

Article

New Fault Recognition Method for Rotary Machinery Based on Information Entropy and a Probabilistic Neural Network

Quansheng Jiang ^{1,2,*} , Yehu Shen ^{1,2}, Hua Li ^{1,2} and Fengyu Xu ³

¹ School of Mechanical Engineering, Suzhou University of Science and Technology, Suzhou 215009, China; yehushen@mail.usts.edu.cn (Y.S.); lihua@mail.usts.edu.cn (H.L.)

² Suzhou Key Laboratory of Precision and Efficient Machining Technology, Suzhou 215009, China

³ College of Automation, Nanjing University of Posts and Telecommunications, Nanjing 210003, China; xufengyu598@163.com

* Correspondence: qschiang@163.com; Tel.: +86-512-6832-0622

Received: 30 November 2017; Accepted: 22 January 2018; Published: 24 January 2018

Abstract: Feature recognition and fault diagnosis plays an important role in equipment safety and stable operation of rotating machinery. In order to cope with the complexity problem of the vibration signal of rotating machinery, a feature fusion model based on information entropy and probabilistic neural network is proposed in this paper. The new method first uses information entropy theory to extract three kinds of characteristics entropy in vibration signals, namely, singular spectrum entropy, power spectrum entropy, and approximate entropy. Then the feature fusion model is constructed to classify and diagnose the fault signals. The proposed approach can combine comprehensive information from different aspects and is more sensitive to the fault features. The experimental results on simulated fault signals verified better performances of our proposed approach. In real two-span rotor data, the fault detection accuracy of the new method is more than 10% higher compared with the methods using three kinds of information entropy separately. The new approach is proved to be an effective fault recognition method for rotating machinery.

Keywords: fault recognition; information entropy; probabilistic neural network; rotary machinery; feature extraction

1. Introduction

Rotating machinery plays a key role in industrial production. Once failure occurs, it may lead to significant downtime losses. The condition monitoring and fault identification diagnosis for rotating machinery is an important guarantee to improve the reliability of mechanical equipment operations [1]. Since the fault feature information of rotating machinery is complex and changeable, it is important for mechanical fault diagnosis to accurately extract the intrinsic characteristics from all kinds of fault signals [2].

Common signal feature extraction methods for rotating machinery include time domain and frequency domain methods [3], e.g., time domain analysis, spectrum analysis, correlation analysis, zoom spectral analysis, independent component analysis, wavelet transform, as well as empirical mode decomposition [4–6]. These methods are effective only for single fault diagnosis.

As modern rotating machinery becomes more complex and intelligent, the state features have non-stationary dynamic and multi-source coupling characteristics [7]. There are many factors that cause faults. A new fast estimator of the spectral correlation (the fast spectral correlation) based on the short-time Fourier transform (STFT) for a cyclostationary signal is proposed [8]. One kind of fault can be described with different characteristic indices. The same symptom is often the result of

several faults, and the fault feature and, as a result, the mapping between the fault feature and the fault source is nonlinear [9]. The complexity of fault diversity, uncertainty, and the connection between various faults constitute the technical difficulties for fault diagnosis. It is difficult to extract effective fault feature information and diagnose faults by means of a single fault feature and diagnosis method. Therefore, it is necessary to use information fusion methods to combine multi-symptom information.

Entropy is a physical quantity that represents the degree of regularity and complexity of the system. It has been applied to the field of mechanical fault diagnosis in recent years [10–12]. Antoni [13] proposed an entropic evidence method to measure the negentropy of the squared envelope (SE) and of the squared envelope spectrum (SES) of the signal, which can capture the signature of repetitive transients in time or frequency domains. Gao et al. [14] proposed an ensemble empirical mode decomposition (EEMD) method-based singular spectrum entropy for signal analysis and fault diagnosis of rotating machinery. A fault detection method based on multi-sensor data fusion for planetary gearboxes is presented in [15]. Compared with the methods based on individual sensors, the proposed method achieves much higher accuracies in detecting planetary gearbox faults. Bai et al. [16] presented a fault quantitative diagnosis method which described the change disciplinary of rotor vibration by the power spectrum entropy with more measurement points and multiple speeds based on information entropy theory. Xu et al. [17] used approximate entropy to describe the irregularity and complexity of mechanical vibration signals and measured the complexity of two-dimensional signals of axis orbits. It also quantitatively evaluated the running status of rotating machinery. An approximate entropy-based method for machine health monitoring from a rolling bearing is proposed in [18]. Gui et al. [19] proposed wavelet packet characteristic entropy and a neural network for dynamic feature extraction and fault detection of a pressure pulsation signal of a turbine draft tube. When a rolling bearing is at the early fault stage, the signal is easily contaminated by the gear and other noise sources, and Sun et al. [20] presented a method of singular point recognition and feature extraction for the incipient bearing fault based on instantaneous envelope scalogram entropy. Jing [21] proposed an adaptive multi-sensor data fusion method based on deep convolutional neural networks (DCNN) for fault diagnosis. It obtained satisfactory diagnosis accuracy.

Staszewski [22] proposed a method of wavelet-based compression and feature selection for periodic, continuous, non-stationary, and transient non-stationary signals analysis. Kahirdeh et al. [23] utilized acoustic entropy to characterize the Materials in the Course of Degradation, and demonstrated that the maximum acoustic entropy that can emanate from materials during the course of degradation remains similar. Cui et al. [24] proposed a method of compressor valve fault diagnosis using information entropy and an SVM, obtaining good performance from experiments. Cabal et al. [25] presented a methodology for induction motor condition monitoring by analyzing a single parameter, which can detect different faults quantitatively.

