

Research Article

Mass Laplacian Discriminant Analysis and Its Application in Gear Fault Diagnosis

Guangbin Wang,¹ Ying Lv ,¹ Tengqiang Wang,² Xiaohui Wang,¹ and Huanke Cheng³

¹College of Mechanical Engineering, Lingnan Normal University, Zhanjiang 524048, China

²Mechanical and Electrical Engineering, Hunan University of Science & Technology, Xiangtan 411210, China

³Hunan Key Laboratory of Health Maintenance for Mechanical Equipment, Hunan University of Science & Technology, Xiangtan 411210, China

Correspondence should be addressed to Ying Lv; lvying1108@163.com

Received 31 May 2020; Revised 19 June 2020; Accepted 29 June 2020; Published 3 August 2020

Academic Editor: Jiawei Xiang

Copyright © 2020 Guangbin Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fault diagnosis is essentially the identification of multiple fault modes. How to extract sensitive features and improve diagnostic accuracy is the key to fault diagnosis. In this paper, a new manifold learning method (Mass Laplacian Discriminant Analysis, MLDA) is proposed. Firstly, it is assumed that each data point in the space is a point with mass, and the mass is defined as the number of data points in a certain area. Then, the idea of universal gravitation is introduced to calculate the virtual universal gravitation between data points. Based on the Laplace eigenmaps algorithm, the gravitational Laplacian matrix between the same kind of data and the heterogeneous data is obtained, and the discriminant function is constructed by the ratio of the virtual gravitation between the heterogeneous data and the virtual gravitation between the similar data; the projection function with the largest discriminant function value is the optimal feature mapping function. Finally, based on the mapping function, the eigenvalues of the training data and the test data are calculated, and the softmax algorithm is used to classify the test data. Experiments on gear fault diagnosis show that this method has higher diagnostic accuracy than other manifold learning algorithms.

1. Introduction

The integration of mechanical equipment such as wind turbines, aeroengines, and rail trains is getting higher and higher, the service time is getting longer and longer, and the system scale is getting bigger and bigger. When the state is monitored and diagnosed, the signals detected by the sensors have the characteristics of large amount of data, high nonlinearity, and strong noise interference. How to excavate the essential characteristics and internal laws of signals from massive and complex raw data has become the key link of mechanical fault diagnosis [1, 2]. For the state monitoring and fault diagnosis of mechanical equipment and its key components, domestic and foreign scholars have proposed many practical and novel algorithms, and good diagnostic results have been obtained [3–5]. As an effective tool for

dimension reduction, data mining, and machine learning, manifold learning seeks to cover the complex data processing mechanism behind the identification behavior, which can achieve the dimension reduction and feature extraction of massive fault data. Therefore, it can be used as a technical means of mechanical fault diagnosis, and the fault category is better identified on the basis of maintaining the essential characteristics of fault data [6, 7]. The feature extraction and fault diagnosis of mechanical equipment based on manifold learning have been a research hot spot in recent years. Many scholars have proposed a variety of new learning methods in combination with the actual mechanical fault diagnosis problems, such as multiscale manifold learning [8, 9], sparse manifold learning [10, 11], time-frequency manifold learning [12–14], fast manifold learning [15], mixed-domain manifold learning [16], and other

improved manifold learning methods [17–20]. These studies based on manifold learning have achieved certain effects, but at present it has not been found that the fault features contained in complex data are extracted to improve the accuracy of fault diagnosis from the perspective of data mass.

In this paper, a new manifold learning method is proposed: Mass Laplacian Discriminant Analysis (MLDA) algorithm. Firstly, the data mass and virtual gravitation study are carried out. Then, based on the Laplace map algorithm, the virtual gravity Laplacian matrix between the same data and the heterogeneous data is calculated, the discriminant function is constructed and the optimal feature mapping function is obtained, and the sensitive feature of the data is extracted. Finally, the characteristic value of the test data is obtained based on the mapping function, and the type of fault is discriminated based on the softmax algorithm. Experiments on gear fault diagnosis show that this method has higher diagnostic accuracy than other manifold learning algorithms.

2. Mass of Data and the Construction of Virtual Gravity

Due to the complicated operating state of the mechanical equipment, the harsh working environment, and long running time, its state signal is high-dimensional, nonlinear, and nonstationary. The data analysis and state decision are extended from Euclidean space to high-dimensional manifold, and the essential characteristics of the data are found through the manifold learning. It has become the main means of signal feature extraction and pattern recognition. The manifold is a high-dimensional space. If it is regarded as a vast space of the universe; then the data distributed on the manifold is the stars in the sky. Based on this idea, assuming that the data on the manifold is a mass point, there is physical gravitation between the pairs of data points.

For u -dimensional data point $x_i (i = 1, 2, \dots, n) \in R^{n \times u}$ distributed on the manifold, the mass m_i of the point x_i is defined as the number q_i of points in the neighborhood of x_i . Assuming the neighborhood distance d_i , the mass m_i is

$$m_i = \frac{q_i}{d_i}. \quad (1)$$

Define the centroid C_i of x_i as the sum of the product of $m_{i,p} (p = 1, 2, \dots, q)$ of each point $x_{i,p} (p = 1, 2, \dots, q)$ in the neighborhood distance d_i of that point and its coordinate vector, that is,

$$C_i = \sum_{p=1}^q m_{i,p} x_{i,p}. \quad (2)$$

It can be seen that the mass of x_i is related to the uniformity of the distribution of the mass point on the manifold and the selection of the neighborhood distance. In terms of material theory, the greater the density of an object, the greater the relative mass. Similarly, the greater the density of the neighborhood of the data, the greater the mass of the data. Conversely, when the mass m_i of the data point increases with the increase of the neighborhood distance d_i ,

if it changes linearly in the forward direction, it can be considered that the data distribution is uniform within the range of the neighborhood distance d_i ; if the mass m_i of the data point decreases with the increase of the neighborhood distance d_i , the data point is considered to be at the boundary of the data concentration area. If the mass m_i of the data point is 0 at a certain neighborhood distance, then the data point can be considered as outlier, and it can be discarded during analysis to reduce the impact of outliers on data operations and analysis and improve the robustness of the analysis algorithm.

When data is given mass and becomes a particle distributed in a high-dimensional space, the relationship between pairs of data points can be expressed by the gravitational attraction in Newtonian mechanics. Because this gravitation is based on the assumption that the data has mass, it is defined here as virtual gravitational $F_{i,j}$:

$$F_{i,j} = G \frac{m_i m_j}{r_{i,j}}, \quad (3)$$

where G is the virtual gravitational coefficient and $r_{i,j}$ is the Euclidean distance between x_i and x_j . Substituting (1) into (3), you can get

$$F_{i,j} = G \frac{q_i q_j}{d_i d_j r_{i,j}}. \quad (4)$$

In the data analysis, in order to make the calculated data mass comparable, the same neighborhood distance is generally used for the same spatial data; that is to say, $d = d_i = d_j$, and (4) becomes

$$F_{i,j} = G \frac{q_i q_j}{d r_{i,j}}. \quad (5)$$

It can be seen from (5) that the virtual gravitational force $F_{i,j}$ is inversely proportional to the distance $r_{i,j}$ between the centroid point pairs and the neighborhood distance d for determining the data mass and within the range d from the neighborhood of the mass point x_i . The amount of data q_i is proportional to the amount of data q_j in the range of the distance d of the particle x_j . When the neighborhood distance d is selected, q_i and q_j are also determined; then, the virtual gravitational force $F_{i,j}$ is inversely proportional to the distance $r_{i,j}$ between the point pairs.

3. Mass Laplacian Discriminant Analysis (MLDA)

Laplace eigenmap algorithm is a classical nonlinear manifold learning algorithm. It assumes that the data points are distributed on the compact Riemannian manifold. The main idea is to describe a manifold with an undirected weight graph and construct the corresponding space embedded objective function through the Beltrami operator to obtain its low-dimensional embedding representation [21]. In short, the graph is mapped from a high-dimensional space to a low-dimensional space while maintaining the local adjacency of the graph. Suppose the Laplacian Beltrami operator is as follows:

$$L(f) = \frac{1}{2} \sum_{i,j} (f_i - f_j)^2 W_{ij}. \quad (6)$$

Then, the minimization problem can be transformed into an eigenfunction problem that is combined with the Lagrange auxiliary function. According to Graph theory, if the data is uniformly sampled from the low-dimensional manifold in the high-dimensional space, the Laplacian Beltrami operator on the manifold is defined as the negative divergence function of the gradient vector on the manifold-cut space, and the discrete approximation [22] can be performed by the previous several feature vectors of the Laplacian matrix on the graph.

