

Article

Correlated EEMD and Effective Feature Extraction for Both Periodic and Irregular Faults Diagnosis in Rotating Machinery

Jiejunyi Liang ^{1,2}, Jian-Hua Zhong ^{1,*}  and Zhi-Xin Yang ³ 

¹ School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350108, China; jiejunyiliang@gmail.com

² School of Electrical, Mechanical and Mechatronic Systems, University of Technology Sydney, Sydney, NSW 2007, Australia

³ Department of Electromechanical Engineering, Faculty of Science and Technology, University of Macau, Macau 999078, China; zxyang@umac.mo

* Correspondence: jhzhong@fzu.edu.cn

Received: 16 September 2017; Accepted: 14 October 2017; Published: 19 October 2017

Abstract: Intelligent fault diagnosis of complex machinery is crucial for industries to reduce the maintenance cost and to improve fault prediction performance. Acoustic signal is an ideal source for diagnosis because of its inherent characteristics in terms of being non-directional and insensitive to structural resonances. However, there are also two main drawbacks of acoustic signal, one of which is the low signal to noise ratio (SNR) caused by its high sensitivity and the other one is the low computational efficiency caused by the huge data size. These would decrease the performance of the fault diagnosis system. Therefore, it is significant to develop a proper feature extraction method to improve computational efficiency and performance in both periodic and irregular fault diagnosis. To enhance SNR of the acquired acoustic signal, the correlation coefficient (CC) method is employed to eliminate the redundant intrinsic mode functions (IMF), which comes from the decomposition procedure of pre-processing known as ensemble empirical mode decomposition (EEMD), because the higher the correlated coefficient of an IMF is, the more significant fault signatures it would contain, and the redundant IMF would compromise both the SNR and the computational cost performance. Singular value decomposition (SVD) and sample Entropy (SampEn) are subsequently used to extract the fault feature, by exploiting their sensitivities to irregular and periodic fault signals, respectively. In addition, the proposed feature extraction method using sparse Bayesian based pairwise coupled extreme learning machine (PC-SBELM) outperforms the existing pairwise-coupling probabilistic neural network (PC-PNN) and pairwise-coupling relevance vector machine (PC-RVM) by 1.8% and 2%, respectively, to achieve an accuracy of 93.9%. The experiments conducted on the periodic and irregular faults in the gears and bearings have demonstrated that the proposed hybrid fault diagnosis system is effective.

Keywords: acoustic signal; ensemble empirical mode decomposition; sparse Bayesian extreme learning machine; sample entropy; singular value decomposition

1. Introduction

Nowadays, many efforts have been devoted to revealing the hidden patterns in fault-related signals. Different possible signal sources, such as vibration signal and acoustic signal, are investigated to provide more accurate and efficient methods [1–3]. Although vibration signal is adopted in most fault diagnosis cases, acoustic signal is more desirable in some implementations due to its unique merits [2]. The first advantage for acoustic signal is the low hardware cost because of its independence

of direction. Only one acoustic sensor is required to achieve the same functionality as that of vibration signal. Second, when compared with vibration signal, acoustic signal is relatively immune to the underlying structural resonances. Third, it is easier for acoustic sensor to capture incipient fault information than the vibration sensor as it is more sensitive [4]. These advantages, in two aspects, fulfill the functional requirements of fault diagnosis. On the one hand, incipient faults would likely generate weak vibration energy, which could hardly cause any significant physical degradation within the solid structure, which makes it impossible for the vibration sensor to capture any useful pattern. On the other hand, acoustic sensor can cover a much more extensive range of frequencies which enables it to be embedded with much more extended types of fault signals [5,6]. However, the disadvantages accompanying the merits are the low signal to noise ratio (SNR) and low computational efficiency. To solve this problem, proposing a suitable signal processing approach for acoustic signal is essential.

To handle different implementation situations, many methods have been developed like fast Fourier transform (FFT) [7], wavelet transform [8], and empirical mode decomposition (EMD) [9,10]. FFT could generate the statistical average characteristics under the circumstance of steady operations, but it cannot deal with time and frequency domain problems. The wavelet transform, as a time-frequency analysis approach, is widely used in fault feature extraction [11], but it could only use certain frequency bands to deal with the acquired signal [12].

EMD is a method proposed recently for analyzing nonlinear and non-stationary signals. The main idea is to use the envelope lines and the mean value of a signal to decompose it into several intrinsic mode functions (IMFs) [13], and by analyzing each of the obtained IMFs, the patterns hiding in the original signal could be accurately and effectively extracted. Although the investigations into EMD have been reported in the literature in the field of fault diagnosis [14], the major disadvantages are the mode mixing issues and redundant IMFs, which degrades the performance of EMD in application. To solve the problems of mode mixing and redundant IMFs, the traditional EMD has been modified, namely ensemble empirical mode decomposition (EEMD) [15], to fix the mode mixing problem [16], and correlation coefficient technology is applied to eliminate the redundant IMFs. The detailed explanation will be introduced in Section 3.2.

Although the proposed EEMD helps to alleviate the problem of mode mixing, it still has to face the problem of the large dataset because fundamentally it decomposes one signal into ten or more IMFs. In general, the large size dataset could compromise the performance of the classifiers. To improve the overall performance, useful information should be further extracted during signal processing to make the decomposed IMFs more meaningful and more representative.

As the generated IMFs could perfectly form a matrix, singular value decomposition (SVD) becomes one of the suitable solutions because it could reduce the size of a matrix and reveal the hidden information in it [17]. As the singular value of a matrix is relatively stable and it could represent its natural characteristic, it is desirable to be used to detect periodic faults. Moreover, it is not necessary to manually choose the reconstruction parameters in SVD because the initial feature vector matrix can be generated by combining the IMFs directly [18]. When considering the periodic impulse could reveal related faults by singular values of the matrixes while applying SVD, the SVD technique based on EEMD is proposed. To further solve the high computational cost problem, sample entropy (SampEn) as a feature selection method can be employed to reduce the input size and dimension for the fault classifiers [19], as it could also select the fault features from irregular faults of each IMF generated by EMD method [20]. As the faults of rotating machinery can usually be classified into two types as periodic and irregular faults, this study proposed a proper feature selection method via the combination of SVD and SampEn to generate a feature matrix that could identify both irregular and periodic faults adequately.