As an intelligent data processing method, neural networks have a strong ability to deal with the nonlinear characteristics in the data. Among them, a probabilistic neural network (PNN) can approximate any nonlinear continuous function with arbitrary precision, and has the ability of self-organization, self-learning, fast learning speed, and parallel processing [26–28]. Deep learning is one of the hotspots in artificial intelligence (AI) techniques [29]. A novel AI method based on a deep belief network (DBN) is proposed for the unsupervised fault diagnosis of a gear transmission chain in [30].

In this paper, for the purpose of quantifying the characteristics of vibration signals, we propose a feature information fusion model based on a probabilistic neural network and information entropy. We extract three kinds of information entropy of vibration signals in the time domain, frequency domain, and signal complexity, respectively, according to information theory. We also conduct a typical fault diagnosis verification experiment for rotating machinery. The remainder of the paper is organized as follows: Section 2 describes three kinds of information entropy features. Then the probabilistic neural network is discussed in Section 3. The information entropy features fusion model is designed and

analyzed in Section 4. The simulation and application experiments analysis for the presented features fusion model are carried out in Section 5. Conclusions of this paper are presented in Section 6.

2. Descriptions of Information Entropy Features

Information entropy is a description of the degree of uncertainty of the system, so we can use it to measure the state change of the rotating machinery. A lower entropy value means less uncertainty of the information. That is to say, there are fewer disorders in the information.

The definition of the entropy within the system is as follows [31]:

Assume M is a Lebesgue space which is a σ -algebra and generated by measurable set S . It has μ measure and $\mu(M) = 1$. M could be described by a finite partition $A = \{A_i\}$ which is incompatible. That is to say: $M = \bigcup_{i=1}^n A_i$ and $A_i \cap A_j = \Phi, \forall i \neq j$. Entropy with regard to partition A is defined as:

$$H(A) = -\sum_{i=1}^n \mu(A_i) \log \mu(A_i) \quad (1)$$

where $\mu(A_i) (i = 1, 2, 3, \dots, n)$ is the measure for set A_i .

According to the theory of entropy, we analyze the energy features of vibration signals in the time and frequency domains. Furthermore, we extract entropy features in vibration signals and generate the following three kinds of features.

2.1. Singular Spectrum Entropy

Singular spectral entropy gives an indicator to measure the complexity or uncertainty of a vibrational signal with multiple spatial distributions on the whole. For a discrete time series $Y_t = [y_1, y_2, \dots, y_N]$ (N is the number of samples) of the vibration signal, we can map the original signal into the embedding space by using the delay inlay technique. Assuming that the length of the embedding is M , we can obtain an $(N - M + 1) \times M$ dimensional matrix:

$$A = \begin{bmatrix} y_1 & y_2 & \cdots & y_M \\ y_2 & y_3 & \cdots & y_{M+1} \\ \vdots & \vdots & \vdots & \vdots \\ y_{N-M} & y_{N-M+1} & y_{N-M+2} & y_N \end{bmatrix} \quad (2)$$

The singular values can be obtained by the SVD decomposition from the matrix A . The definition of singular value decomposition is as follows [32]:

We define Y as a matrix with the dimension $m \times n$, there are two orthogonal matrices $U = [u_1, \dots, u_m] \in R^{m \times m}$, $V = [v_1, \dots, v_n] \in R^{n \times n}$ which satisfy the following equation:

$$Y = UDV^T, D = \text{diag}(\sigma_1, \dots, \sigma_r) \in R^{m \times n}, r = \text{rank}(Y) \quad (3)$$

$\sigma_i = \sqrt{\lambda_i} (i = 1, 2, \dots, r, \dots, m)$ can be called the singular value of the matrix Y , where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r \geq 0, \dots, \lambda_{r+1} = \lambda_{r+2} = \dots = \lambda_m = 0$.

Thus, we decompose A with the SVD decomposition, and obtain all of the singular values $[\sigma_1, \sigma_2, \dots, \sigma_m]$. All of the singular values form the singular spectrum entropy of the vibration signal. We assume k is the number of non-zero singular values, and k stands for the number of different patterns of the column space of matrix A . Singular spectrum entropy can be defined as follows:

$$H_s = -\sum_{i=1}^m p_i \log p_i \quad (4)$$

where $p_i = \frac{\sigma_i}{\sum_{i=1}^m \sigma_i}$ is the relative weight of the i th singular value with regard to all the singular values.

Singular spectrum entropy extracts the intrinsic complexities of the system and describes its status species according to the SVD technology.

2.2. Power Spectrum Entropy

We assume $Y(\omega)$ as the discrete Fourier transform of the vibration signal $Y_t = [y_1, y_2, \dots, y_N]$ (N is the number of samples). Then the power spectrum of Y_t is $S(\omega) = \frac{1}{2\pi N} |Y(\omega)|^2$ [33]. According to the energy conservation law, we set:

$$\sum y^2(t)\Delta t = \sum |Y(\omega)|^2 \Delta \omega \quad (5)$$

As a result, the power spectrum $S = \{S_1, S_2, \dots, S_n\}$ could be considered as a partition of the vibration signal in frequency domain. Thus, we can define the corresponding power spectrum entropy as follows.

$$H_f = -\sum_{i=1}^n q_i \log q_i \quad (6)$$

where $q_i = \frac{S_i}{\sum_{i=1}^n S_i}$ is the proportion of the i th power spectrum with regard to the whole power spectrum.

From the definition of the power spectrum entropy, we know that the power spectrum entropy of the vibration signal depicts the distribution modes of the vibration energy in the frequency domain.