Laplace eigenmap algorithm is mainly to keep the adjacent data points in the high-dimensional space still adjacent when embedded in the low-dimensional space. The core is to construct the neighboring weight matrix $W_{i,j}$, that is, the neighboring weight matrix $W_{i,j}$ to describe the relationship between pairs of data points. Here, if the virtual gravitational force $F_{i,j}$ is normalized, if the two particles are close to each other, or if the particle x_i and the particle x_j are adjacent, then $r_{i,j}$ gets smaller and $F_{i,j}$ gets bigger until it equals 1. If the two particles are further apart or if the particle x_i and the particle x_j are not adjacent, then $r_{i,j}$ gets bigger and $F_{i,j}$ gets smaller until it equals 0. That is, the virtual gravitational force $F_{i,j}$ is inversely proportional to the distance $r_{i,j}$ between point pairs, which is consistent with the value of $W_{i,j}$ using thermonuclear method. Therefore, the virtual gravity matrix F between point pairs can be used to replace the adjacency weight matrix W .

However, since the center of mass is also related to the selection of neighborhood distance d , that is to say, the virtual gravity moment F can also adjust the neighborhood distance d to realize the adjustment of the adjacency weight matrix, the virtual gravity matrix can describe the mutual relationship between data point pairs more accurately. For the normalized virtual gravitation matrix, its degree matrix D is

$$D_{i,j} = \sum_j F_{i,j}. \quad (7)$$

Then, we can obtain the Laplace matrix L of mass transformation, namely, $L = D - F$:

$$L_{i,j} = \sum_j F_{i,j} - F_{i,j}. \quad (8)$$

In mechanical fault diagnosis, each data set contains information on multiple fault categories, and multi-manifold theory thinks that each category corresponds to one manifold, there are many manifolds in the data set. The purpose of multi-manifold learning is to search for a suitable mapping function during the process of mapping high-dimensional data to a low-dimensional space, so that the low-dimensional data set can not only faithfully retain the internal structure of the data but also effectively distinguish the types of data. Using Laplacian Eigenmap

theoretical framework, suppose that \bar{L}^b and \bar{L}^w are used to represent the manifold structure of heterogeneous and homogeneous data and are defined as a qualitative Laplacian matrix of heterogeneous and homogeneous data, which can be expressed by

$$\begin{aligned} \bar{L}^b &= (C^b)^T L^b C^b, \\ \bar{L}^w &= (C^w)^T L^w C^w, \end{aligned} \quad (9)$$

where C^b and L^b are the center of mass and Laplace matrix of the heterogeneous data and C^w and L^w are the center of mass and Laplace matrix of the homogeneous data.

Then, in order to achieve the classification of data in the process of dimensionality reduction, it is necessary to obtain the low-dimensional embedded data of sample data after projection. The ratio of the mass Laplace structure \bar{L}^b of heterogeneous data and the mass Laplace structure \bar{L}^w of similar data is the minimum. The discriminant function $E(C)$ is expressed in

$$E(C) = \frac{\bar{L}^b}{\bar{L}^w} = \frac{(C^b)^T L^b C^b}{(C^w)^T L^w C^w}. \quad (10)$$

So the purpose of mass Laplacian discriminant algorithm is to find the optimal projection vector and minimize the discriminant function $E(C)$. The problem of solving the optimal projection vector is equivalent to the problem of solving the eigenvector corresponding to the minimum eigenvalue of the generalized eigenequation in

$$(C^b)^T L^b C^b \alpha = \lambda (C^w)^T L^w C^w \alpha. \quad (11)$$

Assuming $\alpha_0, \alpha_1, \dots, \alpha_{v-1}$ are the eigenvectors corresponding to the first v minimum eigenvalues $\lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{v-1}$ of (11), then the dimensionality reduction mapping obtained by MLDA algorithm can be expressed as

$$x_i \longrightarrow y_i = A^T x_i. \quad (12)$$

The mapping matrix $A = [\alpha_0, \alpha_1, \dots, \alpha_{v-1}] \in R^{u \times v}$ is the linear transformation matrix of the MLDA algorithm to achieve the optimal classification while maintaining the local structure of the data.

4. Softmax Classifier

Softmax regression is the promotion of logistic regression model in multiclassification problem, which is mainly used to solve the problem of multiclassification. After mass Laplacian discriminant for sampling data, data sets $\{(y^{(1)}, z^{(1)}), (y^{(2)}, z^{(2)}), \dots, (y^{(n)}, z^{(n)})\}$ are obtained, where y is the fault eigenvector and z is the category label $z \in \{1, 2, \dots, k\}$, where k is the number of categories.

For a given test input $y^{(i)} \in R^{n \times v}$, softmax function estimates the probability $p(z = j | y)$ for each category, that is, estimating the probability of each category of y . Suppose the probability function $h_\theta(y^{(i)})$ is as follows:

$$h_{\theta}(y^{(i)}) = \begin{bmatrix} p(z^{(i)} = 1 | y^{(i)}; \theta) \\ p(z^{(i)} = 2 | y^{(i)}; \theta) \\ \vdots \\ p(z^{(i)} = k | y^{(i)}; \theta) \end{bmatrix} \quad (13)$$

$$= \frac{1}{\sum_{j=1}^k e^{\theta(T/j)y^{(i)}}} \begin{bmatrix} e^{\theta_1^T y^{(i)}} \\ e^{\theta_2^T y^{(i)}} \\ \vdots \\ e^{\theta_k^T y^{(i)}} \end{bmatrix} c^2.$$

In (13), $h_{\theta}(y^{(i)})$ is the vector, any element $p(z^{(i)} = k | y^{(i)}; \theta)$ in the vector is the probability that sample $y^{(i)}$ belongs to category k , and the sum of each element in the vector is equal to 1. $\theta_1, \theta_2, \dots, \theta_k$ are the classifier parameters corresponding to each category, $\theta = [\theta_1^T, \theta_2^T, \dots, \theta_k^T]^T$. Then, the cost function of softmax regression algorithm can be defined as

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^k 1\{z^{(i)} = j\} \ln \left(\frac{e^{\theta(T/j)y^{(i)}}}{\sum_{n=1}^k e^{\theta(T/n)y^{(i)}}} \right) \right], \quad (14)$$

where 1 is an indicative function, $1\{\text{true}\} = 1$, and $1\{\text{false}\} = 0$. The value of θ can be obtained by minimizing the cost function $J(\theta)$ of the softmax regression, thereby predicting the category of a new sample. Generally gradient descent algorithm is used to solve the minimization of $J(\theta)$. The gradient formula of $J(\theta)$ is as follows:

$$\nabla_{\theta_j} J(\theta) = \frac{1}{m} \sum_{i=1}^m [y^{(i)} 1\{z^{(i)} = j\} - p(z^{(i)} = j | y^{(i)}; \theta)], \quad (15)$$

where the value of $\nabla_{\theta_j} J(\theta)$ is a vector, which means $J(\theta)$ is a partial derivative of θ_j for the j category.

5. MLDA Fault Diagnosis Model

Based on the above analysis, aiming at the problems in mechanical equipment fault diagnosis, the mass Laplacian discriminant algorithm was used to extract the fault features. Then softmax classifier was built to identify and classify the faults. Thus, the mass Laplacian fault diagnosis algorithm was established. And the algorithm model is shown in Figure 1. The specific steps are as follows.

Step 1. After extracting the monitoring data of the running state of mechanical equipment and preprocessing, the training sample $x_i (i = 1, 2, \dots, n) \in R^{n \times u}$ and the category label vector $z \in \{1, 2, \dots, k\}$ and k are obtained. Here, n is the number of data points, u is the dimension of data, and k is the number of categories of running state.