Conventional pattern recognition methods always generate binary results, which could not meet the demand of modern implementations. As a result, probabilistic neural network (PNN) and relevance vector machine (RVM) [21,22] are proposed to serve as probabilistic classifiers. The prevalence of

PNN is hindered by its complexity and the input-related training time. For RVM, a Bayesian based support vector machine, as it needs to deal with Hessian matrix, which computes attains $O(N^2)$, it can not handle big data problems. Recently, a sparse Bayesian based extreme learning machine (SBELM) emerged as a more advanced classifier [23]. The Sparse Bayesian framework is adopted to modify the probabilistic classifier extreme learning machine (ELM) to reduce the training time and entitle it with the ability to effectively handle big data problems. To make the proposed approach more practical, a pairwise coupling strategy, named one versus one strategy (1vs1), is further incorporated, thus making it possible to deal with multi-classification problems [24]. The new classifier is named pairwise-coupled sparse Bayesian extreme learning machine (PC-SBELM). The main advantage of the pairwise coupling strategy is to include the correlation between different faults that make the probabilistic estimation more accurate. The details of how the proposed 1vs1 strategy works will be discussed in Section 2.2. Although a similar framework is proposed by the authors in [25], the new framework here adopts a more effective technique, correlation coefficient, to select the suitable number of IMF to reduce the redundant information and improve the SNR. Moreover, the number of periodic and irregular faults tested in the experiments is extended from four to six. Besides, to verify the effectiveness of the proposed classifier, pairwise-coupled relevance vector machine (PC-RVM) and pairwise-coupled probabilistic neural network (PC-PNN) are also adopted as comparisons.

The organization of this paper is as follows. The second section introduces the proposed framework and the methods adopted. Then, a test rig is set up and the corresponding signal pre-processing is discussed in Section 3. In the fourth section, experimental results are generated and compared with the aforesaid existing approaches. The last section summarizes the proposed system and conclusions are given.

2. Proposed Frameworks

Figure 1 shows the proposed periodic and irregular fault diagnosis system for rotating machinery. The proposed fault diagnosis system consists of two components, namely, (1) data processing and (2) pattern recognition, which will be thoroughly discussed in the following sub-sections.

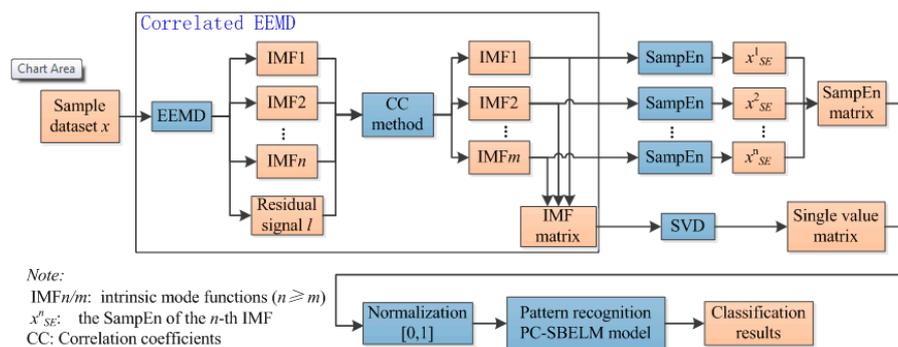


Figure 1. Proposed multiple faults diagnosis system.

In the phase of data processing, EEMD decomposes the acquired acoustic signal into a series of IMFs. To de-noise by reducing the redundant IMFs, correlation coefficient method is applied to select the useful IMFs. After the signal being de-noised, SVD and SampEn would generate a selected feature matrix that contains both irregular and periodic fault features. Finally, the proposed pattern recognition method, PC-SBELM, will be verified using the extracted features in training, validating and testing.

2.1. Data Processing

2.1.1. Ensemble Empirical Mode Decomposition

As conventional EMD suffers from mode mixing, EEMD is adopted in this study. It employs white noises of certain amplitudes to help separate different IMFs more effectively and calculates the ensemble mean of each decomposition to eliminate the white noise. The general rule for the added white noise is that small amplitude white noise should be added to high-frequency components and large amplitude white noise for low-frequency components. In [26], the recommended standard deviation range is from 0.1 to 0.4. The algorithm can be summarized as:

1. Set the ensemble number as E , the white noise (amplitude) and $e = 1$.
2. For e th trail, a white noise with corresponding amplitude will be added to the investigated signal.

$$x_e = x + n_e, \quad (1)$$

where n_e is the e th white noise, and x_e is the resultant signal.

3. Then x_e will be further decomposed by conventional EMD into I IMFs as $c_{i,e}(t)$, where I is the number of IMF and $c_{i,e}$ is the i th IMF of the e th trial.
4. Let $e = e + 1$ and repeat steps 2 and 3 with different white noise until $n = E$.
5. Find the ensemble mean to eliminate the white noise:

$$\bar{c}_i = \frac{1}{E} \sum_{e=1}^E c_{i,e}, \quad i = 1, 2, \dots, I, e = 1, 2, \dots, E. \quad (2)$$

7. Treat the means $\bar{c}_i (i = 1, 2, \dots, I)$ as the final results.

After the decomposition, the dimension of the original signal increased because of the generated IMFs, but the data points in each IMF is the same as the original signal. Depending on the sampling rate, the number of points of a signal can be very large, implying that a very large input vector needs to be constructed for the fault classifiers. Moreover, it is generally believed that high dimensional and large input vector can undesirably compromise the performance of the classifier. Therefore, it is necessary to extract useful features from the pre-processed signal to reduce the data dimension and scale.

2.1.2. Correlation Coefficient Based EEMD

In order to reduce the number of IMFs and improve the overall computational efficiency, the correlation coefficient of an IMF $I_i(t)$ to the original signal $x(t)$ is calculated as:

$$Coe_{x(t), I_i(t)} = \frac{\sum_{i=1}^M (x(t) - \bar{x})(I_i(t) - \bar{I}_i)}{\sqrt{\sum_{i=1}^M (x(t) - \bar{x})^2} \sqrt{\sum_{i=1}^M (I_i(t) - \bar{I}_i)^2}} \quad (3)$$

where the overline denotes the mean value of certain variables and M is the number of IMFs. The larger the coefficient is, the more significant fault information an IMF could contain. To improve the SNR, those IMFs with a small $Coe_{x(t), I_i(t)}$ value are treated as noise and are ignored. The correlation coefficient analysis transforms the raw signals into a set of fault signatures for the downstream feature extraction module, by means of the SVD and SampEn, respectively, as depicted in Section 3.2.

