RESEARCH ON FEATURE EXTRACTION OF MECHANICAL FAULT BASED ON ORTHOGONAL LOCAL FISHER DISCRIMINANT ANALYSIS

Guoqin SONG \textsuperscript{1a}\ and Chun HE \textsuperscript{1b}

ABSTRACT: The basic problem of fault diagnosis is to extract the characteristic parameters and design the decision function according to the running state signals collected by the sensors. In addition, it also can find out the fault states. Due to the complexity of the operation state of the mechanical equipment, the state signal has the characteristics of large amount of data and high degree of nonlinearity. Therefore, it is more difficult for people to control and deal with these large and complex data. In recent years, manifold learning methods have been developed to extend data analysis and state decision from Euclidean space to manifold. It is able to dig out the essential features of the data efficiently and quickly from the data set distributed on the high dimensional manifold. In order to solve the problem, a new method based on iterative orthogonal and Schur orthogonal are proposed. The signal is projected into a high dimensional feature space by nonlinear kernel function, and the orthogonal local discriminant analysis and fault feature extraction are carried out in this space. This method realizes the transformation of linear manifold to the nonlinear method, and achieves better fault diagnosis effect than the linear orthogonal method. The local boundary neighbourhood is used to construct the discriminant function for fault feature extraction and diagnosis. The efficiency of the method is greatly improved by using the boundary point pairs in the neighbourhood space to calculate the local intra class divergence and inter class divergence. In order to avoid the interference of the pseudo boundary points, kernel method is used to realize the change of local fuzzy clustering boundary identification from linear to nonlinear.

KEY WORDS: Fisher discriminant analysis; manifold learning; signal denoising; fault diagnosis.

1 INTRODUCTION

Fault diagnosis of mechanical equipment is a typical data and pattern recognition process, and the related technology involves many fields of mathematics, chemistry, mechanics, machinery and computer science, which is an interdisciplinary comprehensive science. With the rapid development of industrial modernization, scientific theory and technology, the integration degree of mechanical equipment is becoming higher and higher. The direct result of this is that the requirements of the bearing capacity of the equipment system and reliable working speed are more and more strict. The mechanical equipment state monitoring signal processing, the correct amount of useful signal and its characteristic of effective extraction, different fault diagnosis, fault location, fault severity recognition and correct judgment of health maintenance for mechanical equipment is becoming more and more difficult. Therefore, it is necessary to put forward different ideas to improve the diagnostic methods and ideas based on the failure mechanism of different mechanical systems, so as to obtain reliable and effective diagnostic results [1].

The basic process of equipment fault diagnosis is: First of all, a variety of sensor detection devices are used to detect the signal and obtain a variety of physical and chemical properties of the mechanical equipment system operation process. The detected material is converted into a variety of easily processed electrical signals through the hardware circuit and the device, and then into the computer system. The linear and nonlinear analysis is used to find out the best characteristic of the running state of the mechanical system, such as vibration peak value, kurtosis and other time-domain eigenvalue. Secondly, the classifier designed for fault diagnosis should meet the performance of the equipment. For example, nonlinear classifier is usually used for complex machines and general linear classifier is used for simple equipment. Finally, we need to give the equipment maintenance strategy and treatment measures according to the diagnosis results [3]. The essence of fault diagnosis is to make use of the interaction between the components and the fault signal in the process of mechanical system. The operation origin of the mechanical equipment...
system is derived through the theory of the phenomenon backstepping. Finally, the maximum economic and social benefits during the normal operation period of the equipment can be utilized, and the occurrence of equipment accidents and accidents can be prevented [2].

In the fault diagnosis of mechanical equipment, the monitoring and control system is more precise, then the number of sensors is more, which indicates the more index of the state on equipment operation. The data that describe the state with multiple variables are abstracted into high dimensional data. High dimensional data provides a very rich and detailed information about the state of the device, but a significant increase in the data dimension will bring unprecedented difficulties to the subsequent data processing. The research of fault diagnosis based on manifold learning has just started, and a nonlinear time series noise reduction method based on phase reconstruction and mainstream shape recognition is proposed. Through the reconstruction of the phase space of the data, it can reduce the noise of the signal. Meanwhile, the shock characteristics of gearbox fault signal submerged in noise are extracted successfully.

2 METHODS

2.1 The proposal of manifold learning noise reduction problem

When the machine is in operation, the signals collected by the sensors are generally nonlinear and non-stationary, which will inevitably be subject to noise interference. The basic method of signal noise reduction is to decompose the collected time series data into the noise generated by the dynamic behavior of the original system and the noise produced by the external disturbances according to certain objective criteria. The traditional method of signal de-noising is hardware or software filtering technology based on linear smoothing [4].