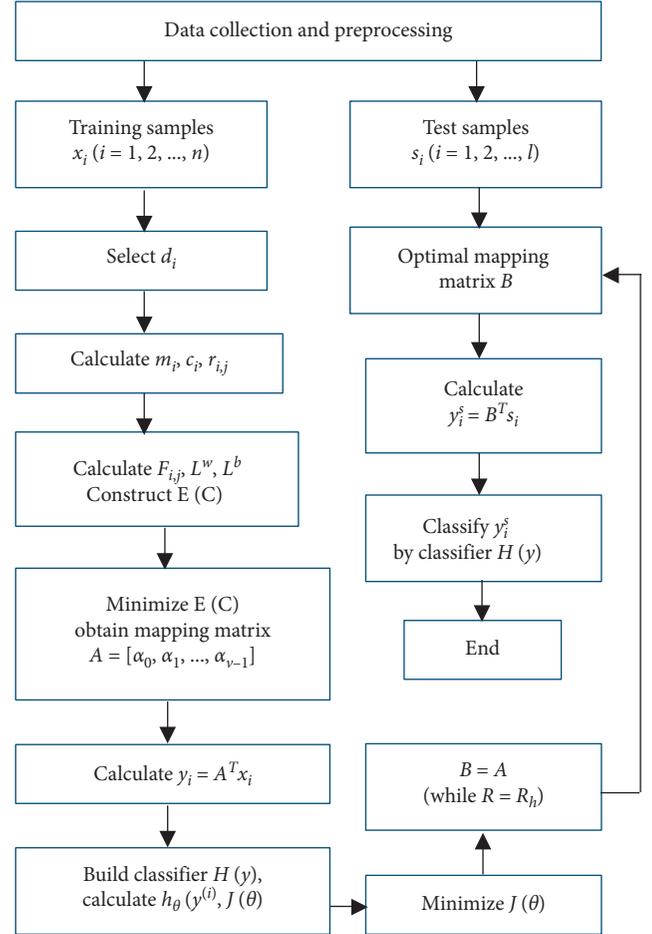


FIGURE 1: MLDA fault diagnosis model.

Step 2. The neighborhood distance d_i of x_i is selected. Suppose \bar{d}_{ij} is the average of the neighborhood distance of the sample point; then,

$$d_i = \varepsilon \bar{d}_{ij}, \quad (16)$$

where ε is the distance adjustment coefficient. When the data is evenly distributed, $\varepsilon = 1$, this is $d_i = \bar{d}_{ij}$; if the data distribution is uneven, then ε takes value around 1, and the data quality error caused by this unevenness can be adjusted.

Step 3. According to (1), the mass m_i of each point x_i is calculated, and the Euclidean distance $r_{i,j}$ between the sample point pairs is calculated. Equation (3) is used to calculate the virtual universal gravitation between x_i and x_j at each point. In this case, the gravitational coefficient $G = 1$.

Step 4. According to the category label of the training sample, classify heterogeneous data sets and homogeneous data sets, according to (2) and (9). The centroid C^b and C^w of each point in the heterogeneous data set and homogeneous data set and mass Laplace matrices L^b and L^w were calculated, respectively.

Step 5. According to (12), the discriminant function $E(C)$ is constructed and transformed into the minimization problem of the characteristic equation $(C^b)^T L^b C^b \alpha = \lambda (C^w)^T L^w C^w \alpha$, and the classification mapping matrix $A = [\alpha_0, \alpha_1, \dots, \alpha_{v-1}] \in R^{u \times v}$ is constructed by taking the eigenvectors corresponding to the first v minimum eigenvalues $\lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{v-1}$.

Step 6. Low-dimensional embedded y_i of training sample x_i was calculated according to (12), and softmax classifier $H(y)$ was constructed by combining category label vector $z \in \{1, 2, \dots, k\}$. Calculate probability function $h_\theta(y^{(i)})$ and softmax regression cost function $J(\theta)$ according to formulas (15) and (16), and obtain classification accuracy $R(y)$ by minimizing the predicted sample category of $J(\theta)$.

Step 7. Choose the mapping matrix A with the highest recognition rate R_{li} as the optimal mapping matrix $B = A(R = R_{li}) \in R^{u \times v}$.

Step 8. The test data $s_i (i = 1, 2, \dots, l) \in R^{l \times u}$ of the running state was obtained by the preprocessed mechanical equipment. Its low-dimensional embedded $y_i^s \in R^{l \times v}$ was obtained through $y_i^s = B^T s_i$. Softmax classifier was used to classify y_i^s , and the failure type of test data was obtained to realize the classification and recognition of mechanical fault.

6. Simulation and Analysis

According to the vibration theory of the gear, the normal gear mainly represents the shaft speed frequency and frequency multiplication and meshing frequency and its frequency multiplication. Wear characteristic frequency mainly on the tooth surface shows the meshing frequency and its harmonic components, and the fractional harmonic occurs when the wear is intensified. The characteristic frequency of the broken teeth shows a sideband near the meshing frequency and its higher harmonics, and the sideband is the rotational frequency and its harmonic frequency. Based on these, the vibration simulation signals x_1, x_2, x_3 of normal gear, worn gear, and broken gear can be constructed as shown in equations (17)–(19). Here, shaft speed is $f_0 = 10$ Hz, the number of gear teeth is $z = 23$, the meshing frequency is $f_1 = 460$ Hz, the sampling frequency is $F_s = 10$ kHz, a, b, c is constant coefficient, and s is interference and noise:

$$\begin{aligned} x_1 = & a_{11} \cos(2\pi f_0 t) + a_{12} \cos(4\pi f_0 t) \\ & + b_{11} \cos(2\pi f_1 t) + b_{12} \cos(4\pi f_0 t) + \dots \\ & + b_{14} \cos(8\pi f_0 t) + c_1 s, \end{aligned} \quad (17)$$

$$\begin{aligned} x_2 = & a_{21} \cos(2\pi f_0 t) + a_{22} \cos(4\pi f_0 t) + b_{21} \cos(2\pi f_1 t) \\ & + b_{22} \cos(4\pi f_0 t) + \dots + b_{24} \cos(8\pi f_0 t) + b_{25} \cos(\pi f_1 t) \\ & + b_{26} \cos(0.5\pi f_0 t) + c_2 s, \end{aligned} \quad (18)$$

$$\begin{aligned} x_3 = & a_{31} \cos(2\pi f_0 t) + a_{32} \cos(4\pi f_0 t) \\ & + b_{31} \cos(2\pi f_1 t) + b_{32} \cos(2\pi(f_0 - f_1)t) \\ & + b_{33} \cos[2\pi(f_0 + f_1)t] \\ & + b_{41} \cos(4\pi f_1 t) + b_{42} [4\pi(f_0 - f_1)t] \\ & + b_{43} \cos[4\pi(f_0 + f_1)t] \\ & + b_{51} \cos(6\pi f_1 t) + b_{52} [6\pi(f_0 - f_1)t] \\ & + b_{53} \cos[6\pi(f_0 + f_1)t] + c_3 s. \end{aligned} \quad (19)$$

The time domain and frequency domain diagrams of three kinds of fault gear simulation signals are constructed as shown in Figures 2–4.

During signal simulation, the different fault statuses are simulated by changing the magnitude of the amplitude without changing the characteristic frequency. Each group is 0.5 seconds (120 data points) and there are a total of 240 groups data (80 groups data for each fault). In order to study the impact of noise on the algorithm, adding different degrees of noise to each fault (none noise, signal-to-noise ratio (SNR) is 40, 20, 0, -20, -30, and -40), 180 groups (60 groups for each fault) were taken as training data, and the remaining 60 groups (20 groups for each fault) were used as test data to analyze the MLDA algorithm. When the distance coefficient ε is 0.7 to 1.32, the training data can correctly identify the three fault states, and the correct recognition rate can reach 100%. At this time, the mapping matrix is used to identify the fault state of the test data.

Figure 5 and Table 1 show the recognition results of fault data under different noise levels and distance coefficients. It can be seen that as SNR decreases, the correct recognition rate also decreases, but from noise-free to signal-to-noise ratio (40, 20, 0, -20), the correct rate is stable more than 90%. When the signal-to-noise ratio is -30, the correct recognition rate is between 83.33% and 91.67%. When the SNR is -40, the recognition rate suddenly decreases to 50%. In the actual engineering condition, the SNR of the vibration signal collected by sensor is generally not less than 10 dB. That is to say, as long as the appropriate distance coefficient is selected and the optimal mapping matrix is obtained through the training data, the sensitivity of the MLDA algorithm to noise is very low, and the robustness of the algorithm is very good.

Figures 6 and 7 show two extreme noise conditions (no noise and SNR = -40). The three-dimensional manifold of the three fault states obtained through training data shows that three types of faults can be well identified. Figures 8–13 show the classification results of the test data obtained from the training data with six different noise levels (no noise, signal-to-noise ratio of 40, 20, -20, -30, and -40). When the distance coefficient is 0.96 and SNR is above -20, the fault state can be better recognized and the recognition rate is above 91.67%. When SNR is -30, the fault can be identified, but the recognition rate is reduced to 83.33%. When SNR is -40, the recognition rate is lower than 50%, and the fault state is completely unrecognizable.

