Roller Bearing Fault Diagnosis Based on Empirical Mode Decomposition and Targeting Feature Selection

Xiaoyue chen¹²*, Dang Ge¹, Xiong Liu¹, and Mengchao Liu¹

¹ School of Electrical Engineering and Automation, East China Jiaotong University, Nanchang 330013, China
² Huazhong University of Science and Technology, Wuhan, Hubei 430074, China

*976932632@qq.com; wdmxdefeng@126.com; 997921925@qq.com; 601075460@qq.com.

Abstract. At present, the feature extraction of frequency signal based on empirical mode decomposition (EMD) has been widely studied and applied in fault diagnosis of rolling bearings. However, there are still some shortcomings in fault diagnosis based on EMD. Therefore, a fault diagnosis method based on the combination of EMD and target feature selection (TFS) is proposed in this paper. The method firstly analyzes the fault signal through EMD and extracts the fault features. Then, it removes the redundant features and optimizes the feature subsets by using TFS. TFS selects the most effective feature for each target sample space through filtering evaluation criteria and heuristic search strategy, thereby effectively improving the accuracy and efficiency of fault diagnosis.

1. Introduction
Rolling bearing is a key component in rotating machinery such as industrial machinery, spaceflight and navigation. It needs high reliability and safe operation. However, due to heavy workload in a long and bad environment, mechanical fault often occur, affecting industrial production, increasing cost and maintenance time. Therefore, the safety and reliability of rolling bearings are particularly important for mechanical equipment [1]. The timely diagnosis of rolling bearing fault is the key to ensure safety and reliability, so fault diagnosis has great research value and significance [2].

In the past few years, a variety of signal processing methods have been proposed to detect and diagnose bearing defects. These methods include vibration and acoustic measurements, temperature measurement and wear particle analysis [1,5], in which bearing vibration signals are still the most direct, simplest and most effective information for the analysis of bearing defects [2,5]. Currently, non-linear, non-Guess, and non-stationary processing methods in vibration signals include short-time Fourier transform, Local feature-scale decomposition (LCD), Wavelet transform(WT), etc. These methods are widely used in feature extraction of fault signals of rotating machinery [3-5]. Compared with these methods, the EMD proposed by Hilbert-Huang can accurately give the combination of the frequency and time distribution of the signal energy. The local features of the signal time and frequency domain are accurately expressed by the instantaneous frequency, so this method is more effective in extracting the fault feature signal [3,6]. Based on EMD, some signal processing methods such as Ensemble Empirical Mode Decomposition (EEMD) [7] and Generalized Empirical Mode Decomposition (GEMD) [8] are proposed.
After extracting bearing signal features, there are still invalid features that affect the accuracy of feature subset characterization. In order to solve this problem, effective feature optimization and selection are usually used. At present, the feature selection method can be divided into three categories according to the feature evaluation criteria: The Filter, the Wrapper, and the Embedded [9]. Based on these, using Dynamic Relevance and Joint Mutual Information Maximization (DRJMMI), the redundant features are combined with dynamic weights to reduce the possibility of selection of real redundant features [10]; Based on Affinity Propagation (AP) clustering, RReliefF and Sequential Forward Search, the ARSFS is proposed to select the effective subset of the original feature set, reducing the feature space dimension and multiple correlation [11]. In addition, Ben Ali proposes a feature selection method combining Linear Discriminant Analysis (LDA) with Principal Component Analysis (PCA) [5], and successfully applied to fault diagnosis. These feature selection methods have achieved good results in eliminating invalid features, but there are still some shortcomings: 1) The combination of clustering and search is generally adopted, which results in the complexity of feature optimization. 2) For a variety of fault categories, effective feature selection is more difficult. Therefore, in order to solve this problem, this paper proposes a target feature selection method based on the empirical mode decomposition. The method is based on the transformation of multi-classifications into two classifications. After the fault defect is analyzed by EMD, finding the optimal feature between two kinds of fault categories, and the nonoptimal feature is eliminated. Through experiments, it can be seen that this method not only has a good diagnostic effect on various fault categories, but also has a simple feature optimization process.