2.3. Approximate Entropy

Approximate entropy is a new metric to evaluate the complexity of a series. It estimates the probability of new mode generation according to the analysis of the complexity of a time series [34]. A non-negative value is applied to describe this complexity. We show an approximate entropy algorithm as follows:

We assume $\{u(i)\} (i = 1, 2, 3 \dots N)$ (N is the number of samples) is the given time series, m is the predefined number of dimensions of the patterns, and r is the predefined similarity threshold. The algorithm is as follows:

- (1) Form the series of data $\{u(i)\} (i = 1, 2, 3 \dots N)$ into an m dimensional vector $X(i)$;
- (2) Calculate the distance between $X(i)$ and $X(j)$:

$$d[X(i), X(j)] = \max_{k=0:m-1} |u(i+k) - u(j+k)| (i = 1 : N - m + 1, j = 1 : N - m + 1, i \neq j)$$

- (3) Count the number of distances that satisfy $d[X(i), X(j)] < r$ and set it as num . Then we can calculate the ratio between num and the total number of vectors as follows:

$$C_i^m(r) = \{d[X(i), X(j)] < r\} / (N - m + 1) \quad (7)$$

- (4) Take the logarithm for each $C_i^m(r)$ and obtain their mean:

$$\phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln C_i^m(r) \quad (8)$$

- (5) Set the dimension to be $m + 1$ and repeat the procedures from Step 1 to Step 4. We can obtain $\phi^{m+1}(r)$.
- (6) The approximate entropy of the series $\{u(i)\} (i = 1, 2, 3 \dots N)$ can be calculated with the following equation:

$$ApEn(m, r, N) = \phi^m(r) - \phi^{m+1}(r) \quad (9)$$

According to the definition of approximate entropy, we know that it is a function with respect to the dimension of the pattern, similarity threshold, and the number of samples. It estimates the probability of new pattern generation for a time series when the dimension changes. It is an effective non-linear analysis method since it only needs data with small length and could be applied to both determined and stochastic processes.

3. Probabilistic Neural Network (PNN)

The basic idea of a probabilistic neural network is to generate a decision space in a high-dimensional input space based on Bayesian rules, which is to say to minimize the expected risk of erroneous classification. A probabilistic neural network is a kind of artificial neural network which is based on statistics. More specifically, it is a kind of feed-forward network whose activation function is the Parzen window function. A PNN contains the strengths of both the Radial Basis Function (RBF) neural network and the classical probabilistic density estimation method. It performs better in pattern classification compared with traditional feed-forward neural networks.

We map the sample space to pattern space with the help of the PNN. As a result, we can obtain a network system which is robust and has an adaptive structuring ability. The structure of the PNN is similar with that of an RBF neural network. As shown in Figure 1, the main structure contains four layers: input layer, model layer, summation layer, and output layer.

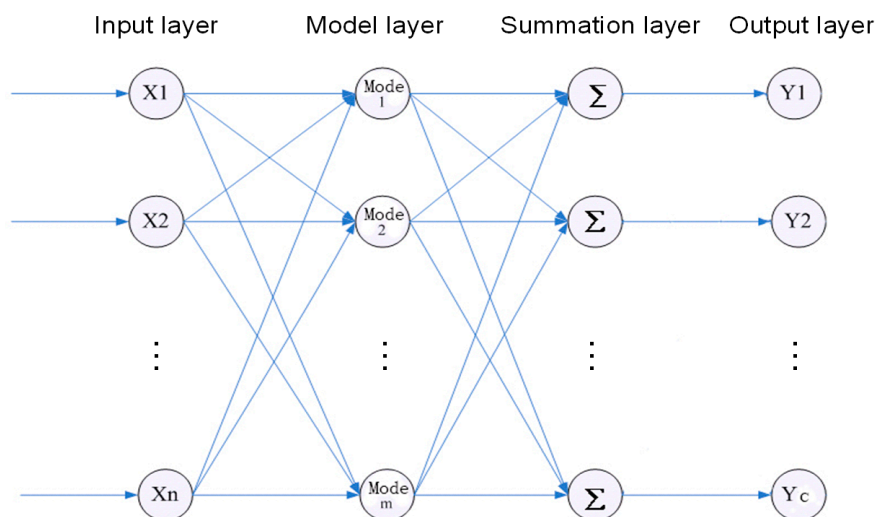


Figure 1. Structure diagram of the PNN.

The input layer receives the samples (the input dimension is n). The model layer is responsible for calculating the pattern correspondence between the input feature vector and the training set. It will generate a non-linear function, i.e., $\exp[-(n_k - 1)/\sigma^2]$, ($n_k = w_k^T x$). Here w_k is the connection weight between the input layer and the model layer. We can obtain the distance p_j^i ($i = 1, 2, \dots, m$, where m is the number of models in the sample) between the input vector and the weight vector from every model layer unit. Finally, we can obtain the conditional probabilistic density in the summation layer:

$$P_k = \sum_{j=1}^{m_k} \pi_j^i p_j^i \quad (10)$$

The output layer estimates the maximum probability R_k of a test sample, which belongs to a specific class based on Bayesian minimal risk estimation theory:

$$R_k = \sum_{l=1}^c v_l^k \alpha_l P_l \quad (11)$$

Thus, we can obtain the expected class of the test sample.

Since the weights between the input layer and the model layer could be adaptively chosen according to the training samples, the PNN is simple, robust, and easy to train.

4. Information Entropy Features Fusion Model

Information fusion is a signal processing procedure. It could manipulate complex multi-source information from different scales and aspects.

The fusion process could be divided into three levels according to the relationships among multi-source information: data-level fusion, feature-level fusion, and decision-level fusion. The main procedures for feature-level fusion are as follows: Firstly, we transform the raw data from every sensor into a feature vector. Then we fuse all these feature vectors. We obtain our final decision based on the fused results. The fused results compressed the raw data and extracted the key information. As a result, it could reduce the computational complexity.

