversely proportional to the value of within-class scatter S_w in other words, the larger the value of between-class scatter S_b and synthesized within-class scatter SS, but the smaller the value of within-class scatter S_w , the better the clustering performance of the first KPCA-based 3D reduction features of features sets originating from four different feature score-weighting methods, as well as that of the original unweighted-feature set.

3.1.2. Fault-Diagnosis Model Based on SVM

The SVM is a machine learning method based on statistical theory and structural risk minimization principle. It is a large edge classifier suitable for small samples. The thought of the SVM is to find an optimal hyperplane that can be classified, which can not only guarantee the classification accuracy but also maximize the blank space on both sides of the hyperplane. Theoretically, SVM can achieve the optimal classification of linear classification data, and the classification of complex nonlinear data in high-dimensional Hilbert space can be achieved through nonlinear projection.

Given a nonlinear classification sample set $\{(x_i, y_i)\}_{i=1}^N$, where $x_i \in \mathbb{R}^d$, -dimensional input vector, $y_i[\pm 1]$ is the corresponding class label of x_i , N is the number of samples in the sample set. Through the nonlinear mapping function $\varphi(\cdot)$, the input sample data is mapped from the original space to

the high-dimensional feature space, and the optimal classification hyperplane is constructed in the high-dimensional feature space:

$$\begin{cases} (\boldsymbol{\omega} \cdot \boldsymbol{\varphi}(\mathbf{x})) + \boldsymbol{b} = 0\\ s.t.y_i((\boldsymbol{\omega} \cdot \boldsymbol{\varphi}(\mathbf{x})) + \boldsymbol{b}) + \varepsilon_i \ge 1, i = 1, 2, \cdots, n \end{cases}$$
(13)

where ω is the weight vector; b is bias; and $\varepsilon_i \ge 0$ is a soft variable that allows a certain degree of misclassification for some points around the decision boundary. ω and b determine the location of the classification hyperplane. The determination of the optimal classification hyperplane can be achieved by solving the following convex quadratic programming optimization problem:

$$\min\varphi(\omega,\varepsilon) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n \varepsilon_i$$
(14)

where *C* is a penalty factor. By changing the penalty factor, the generalization ability and misclassification rate of the classifier can be balanced.

The kernel function plays an important role in the SVM classifier. It can calculate the dot product result of the high space in low-dimensional space and use it as a means of classification. Now, the kernel function has been fully developed. Radial-basis functions (RBF) are the most widely used and have a wide convergence domain. It is an ideal classification function and can be defined as:

$$K(x_i, y_j) = \exp(-||x_i - y_j||^2 / 2\sigma^2)$$
(15)

where σ represents nuclear parameters. We use an RBF constructed SVM, where penalty factor *C* and kernel parameter σ directly affect the classification accuracy of the SVM classifier. Therefore, in this paper, different input feature sets adopt a SVM classification model with the same parameter value, thus reflecting the fairness and objectivity of the program comparison.

3.2. Fault-Diagnosis Process Based on Multimeasurement Hybrid-Feature Evaluation

In order to prove the model presented in this paper is better than a traditional single measure model, two different diagnostic models were used for analysis, respectively, and the specific process is shown in Figure 2.

(1) Vibration signals of rolling-bearing faults are statistically analyzed relating to the time and frequency domains, building $T_1 = \{P_1, P_2, \dots, P_{15}\}$, the time and frequency domains of the original feature set; and EMD analysis is carried out on the vibration signal, extraction, and original signal correlation based on correlation analysis of the top five largest energy and Lempel–Ziv complexity characteristics of the IMF, building the original frequency-domain feature set $T_2 = \{E_1 \sim P_2, L_1 \sim L_5\}$, a combination of T_1 and T_2 building 1 × 25 D multidomain fault-feature vector $T = \{t_1, t_2, \dots, t_{25}\}$.

(2) Using the proposed fault-diagnosis method to evaluate each feature parameter t_i in the original fault-feature set T, the comprehensive score value HFS_i of each feature is obtained, and new fault-feature set T is formed by combining the corresponding feature weight and feature value.

(3) The new fault-feature set in Step (2) is input in the category-divergence fault-diagnosis model based on KPCA and the fault-diagnosis model based on SVM. This obtains category divergence and classification accuracy, respectively.





Figure 2. Flowchart of diagnosis based on multi-measurement hybrid-feature evaluation.