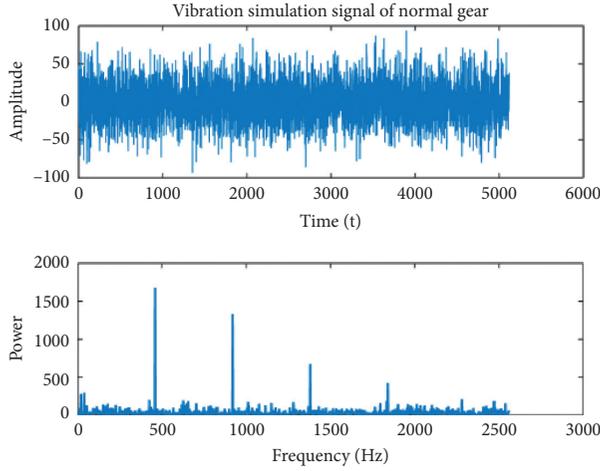


FIGURE 2: Vibration simulation signal of normal gear.

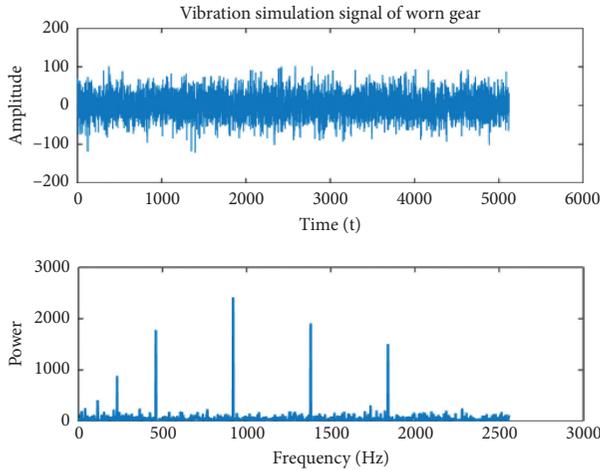


FIGURE 3: Vibration simulation signal of worn gear.

7. Experiment and Analysis

7.1. Experiment Platform and Conditions. The gear fault diagnosis experiment was carried out in the fault simulation bench of Spectra Quest, as shown in Figure 14. Figure 15 shows internal structure of the gearbox. The experiment bench is powered by the motor, and power is transmitted to the input shaft of the gearbox through the coupling, the drive shaft, and the pulley. The gear on the input shaft drives the intermediate shaft to mesh the gear, and the other gear on the intermediate shaft transmits the power through the output shaft, at the same time connecting the brake at the end of the gearbox output shaft. The experiment uses the vibration acceleration sensor of B&K Company to collect the signal of the multipoint on the gearbox.

Figure 16 shows sensor distribution: six sensors are arranged in the horizontal direction and two sensors are arranged in vertical directions. The sensor vibration signal was the most obvious in vertical direction at the input shaft end, so these three vibration signals are taken as the analysis signals. In the experiment, the current of the electromagnetic

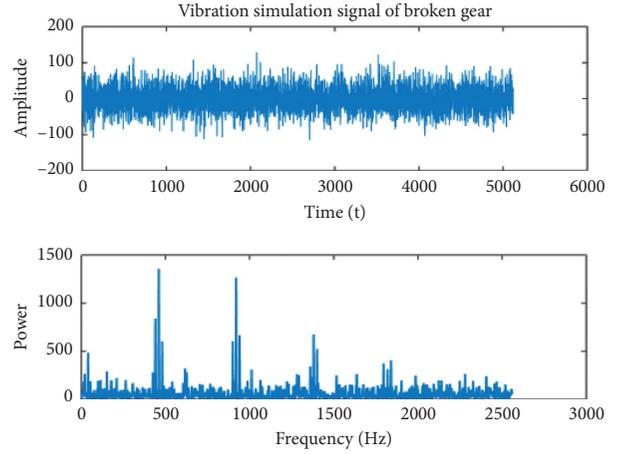


FIGURE 4: Vibration simulation signal of broken gear.

brake is set to 55 mA, the braking torque is 1790 N·m, the sampling frequency of the acquisition system is 16 kHz, and the transmission ratio of the pulley is 3.56.

In the experiment, four degrees of pitting fault (upper part of Figure 17) and five degrees of broken fault gears (lower part of Figure 17) were made on the pair of meshing gears. The sensor signal was the most obvious to vertical direction in the input shaft, so the three vibration acceleration signals here are taken as the analysis signals. The experiment simulates four kinds of working conditions under 30 Hz motor speed (normal gear, pinion pitting fault, large gear, and broken tooth fault). Collect 240 sets of data for each condition, 120 sets of which were used as training samples and the others were used as test samples. The data dimension is the data collected by the acceleration sensor within 0.25 seconds; it is $16\text{ K} \times 0.25 = 4\text{ K}$. The specific working condition data combination is shown in Table 2.

For a pair of meshing gears at the input shaft end, the number of teeth of the small gear is 36, the number of teeth of the large gear is 101, and the meshing frequency is 303.48 Hz when the motor speed is 30 Hz. Since the gear is driven by the motor through the pulley drive input shaft, there is a certain fluctuation in the transmission ratio, so the meshing frequency changes around 303 Hz. Figures 18–21 are the time domain and frequency domain diagram of the vibration signal in four fault states.

It can be seen that the characteristic frequencies of the normal gears are the meshing frequency and its multiplication, the characteristic frequencies of pitting gears are the meshing frequency and its division frequency, and the characteristic frequency of the broken tooth gear is the meshing frequency and its multiplication frequency, and there are the side frequency bands near these frequencies. The characteristic frequency component of the composite fault gear is more complicated, including division frequency, multiplication frequency of the meshing frequency, and the side frequency bands near these frequencies.

7.2. Analysis and Discussion. According to algorithm flow of the mass Laplacian fault diagnosis, four different state gear

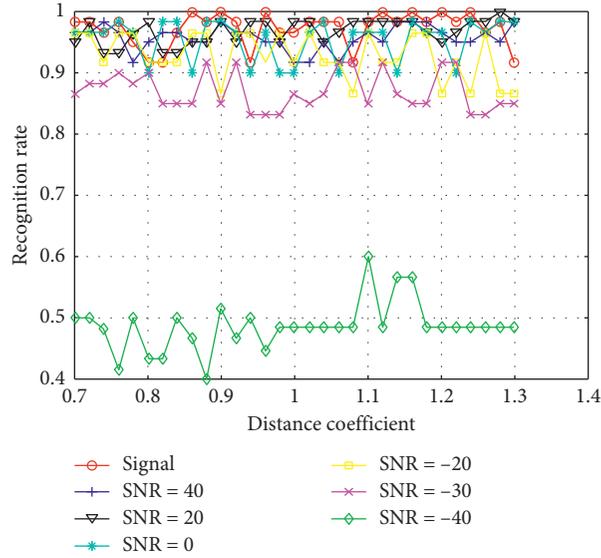


FIGURE 5: Fault recognition rate under the different noise levels and distance coefficients.

TABLE 1: Correct recognition rate of test data under different noise levels and distance coefficients (%).

ε	No noise	SNR = 40	SNR = 20	SNR = 0	SNR = -20	SNR = -30	SNR = -40
0.7	98.33	96.67	95.00	96.67	96.67	86.67	50.00
0.72	98.33	96.67	98.33	96.67	96.67	88.33	50.00
0.74	96.67	98.33	93.33	96.67	91.67	88.33	48.33
0.76	98.33	96.67	93.33	98.33	96.67	90.00	41.67
0.78	95.00	91.67	96.67	96.67	96.67	88.33	50.00
0.8	91.67	95.00	98.33	90.00	91.67	90.00	43.33
0.82	91.67	96.67	93.33	98.33	91.67	85.00	43.33
0.84	96.67	96.67	93.33	98.33	91.67	85.00	50.00
0.86	100.00	95.00	95.00	90.00	96.67	85.00	46.67
0.88	98.33	95.00	95.00	98.33	96.67	91.67	40.00
0.9	100.00	98.33	98.33	98.33	86.67	85.00	51.67
0.92	98.33	96.67	95.00	96.67	96.67	91.67	46.67
0.94	91.67	96.67	98.33	90.00	96.67	83.33	50.00
0.96	100.00	95.00	98.33	96.67	91.67	83.33	44.67
0.98	96.67	95.00	95.00	90.00	96.67	83.33	48.33
1.00	96.67	91.67	98.33	90.00	91.67	86.67	48.33
1.02	98.33	91.67	98.33	96.67	96.67	85.00	48.33
1.04	98.33	95.00	95.00	98.33	91.67	86.67	48.33
1.06	98.33	91.67	96.67	90.00	91.67	91.67	48.33
1.08	91.67	95.00	98.33	96.67	86.67	91.67	48.33
1.1	98.33	96.67	98.33	96.67	96.67	85.00	60.00
1.12	100.00	95.00	98.33	96.67	91.67	91.67	48.33
1.14	98.33	98.33	98.33	90.00	91.67	86.67	56.67
1.16	100.00	98.33	98.33	98.33	96.67	85.00	56.67
1.18	98.33	98.33	96.67	96.67	96.67	85.00	48.33
1.2	100.00	96.67	95.00	96.67	86.67	91.67	48.33
1.22	98.33	95.00	96.67	90.00	91.67	91.67	48.33
1.24	100.00	95.00	98.33	98.33	86.67	83.33	48.33
1.26	96.67	96.67	98.33	96.67	96.67	83.33	48.33
1.28	98.33	95.00	100.00	98.33	86.67	85.00	48.33
1.30	91.67	98.33	98.33	98.33	86.67	85.00	48.33