2.1.3. Sample Entropy Approach

SampEn could generate a family of statistical variables that measure the complexity of a signal and is designed to solve nonlinear problems [20]. To solve the bias problem of self-matching, SampEn takes data samples in time domain from a continuous process [27]. It can be calculated as follows:

$$\text{SampEn}(m, r, N) = \ln\left[\frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} C_r^m(i) - \frac{1}{N - (m + 1) + 1} \sum_{i=1}^{N-(m+1)+1} C_r^{m+1}(i)\right] \quad (4)$$

where N denotes the total number of sample points and m is the number of points that match each other, and $C_r^m(i)$ is the matching probability.

The advantages of SampEn are as follows [28]:

- (1) Only a short dataset is required to give a relatively robust estimation.
- (2) It is immune to strong and short disturbance.
- (3) Noise effect could be avoided if noise filter parameter r could be properly chosen.

In this study, the optimal parameters m and r of SampEn are selected from 1 to 8 and 0.1 to 0.8, respectively. The experimental result shows that the parameters, m should be set as 2 and r should be defined as 0.2. After pre-processing and feature extraction by SampEn, the SampEn of each chosen IMF will be generated and form a feature vector $x_{SE} = [\text{SampEn}_1, \text{SampEn}_2, \dots, \text{SampEn}_n]$.

2.1.4. Singular Value Decomposition Approach

SVD could be used to analyze multivariate data and is especially effective when dealing with low energy noisy data.

If X is a $N \times M$ real matrix, then there exists:

$$X = U\Lambda V^T \quad (5)$$

where U and V are orthogonal matrices, $U = [u_1, \dots, u_N] \in R^{N \times N}$, $U^T U = I$, $V = [v_1, \dots, v_M] \in R^{M \times M}$, $V^T V = I$; S is a diagonal matrix, $\Lambda \in R^{N \times M}$, $\Lambda = [\text{diag}\{\Lambda_1, \dots, \Lambda_p\} : 0]$; $p = \min(N, M)$; $\Lambda_1 \geq \dots \geq \Lambda_p \geq 0$. The values of Λ_i are the singular values of matrix X and the vectors u_i and v_i are respectively the i -th left and right singular vectors.

To form the initial feature vector matrices, EEMD is applied to the original signal and a matrix of IMFs is obtained. IMFs are sorted by frequencies from high to low and are orthogonal, which can be used as the initial matrix. As the IMFs could represent the natural oscillatory mode embedded in the signal, it can reveal the nature characteristics of the potential fault. Before applying SVD to the matrices formed by IMFs, N IMFs should be obtained by EEMD, which is denoted by n_1, n_2, \dots, n_N , and $n = \max(n_1, n_2, \dots, n_N)$. When SVD is conducted with the pre-processed initial matrix, the singular value $\Lambda_{A,j}$ can be represented as:

$$\Lambda_{A,j} = [\Lambda_{A,j}^1, \Lambda_{A,j}^2, \dots, \Lambda_{A,j}^n] \quad (6)$$

where $\Lambda_{A,j}^1 \geq \Lambda_{A,j}^2 \geq \dots \geq \Lambda_{A,j}^n$, $j = 1, 2, \dots, m$ denotes the total number of possible faults. Then, the generated $\Lambda_{A,j}$ could be sent to the classifier for further pattern recognition.

2.2. Pattern Recognition Based on Pairwise-Coupling Sparse Bayesian Extreme Learning Machine

To solve the problem embedded in conventional extreme learning machine, sparse Bayesian mechanism is proposed to calculate the output weights \mathbf{w} instead of only calculating the hidden layer output \mathbf{H} , in which $\mathbf{H} = [\mathbf{h}_1(\mathbf{x}_1), \dots, \mathbf{h}_1(\mathbf{x}_N)]^T \in R^{N \times (L+1)}$ an $\mathbf{h}_i(\mathbf{x}_i) = [1, g_1(\theta_1 \cdot \mathbf{x}_i + b_1), \dots, g_L(\theta_L \cdot \mathbf{x}_i + b_L)]$, where $i = 1, \dots, N$, θ is the weight factor between the input and hidden nodes, $g(\cdot)$ is used to depict the hidden layer and b serves as the hidden node's threshold. In a classification event where there are only two classes, a Bernoulli event $P(t|\mathbf{x})$ can be used to represent each of the training samples and the likelihood can be as follows:

$$P(\mathbf{t}|\mathbf{w}, \mathbf{h}) = \prod_{i=1}^N \sigma[y(\mathbf{h}_i; \mathbf{w})]^{t_i} \{1 - \sigma[y(\mathbf{h}_i; \mathbf{w})]\}^{1-t_i} \tag{7}$$

where $\sigma(\cdot)$ is sigmoid function $\sigma[y(\mathbf{h}; \mathbf{w})] = \frac{1}{1+e^{-y(\mathbf{h}; \mathbf{w})}}$, $y(\mathbf{h}; \mathbf{w}) = \mathbf{h}\mathbf{w}$, $\mathbf{t} = (t_1 \dots t_N)^T$, $t_i = \{0, 1\}$ and $\mathbf{w} = (w_0, w_1, \dots, w_L)^T$.

Traditionally, most of the existing classifiers are based on binary classification, which cannot adequately solve practical problems as real implementations always include multi-classification problems. Modifications should be made to the binary classifiers to entitle them with the probabilistic ability. The one-versus-all strategy is one of the solutions, which constructs a group of classifiers $I_{\text{class}} = [C_1, C_2, \dots, C_d]$ in a d -label classification problem. However, the classification accuracy is relatively poor because it employs only one threshold to separate the target group from the others which leaves large indecisive region, shown in Figure 2. As a result, the 1vs1 strategy is adopted here, whose detailed definition is explained in reference [25].

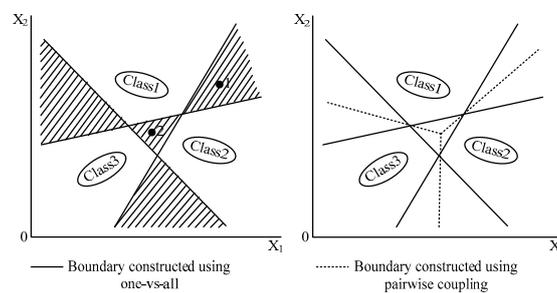


Figure 2. Comparisons between one versus all and one versus one strategy.