4. Experiment Analysis

4.1. Experiment Depictions

In order to evaluate the effectiveness of the proposed fault-diagnosis method for rotating machinery, two real-world experiments involving corrosion faults of rolling-element bearings and gearbox failures were employed in this paper. The specific information of the two cases is as follows.

The first experiment, Case I, was fault diagnosis of rolling-element bearing where fault data originated from the Case Western Reserve University [31]. The layout diagram of this experiment is shown in Figure 3.



Figure 3. Schematic diagram of rolling-element-bearing fault-simulation experiment (Case I).

The test bench comprised a 2 HP reliance electric motor (left), a torque transducer (center), a dynamometer (right), and control electronics. In the experiment process, nine kinds of corrosion faults of the drive-end bearing in the motor were imitated by respectively seeding single points of 0.007, 0.014, and 0.021 inches in diameter on the rolling element, and inner and outer raceway with the electro

discharge machine. Bearing-vibration data were collected by the accelerometer attached to the housing with magnetic bases. Here, 10 kinds of bearing conditions—normal and the nine aforementioned types of corrosion faults—were investigated. Sampling frequency was 12 kHz, and 50 samples were selected in each condition. The experiment data is shown in Table 2.

Motor Speed (rpm)	Motor Load (HP)	Fault location	Fault Diameter (inch)	Name of Setting
1750	2	normal	0	normal
1750	2	Rolling	0.007	B/0.007
1750	2	Rolling	0.014	B/0.014
1750	2	Rolling	0.021	B/0.021
1750	2	Inner-race	0.007	IR/0.07
1750	2	Inner-race	0.014	IR/0.014
1750	2	Inner-race	0.021	IR/0.021
1750	2	Outer-race	0.007	OR/0.007
1750	2	Outer-race	0.014	OR/0.014
1750	2	Outer-race	0.021	OR/0.021

Table 2. Description of dataset in case I.

The second experiment (Case II) was conducted on the gearbox-fault simulation-test rig shown in Figure 4 [32].



Figure 4. Schematic diagram of gearbox-fault experiment (Case II).

The tested bearing type was double-row ball bearing, main shaft frequency was 10 Hz (600 rpm), the bearing-fault vibration signal was measured by the vibration-acceleration sensor installed on the gearbox housing at the upper end of the bearing under test, and sampling frequency of the signal was 48 KHz. Four bearing states were normal; inner and outer race, 0.8 mm wide and 0.3 mm deep notch failure; and ball, 0.5 mm wide, 0.5 mm deep notch failure. The experiment data is shown in Table 3.

Motor Speed(rpm)	Fault Location	Fault Diameter (mm)	Name of Setting
600	normal	0	Normal state
600	Inner-race	0.8	Inner-race failure
600	Outer-race	0.3	Outer-race failure
600	Ball	0.5	Ball failure

4.2. Validation and Comparisons of New Fault-Feature Set

For the two groups of experiment data, the data length of each single sample was set to 2048, with 50 samples for each state, 30 of which were used as training samples and the remaining 20 samples were used as test samples.

First, time-frequency and EMD-based time-frequency domain analysis was performed on the sample data of each type in each experiment to construct eigenvectors $T = [T_i]_{1\times 25}$. Two primitives of different dimensions were constructed according to the number of samples in the two groups of experiment data, respectively. Characteristic matrix $C_1 \in R^{500\times 25}$ and $C_2 \in R^{200\times 25}$.

Second, by using the hybrid-feature evaluation model proposed in this paper, we respectively scored the feature parameters corresponding to matrices C_1 and C_2 of each column vector to obtain the sum of scores corresponding to two matrix feature parameters $s_1 = [HFS_1, HFS_2, \dots, HFS_{25}]$ and s_2 , and then constructed a row vector of s_1 or s_2 . The number of rows was equal to the number of sample-weight matrixes s_1 and s_2 , respective to original feature matrix C_1 and C_2 dot multiplication. Two sets of experiment data corresponding to the hybrid-feature score weighted feature matrix sum were WW_1^H and WW_2^H .