vibration signals are identified and analyzed. Because the MLDA algorithm is relatively sensitive to noise, the digital filter is used to reduce noise before performing data analysis.

Compared with FIR filters, IIR filters have higher accuracy in amplitude-frequency characteristics. Compared with Butterworth filters, Chebyshev filters, and other IIR filters,

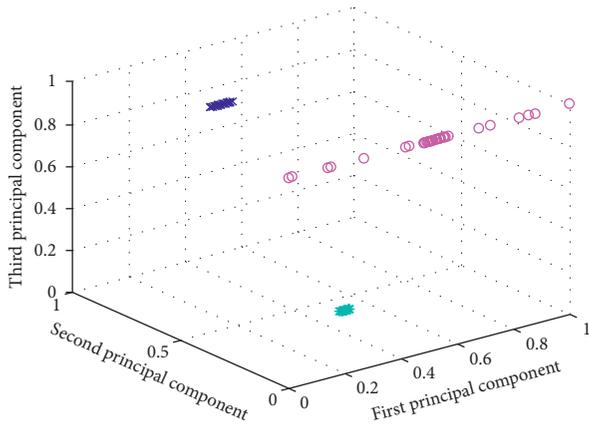


FIGURE 6: Recognition rate (100%) of training data.

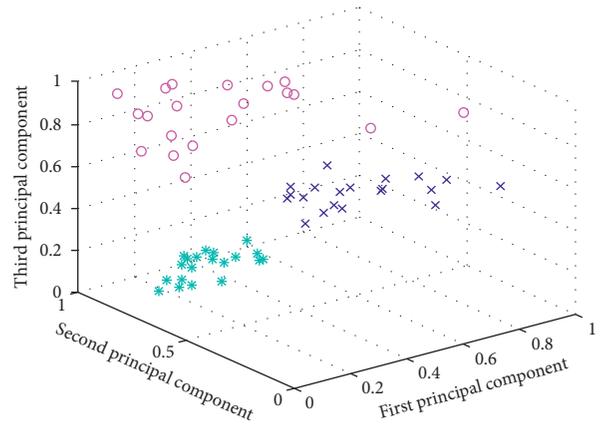


FIGURE 9: Recognition rate (95%) of test data.

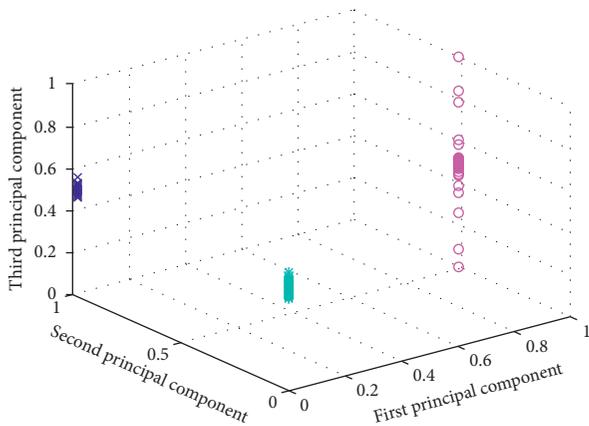


FIGURE 7: Recognition rate (100%) of training data.

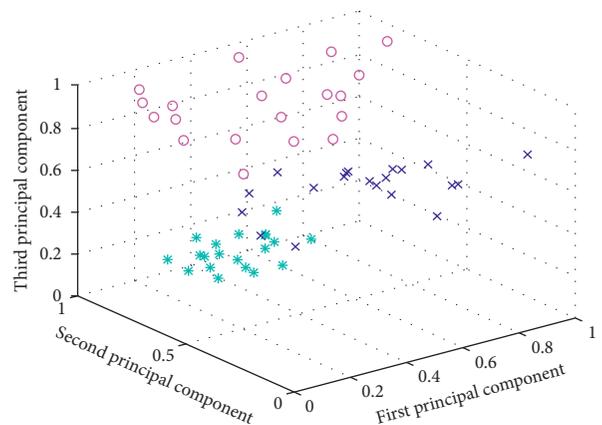


FIGURE 10: Recognition rate (98.33%) of test data.

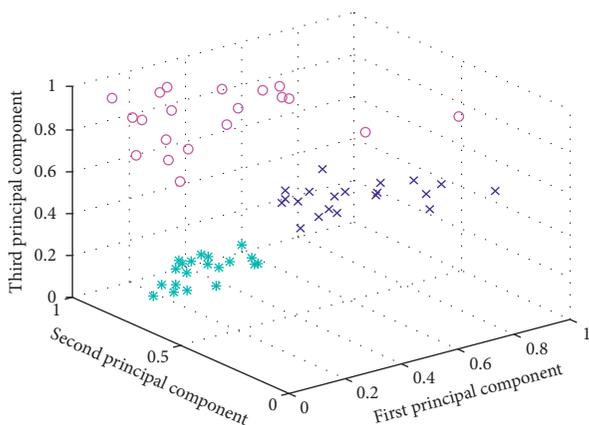


FIGURE 8: Recognition rate (100%) of test data.

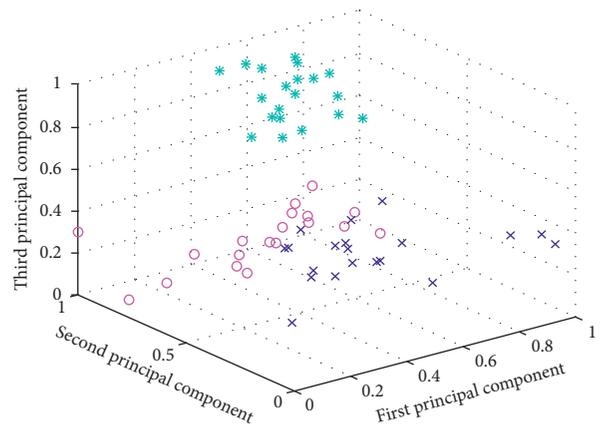


FIGURE 11: Recognition rate (91.67%) of test data.

elliptic filters have the smallest passband and stopband fluctuations, and the filtering effect is more stable, so the elliptical IIR filter is used in this experiment.

It can be seen from Figures 20 and 21 that the characteristic frequency of the gear vibration signal is around the

meshing frequency of 303.48 Hz and 2 times the frequency, plus the shaft speed side frequency of 30 Hz and its low-order frequency multiplication (generally lower than 5 times the shaft speed frequency); the filter cut-off frequency to 800 Hz can contain the fault characteristic frequency. Here take 900 Hz as the passband cut-off frequency, the stopband start frequency is 1200 Hz, the passband ripple is 0.3, and the stopband minimum attenuation is 40. The parameters of the

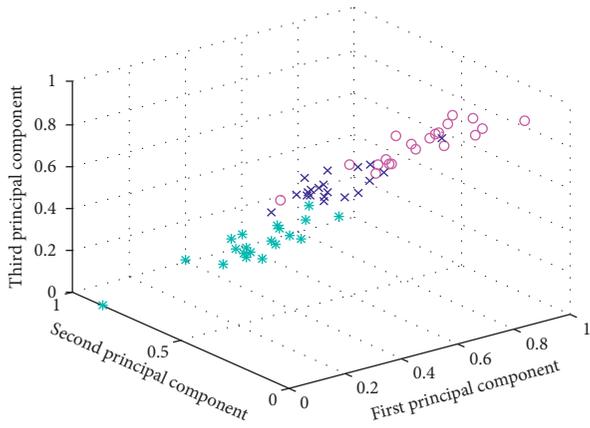


FIGURE 12: Recognition rate (83.33%) of test data.