In a fault diagnosis process where there are d possible faults, I_{class} contains a set of pairwise coupling strategy C_{ij} as shown in Figure 3. Then, the extracted feature vector x will be processed and a probability vector $\rho = [\rho_1, \rho_2, \dots, \rho_d]$ could be generated, Here, $\rho_j = \Pr(j|x) \in [0,1]$ denotes the probability and x belongs to the j th label for $j = 1$ to d . Therefore, ρ_j is an independent probability and $\sum \rho_i \geq 1$. The following pairwise coupling strategy for multi-classification is proposed. The independent probability ρ_i is calculated as

$$\rho_i = C_i(\mathbf{x}) = \frac{\sum_{j=1:i \neq j}^d n_{ij} C_{ij}(\mathbf{x})}{\sum_{j=1:i \neq j}^d n_{ij}} = \frac{\sum_{j=1:i \neq j}^d n_{ij} \rho_{ij}}{\sum_{j=1:i \neq j}^d n_{ij}} \tag{8}$$

where n_{ij} is the number of feature vectors.

Hence, the probability can be more accurately estimated from $\rho_{ij} = C_{ij}(\mathbf{x})$ as the indecisive region between different thresholds in one versus all strategy is eliminated. For simplicity, the notation SBELM since then represented the proposed pairwise-coupled SBELM.

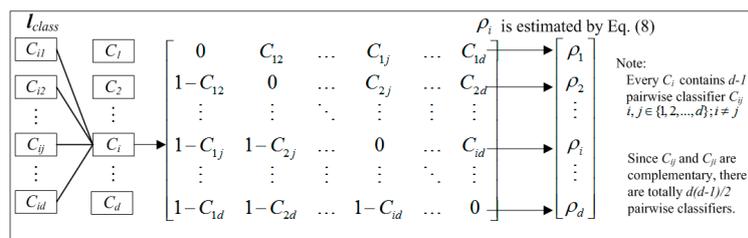


Figure 3. The proposed pairwise coupling strategy.

3. Experimental Setup and Data Analysis

To evaluate the performance of the proposed system, experiments were conducted and the details about the test are discussed in the first sub-section, which is followed by the corresponding results and comparisons. All of the experiments were executed on a PC with an i3-3217U/1.8 GHz/4 GB RAM and were implemented by using Matlab R2013a.

3.1. Experiment Setup and Data Acquisition

Figure 4 shows the test rig where the experiments were conducted, it can simulate common periodic faults and irregular faults of the gears and bearings in rotating machinery. The specifications of test gearbox are shown in Table 1. The required acoustic signals are collected by a microphone which is mounted on the gearbox and connected to a signal acquisition module (NI 9234). In order to make the experiments more representative, the faults were sampled under two conditions (no load and the load of 35 Nm). In this case study, the load system is offered by the resistance box that is connected to the generator of the test rig, and the torque of the generator can be adjusted by changing the resistance. A total of 7 cases, including one normal condition and six periodic and irregular faults are simulated. The details are listed in Table 2. Figure 5 shows the gear faults that include 1/4 chipped tooth, broken full tooth, and 5 mm gear crack; moreover, the rolling elements and the races (inner and outer) are manually worn, in which the diameter of medium wear is from 0.18 mm to 2.5 mm and the depth is 0.18 mm to 0.36 mm.

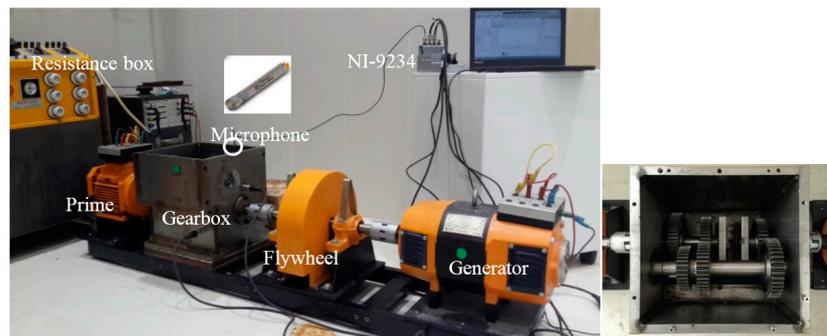


Figure 4. Established fault diagnosis test rig.

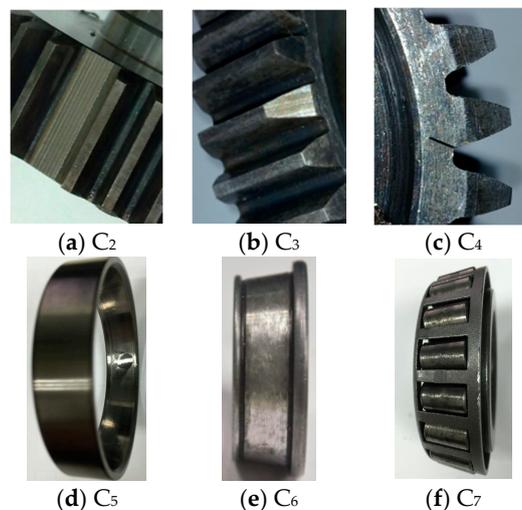


Figure 5. (a) Broken tooth; (b) Chipped tooth; (c) Gear crack; (d) Wear of outer race; (e) Wear of inner race; and, (f) Wear of the rolling elements.

Table 1. Specifications of test gearbox.

Type	Two-Speed Gearbox
First stage	Number of teeth: 52/44
Second stage	Number of teeth: 36/60
Gear ratio	First stage: 1.18 Second stage: 0.6
Tapered roller bearing	No. 32305/30206(SKF)
PRM range	0 to 2500 rpm variable speed
Total weight	50 kg
Oil capacity	About 2.5 L

Table 2. The manually created six faults.