Considering that the eigenvectors of a single sample in this weighting-feature matrix had high dimensionality, which is unfavorable for the visual display of samples, the KPCA, a typical nonlinear feature-dimensionality reduction method, was used for matrix WW₁^H and WW₂^H dimension reduction, so as to obtain the two- or three-dimensional display output of sample data, as shown in Figures 5a and 6a. At the same time, in order to realize new fault-feature matrices WW_1^H and WW_2^H based on the multi-measurement hybrid-feature evaluation proposed in this paper, comparative analysis of the sample-clustering performance for the weighted-feature set based on the four traditional single-measurement features (WP, weighted eigenmatrix based on Pearson correlation coefficient model; WD, weighted eigenmatrix based on category distance model; WI, weighted eigenmatrix based on mutual information model; and WL, weighted feature matrix based on Laplacian scoring model) and original unweighted feature matrices C_1 , C_2 . The KCPA method was applied to the five different feature matrices above that were, respectively, dimensionally reduced to obtain the two-dimensional output results of the five feature matrices, as shown in Figures 5b-f and 6b-f. In Figures 5 and 6, PC1 PC2, PC3 stand for the first, second, third dimensional principal components after KPCA dimensionality reduction, respectively. We can see two different groups of experiment data of rolling-bearing failure. The distinguishing effects of the feature samples were obviously better than those of the four other weighted-feature samples and the original unweighted-feature samples based on the single-measurement evaluation model, which not only made between-class scatter clustering of each type of fault-sample data more concentrated, but also greatly improved existing overlap between different types of samples in five feature-sample sets.

At the same time, in order to realize quantitative evaluation of the effect of the method of feature weighting based on the multi-measurement hybrid-feature evaluation proposed in this paper on sample clustering and classification-performance improvement, according to the two sets of the characteristic-weighting scheme-evaluation system proposed in this paper, the six different failure-characteristic sets in the two different rolling-bearing failure experiments above were calculated and analyzed as necessary. Between them, the first evaluation method aimed to quantitatively evaluate the clustering performance of the feature set by comprehensively calculating the category divergence values of the sample data of the fault feature sets in each feature parameter dimension. At the same time, taking into account the large amount of computation caused by the high dimension of the feature set and the nonlinearity between the feature parameters in the feature set, this paper used KPCA to reduce the dimensionality of different feature sets; the result of the divergence and the mean of the first two principal components were an indirect reflection of the overall divergence index of the feature set. Therefore, with feature-divergence calculation according to Equations (12)–(14), the clustering performance of six different weighted feature sets in the two experiment sets were comprehensively



evaluated. The results of the divergence-value calculation are shown in Table 4 and Figure 7. In Figure 7, F1–F6 stand for feature set WW_2^H , WP_2 , WD_2 , WI_2 , WL_2 , C_2 , respectively.

Figure 5. Experiment data for Case I of six kinds of KPCA result feature matrix: (**a**) WW_1^H ; (**b**) WP_1 ; (**c**) WD_1 ; (**d**) WI_1 ; (**e**) WL_1 ; (**f**) C_1 .



Figure 6. Experiment data for Case II of six kinds of different KPCA result feature matrix: (**a**) WW_2^H ; (**b**) WP_2 ; (**c**) WD_2 ; (**d**) WI_2 ; (**e**) WL_2 ; (**f**) C_2 .

Feature Sets –	First Main Element			Second Main Element			Mean Value		
	S_w	Sb	SS	S_w	Sb	SS	S_w	Sb	SS
WW_1^H	1.26×10^5	3.78×10^8	3.01×10^3	5.82×10^3	8.44×10^{6}	1.45×10^3	$6.59 imes 10^4$	1.93×10^8	2.23×10^3
WP_1	1.25×10^4	7.18×10^{6}	575.19	3.26×10^{3}	1.29×10^{6}	396.76	7.87×10^{3}	4.24×10^{6}	485.98
WD_1	8.59×10^{3}	2.16×10^{7}	2.51×10^{3}	3.25×10^{3}	2.01×10^{6}	616.76	5.92×10^{3}	$1.18 imes 10^7$	1.56×10^{3}
WI_1	8.53×10^{3}	6.43×10^{6}	753.15	2.87×10^{3}	1.17×10^{6}	407.11	5.70×10^{3}	3.79×10^{6}	580.13
WL_1	1.24×10^4	1.91×10^{7}	1.54×10^{3}	2.75×10^{3}	1.71×10^{6}	622.50	7.57×10^{3}	1.04×10^{7}	1.08×10^{3}
C_1	474.22	3.57×10^5	753.15	159.75	6.50×10^4	407.11	316.98	2.11×10^5	580.13





Figure 7. Divergence-calculation results of six different fault-feature sets in experiment data 2. (a) Within-class scatter values corresponding to different measurement methods; (b) between-class scatter values corresponding to different measure methods; and (c) within-class–between-class scatter values corresponding to different measure methods.