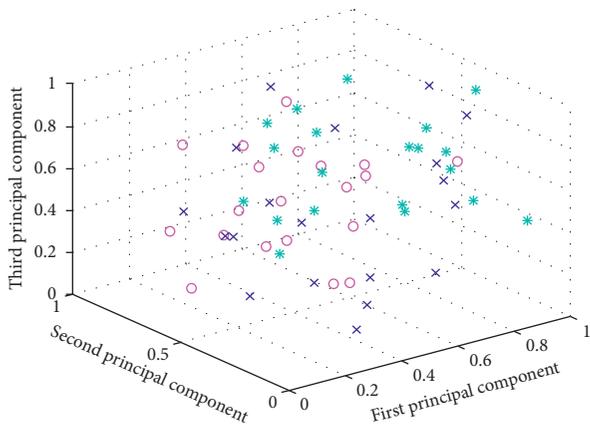


FIGURE 13: Recognition rate (44.67%) of test data.



FIGURE 14: Gear fault experiment platform.

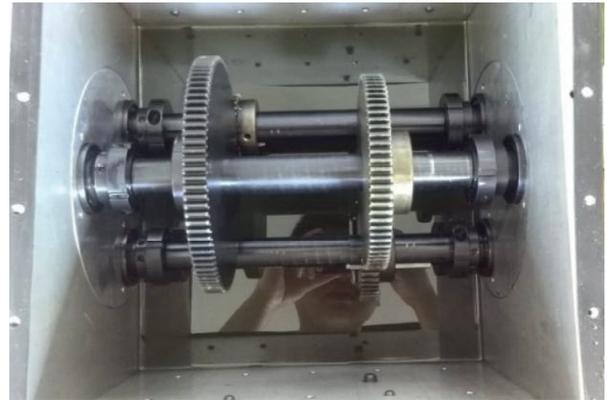


FIGURE 15: Internal structure of the gearbox.

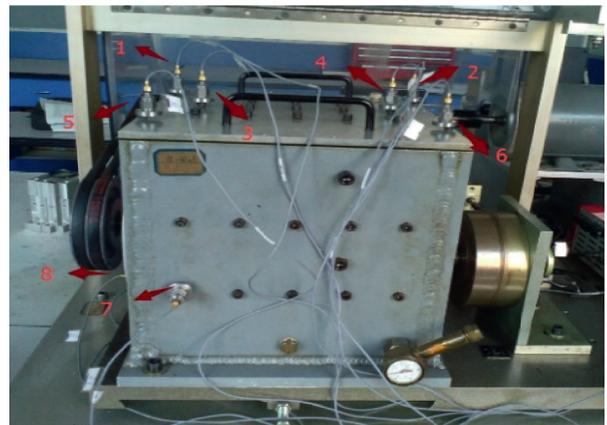


FIGURE 16: Sensor distribution.

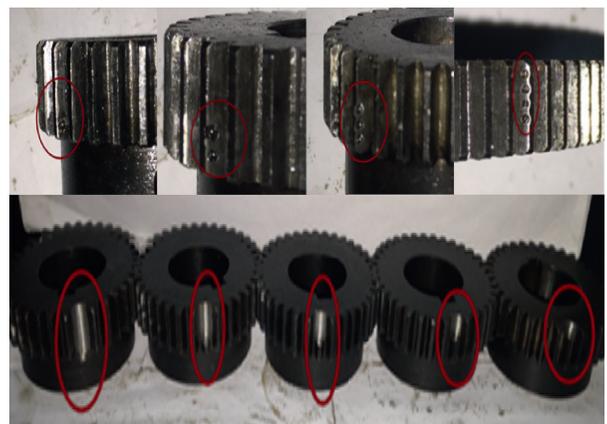


FIGURE 17: Fault gears (pitting and broken tooth).

softmax classifier are as follows: the weight attenuation coefficient is 0.05, the learning rate is 0.01, and the number of iterations is 5000.

Figure 22 and Table 3 show the recognition rates of the four fault status signals before and after noise reduction when the distance coefficient is 0.77 to 1.15. It can be seen that the noise has a great influence on the effect of fault

recognition. After the noise reduction, the recognition rate of the signal is increased by about 20%, and the highest recognition rate is over 90%, and the original sampling signal has a maximum recognition rate of only 70%. At the same time, the distance coefficient also has a great influence on the recognition. When the distance coefficient is 0.87 to 1.07, the recognition rate is steadily rising, and when the distance is over 1.07, the recognition rate starts to decrease

TABLE 2: Data sampling strategy for different fault levels on four states.

Fault status	Data grouping strategy based on working conditions				
Normal gear	240 sets of data (3 sensor signals \times 80 sampling periods)				
Pinion pitting	60 groups (1 pitting pit)	60 groups (2 pitting pits)	60 groups (3 pitting pits)	60 groups (4 pitting pits)	
Large gear broken teeth	48 groups (1/5 broken)	48 groups (2/5 broken)	48 groups (3/5 broken)	48 groups (4/5 broken)	48 groups (all broken)
Large gear broken + pitting	48 groups (1/5 broken + 1 pitting pit)	48 groups (2/5 broken + 2 pitting pits)	48 groups (1 3/5 broken + 3 pitting pits)	48 groups (4/5 broken + 4 pitting pits)	48 groups (all broken + 4 pitting pits)

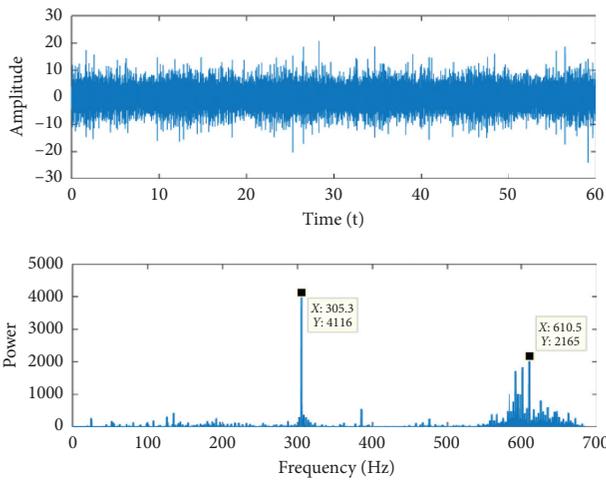


FIGURE 18: Time domain and frequency domain diagram for signal of normal state gear.

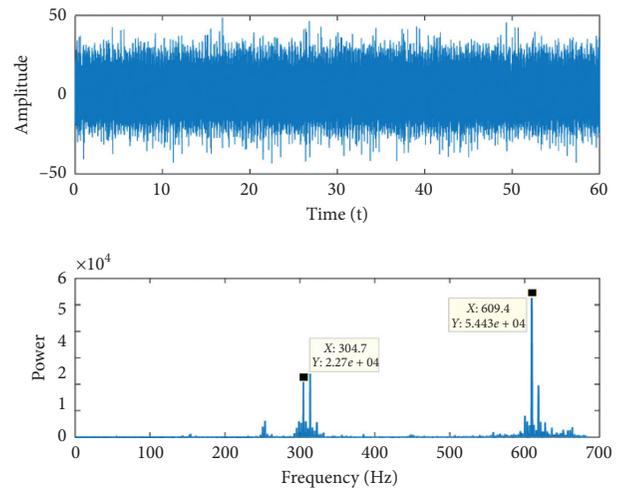


FIGURE 20: Time domain and frequency domain diagram for signal of broken tooth gear.

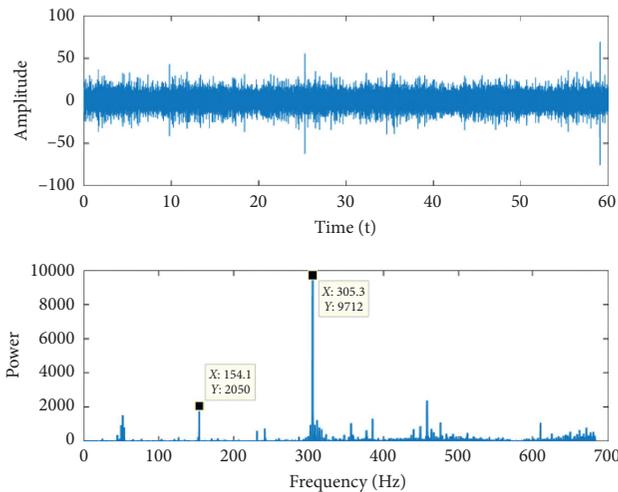


FIGURE 19: Time domain and frequency domain diagram for signal of pitting gear tooth.