In the process of fault diagnosis, an accurate classifier is very important. The accuracy of classifier recognition is determined by the quantity and quality of the input features. Therefore, the classifier should be selected according to the characteristics of the input features. Although the KNN classifier is simple in calculation, it has poor ability to reject invalid features [12-15]. Although the SVM classifier has an inhibitory effect on the feature of invalid redundancy, the structure is too complex and time-consuming, and it is not suitable for a large number of feature recognition [12-15]. Considering comprehensively, a Probabilistic Neural Network (PNN) with a simple structure and stable classification effect is used as a fault classifier [5-16].

The main work of this paper is as follows: 1) extract the vibration signal features of rolling bearing based on EMD. 2) Selecting effective feature sets from all feature sets based on TFS. 3) Using PNN to complete the diagnosis of rolling bearing.

2. Target Feature Selection (TFS)

Through the in-depth analysis of 9 types of faults of the rolling bearing, as long as we can accurately distinguish any two kinds of faults, we can accurately achieve accurate diagnosis of multiple faults. As shown in figure 1, identifying four different color samples of red, yellow, black and green, as long as any two different color samples are distinguished from each other, that is, to complete the six different color-resolving tasks in the figure 1, four kinds of color samples are effectively distinguished. Therefore, the multi-class identification problem can be transformed into multiple two-class identification problems. The two-classification problem is simple, and it can quickly lock the optimal features for the classification and remove the invalid features as much as possible. If each of the two classifications can find the optimal effective features, the feature subsets of these features can be very well identification all faults. According to the above analysis, a feature selection method (TFS) is proposed to ensure that feature subset have excellent distinguishing power for any two kinds of faults.

As a new feature selection strategy, the target feature selection method has the following characteristics: 1) A heuristic search strategy [16]. 2) Feature subset identification performance complements each other, which ensures the global identification ability of feature subsets. 3) Feature identification capability guides feature selection. Feature identification capability evaluation uses a filter-based evaluation criterion [16-18]. 4) High computational efficiency, small number of feature subsets, and good identification.
The TFS structure is shown in Figure 2. Target setting: decomposing the $N$ item classification into the $C^2_n$ item two classification task, each of the two classification tasks is called a Target Group (TG). Target feature selection: For each TG, a filter-based feature evaluation criterion is used to find the most effective feature to distinguish the target, and the target's optimal feature is called the Target Feature (TF). Target feature fusion: With a reasonable fusion of the target features of the entire target group, a new subset of features can be obtained, which is an optimized feature subset of the $N$ item classification problem.

Step1: First of all, according to the target structure, for all the fault samples, two different fault samples constitute a target sample space.

Step2: In each target sample space, according to the feature evaluation criteria, all features are evaluated and selected to get the best identification feature of the target, that is, the target feature of the target.

Step3: Each target sample space is performed by step 2, and the target feature of each target are obtained, and then the target features are reasonably fused to form the optimized feature subset.

3. Evaluation criteria

Figure 1. Identification process of multiple color samples

Figure 2. Target feature selection structure
In pattern characterization and pattern identification, feature optimal selection is advantageous to pattern characterization and identification, and feature selection depends on feature evaluation. The feature evaluation criterion proposed in this paper mainly considers the dispersion between categories and the compactness within categories [16-18]. However, for feature dispersiveness between fault categories, not all features have good dispersion among the fault categories, and there are also some features of overlapped overlaps between the faults cases. In this case, the fault identification ability of the feature is determined by the number of samples in the mixed region of the fault feature range. The feature evaluation criterion is determined by the degree of fault sample aliasing.