From Table 4 we can see that, for the first set of rolling-bearing-fault experiment data, the multi-measurement hybrid-evaluation weighted-feature matrix achieved the maximum between-class and within-class-between-class scatter at the expense of a certain degree of within-class scatter corresponding to it; the four other groups of weighted feature sets based on the single-measurement evaluation of class-divergence between-class and within-class-between-class scatter were second; and the original unweighted-feature between-class and within-class-between-class scatter was minimal. It can be seen from Figure 7 that, in the experiment data of the second group of rolling-bearing failures, the minimum within-class-between-class scatter, the largest between-class scatter, and the within-class-between-class scatter were obtained based on the new fault-feature set of the multi-measurement hybrid-evaluation model proposed in this paper. In addition, the four other weighted feature sets based on the single-measurement evaluation model had higher between-class and within-class-between-class scatter than the original unweighted-feature sets.

To sum up, the weighted method based on multi-measurement hybrid evaluation proposed in this paper has more obvious advantages for improving clustering performance of the original fault-feature set of the rolling bearing than the existing weighted method based on a single measurement.

Another method for evaluating the weighting scheme is to use the classification accuracy of differently weighted feature sets in the same classifier model as a measurement of quality evaluation and weighting scheme corresponding to the new fault-feature set. We used an SVM, a classic small-sample large-edge classifier, as the common classification model for the six different feature sets in the above experiment data; structural parameters of the value were set to C = 2, $\sigma = 1$. SVM model classification accuracy corresponding to each feature set is shown in Figure 8. It can be seen from Figure 8 that, for experiment data 1, the new fault-feature matrix weighted by the score of each feature parameter in the feature set was weighted by the original feature matrix and had higher classification accuracy than the original feature matrix. The weighted-feature matrix based on the hybrid-model scores proposed in this paper overcame the one-sidedness of the traditional single-model evaluation and achieved the highest classification accuracy.



Figure 8. Support vector machine (SVM) test-classification results of six different feature sets in experiment data 1.

The classification results of six different fault-feature sets in experiment data 2 based on the SVM classification model are shown in Table 5. With the help of the same classification model, the new fault-feature set based on the multi-measurement hybrid-evaluation model achieved the highest classification accuracy of 95.0% (cumulative number of misclassified samples was four). In contrast, the four other weighted-feature sets based on the single-measurement evaluation model achieved 90.0% (cumulative number of misclassified samples was eight), 87.5% (cumulative number of misclassified samples was 10), 81.25% (cumulative number of misclassified samples was 15), and 87.5% (cumulative number of misclassified samples was 10) classification accuracy, respectively; the unweighted-feature set had the lowest classification accuracy of only 78.75% (cumulative number of misclassified samples was 17), the multi-measurement hybrid-evaluation model had the advantages of quality features and reduce interference of inferior features. Classification accuracy of the single-measurement evaluation model was not as high as that of the multi-measurement hybrid-evaluation model. The total number of sample misclassifications was not as low as that of the multi-measurement hybrid-evaluation model. The experiment data 2 shows that the proposed method based on the multi-measurement hybrid-evaluation model could achieve higher classification performance than the traditional correlation based on distance, information, and Laplacian score.

Feature Sets	Cumulative Number of	Normal State	Inner Race Failure	Outer Race Failure	Ball Failure	Classification Accuracy (%)
	Misclassified Samples	Misclassified Sample Number	Misclassified Sample Number	Misclassified Sample Number	Misclassified Sample Number	
WW_2^H	4	2	2	0	0	95.0
WP_2	8	0	3	0	5	90.0
WD_2	10	0	10	0	0	87.5
WI_2	15	3	10	2	0	81.25
WL_2	10	2	5	0	3	87.5
C ₂	17	2	5	3	7	78.75

Table 5. Comparison of test-classification results of six different fault feature sets in experiment data 2based on SVM model.

To sum up, for the two sets of experiment data of rolling-bearing failure simulation, the new fault-feature set based on multi-measurement hybrid evaluation achieved better performance than the traditional weighted-feature set based on single measurement, and the original unweighted-feature set higher classification accuracy, thus proving the advantage of feature-based classification based on multi-measurement hybrid evaluation proposed in this paper.