This paper proposes a feature fusion model based on information entropy and probabilistic neural networks for rotary machine fault identification. The structure of the fusion model is demonstrated in Figure 2.

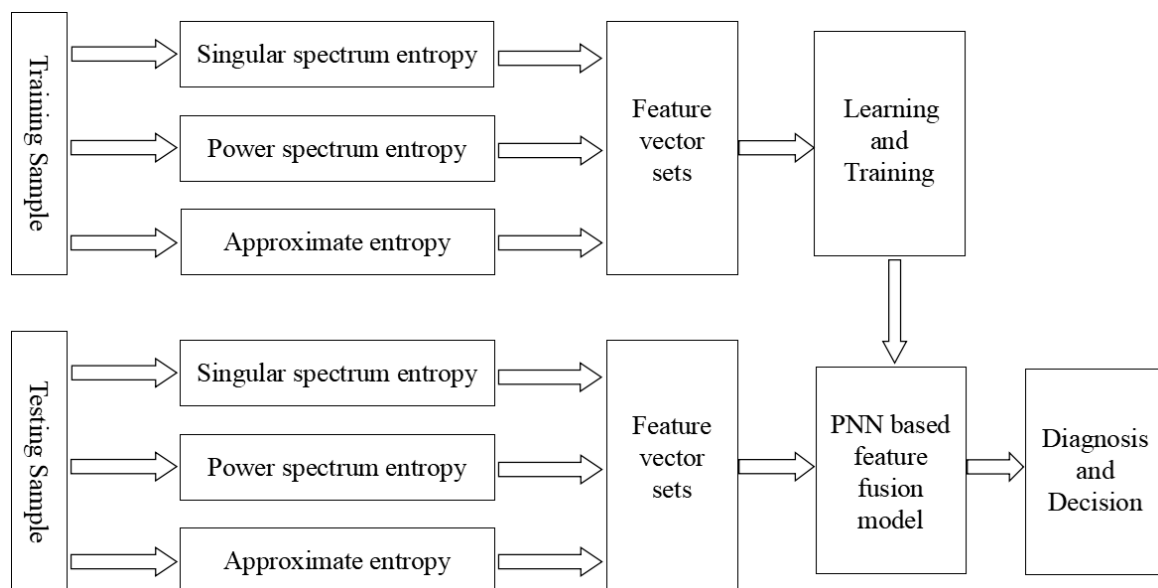


Figure 2. The structure of the feature fusion model based on information entropy and the PNN.

According to Figure 2, the main steps of the feature fusion model based on information entropy and the PNN are as follows.

- (1) The input data are collected from all the sensor outputs for fault identification. We calculated singular spectrum entropy, power spectrum entropy, and approximate entropy from the training data and construct the feature vectors. We feed the feature vectors into the PNN as training samples and train the model based on these samples.
- (2) We calculate the corresponding feature vector for each test sample and feed the vector to the PNN which has already trained in Step 1. We can obtain the fault identification result from the output layer of PNN.
- (3) For each test sample output, if its classification result is correct, we could add it into the training set. We retrain the whole model with this new training set in order to enhance the performance of the model.

5. Simulation and Application Experiments

5.1. Simulated Fault Signal

In this section, we generated three kinds of simulated fault signal of rotary machines as training and test samples.

For all the simulated signals, Class A stands for the type of unbalanced fault, Class B stands for the type of coupling misalignment fault, Class C stands for the type of rubbing fault. We set the sampling number to be 1024 and the rotation speed of the rotor is 3000 rpm. Thus, the working frequency of the rotor is $f_1 = 50$ Hz, double the fundamental frequency is $f_2 = 100$ Hz, high frequency is $f_3 = 200$ Hz, half the fundamental frequency is $f_4 = 25$ Hz, and extremely high frequency is $f_4 = 500$ Hz. We construct the simulated signals according to the different constitutions of frequencies for different faults.

Class A consists of 90% of frequency f_1 , and 5% of frequency f_2 and f_3 , respectively. Thus, the signal can be represented as follows:

$$X_1(t) = 0.9 \cos(2\pi f_1 t) + 0.05 \cos(2\pi f_2 t) + 0.05 \cos(2\pi f_3 t) + \varepsilon(t) \quad (12)$$

where $\varepsilon(t)$ is Gaussian noise whose mean is 0 and standard deviation σ is 0.1.

Class B consists of 50% of frequency f_2 , 40% of frequency f_1 , and 10% of frequency f_3 . Thus, the signal can be represented as follows:

$$X_2(t) = 0.4 \cos(2\pi f_1 t) + 0.5 \cos(2\pi f_2 t) + 0.1 \cos(2\pi f_3 t) + \varepsilon(t) \quad (13)$$

Class C consists of 40% of frequency f_1 , 20% of frequency f_2 , 10% of frequency f_3 , 20% of frequency f_4 , and 10% of frequency f_5 . Thus, the signal could be represented as follows:

$$X_3(t) = 0.4 \cos(2\pi f_1 t) + 0.2 \cos(2\pi f_2 t) + 0.1 \cos(2\pi f_3 t) + 0.2 \cos(2\pi f_4 t) + 0.1 \cos(2\pi f_5 t) + \varepsilon(t) \quad (14)$$

In Equations (12)–(14), $\varepsilon(t)$ is Gaussian white noise whose mean is 0 and standard deviation σ is 0.1.

We generate 100 groups of data from the above-mentioned three classes, respectively. We calculate the singular spectrum entropy, power spectrum entropy, and approximate entropy, respectively, from the training data and form them into feature vectors.