3.1 Single target feature evaluation criteria 1

The feature evaluation criteria are mainly determined by the feature dispersion between the fault categories and the compactness of the features within the category. This paper uses the difference method to describe the compactness of features within a fault category, as Eq1. $d^i_m$ is the compactness of the feature $i$ in the category of fault $m$; $f^i_y_m$ and $f^i_x_m$ are the maximum and minimum of the feature $i$ in the fault category $m$.

$$d^i_m = f^i_y_m - f^i_x_m$$

(1)

The dispersion of features between fault categories is another important factor in the feature evaluation criteria. The boundary difference method is used to measure the dispersion of the feature between the fault categories, as equation (2). $H^i_{m,n}$ is the dispersion of the feature $i$ between fault category $m$ and fault category $n$.

Combining the compactness and dispersion of feature, the average Euler distance is used to evaluate the feature’s fault identification ability [14-21]. The larger the average Euler distance of the feature, the stronger the fault identification capability, as equation (3). $D^i_{m,n}$ is the average Euler distance of the feature $i$ between fault category $m$ and fault category $n$.

$$H^i_{m,n} = \begin{cases} f^i_x_n - f^i_y_n, & f^i_x_n \geq f^i_y_n \\ f^i_x_m - f^i_y_m, & f^i_x_m \geq f^i_y_m \end{cases}$$

(2)

$$D^i_{m,n} = \frac{2 \times H^i_{m,n}}{d^i_m + d^i_n}$$

(3)

3.2 Single target feature evaluation criteria 2

For the cases where some of features ranges mixed overlap between the faults, the fault identification ability of the feature is determined by the number of samples in the mixed region of the fault feature range. The feature evaluation criterion is determined by the degree of fault sample aliasing, as equation (4). The greater the degree of aliasing of features, the weaker the fault identification capability.

$$P^i_{m,n} = \frac{N^i_m + N^i_n}{2 \times N^i}$$

(4)

$P^i_{m,n}$ is the degree of aliasing of features $i$ between fault category $m$ and fault category $n$; $N^i$ is the total number of each category of fault samples; $N^i_m$ is the number of the fault category sample $m$ belongs to the aliased region.

4. Applications in fault diagnosis

The experiment and engineering application based on the above theory are introduced in this section.
The experiment original data includes 9 fault categories, each with 118 samples. The application experiments are conducted according to figure 3. According to TFS, \((x_{1}, x_{2}, x_{6}, x_{12}, x_{19}, x_{32}, x_{42}, x_{53}, x_{63}, x_{76}, x_{79}, x_{83}, x_{85}, x_{90})\) is selected as the feature set to identify the faults. Fault identification ability value of target feature are showed in table 1. The comparison of target features with other features in same TG are listed in figure 4. The results of 9 faults diagnosis are showed in table 2, and the comparison of different methods in table 3. These experiments are carried out under the same condition with the same data.

In the same target sample space, compared with other features, the target feature has good distinctiveness and can classify an unknown sample by this distinction., as Figure 4 shown. This confirms that the feature set selected by TFS is much more prominent in fault diagnosis. It can be seen from table 3 that this method not only guarantees higher accuracy of fault diagnosis, but also has certain advantages for multi-category fault diagnosis, compared to several other feature optimization and selection methods.

![Flowchart](image)

**Figure 3.** the flow of experiment

![Comparison of features](image)