5. Conclusions

In order to make up for the defect of the single-measurement model and improve the accuracy of fault diagnosis, this paper proposed a rolling-bearing fault-diagnosis method based

on multi-measurement hybrid-feature evaluation. In the feature-evaluation step, the category distance, Pearson correlation coefficient, regularization-information gain, and Laplacian score were used and combined with a feature-weight vector. In the fault-diagnosis step, the KPCA was used to calculate the clustering effect, and SVM was used to calculate classification accuracy. This method was evaluated with two different sets of rolling-bearing-fault experiment data. The experiment results showed that the clustering effect of different fault categories was more obvious, and the classification accuracy was improved.

Author Contributions: Conceptualization, J.G.; methodology, J.G. and F.W.; software, G.Y.; validation, G.Y., D.X.; formal analysis, J.G.; investigation, G.Y.; resources, J.G.; data curation, G.Y.; writing—original draft preparation, G.Y.; writing—review and editing, D.X.; visualization, Y.W.; supervision, J.G.; project administration, J.G.; funding acquisition, J.G.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 51575143, and supported by Heilongjiang Provincial Natural Science Foundation of China grant number E2018046.

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

References

- 1. Randall, R.B.; Antoni, J. Rolling element bearing diagnostics—A tutorial. *Mech. Syst. Signal Process.* **2011**, 25, 485–520. [CrossRef]
- Li, Z.; Outbib, R.; Giurgea, S.; Hissel, D. Diagnosis for PEMFC systems: A data-driven approach with the capabilities of online adaptation and novel fault detection. *IEEE Trans. Ind. Electron.* 2015, 62, 5164–5174. [CrossRef]
- 3. Ju, B.; Zhang, H.; Liu, Y.; Liu, F.; Lu, S.; Dai, Z. A feature extraction method using improved multi-scale entropy for rolling bearing fault diagnosis. *Entropy* **2018**, *20*, 212. [CrossRef]
- 4. Xiao, Y.; Wang, Y.; Mu, H.; Kang, N. Research on misalignment fault isolation of wind turbines based on the mixed-domain features. *Algorithms* **2017**, *10*, 67. [CrossRef]
- 5. Guo, X.; Shen, C.; Chen, L. Deep fault recognizer: An integrated model to denoise and extract features for fault diagnosis in rotating machinery. *Appl. Sci.* **2017**, *7*, 41. [CrossRef]
- 6. Alías, F.; Socoró, J.C.; Sevillano, X. A review of physical and perceptual feature extraction techniques for speech, music and environmental sounds. *Appl. Sci.* **2016**, *6*, 143. [CrossRef]
- 7. Uddin, M.; Lee, J.; Rizvi, S.; Hamada, S. Proposing enhanced feature engineering and a selection model for machine learning processes. *Appl. Sci.* **2018**, *8*, 646. [CrossRef]
- 8. Vakharia, V.; Gupta, V.; Kankar, P. A comparison of feature ranking techniques for fault diagnosis of ball bearing. *Soft Comput.* **2016**, *20*, 1601–1619. [CrossRef]
- 9. Islam, M.R.; Islam, M.M.; Kim, J.-M. Feature Selection Techniques for Increasing Reliability of Fault Diagnosis of Bearings. In Proceedings of the 2016 9th International Conference on Electrical and Computer Engineering (ICECE), Dhaka, Bangladesh, 20–22 December 2016; pp. 396–399.
- 10. Chen, X.; Zhang, X.; Zhou, J.; Zhou, K. Rolling Bearings Fault Diagnosis Based on Tree Heuristic Feature Selection and the Dependent Feature Vector Combined with Rough Sets. *Appl. Sci.* **2019**, *9*, 1161. [CrossRef]
- 11. He, C.; Niu, P.; Yang, R.; Wang, C.; Li, Z.; Li, H. Incipient rolling element bearing weak fault feature extraction based on adaptive second-order stochastic resonance incorporated by mode decomposition. *Measurement* **2019**, 145, 687–701. [CrossRef]
- 12. Zhu, K.; Chen, L.; Hu, X. A Multi-scale Fuzzy Measure Entropy and Infinite Feature Selection Based Approach for Rolling Bearing Fault Diagnosis. *J. Nondestruct. Eval.* **2019**, *38*, 90. [CrossRef]
- 13. Li, J.; Wang, H.; Song, L.; Cui, L. A novel feature extraction method for roller bearing using sparse decomposition based on self-Adaptive complete dictionary. *Measurement* **2019**, *148*, 106934. [CrossRef]
- 14. Jiang, Q.; Yan, X.; Huang, B. Performance-driven distributed PCA process monitoring based on fault-relevant variable selection and Bayesian inference. *IEEE Trans. Ind. Electron.* **2015**, *63*, 377–386. [CrossRef]
- 15. Lu, F.; Jiang, J.; Huang, J.; Qiu, X. An Iterative Reduced KPCA Hidden Markov Model for Gas Turbine Performance Fault Diagnosis. *Energies* **2018**, *11*, 1807. [CrossRef]
- 16. Lei, Y.; He, Z.; Zi, Y. Fault diagnosis based on novel hybrid intelligent model. *Chin. J. Mech. Eng.* **2008**, *44*, 112–117. [CrossRef]