	Case No.	Condition
Normal	C ₁	Normal
Periodic faults (gear)	C ₂	Broken tooth
	C ₃	Gear crack
	C ₄	Chipped tooth
Irregular faults (bearing)	C ₅	Wear of outer race
	C ₆	Wear of inner race
	C ₇	Wear of the rolling elements

In this study, each fault was kept running for several minutes, 400 datasets are selected randomly under 25.6 kHz sampling rate. For each simulated case, there were 2800 sample data (i.e., (normal condition, gear crack, chipped tooth, broken tooth, wear of inner & outer race of bearing, wear of the rolling elements) \times 400 samples). As required by the proposed system, the sample data was divided into validation, training, and test datasets, as shown in Table 3.

Table 3. Division of the acquired data.

	Type of Dataset	(2800)
Feature extraction (EEMD+SVD/SampEn)	Validation dataset	D_{Proc_Valid} (700)
	Training dataset	D_{Proc_Train} (1400)
	Test dataset	D_{Proc_Test} (700)

3.2. Feature Extraction

In this case study, the ensemble number of EEMD is set as 100 and the corresponding amplitude of the added white noise is 0.3 times of the standard deviation. Take the sound signal of fault C₃ as an example; it is decomposed into ten IMFs, in which the last one is the residual, as shown in Figure 6. It can be seen in the last diagram of Figure 7 that the correlation coefficients for IMF8 and IMF9 are much smaller than those of IMF1 to IMF7. By theory, a large value of correlation coefficient means more information. As a result, in order to improve the SNR and reduce the data scale, IMF8 and IMF9 are neglected. After taking many trials of correlation coefficients, it is found that for C₁ to C₇ datasets, the main features of the periodic and irregular faults in the gearbox are also contained mostly in the first seven IMFs. Hence, only the SVD/SampEn of IMF1 to IMF7 are retained to represent the signal features for C₁ to C₇, which is shown in Table 4.

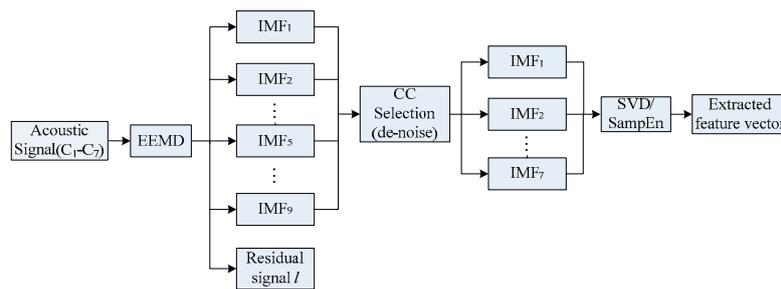


Figure 6. Procedures for performing feature extraction.

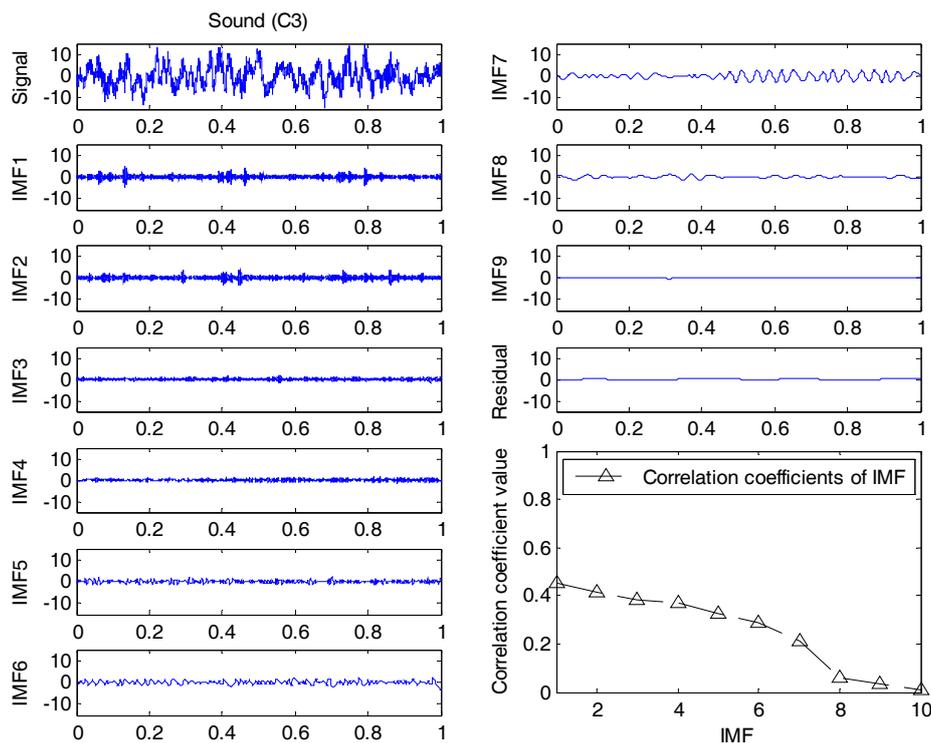


Figure 7. Example of ensemble empirical mode decomposition (EEMD) for decomposing sound signal of C_3 and correlation coefficient of each intrinsic mode functions (IMF).

Table 4. Number of extracted features in C_1 to C_6 signal.

	Raw Data	EEMD	SampEn	SVD	×Five Loading Condition	Total Features
Size of the data matrix	$1 \times 25,600$	$7 \times 25,600$	7	7	2	$(7 + 7) \times 2 = 28$

4. Results and Discussion

4.1. Results Analysis of EEMD Based SVD

By using PC-SBELM, every possible fault could get a corresponding fault probability, which would help to identify the faults and find out the potential fault when the system could not make the right decision. It is worth noting that in each PC-SBELM, there are totally $7 \times (7 - 1)/2 = 21$ SBELM fault classifiers trained as there are totally seven types of gearbox faults in this application. The initial number of hidden nodes in each PC-SBELM was set to 20.

Singular value decomposition is a factorization of a real or complex matrix; it can extract the values that can represent the main characteristic of the signal. The initial feature vector matrices

could be formed automatically by the IMFs components. The singular values of the matrix gained by utilizing SVD to the original feature vector matrices could be used as the fault feature vectors for the classifier. Figure 8 presents the results based on SVD alone in which the accuracy of periodic faults is almost 6% higher than irregular faults. It can be easily seen that SVD could adequately solve the problem which SampEn encounters. To verify the explanation, Table 5 shows the results of 11 sample cases of wear of outer race which are improperly classified using SVD alone.