We randomly split the generated data into two halves. We use one half as the training set and train the PNN. The other half is used as the test set. Table 1 compares the results of the four methods. The first three methods use singular spectrum entropy, power spectrum entropy, and approximate entropy separately. The last method fuses all three kinds of entropies with the PNN.

Table 1. Results of four fault identification methods on simulated data.

Methods	Singular Spectrum Entropy	Power Spectrum Entropy	Approximate Entropy	Feature Fusion Model
Fault identification accuracy	91.7%	93.5%	68.3%	95.2%

From Table 1 we can conclude that the fault identification accuracy becomes much higher while fusing all three entropies when compared to using the three kinds of information entropy separately. This reflects the better performances and effectiveness of the proposed method based on information entropy and a probabilistic neural network in fault diagnosis.

5.2. Rotor Test Platform Fault Experiments

In this section, we collect signals from the rotor test platform, and calculate three kinds of information entropy from the collected signals.

The used rotor test rig is an experimental device to simulate the vibration condition of rotating machinery, which can effectively reproduce many kinds of vibration phenomena generated by rotating machinery. The rotor rig can simulate the running state of the machine by changing the rotor speed, shaft stiffness, mass unbalance, bearing friction, coupling form, or impact condition through different choices.

The working principle diagram of the rotor test rig is shown in Figure 3.

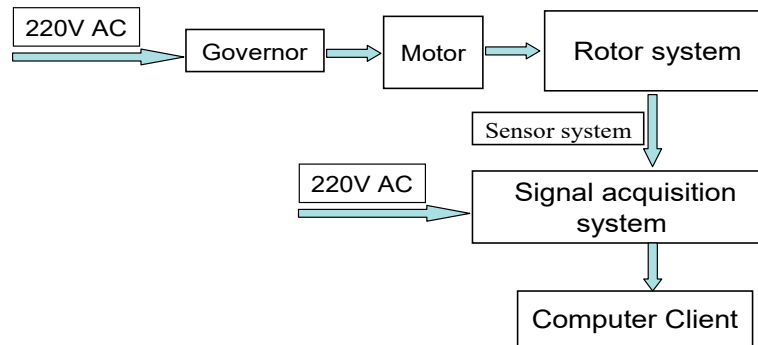


Figure 3. The working principle diagram of the rotor test rig.

We conduct experiments for identifying unbalance fault, coupling misalignment fault and rubbing fault with the propose feature fusion model.

5.2.1. Identifying Unbalance and Rubbing Faults for a Single-Span Rotor

We design a PNN model with three input units and two output units. The three input units correspond to the above-mentioned three kinds of information entropy and the two output units correspond to unbalance and rubbing faults, respectively.

We tested at three different rotation speeds, which are 1200 rpm, 2400 rpm, and 3600 rpm, for a single-span rotor. We collected vibration fault signals at the three rotation speeds. For each rotation speed, we collected 100 groups of data for training and 20 groups of data for testing.

For the single-span rotor, the calculated singular spectrum entropy of each vibration signal in different rotation speeds is shown in Figures 4 and 5.

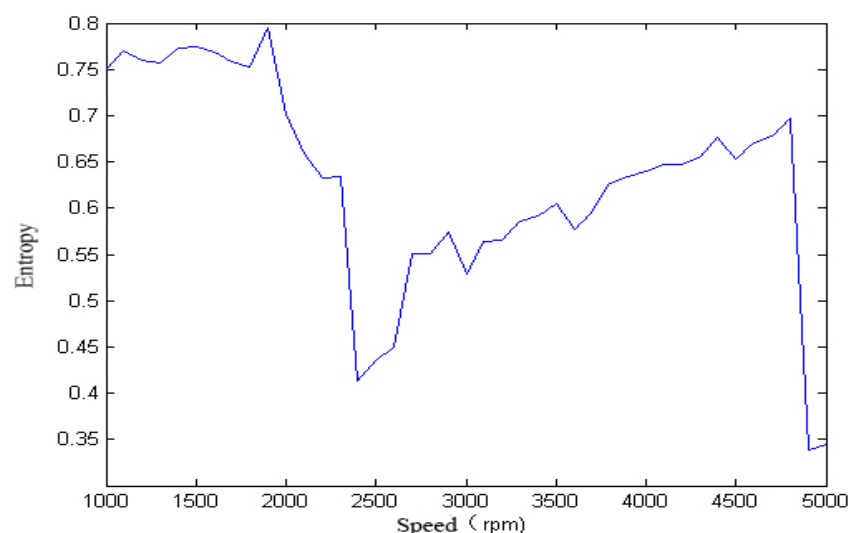


Figure 4. Singular spectrum entropy of a single-span rotor under the unbalance fault.

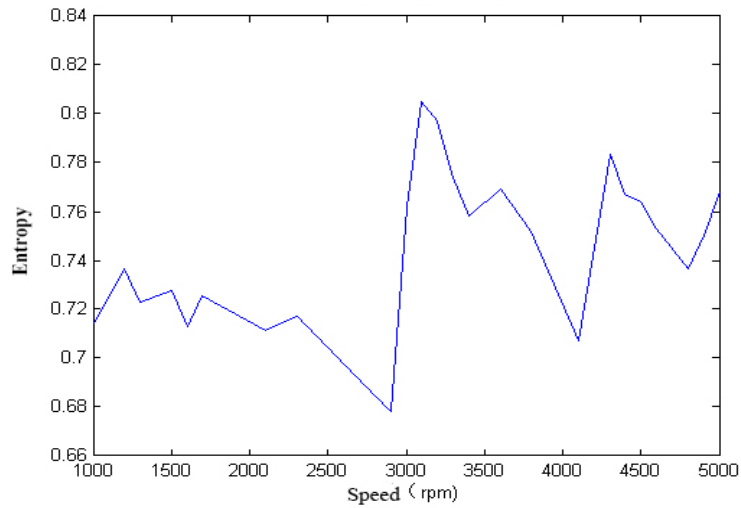


Figure 5. Singular spectrum entropy of a single-span rotor under the rubbing fault.