- 17. Guyon, I.; Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.
- 18. Li, J.; Cheng, K.; Wang, S.; Morstatter, F.; Trevino, R.P.; Tang, J.; Liu, H. Feature selection: A data perspective. *ACM Comput. Surv.* (*CSUR*) **2018**, *50*, 94. [CrossRef]
- 19. Zheng, K.; Wang, X. Feature selection method with joint maximal information entropy between features and class. *Pattern Recognit.* **2018**, 77, 20–29. [CrossRef]
- 20. Gu, Q.; Li, Z.; Han, J. Generalized Fisher Score for Feature Selection. arXiv 2012, arXiv:1202.3725.
- 21. He, X.; Cai, D.; Niyogi, P. Laplacian Score for Feature Selection. In *Advances in Nerual Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2005.
- 22. Si, L.; Wang, Z.; Tan, C.; Liu, X. Vibration-based signal analysis for shearer cutting status recognition based on local mean decomposition and fuzzy C-means clustering. *Appl. Sci.* **2017**, *7*, 164. [CrossRef]
- 23. Widodo, A.; Yang, B.-S. Support vector machine in machine condition monitoring and fault diagnosis. *Mech. Syst. Signal Process.* **2007**, *21*, 2560–2574. [CrossRef]
- 24. Chen, F.; Tang, B.; Song, T.; Li, L. Multi-fault diagnosis study on roller bearing based on multi-kernel support vector machine with chaotic particle swarm optimization. *Measurement* **2014**, 47, 576–590. [CrossRef]
- 25. Bordoloi, D.; Tiwari, R. Support vector machine based optimization of multi-fault classification of gears with evolutionary algorithms from time-frequency vibration data. *Measurement* **2014**, *55*, 1–14. [CrossRef]
- Liu, Z.; Cao, H.; Chen, X.; He, Z.; Shen, Z. Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings. *Neurocomputing* 2013, *99*, 399–410. [CrossRef]
- 27. Jiang, Y.; Cheng, G.; Kan, J.; Xuan, Z.; Ma, J.; Zhang, Z. Rolling bearing fault diagnosis based on NGA optimized SVM. *Chin. J. Sci. Instrum.* **2013**, *34*, 2684–2689.
- Liu, R.; Yang, B.; Zio, E.; Chen, X. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mech. Syst. Signal Process.* 2018, 108, 33–47. [CrossRef]
- 29. Zeng, L.; Long, W.; Li, Y. A Novel Method for Gas Turbine Condition Monitoring Based on KPCA and Analysis of Statistics T2 and SPE. *Processes* **2019**, *7*, 124. [CrossRef]
- 30. Liu, Y.-H.; Wang, S.-H.; Hu, M.-R. A self-paced P300 healthcare brain-computer interface system with SSVEP-based switching control and kernel FDA + SVM-based detector. *Appl. Sci.* **2016**, *6*, 142. [CrossRef]
- Loparo, K. Case Western Reserve University Bearing Data Center. Available online: https://csegroups. case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website (accessed on 15 September 2019).
- 32. Sawalhi, N.; Randall, R. Simulating gear and bearing interactions in the presence of faults: Part I. The combined gear bearing dynamic model and the simulation of localised bearing faults. *Mech. Syst. Signal Process.* **2008**, *22*, 1924–1951. [CrossRef]



© 2019 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/).