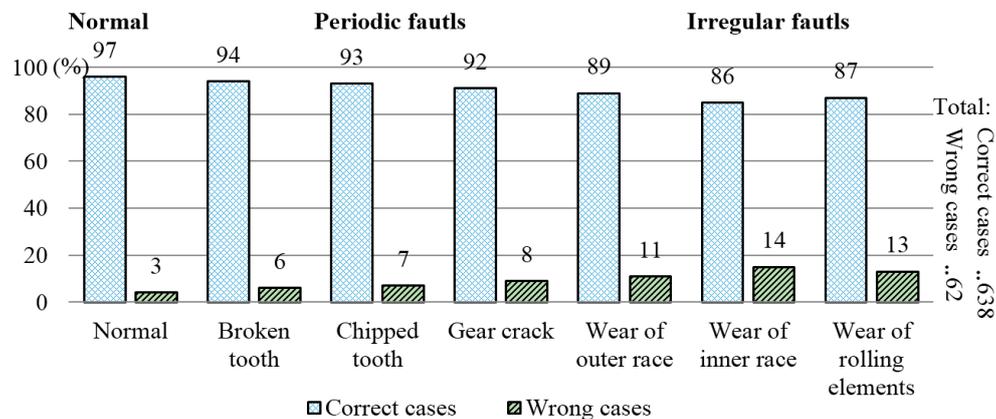


Figure 8. The experiment outcome based on singular value decomposition (SVD). Remark: Pattern recognition is based on the proposed method pairwise coupled extreme learning machine (PC-SBELM).

Table 5. Results of 11 improperly classified cases using SVD.

	Periodic Faults			Irregular Faults		
	Broken Tooth	Chipped Tooth	Gear Crack	Wear of Outer Race	Wear of Inner Race	Wear of Rolling Elements
Case 1	0.1892	0.2438	0.2634	0.5548	0.7142	0.6794
Case 2	0.1345	0.2765	0.2890	0.5074	0.7089	0.6246
Case 3	0.1238	0.3489	0.3720	0.5523	0.6927	0.6746
Case 4	0.1228	0.2850	0.2672	0.4850	0.6852	0.7256
Case 5	0.1254	0.6365	0.5779	0.4560	0.7257	0.4022
Case 6	0.2617	0.6156	0.5643	0.6072	0.6932	0.3291
Case 7	0.1243	0.3173	0.3412	0.4720	0.5636	0.7270
Case 8	0.1470	0.3010	0.2863	0.3172	0.6558	0.6940
Case 9	0.1244	0.2349	0.5335	0.4125	0.4279	0.4562
Case 10	0.3160	0.5177	0.3878	0.3742	0.4640	0.3705
Case 11	0.2579	0.6043	0.6150	0.4781	0.6898	0.3074

Table 5 also shows that, except for case 9 and case 10, the other nine wrongly cases are classified into the wear of inner race or rolling elements fault, which complies with that the SVD of irregular faults, such as wear of races (inner and outer) and that of rolling elements are similar, which are hardly able to be separated. By using PC-SBELM, it is adequately illustrated that SVD performs better on periodic faults rather than irregular faults which could often be misclassified.

4.2. Results Analysis of EEMD Based SampEn

To verify that the SampEn is sensitive to the irregular faults in rotating machinery, the values of SampEn for each IMF are shown in Table 6. As can be seen in Table 6, the wear of the rolling elements has the largest value of SampEn. It accords with the reality that the wear of rolling elements' acoustic signal is the most irregular because of the changing of the contacting area between rolling elements and races (inner and outer). What is more, the rolling elements will rotate while rolling with the rage. Moreover, the difference of SampEn between irregular faults is more obvious than that of

the periodic faults, which contributes to improving the performance of classifier when classifying these irregular faults. The experimental results in Figure 9 prove that the classification accuracy of irregular faults is almost 7% higher than periodic faults. The main reason is that although the impulses of different tooth faults are different, the regularities of them are similar as they occur at the same periodical frequency, which makes the difference between the signals too small to be identified correctly. To verify the explanation, Table 7 shows the results of 12 wrongly classified cases of chipped tooth using SampEn alone.

Table 6. The values of SampEn of each IMF of each fault type.

SampEn of Each IMF	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7
Broken tooth	0.06	0.42	0.46	0.25	0.12	0.04	0.02
Chipped tooth	0.07	0.43	0.47	0.25	0.13	0.05	0.03
Gear crack	0.07	0.40	0.44	0.24	0.10	0.04	0.02
Wear of inner race	0.12	0.45	0.48	0.30	0.15	0.08	0.02
Wear of outer race	0.16	0.51	0.52	0.33	0.19	0.11	0.02
Wear of rolling elements	0.19	0.50	0.57	0.38	0.23	0.15	0.03

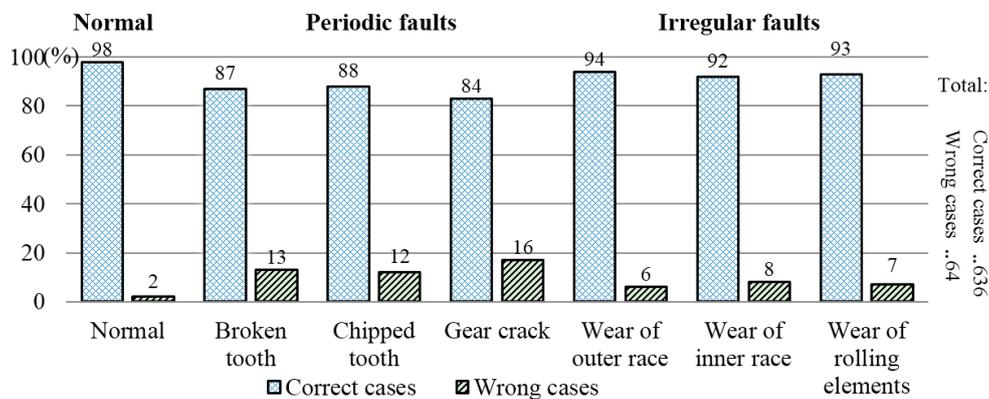


Figure 9. The experiment outcome based on SampEn. Remark: Pattern recognition is based on the proposed method PC-SBELM.

Table 7. Results of 12 wrongly classified cases of chipped tooth using SampEn.