We calculated the singular spectrum entropy of each vibration signal in the single-span rotor at different rotation speeds, and the result is shown in Figures 6 and 7.

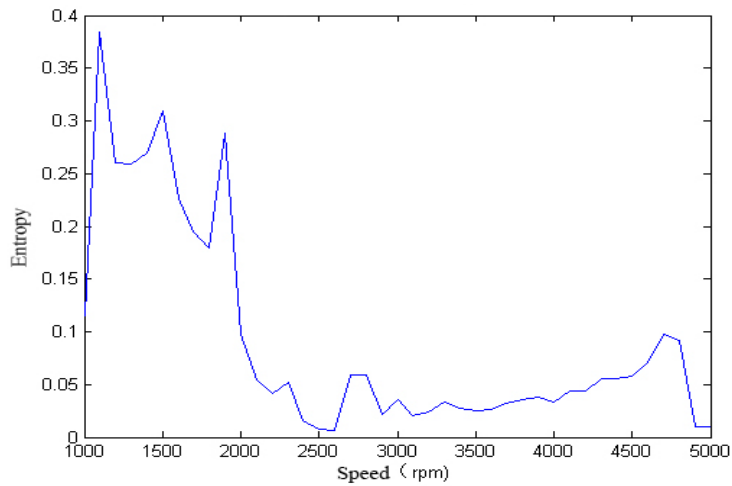


Figure 6. Power spectrum entropy of a single-span rotor under the unbalance fault.

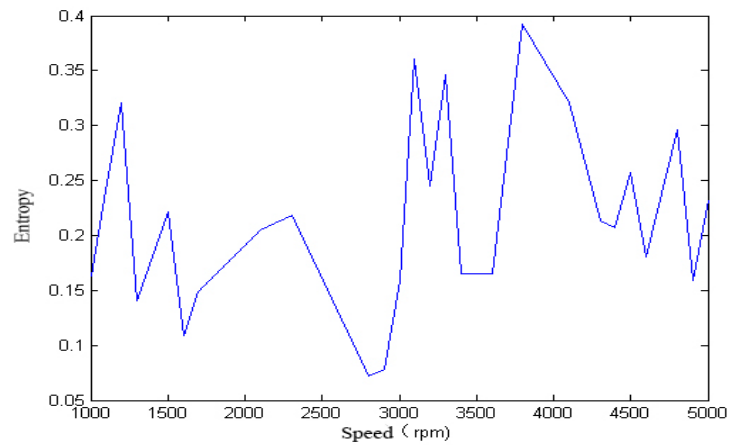


Figure 7. Power spectrum entropy of a single-span rotor under the rubbing fault.

We calculated the approximate entropy of each vibration signal in the single-span rotor at different rotation speeds, and the result is shown in Figures 8 and 9.

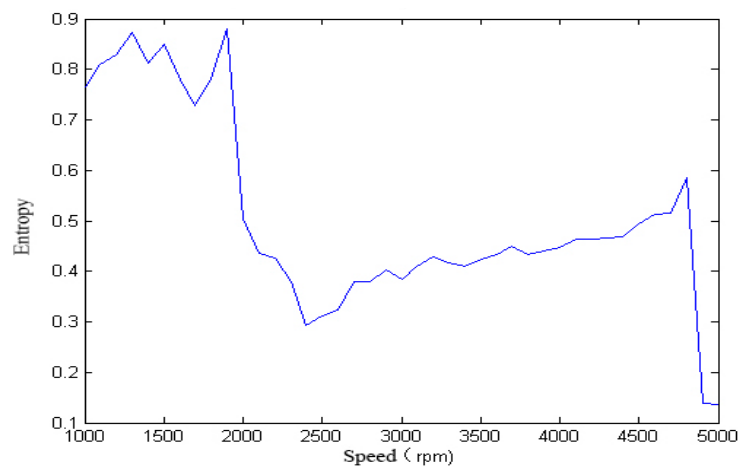


Figure 8. Approximate entropy of a single-span rotor under the unbalance fault.

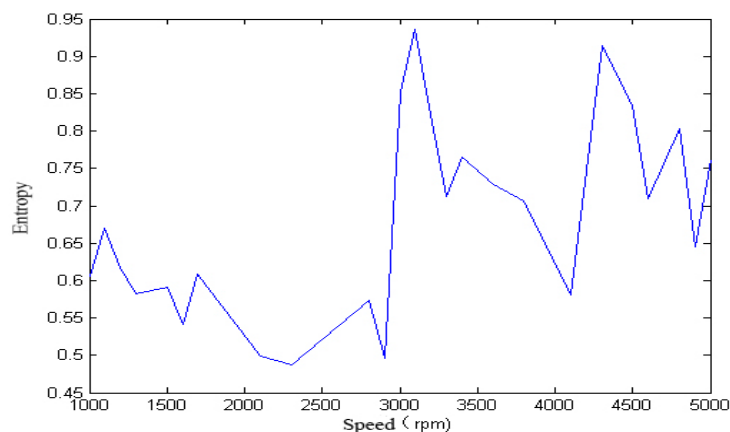


Figure 9. Approximate entropy of a single-span rotor under the rubbing fault.