**Figure 4.** the comparison of features

| Target group sample | Target feature | Average Euler distance | Aliasing degree | Target group sample | Target feature | Average Euler distance | Aliasing degree |
|---------------------|----------------|------------------------|-----------------|---------------------|----------------|------------------------|-----------------|
| F1, F2              | \(x_{6}\)       | 3.581                  | --              | F3, F7              | \(x_{96}\)     | 0.049                  | --              |
| F1, F3              | \(x_{76}\)      | 0.709                  | --              | F3, F8              | \(x_{1}\)      | --                     | 0.093           |
| F1, F4              | \(x_{83}\)      | 7.703                  | --              | F3, F9              | \(x_{76}\)     | 1.059                  | --              |
| F1, F5              | \(x_{2}\)       | 4.504                  | --              | F4, F5              | \(x_{19}\)     | 2.081                  | --              |
| F1, F6              | \(x_{19}\)      | 1.550                  | --              | F4, F6              | \(x_{32}\)     | 0.955                  | --              |
| F1, F7              | \(x_{42}\)      | 0.247                  | --              | F4, F5              | \(x_{63}\)     | 1.183                  | --              |
| F1, F8              | \(x_{19}\)      | 0.715                  | --              | F4, F8              | \(x_{53}\)     | 3.475                  | --              |
| F1, F9              | \(x_{1}\)       | --                     | 0.076           | F4, F9              | \(x_{53}\)     | 2.632                  | --              |
| F2, F3              | \(x_{95}\)      | 1.149                  | --              | F5, F6              | \(x_{19}\)     | 1.463                  | --              |
| F2, F4              | \(x_{12}\)      | 7.031                  | --              | F5, F7              | \(x_{70}\)     | 0.791                  | --              |
| F2, F5              | \(x_{12}\)      | 3.891                  | --              | F5, F8              | \(x_{3}\)      | 1.609                  | --              |
| F2, F6              | \(x_{12}\)      | 0.869                  | --              | F5, F9              | \(x_{32}\)     | 1.997                  | --              |
| F2, F7              | \(x_{83}\)      | 1.088                  | --              | F6, F7              | \(x_{32}\)     | 0.657                  | --              |
| F2, F8              | \(x_{83}\)      | 2.201                  | --              | F6, F8              | \(x_{32}\)     | 1.275                  | --              |
| F2, F9              | \(x_{83}\)      | 3.106                  | --              | F6, F9              | \(x_{32}\)     | 1.419                  | --              |
| F3, F4              | \(x_{6}\)       | 0.883                  | --              | F7, F8              | \(x_{79}\)     | --                     | 0.016           |
| F3, F5              | \(x_{85}\)      | 0.632                  | --              | F7, F9              | \(x_{96}\)     | 0.224                  | --              |
F3, F6 \( x_{32} \) 0.562 -- F8, F9 \( x_{76} \) 0.397 --

| Fault category | F1   | F2   | F3   | F4   | F5   | F6   | F7   | F8   | F9   |
|----------------|------|------|------|------|------|------|------|------|------|
| Diagnostic accuracy | 100% | 100% | 89.1%| 100% | 100% | 100% | 90.8%| 99.7%| 100% |

| Feature selection method | Classification | Number of fault | Diagnostic accuracy |
|--------------------------|-----------------|-----------------|---------------------|
| PCA                      | SVM             | 4               | 98.75               |
| MPE                      | ANN             | 4               | 96.53%              |
| DET-PSO                  | PNN             | 6               | 97.23%              |
| TFS                      | PNN             | 9               | 97.74%              |

MPE (Multiscale permutation entropy) [20]; DET-PSO (Distance evaluation technique particle swarm optimization )[12]; LDA (Independent component analysis) [21];

5. Conclusions
A feature selection method based on EMD and TFS is presented and discussed for fault diagnosis in this paper. On the basis of filter evaluation criterion and heuristic search strategy, this novel feature selection method selects the optimal feature for each target sample space. It not only could remove the redundant information to the hilt, but also retains the most obvious difference between samples. Moreover, TFS is easy to design and could complete the calculation with great rapidity. Therefore, TFS effectively improves the accuracy of fault diagnosis, and it is a very useful feature selection method for multi fault diagnosis.