	Periodic Faults			Irregular Faults		
	Broken Tooth	Chipped Tooth	Gear Crack	Wear of Outer Race	Wear of Inner Race	Wear of Rolling Elements
Case 1	0.7160	0.5036	0.4835	0.2835	0.1852	0.1282
Case 2	0.7234	0.4350	0.5145	0.3436	0.2457	0.1086
Case 3	0.2413	0.4647	0.4265	0.6313	0.3256	0.1288
Case 4	0.7056	0.5103	0.5163	0.2487	0.2684	0.0943
Case 5	0.6916	0.5456	0.5467	0.2607	0.2324	0.0889
Case 6	0.6341	0.4879	0.6878	0.3898	0.3456	0.1078
Case 7	0.7086	0.5263	0.7266	0.2854	0.2545	0.1356
Case 8	0.7456	0.4945	0.4834	0.3172	0.2785	0.1873
Case 9	0.2234	0.5758	0.6764	0.4032	0.4122	0.0945
Case 10	0.5568	0.5367	0.6565	0.3634	0.3653	0.1116
Case 11	0.6032	0.5236	0.4831	0.4681	0.4455	0.1240
Case 12	0.3022	0.4780	0.4556	0.6050	0.6785	0.1323

Table 7 shows that, except for case 3 and case 12, other 10 wrongly cases are classified into the broken tooth and gear crack, which complies with that the SampEn of tooth faults are almost

the same, which could hardly be divided. Meanwhile, it can be found that the wear of the rolling elements and the chipped tooth have the smallest odds to be misclassified since the SampEn of them are obviously different.

4.3. Results and Discussion of PC-SBELM Classifier

For the purpose of demonstrating the effectiveness of the proposed approach, based on the pairwise-coupling strategy methods, PC-PNN and PC-RVM, are compared with the proposed framework. A fivefold cross-validation method is used to process the experimental results. The value of spread s , a hyperparameter of PC-PNN, is set to 0.65 which is examined from 1 to 3, at an interval of 0.5, and width w , a hyperparameter of PC-RVM, is defined to 0.75, which is selected from 1 to 8 at an interval of 0.5.

The experimental results in Table 8 show that PC-SBELM classifier based on SVD and SampEn both achieve the best accuracy of 93.9%, which outperforms the methods based on SVD or SampEn alone by about 2.9% and 3.3%, respectively, under six common periodic and irregular faults in the gear and bearing. It proves that combining the SVD and SampEn method can overcome the weaknesses in each of them, which is conducive to improving the performance of the classifier. Moreover, the proposed framework, PC-SBELM based on SVD and SampEn, is better than traditional probabilistic classifier PC-PNN and PC-RVM outperforming by 1.8% and 2%, respectively. In summary, the proposed framework is a powerful tool to fix the aforesaid problems during diagnosing fault in rotating machinery.

Table 8. Evaluation results of proposed PC-SBELM with pairwise-coupled probabilistic neural network (PC-PNN) and pairwise-coupled relevance vector machine (PC-RVM)

Classification Techniques	Condition Indicators	Normal	Accuracies for Test Cases (%)						Average Accuracy
			Broken Tooth	Chipped Tooth	Gear Crack	Wear of Outer Race	Wear of Inner Race	Wear of Rolling Elements	
PC-PNN	SVD&SampEn	96 ± 0.1	91 ± 0.2	92 ± 0.2	91 ± 0.2	92 ± 0.3	91 ± 0.3	92 ± 0.2	92.1 ± 0.2
PC-RVM	SVD&SampEn	97 ± 0.0	90 ± 0.1	91 ± 0.2	91 ± 0.2	91 ± 0.2	92 ± 0.3	91 ± 0.2	91.9 ± 0.2
PC-SBELM	SVD	98 ± 0.0	94 ± 0.1	93 ± 0.0	91 ± 0.1	89 ± 0.0	85 ± 0.0	87 ± 0.0	91.0 ± 0.0
	SampEn	97 ± 0.0	87 ± 0.1	88 ± 0.1	83 ± 0.1	94 ± 0.1	92 ± 0.1	93 ± 0.1	90.6 ± 0.1
	SVD&SampEn	98 ± 0.1	93 ± 0.2	93 ± 0.2	92 ± 0.3	94 ± 0.3	93 ± 0.3	94 ± 0.2	93.9 ± 0.2

5. Conclusions

In modern industry, accuracy and fast fault diagnosis are difficult due to different representation forms, simultaneously occurring, and noisy working environments, for signal acquiring. We addressed these problems by first proposing to adopt the acoustic signals to the periodic and irregular fault diagnosis, for its inherent characteristics in terms of non-directional and insensitive to structural resonances. In order to solve low SNR and high computational difficulties of acoustic signals, our second effort is to propose an effective signal processing method. The correlation coefficient (CC) analysis is applied to select the suitable IMFs, which improves the SNR and enhances the effectiveness of signal decomposition. Moreover, a feature selection method via the combination of SVD and SampEn are used to build the feature extraction system, which could compress the dimension of the data to improve the computational efficiency in terms of periodic and irregular signal processing. Figures 8 and 9 show that the combined method EEMD + CC + SVD and EEMD + CC + SampEn are only effective in periodic and irregular fault diagnosis, respectively. In this case study, to verify the effectiveness of the proposed framework, the PC-PNN and PC-RVM are adopted as comparison sets. The results from experiment indicate that the proposed feature selection (SVD + SampEn) method combining with PC-SBELM as classifier outperforms the PC-PNN based and PC-RVM based methods by 1.8% and 2%, respectively, and achieves the best accuracy of 93.9%. Therefore, the proposed framework is a potential way to detect periodic and irregular faults precisely for rotating machinery,

which is successfully demonstrated in the experiments on the test rig gearbox system. Although the proposed method has been evaluated under experimental environment, as there are much more unforeseeable and uncontrollable factors in real industrial settings, its effectiveness still needs to be investigated in industrial ambience.

Acknowledgments: The authors would like to thank the funding support by the Fuzhou University, Grant number: (GXRC-17029-510499); and University of Macau, Grant numbers: MYRG2016-00160-FST, and MYRG2015-00077-FST.

Author Contributions: Jiejunyi Liang and Jian-Hua Zhong conceived and designed the experiments; Zhi-Xin Yang performed the experiments; Jiejunyi Liang and Jian-Hua Zhong analyzed the data; Zhi-Xin Yang contributed reagents/materials/analysis tools; Jiejunyi Liang and Jian-Hua Zhong wrote the paper.