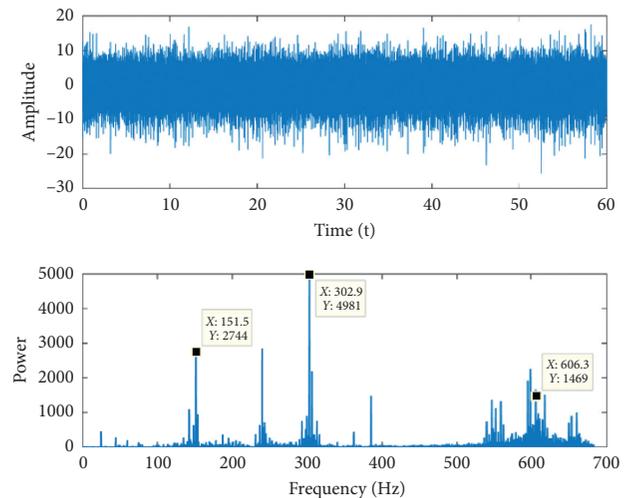


FIGURE 21: Time domain and frequency domain diagram for signal of compound fault gear.

greatly. When the distance is lower than 0.87, the recognition rate fluctuates obviously.

Figure 23 shows the recognition result of the original signals in the four fault states, and the distance coefficient is 1.03. The recognition rate is 76.04%. Figure 24 shows the

recognition effect of the four fault status signals after noise reduction, and the distance coefficient is 0.97. The recognition rate is 91.94%. In the figure, red \diamond is the normal signal, purple \circ is the pinion pitting fault, blue \times is the big gear broken tooth fault, and cyan $*$ is the small gear pitting

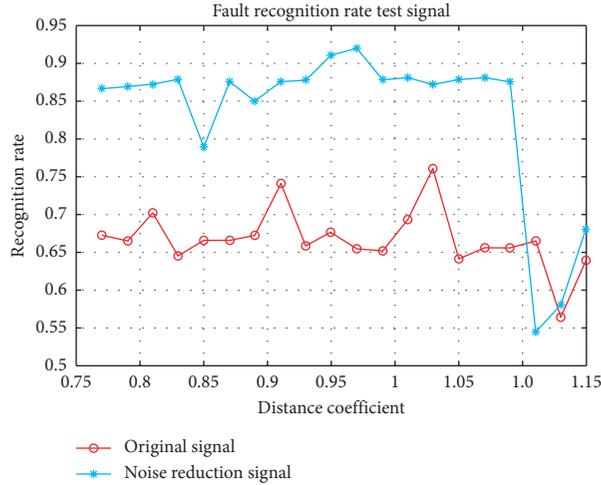


FIGURE 22: Recognition rate of four fault status signals before and after noise reduction.

TABLE 3: Recognition rate under different distance coefficient ϵ .

Different coefficient ϵ	0.77	0.79	0.81	0.83	0.85	0.87	0.89	0.91	0.93	0.95
Original signal recognition rate (%)	67.29	66.46	70.21	64.58	66.67	66.67	67.29	74.17	65.83	67.71
Signal recognition rate after noise reduction (%)	86.67	86.94	87.22	87.78	78.89	87.5	85	87.5	87.78	91.11
Distance coefficient ϵ	0.97	0.99	1.01	1.03	1.05	1.07	1.09	1.11	1.13	1.15
Original signal recognition rate (%)	65.42	65.21	69.37	76.04	64.17	65.63	65.63	66.46	56.46	63.96
Signal recognition rate after noise reduction (%)	91.94	87.78	88.06	87.22	87.78	88.06	87.5	54.44	58.06	68.06

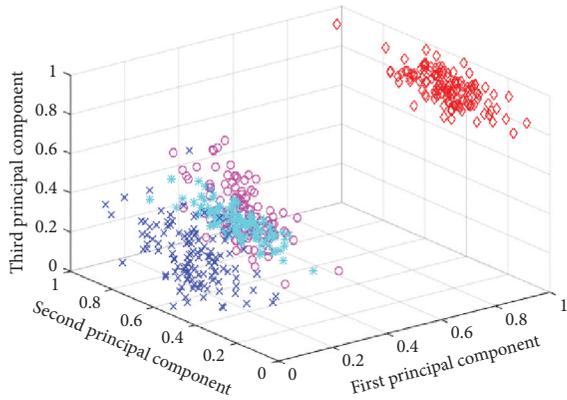


FIGURE 23: Recognition result of original signal ($\epsilon = 1.03$).

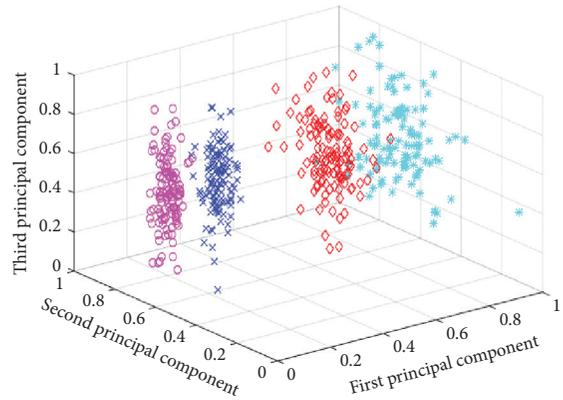


FIGURE 24: Recognition result of signal after noise reduction ($\epsilon = 0.97$).

and large gear broken tooth composite fault. It can be seen that the four fault status signals can be well identified after noise reduction, and only some faults can be identified before noise reduction. In this example, the normal signals are clearly identified, but the composite faults are mixed in two faults; thus, the recognition rate is greatly reduced.

In order to compare the effect of the algorithm, the LLE (local linear embedding), LPP (locally preserved projection), and LE (Laplace eigenmaps) are used for the vibration signals after noise reduction in the normal fault, pitting gear, worn gear, and composite fault gear to identify faults in this experiment. After taking the appropriate parameter

(number of neighborhoods k and distance coefficients), the highest recognition rate is shown in Table 4.

Figures 25–27 show recognition result for test signal after noise reduction based on three common manifold learning algorithms (LLE, LPP, and LE). It is visible for the sensor vibration acceleration signal, even after noise reduction and selection of appropriate parameters. The common manifold learning does not completely distinguish the four states, and the accuracy of fault recognition is also low.

TABLE 4: Highest recognition rate under different distance manifold learning methods.

Methods	LLE	LPP	LE	MLDA
Highest recognition rate (%)	71.67 ($k = 10$)	47.92 ($k = 10$)	60.21 ($k = 15$)	91.94 ($\epsilon = 0.97$)

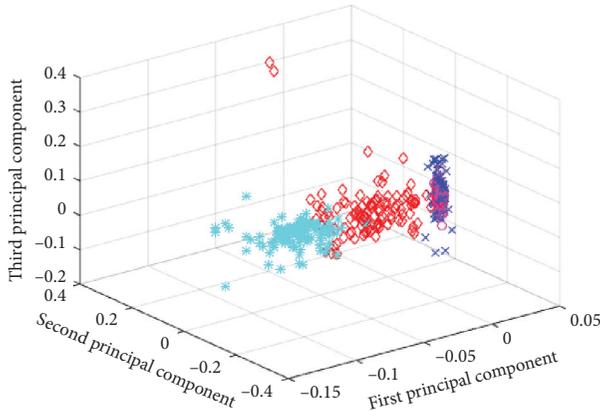


FIGURE 25: Recognition result of test signal based on LLE.

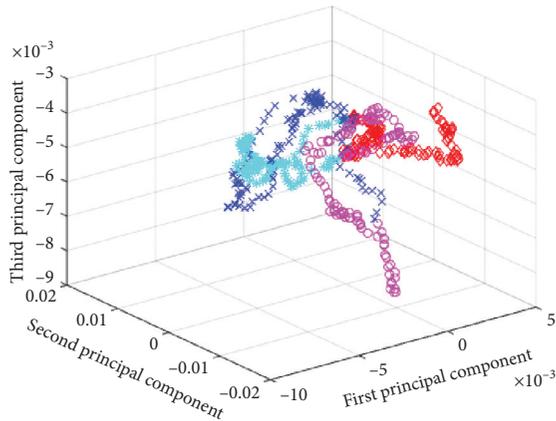


FIGURE 26: Recognition result of test signal based on LPP.