References
[1] Nandi S, Toliyat H A, Li X. Condition Monitoring and Fault Diagnosis of Electrical Motors—A Review[J]. IEEE Transactions on Energy Conversion, 2005, 20(4):719-729.
[2] Choi S, Akin B, Rahimian M M, et al. Implementation of a Fault-Diagnosis Algorithm for Induction Machines Based on Advanced Digital-Signal-Processing Techniques[J]. IEEE Transactions on Industrial Electronics, 2011, 58(3):937-948.
[3] Zhang W F, Yan H. Exon prediction using empirical mode decomposition and Fourier transform of structural profiles of DNA sequences[J]. Pattern Recognition, 2012, 45(3):947-955.
[4] Cherif L H, Debbal S M, Bereksi-Reguig F. Choice of the wavelet analyzing in the phonocardiogram signal analysis using the discrete and the packet wavelet transform[J]. Expert Systems with Applications, 2010, 37(2):913-918.
[5] Ali J B, Saidi L, Mouelhi A, et al. Linear feature selection and classification using PNN and SFAM neural networks for a nearly online diagnosis of bearing naturally progressing degradations[J]. Engineering Applications of Artificial Intelligence, 2015, 42(C):67-81.
[6] Huang … [et. al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time seri[M]. The Royal Society, 1998.
[7] Amirat Y, Benbouzid M E H, Wang T, et al. EEMD-based notch filter for induction machine bearing faults detection[J]. Applied Acoustics, 2018, 133:202-209.
[8] Zheng J, Cheng J, Yang Y. Generalized empirical mode decomposition and its applications to rolling element bearing fault diagnosis[J]. Mechanical Systems & Signal Processing, 2013, 40(1):136-153.
[9] Piao Y, Park H W, Jin C H, et al. Ensemble method for classification of high-dimensional data[C]// International Conference on Big Data and Smart Computing. IEEE, 2014:245-249.
[10] Hu L, Gao W, Zhao K, et al. Feature selection considering two types of feature relevancy and feature interdependency[J]. Expert Systems with Applications, 2018, 93.
[11] Tang T, Bo L, Liu X, et al. Variable predictive model class discrimination using novel predictive models and adaptive feature selection for bearing fault identification[J]. Journal of Sound & Vibration, 2018, 425:137–148.
[12] Van M, Kang H J. Bearing-fault diagnosis using non-local means algorithm and empirical mode decomposition-based feature extraction and two-stage feature selection[J]. Science Measurement & Technology Lett, 2015, 9(6):671-680.
[13] Babu N R, Mohan B J. Fault classification in power systems using EMD and SVM[J]. Ain Shams Engineering Journal, 2015, 8(2).
[14] Kumar I, Virmani J, Bhadauria H S, et al. Classification of Breast Density Patterns Using PNN, NFC, and SVM Classifiers[M]// Soft Computing Based Medical Image Analysis. 2018.
[15] Tomar D, Agarwal S. Feature Selection based Least Square Twin Support Vector Machine for diagnosis of Heart Disease[J]. International Journal of Bio-Science and Bio-Technology, 2014, 6:69-82.
[16] Chen X, Zhou J, Xiao J, et al. Fault diagnosis based on independent feature vector and probability neural network for rolling element bearings[J]. Applied Mathematics & Computation, 2014, 247(C):835-847.
[17] Chen X, Zhou J, Xiao H, et al. Fault diagnosis based on comprehensive geometric characteristic and probability neural network[J]. Applied Mathematics & Computation, 2014, 230(3):542-554.
[18] K. Fujarewicz, M. Wiench. Selecting differentially expressed genes for colon tumor classification[C]// Int. J. Appl. Math. Comput. Sci. 2003:327-335.
[19] Chen X Y, Zhou J Z, Xu X M, et al. A Hierarchical Feature Selection Method Based on Classification Tree for HGU Fault Diagnosis[J]. Advanced Materials Research, 2014, 1037:398-403.
[20] Vakharia V, Gupta V K, Kankar P K. A multiscale permutation entropy based approach to select wavelet for fault diagnosis of ball bearings[J]. Journal of Vibration & Control, 2014, 21(16).
[21] Widodo A, Yang B S, Han T. Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors[J]. Expert Systems with Applications, 2007, 32(2):299-312.