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

References

1. Jiang, W.; Hu, W.; Xie, C. A New Engine Fault Diagnosis Method Based on Multi-Sensor Data Fusion. *Appl. Sci.* **2017**, *7*, 280. [[CrossRef](#)]
2. Tan, C.K.; Irving, P.; Mba, D. A comparative experimental study on the diagnostic and prognostic capabilities of acoustics emission, vibration and spectrometric oil analysis for spur gears. *Mech. Syst. Signal Process.* **2007**, *21*, 208–233. [[CrossRef](#)]
3. Amar, M.; Gondal, I.; Wilson, C. Vibration Spectrum Imaging: A Novel Bearing Fault Classification Approach. *IEEE Trans. Ind. Electron.* **2015**, *62*, 494–502. [[CrossRef](#)]
4. Tandon, N.; Mata, S. Detection of defects in gears by acoustic emission measurements. *J. Acoust. Emiss.* **1999**, *17*, 23–27.
5. Loutas, T.; Roulias, D.; Pauly, E.; Kostopoulos, V. The combined use of vibration, acoustic emission and oil debris on-line monitoring towards a more effective condition monitoring of rotating machinery. *Mech. Syst. Signal Process.* **2011**, *25*, 1339–1352. [[CrossRef](#)]
6. Gómez Muñoz, C.; García Márquez, F. A New Fault Location Approach for Acoustic Emission Techniques in Wind Turbines. *Energies* **2016**, *9*, 40. [[CrossRef](#)]
7. Yang, T.; Pen, H.; Wang, Z.; Chang, C.S. Feature Knowledge Based Fault Detection of Induction Motors Through the Analysis of Stator Current Data. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 549–558. [[CrossRef](#)]
8. Jong-Hyo, A.; Dae-Ho, K.; Bong-Hwan, K. Fault Detection of a Roller-Bearing System through the EMD of a Wavelet Denoised Signal. *Sensors* **2014**, *14*, 15022–15038.
9. El Hachemi Benbouzid, M. A review of induction motors signature analysis as a medium for faults detection. *IEEE Trans. Ind. Electron.* **2000**, *47*, 984–993. [[CrossRef](#)]
10. Zou, J.; Chen, J. A comparative study on time-frequency feature of cracked rotor by Wigner-Ville distribution and wavelet transform. *J. Sound Vib.* **2004**, *276*, 1–11. [[CrossRef](#)]
11. Meltzer, G.; Dien, N.P. Fault diagnosis in gears operating under non-stationary rotational speed using polar wavelet amplitude maps. *Mech. Syst. Signal Process.* **2004**, *18*, 985–992. [[CrossRef](#)]
12. Vincent, H.; Hu, S.; Hou, Z. Damage detection using empirical mode decomposition method and a comparison with wavelet analysis. In Proceedings of the Second International Workshop on Structural Health Monitoring, Stanford, CA, USA, 8–10 September 1999; pp. 891–900.
13. Tsakalozos, N.; Drakakis, K.; Rickard, S. A formal study of the nonlinearity and consistency of the empirical mode decomposition. *Signal Process.* **2012**, *92*, 1961–1969. [[CrossRef](#)]
14. Cai, J.H. Fault diagnosis of rolling bearing based on empirical mode decomposition and higher order statistics. *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2015**, *229*, 1630–1638. [[CrossRef](#)]
15. Chen, H.; Chen, P.; Chen, W.; Wu, C.; Li, J.; Wu, J. Wind Turbine Gearbox Fault Diagnosis Based on Improved EEMD and Hilbert Square Demodulation. *Appl. Sci.* **2017**, *7*, 128. [[CrossRef](#)]
16. Lei, Y.; He, Z.; Zi, Y. Application of the EEMD method to rotor fault diagnosis of rotating machinery. *Mech. Syst. Signal Process.* **2009**, *23*, 1327–1338. [[CrossRef](#)]
17. De Lathauwer, L.; De Moor, B.; Vandewalle, J. A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.* **2000**, *21*, 1253–1278. [[CrossRef](#)]
18. Wang, W.; Lin, R. The application of pseudo-phase portrait in machine condition monitoring. *J. Sound Vib.* **2003**, *259*, 1–16. [[CrossRef](#)]

19. Song, Y.; Starkov, P.A. A new approach for epileptic seizure detection: Sample entropy based feature extraction and extreme learning machine. *J. Biomed. Sci. Eng.* **2010**, *3*, 556–567. [[CrossRef](#)]
20. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.-Heart Circ. Physiol.* **2000**, *278*, H2039–H2049. [[PubMed](#)]
21. Refregier, P.; Vallet, F. Probabilistic approach for multiclass classification with neural network. *Artif. Neural Netw.* **1991**, *10*, 1003–1006.
22. Ciarelli, P.M.; Oliveira, E.; Badue, C.; Souza, A.F.D. Multi-Label Text Categorization Using a Probabilistic Neural Network. *Int. J. Comput. Inf. Syst. Ind. Manag. Appl.* **2009**, *1*, 133–144.
23. Widodo, A.; Kim, E.Y.; Son, J.-D.; Yang, B.-S.; Tan, A.C.; Gu, D.-S. Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine. *Expert Syst. Appl.* **2009**, *36*, 7252–7261. [[CrossRef](#)]
24. Abe, S. *Support Vector Machines for Pattern Classification*; Springer: Longdon, UK, 2010.
25. Yang, Z.-X.; Zhong, J.-H. A Hybrid EEMD-Based SampEn and SVD for Acoustic Signal Processing and Fault Diagnosis. *Entropy* **2016**, *18*, 112. [[CrossRef](#)]
26. Wu, Z.; Huang, N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **2009**, *1*. [[CrossRef](#)]
27. Rafiee, J.; Arvani, F.; Harifi, A.; Sadeghi, M. Intelligent condition monitoring of a gearbox using artificial neural network. *Mech. Syst. Signal Process.* **2007**, *21*, 1746–1754. [[CrossRef](#)]
28. Fusheng, Y.; Bo, H.; Qingyu, T. Approximate entropy and its application to biosignal analysis. *Nonlinear Biomed. Signal Proc. Dyn. Anal. Model.* **2001**, *2*, 72–91.



© 2017 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/>).