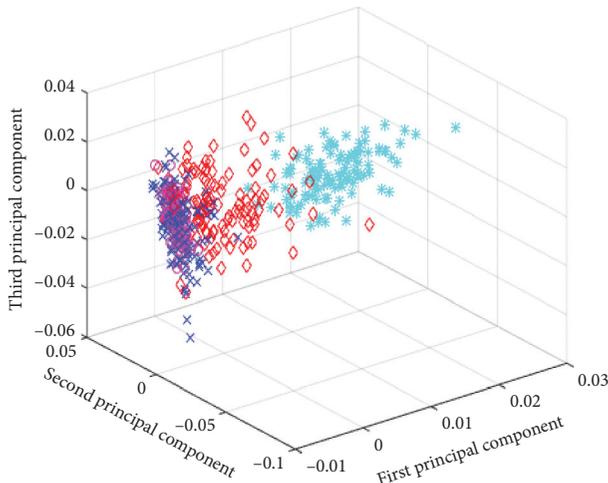


FIGURE 27: Recognition result of test signal based on LE.

8. Conclusions

The manifold learning algorithm is good for mechanical fault diagnosis based on data eigenvalues, but if the original vibration signal is directly used for fault diagnosis, the correct recognition rate is greatly reduced. To overcome this shortcoming, MLDA algorithm is proposed. In the paper, related theories and models are proposed, algorithm steps are given, and algorithm verification is carried out through simulation and experiment [23]:

- (1) The paper has theoretical innovation, puts forward the concept that the data points have mass, and uses the idea of gravitation between the particles to construct the virtual gravitation Laplacian matrix of the data point set. This concept provides a feasible way for conducting data-driven fault diagnosis analysis
- (2) The paper gives the Laplacian discriminant projection analysis method to find the optimal mapping function that can be classified; that is, the projection function is the optimal mapping function when the virtual gravity between different data sets is the largest and the virtual gravity between the similar data is the smallest. Feature extraction of data and fault classification using softmax algorithm. From the gear failure simulation and experiment, it can be seen that the extracted faults can distinguish the fault types.
- (3) From the simulation analysis results, the distance coefficient and noise level have a certain influence on the classification effect of the algorithm. When the appropriate distance coefficient is taken and the signal-to-noise ratio is good, the effect of the algorithm on the classification effect is not very large, but when the noise increases to a certain extent, even if the appropriate distance coefficient is selected, the classification effect will drop sharply. Therefore, it is necessary to further consider how to improve the antinoise performance of the algorithm in the future
- (4) From the experimental results, for the nonfeature value of the vibration fault signal, compared with the commonly used manifold learning algorithms such as LLE, LE, and LPP, the classification of the Mass Laplacian Discriminant Analysis algorithm is the best

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (grant no. 51575178), Hunan Natural Science Foundation of China (grant no. 2018JJ2120), and the

Non-Funded Science and Technology Public Relations Plan Project Foundation of Zhanjiang (grant no. 2017B01092).

References

- [1] M. N. Rahman, A. Esmailpour, and J. Zhao, "Machine learning with big data an efficient electricity generation forecasting system," *Big Data Research*, vol. 5, pp. 9–15, 2016.
- [2] L. Cui, N. Wu, C. Ma, C. Ma, and H. Wang, "Quantitative fault analysis of roller bearings based on a novel matching pursuit method with a new step-impulse dictionary," *Mechanical Systems and Signal Processing*, vol. 68-69, pp. 34–43, 2016.
- [3] W. Wang, Y. Zi, B. Chen, F. Wu, and Z. He, "Automatic fault feature extraction of mechanical anomaly on induction motor bearing using ensemble super-wavelet transform," *Mechanical Systems and Signal Processing*, vol. 54-55, pp. 457–480, 2015.
- [4] W. Wu, C. Little, and R. Court, "S-Transform and its contribution to wind turbine condition monitoring," *Renewable Energy*, vol. 62, pp. 137–146, 2014.
- [5] F. Gu, T. Wang, A. Alwodai, X. Tian, Y. Shao, and A. D. Ball, "A new method of accurate broken rotor bar diagnosis based on modulation signal bispectrum analysis of motor current signals," *Mechanical Systems and Signal Processing*, vol. 50-51, pp. 400–413, 2015.
- [6] F. Wang, B. Dun, X. Liu, Y. Xue, H. Li, and Q. Han, "An enhancement deep feature extraction method for bearing fault diagnosis based on kernel function and autoencoder," *Shock and Vibration*, vol. 2018, Article ID 6024874, 12 pages, 2018.
- [7] G. Wang, J. Luo, Y. He, and Q. Chen, "Fault diagnosis of supervision and homogenization distance based on local linear embedding algorithm," *Mathematical Problems in Engineering*, vol. 2015, Article ID 981598, 8 pages, 2015.
- [8] C. Sun, Z. Zhang, Z. He, Z. Shen, and B. Chen, "Manifold learning-based subspace distance for machinery damage assessment," *Mechanical Systems and Signal Processing*, vol. 70-71, pp. 637–649, 2016.
- [9] Y. Feng, B. Lu, and D. Zhang, "Multiscale singular value manifold for rotating machinery fault diagnosis," *Journal of Mechanical Science and Technology*, vol. 31, no. 1, pp. 99–109, 2017.
- [10] W. Guangbin, D. Moujun, H. Liangpei, and L. Long, "Prediction of bearing damage in wind turbines based on the quadratic root mean square of sub-band manifold," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 232, no. 18, pp. 3213–3223, 2017.
- [11] C. Sun, P. Wang, R. Yan, R. X. Gao, and X. Chen, "Machine health monitoring based on locally linear embedding with kernel sparse representation for neighborhood optimization," *Mechanical Systems and Signal Processing*, vol. 114, pp. 25–34, 2019.
- [12] G. Gao, F. Liu, and W. Huang, "Sparse discriminant manifold projections for bearing fault diagnosis," *Journal of Sound and Vibration*, vol. 399, no. 7, pp. 330–344, 2017.
- [13] Q. He and X. Wang, "Time-frequency manifold correlation matching for periodic fault identification in rotating machines," *Journal of Sound and Vibration*, vol. 332, no. 10, pp. 2611–2626, 2013.
- [14] Q. He, "Time-frequency manifold for nonlinear feature extraction in machinery fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 200–218, 2013.
- [15] Y. Wang, G. Xu, L. Liang, and K. Jiang, "Detection of weak transient signals based on wavelet packet transform and manifold learning for rolling element bearing fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 54-55, pp. 259–276, 2015.
- [16] P. Jiang and A. Elgammal, "A fast algorithm for manifold learning by posing it as a symmetric diagonally dominant linear system," *Applied and Computational Harmonic Analysis*, vol. 40, no. 3, pp. 622–628, 2016.
- [17] R. Chen, S. Chen, L. Yang, J. Wang, X. Xu, and T. Luo, "Looseness diagnosis method for connecting bolt of fan foundation based on sensitive mixed-domain features of excitation-response and manifold learning," *Neurocomputing*, vol. 219, pp. 376–388, 2017.
- [18] B. Wang, T. Song, F. Li, and L. Deng, "Fault diagnosis for a wind turbine transmission system based on manifold learning and Shannon wavelet support vector machine," *Renewable Energy*, vol. 62, pp. 1–9, 2017.
- [19] J. Yu, M. Liu, and H. Wu, "Local preserving projections-based feature selection and Gaussian mixture model for machine health assessment," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 225, no. 7, pp. 1703–1717, 2011.
- [20] W. Yang, P. Zhang, and D. Wu, "Fault diagnosis based on improved semi-supervised locally preserving projections," *Journal of Central South University*, vol. 46, no. 6, pp. 2059–2064, 2015.
- [21] Y. Dong and H. Gu, "Wind turbine fault early warning method based on insurance projection and self-organizing map," *Journal of Solar Energy*, vol. 66, no. 5, pp. 1123–1128, 2015.
- [22] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation*, vol. 15, no. 6, pp. 1373–1396, 2003.
- [23] V. d. Silva and J. B. Tenenbaum, "Global versus local methods in nonlinear dimensionality reduction," in *Proceedings of the Conference on Neural Information Processing Systems*, vol. 15, pp. 705–712, Piscataway, NJ, USA, December 2003.