It can be seen from Figures 5–9 that, with the change of speed, the variation law of entropy of a single-span rotor in different failure states presents different trends, and the range of variation is wide. With the increase of speed, the singular spectrum entropy, power spectrum entropy, and approximate entropy in the unbalance fault show a generally decreasing trend, while the singular spectrum entropy, power spectrum entropy, and approximate entropy of the rubbing fault show a generally increasing trend. Table 2 compares the fault identification accuracy rate of the four methods.

Table 2. Results of four fault identification methods on real data of a single-span rotor.

Rotation Speed	Singular Spectrum Entropy	Power Spectrum Entropy	Approximate Entropy	Feature Fusion Model
1200 rpm	71.2%	78.9%	78.9%	83.1%
2400 rpm	81.6%	92.2%	86.9%	94.8%
3600 rpm	93.9%	64.3%	96.2%	98.7%

From Table 2 we can conclude that there is no clear winner for any single kind of entropy from singular spectrum entropy, power spectrum entropy, and approximate entropy. Their accuracies are much lower than that of using the feature fusion model. This is because the above three kinds of

entropy describe the characteristics of signals from the time domain, frequency domain, and complexity domains, respectively. The proposed feature fusion model could combine information from different characteristics and, as a result, it is more sensitive to the fault features and could boost the performance.

5.2.2. Identifying Unbalance, Coupling Misalignment Faults, and Rubbing Faults for Two-Span Rotors

We designed a PNN model with three input units and three output units. The three input units correspond to the above-mentioned three kinds of information entropy and the three output units correspond to the unbalance, coupling misalignment fault and rubbing faults, respectively.

Similar with the above experiment, we test at three different rotation speeds, which are 1000 rpm, 3000 rpm, and 5000 rpm. We collected vibration fault signals at the three rotation speeds. For each rotation speed, we collected 100 groups of data for training and 20 groups of data for testing. Table 3 compares the results of the four methods.

Table 3. Results of four fault identification methods on real data of two-span rotors.

Rotation Speed	Singular Spectrum Entropy	Power Spectrum Entropy	Approximate Entropy	Feature Fusion Model
1000 rpm	67.74%	64.51%	66.13%	80.64%
3000 rpm	73.69%	81.03%	79.52%	92.17%
5000 rpm	82.85%	75.97%	85.21%	95.38%

From Table 3 we can see that, when the singular spectrum entropy, power spectrum entropy, and approximate entropy are used separately to classify the faults of two-span rotors, the classification accuracy is not satisfactory, and the fault characteristics cannot be better reflected. However, for the proposed feature fusion model method, the fault classification accuracy is significantly higher than that of any single feature fault classification methods. This proves the effectiveness of the proposed approach.

6. Conclusions

In order to cope with the difficulty of identifying complex faults for rotary machinery, this paper proposes a feature fusion model based on information entropy and a probability neural network. In order to obtain the final fault identification result, we implement the characteristic measurement of a vibration signal combining singular spectrum entropy, power spectrum entropy, and approximate entropy, and fuse them with a PNN model to classify and diagnose the fault signals. Finally, the fault detection accuracy of our proposed method is more than 10% higher compared with the methods using three kinds of information entropy separately in real two-span rotor data, and proved to be an effective fault diagnosis method.

Acknowledgments: The authors would like to thank all the reviews for their constructive comments. This work was supported by the National Natural Science Foundation of China (no. 61501451, 51775284), and the Natural Science Foundation of Jiangsu Province (BK20151199).

Author Contributions: Quansheng Jiang and Yehu Shen conceived and designed the experiments; Quansheng Jiang, Hua Li and Fengyu Xu analyzed the data; and Quansheng Jiang wrote the paper.

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

References

1. Lei, Y.; Lin, J.; He, Z.; Zi, Y. Application of an improved kurtogram method for fault diagnosis of rolling element bearings. *Mech. Syst. Signal Process.* **2011**, *25*, 1738–1749. [[CrossRef](#)]
2. 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](#)]