Manifold learning is a nonlinear dimensionality reduction method, which is a method to reduce the dimension of data on the basis of keeping its main features. The study of noise manifold learning is mainly about the influence of noise in high dimensional data on the stability and robustness of manifold learning algorithm. In this paper, the noise reduction algorithm based on manifold learning is to solve the problem of how to design an effective noise reduction algorithm [5]. The main purpose is to reduce the noise in the signal and to consider the robustness of the noise. In the local tangent space alignment (LTSA) algorithm, the local geometry of low dimensional manifold is constructed by approximating the tangent space of each sample point.

1.1 Signal phase space reconstruction

The research of phase space reconstruction is a hot spot in nonlinear time series analysis, and researchers have proposed many methods at the same time. The traditional methods, such as the MMI/FNN method based on the pseudo neighbor method and the minimum mutual information method, is rebuilt through analytical method. Modern algorithms are mainly phase space reconstruction method based on wavelet transform and phase space reconstruction method based on wavelet neural network. Regardless of the method, the selection of embedding dimension and delay time is the key to the reconstruction of the phase space. We assume that the one-dimensional time series is \( x_i(i = 1, 2, \ldots n) \), then the reconstructed phase space can be represented as follows:

\[
X_n = (x_{n-(m-1)\tau}, x_{n-(m-2)\tau}, \ldots x_n)
\]

(1)

\( X_n \) is the N phase point in phase space and represents a state in m dimensional phase space. Takens proved that there exists diffeomorphism between the trajectories and dynamical systems in the reconstructed Rm space as long as to find a suitable embedding dimension m from mathematics, and delay dimension \( m \geq 2d + 1 \) coordinates (d is the intrinsic dimension of power system), then reconstructs the trajectory in the Rm space and the original dynamic system to maintain a diffeomorphism. There is a suitable range about the selection of delay time \( \tau \), which is not too small or too large. If \( \tau \) is too small, then the difference between the adjacent delay coordinates of the delay vector is too small. That is, the redundancy is larger, and the information of the original attractor contained in the sample points in the reconstructed phase space is small. It is manifested in the phase space morphology, which is the signal trajectory to the main axis of the phase space compression. If the difference is too large, the mutual information between the coordinate elements in the phase space is lost, and the signal path may be folded.
2.2 Local tangent space permutation noise reduction based on intrinsic dimension

The intrinsic dimension (ID) of the data set is also called the topological dimension, which is the minimum number of independent parameters to describe all the data in the data set. It is also an inherent property of a data set [6]. The intrinsic dimension of a signal is the minimum number of independent parameters that describe the instantaneous state of the signal system, that is, the minimum dimension required for the phase space of the system.

The intrinsic dimension of the signal determines the distribution properties of the phase points in the phase space, so it can be determined by analyzing the subspaces of the points in the phase space. At present, there are three kinds of methods to estimate the intrinsic dimensionality, which are feature mapping, geometric learning and statistical learning. The most representative algorithm of feature mapping is the method based on principal component analysis (PCA) and its improved algorithm [7]. This kind of algorithm requires the user to determine the closed value of eigenvalues, and the result of dimension estimation depends on the accurate selection of local region scale.

The idea of local PCA estimation on the intrinsic dimension is to find out the number of main features, and then reduce the size of the local region up to a certain limit dimension. The algorithm flow is as follows:

Step 1. The data is set to $S = \{X_1, X_2, ..., X_n\} \subset \mathbb{R}^D$. The threshold value $t$ and the region size (such as the radius of the sphere and k-neighbor-domain) are given.

Step 2. The k value is set to $k=1$, and there are $n_k$ points in the local area of $X_k$, remarking as $Y = \{X_1^{(k)}, ..., X_{n_k}^{(k)}\}D_{nk}$. The eigenvalues $\lambda_1, \lambda_2, ..., \lambda_{n_k}$ (all features are eliminated by the largest eigenvalue and arranged in order from small to large).

Step 3. $\sum = \sum(Y - \bar{Y})(Y - \bar{Y})$, $1, \lambda_2^{(k)}, ..., \lambda_{n_k}^{(k)}$ are the largest eigenvalues of the local domain.

Step 4. If $k=k+1$, the result of the step 3 is repeated. Otherwise, the step 5 is used.

Step 5. The dim value is remarked as $\dim = \{d_1, ..., d_N\}$, the distribution of each dimension of $\dim$ is calculated and a consistent estimate of the intrinsic dimension $\hat{d}(\delta)$ is determined.