3. Liu, W.Y.; Tang, B.P.; Han, J.G.; Lu, X.N.; Hu, N.N.; He, Z.Z. The structure healthy condition monitoring and fault diagnosis methods in wind turbines: A review. *Renew. Sustain. Energy Rev.* **2015**, *44*, 466–472. [[CrossRef](#)]
4. Lu, C.; Wang, S.; Zhang, C. Fault diagnosis of hydraulic piston pumps based on a two-step EMD method and fuzzy C-means clustering. *Proc. IMechE Part C* **2016**, *230*, 2913–2928. [[CrossRef](#)]
5. Chen, J.; Lia, Z.; Pan, J.; Chen, G.; Zi, Y.; Yuan, J.; Chen, B.; He, Z. Wavelet transform based on inner product in fault diagnosis of rotating machinery: A review. *Mech. Syst. Signal Process.* **2016**, *70–71*, 1–35. [[CrossRef](#)]
6. He, Q. Time–Frequency manifold for nonlinear feature extraction in machinery fault diagnosis. *Mech. Syst. Signal Process.* **2013**, *35*, 200–218. [[CrossRef](#)]
7. Meng, Z.; Hu, M.; Gu, W.; Zhao, D. Rolling bearing fault diagnosis method based on LMD Multi-scale entropy and probabilistic neural network. *China Mech. Eng.* **2016**, *27*, 433–437.
8. Antoni, J.; Xin, G.; Hamzaoui, N. Fast computation of the spectral correlation. *Mech. Syst. Signal Process.* **2017**, *92*, 248–277. [[CrossRef](#)]
9. Wang, C.; Zhou, J.; Qin, H.; Li, C.; Zhang, Y. Fault diagnosis based on pulse coupled neural network and probability neural network. *Expert Syst. Appl.* **2011**, *38*, 14307–14313. [[CrossRef](#)]
10. Niu, X.; Zhao, H.; Chen, X. Target recognition in naval battle field based on information entropy and PNN. *Electron. Opt. Control* **2010**, *17*, 83–86.
11. Wang, Y.; Xu, G.; Liang, L.; Jiang, K. Detection of weak transient signals based on wavelet packet transform and manifold learning for rolling element bearing fault diagnosis. *Mech. Syst. Signal Process.* **2015**, *54–55*, 259–276. [[CrossRef](#)]
12. Liu, H.; Han, M. A fault diagnosis method based on local mean decomposition and multi-scale entropy for roller bearings. *Mech. Mach. Theory* **2014**, *75*, 67–78. [[CrossRef](#)]
13. Antoni, J. The infogram: Entropic evidence of the signature of repetitive transients. *Mech. Syst. Signal Process.* **2016**, *74*, 73–94. [[CrossRef](#)]
14. Gao, Q.; Jia, M. EEMD method based singular value spectral entropy in fault diagnosis of rotating machinery. *J. Southeast Univ.* **2011**, *41*, 998–1001.
15. Lei, Y.; Lin, J.; He, Z.; Kong, D. A method based on multi-sensor data fusion for fault detection of planetary gearboxes. *Sensors* **2012**, *12*, 2005–2017. [[CrossRef](#)] [[PubMed](#)]
16. Bai, B.; Bai, G.; Li, C. Application of process power spectrum entropy in rotor vibration quantitative diagnosis. *Aeroengine* **2015**, *41*, 27–31.
17. Xu, Y.; Li, L.; He, Z. Approximate entropy and its applications in mechanical fault diagnosis. *Inf. Control* **2002**, *31*, 547–551.
18. Yan, R.; Gao, R.X. Approximate entropy as a diagnostic tool for machine health monitoring. *Mech. Syst. Signal Process.* **2007**, *21*, 824–839. [[CrossRef](#)]
19. Gui, Z.; Han, F. Neural network based on wavelet packet-characteristic entropy for fault diagnosis of draft tube. *Proc. CSEE* **2005**, *4*, 101–104.
20. Sun, X.; Liu, H.; Zhao, X. Singular point recognition and feature extraction for incipient bearing fault based on instantaneous envelope scalogram entropy. *J. Chin. Mech. Eng.* **2017**, *53*, 73–80. [[CrossRef](#)]
21. Jing, L.; Wang, T.; Zhao, M.; Wang, P. An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox. *Sensors* **2017**, *17*, 414. [[CrossRef](#)] [[PubMed](#)]
22. Staszewski, W.J. Wavelet based compression and feature selection for vibration analysis. *J. Sound Vib.* **1998**, *211*, 735–760. [[CrossRef](#)]
23. Kahirdeh, A.; Khonsari, M. Acoustic entropy of the materials in the course of degradation. *Entropy* **2016**, *18*, 280. [[CrossRef](#)]
24. Cui, H.; Zhang, L.; Kang, R.; Lan, X. Research on fault diagnosis for reciprocating compressor valve using information entropy and SVM method. *J. Loss Prev. Process Ind.* **2009**, *22*, 864–867. [[CrossRef](#)]
25. Cabal-Yepez, E.; Romero-Troncoso, R. J.; Garcia-Perez, A.; Osornio-Rios, R.A. Single-parameter fault identification through information entropy analysis at the startup-transient current in induction motors. *Electric Power Syst. Res.* **2012**, *89*, 64–69. [[CrossRef](#)]
26. Chen, X.; Zhou, J.; Xiao, H.; Wang, E.; Xiao, J.; Zhang, H. Fault diagnosis based on comprehensive geometric characteristic and probability neural network. *Appl. Math. Comput.* **2014**, *230*, 542–554. [[CrossRef](#)]

27. Xu, J.; Wang, Z.; Tan, C.; Si, L.; Liu, X. A cutting pattern recognition method for shearers based on improved ensemble empirical mode decomposition and a probabilistic neural network. *Sensors* **2015**, *15*, 27721–27737. [[CrossRef](#)] [[PubMed](#)]
28. Saritha, M.; Joseph, K.P.; Mathew, A.T. Classification of MRI brain images using combined wavelet entropy based spider web plots and probabilistic neural network. *Pattern Recognit. Lett.* **2013**, *34*, 2151–2156. [[CrossRef](#)]
29. Zhang, W.; Peng, G.; Li, C.; Chen, Y.; Zhang, Z. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals. *Sensors* **2017**, *17*, 425. [[CrossRef](#)] [[PubMed](#)]
30. He, J.; Yang, S.; Gan, C. Unsupervised fault diagnosis of a gear transmission chain using a deep belief network. *Sensors* **2017**, *17*, 1564. [[CrossRef](#)] [[PubMed](#)]
31. Zhang, Q.S.; Jiang, S.Y. A note on information entropy measures for vague sets and its applications. *Inf. Sci.* **2008**, *178*, 4184–4191. [[CrossRef](#)]
32. Lu, Y.; Li, Y.; Xiao, H.; Wang, Z. A Sliding Singular Spectrum Entropy Method and Its Application to Gear Fault Diagnosis. *Int. Conf. Intell. Comput.* **2008**, *5226*, 669–678.
33. Kojima, M.; Obuchi, S.; Mizuno, K.; Henmi, O.; Ikeda, N. Power spectrum entropy of acceleration time-series during movement as an indicator of smoothness of movement. *J. Physiol. Anthropol.* **2008**, *27*, 193–200. [[CrossRef](#)] [[PubMed](#)]
34. Pincus, S.M. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2297–2301. [[CrossRef](#)] [[PubMed](#)]



© 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/>).