3 FAULT FEATURE EXTRACTION AND DIAGNOSIS BASED ON ORTHOGONAL LOCAL FISHER DISCRIMINANT

In general, vibration, sound, light, magnetism, radiation and other means are used to make fault diagnosis of rotating machinery. The running state data represented by various physical and chemical phenomena accompanying the fault is collected. The change rules and characteristics are observed and the hidden state of the data is analyzed. In addition, the normal and abnormal (fault) state are identified and distinguished. Based on the in-depth research of local Fisher discriminant analysis (LFDA), the iterative orthogonal local Fisher discriminant analysis (IOLFDA) and schur orthogonal local Fisher discriminant analysis (SOLFDA) are proposed. These are two orthogonal supervised manifold learning methods. The local Fisher discriminant function of the rotor system is established by calculating the local intra class and inter class divergence of the training samples. The optimal projection vector is solved by orthogonal iteration, and the dominant eigenvector of new data is obtained by the projection of test data. The results show that the orthogonal LFDA algorithm such as IOLFDA and SOLFDA is better than LPP, FDA, LFDA algorithm [8].

3.1 Local Fisher discriminant analysis

(1) Fisher discriminant analysis

In 1936, Fisher first putted forward the linear discriminant function of different characteristic variables in his classic paper. Since then, Fisher discriminant analysis FDA method has been widely used in multivariate statistics, pattern recognition, information retrieval and classification. The basic idea of FDA is to find the maximum data of the same data set, while the minimum projection direction of different types of data from divergence also is found. The experiment separates the different classes of samples as much as possible after projective transformation and gathers the same types as much as possible. FDA is a linear global supervised learning algorithm. LPP is an unsupervised learning algorithm which can preserve the local features of data. When the dimension...
reduction and information classification, the different data sets represent different results. LFDA combines the advantages of the two algorithms, and can effectively maintain the local structure of data in the process of data feature extraction.

We suppose that there are \( n \) \( m \) dimensional data samples \( x_i \in R^m(i = 1,2...n) \), and each data sample is a class of \( C \). The sample matrix is remarked as \( X = [x_1,x_2,...x_n] = [X_1,X_2,...X_n] \).

\( X_i \) represents the \( n_j \) sample set in the \( i \) category. The scatter of data samples can be represented by the matrices: \( S^y \) represents the within-class scatter matrix of the samples.

\[
S^y = \frac{1}{n} \sum_{j=1}^{c} \sum_{i \in X_j} (x_i - \mu_j)(x_i - \mu_j)^T
\]  

(2)

In the formula, \( \mu_j \) is the mean value of \( j \) sample and \( \mu \) is the mean value of total sample.

(1) Locality preserving projection

The locality preserving projection LPP algorithm not only solves the shortcomings of the traditional linear methods such as PCA, but also solves the nonlinear manifold of the original data. The following discusses the basic principles of LPP algorithm.

For a given data set \( X \), LPP establishes the weight graph based on the nearest neighbor relation of each data point \( G \). In order to ensure that the two points in the original data set \( X \) are adjacent to each other, they can be connected to each other in the weighted graph \( G \). we suppose that \( Y = [y_1,y_2,...,y_d] \in R^{ld} \).

It is the mapping of the data set \( X \) after the transformation matrix \( A \), then a reasonable criterion for selecting the optimal mapping of LPP is to minimize the objective function under the appropriate constraints. \( S_q \) is the neighborhood weight matrix.

\[
\sum_{i=1}^{c} (y_i - y_j)S_q(y_i - y_j)^T
\]  

(3)

3.2 Iterative orthogonal local Fisher discriminant method

As mentioned earlier, the projection base vector obtained by LFDA algorithm is not necessarily mutually orthogonal, which may cause the projection base vectors. It also leads to the fact that the discriminant function cannot be used to diagnose the fault class features in the case of complex data types. In this paper, an orthogonal Laplasse manifold learning algorithm is proposed to solve the problem of non-orthogonal vector in LPP algorithm, which ensures the orthogonality between the base vectors and obtains a lot of face recognition results. Some scholars also propose a neighborhood preserving embedding algorithm based on orthogonal projection. The core of the two method is to introduce the idea of iterative orthogonal to the corresponding learning algorithm, so as to improve the algorithm and improve the effect of pattern recognition. In this section, an iterative orthogonal method is introduced into the LFDA algorithm, and a fault feature extraction method based on iterative orthogonal local Fisher discriminant analysis (IOLFDA) is proposed. The IOLFDA algorithm uses the iterative orthogonal method to construct the orthogonal basis function, which can effectively preserve the structure information related to the distance in the manifold space of the fault signal. Orthogonal effects of projection vectors are as follows: we assume that \( T_m = [a_1,a_2,...a_m] \) is orthogonal projection matrix, then the Euclidean distance between two points in the projection space is:

\[
\text{dist}(y_i - y_j) = \|y_i - y_j\| = \|T^T x_i - T^T x_j\| 
\]  

It is shown that the spatial structure of the original space can be preserved completely after the orthogonal projection because of \( T^TT = 1 \). We assume that \( x_i \in R^d(i = 1,2,...n) \) is the \( d \) dimension training sample with class labels. \( Z_j \in R^d(j = 1,2,...) \) is the test sample, and \( y_i \in R^m(i = 1,2,...n) \) is the \( m \) dimension data after the projection.

3.3 Rotor fault feature extraction and diagnosis

In order to test the effect of two orthogonal LFDA (IOLFDA and SOLFDA) algorithms in this section, the rotor fault diagnosis experiment is carried out on the rotor test bench. The experimental platform is based on the model of vibration and fault simulation of the QPZZ-II type rotating machinery in Jiangsu.

The sensor is made by the United States Langsi Testing Technology Co. Ltd CA-YD-1182 piezoelectric acceleration sensor. DEWE-201 data acquisition system is used to collect the rotor
vibration signal, and the sampling frequency is set to 5000HZ. L1-L4 is 4 sensors mounted on the vertical and horizontal directions of the two rolling bearing housings. The experiment collects three kinds of fault, which is vibration acceleration signal of rotor without fault, bearing inner ring crack fault running vibration acceleration signal and vibration acceleration signal of rotor bearing pedestal looseness.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>The number of data samples collected by each fault state</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal type</td>
<td>Speed 10 HZ</td>
</tr>
<tr>
<td></td>
<td>Small load</td>
</tr>
<tr>
<td>No fault</td>
<td>12</td>
</tr>
<tr>
<td>Bearing fault</td>
<td>12</td>
</tr>
<tr>
<td>Looseness fault</td>
<td>12</td>
</tr>
</tbody>
</table>

Each fault is collected in 9 kinds of operating conditions, which is small load, medium load and large load state with the speed of 10 HZ, 20 HZ and 30 HZ. The state of the data collected is shown in Table 1. Each group is sampled for 2 seconds with a value of 10000. Figure 2 is the time domain waveform of the L1 sensor signal when the speed is 20 HZ without load.

### 3.4 Analysis of feature extraction and fault diagnosis

The experimental data of the 288 groups are divided into an average of two parts, and each part contains three kinds of faults. 144 sets of data are used as training samples, and the others groups are used as test samples. Firstly, the intrinsic dimension of the rotor vibration signal is estimated, and the estimated principal component analysis is 1.2892. The correlation dimension estimated by the fractal geometry method is 1.8864. The maximum likelihood estimation is about 2, while the intrinsic dimension of the signal is 2. That is, each group of 32 dimensional data samples are reduced to 2 dimension. For the sake of comparison, the reduction results are normalized. Figure 3 is the result of feature extraction and fault diagnosis of training data using IOLFDA and SOLFDA manifold learning methods. In the figure, the abscissa indicates that the principal component is 1, and the ordinate indicates that the principal characteristic component is about 2. The ‘*’ represents the normal state, and ‘O’ represents the pedestal looseness fault. The ‘+’ represents bearing inner ring failure. It can be seen that the fault features extracted by SOLFDA and IOLFDA can well distinguish three different fault states.

![IOLFDA fault identification results](image1)

![SOLFDA fault identification results](image2)

**Figure 1** Results of IOLFDA and SOLFDA fault pattern recognition
4 CONCLUSIONS
In order to solve the problem of fault diagnosis of rotating machinery, this paper proposes an orthogonal local Fisher discriminant method based on manifold learning noise reduction, fault feature extraction and diagnosis. The main innovation lies in the introduction of iterative orthogonal and Schur orthogonal decomposition method into the local Fisher discriminant fault diagnosis, which solves the problem that the data obtained by the LFDA method is not orthogonal to the maximum eigenvalue of the asymmetric characteristic equation. Fault diagnosis based on orthogonal local Fisher discriminant is a kind of supervised manifold learning method. The orthogonal basis function can be constructed by the method of iterative orthogonal or Schur orthogonal decomposition, which can effectively preserve the structural information related to the nearest neighbor distance in the manifold space of the fault signal. In addition, it can retain the category information in the process of the main feature extraction. That is to say, the extraction of the main features of the amount of energy can maintain or even reduce the intra class divergence as far as possible, so that the distance between the signal characteristics of the class as far away as possible. Therefore, the fault classification is realized.

5 REFERENCE