

Research Article

Stability Evaluation of Fault Diagnosis Model Based on Elliptic Fourier Descriptor

YanZhu Hu , **Yu Hu** , **XinBo Ai**, **HuiYang Zhao**,
Zhen Meng, **WenJia Tian**, and **Jiao Wang** 

School of Automation, Beijing University of Posts and Telecommunications, No. 10 Xi Tu Cheng Road, Beijing/100876, China

Correspondence should be addressed to Yu Hu; hu_yu@bupt.edu.cn

Received 1 June 2018; Accepted 24 July 2018; Published 15 November 2018

Academic Editor: Yong Chen

Copyright © 2018 YanZhu Hu 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.

The performance evaluation of fault diagnosis algorithm is an indispensable link in the development and acceptance of the fault diagnosis system. Aiming at the stability evaluation of the fault diagnosis model based on the characteristic clustering, an image edge detection method based on the Elliptic Fourier Descriptor (EFDSE) is proposed to evaluate the stability of the fault diagnosis model, which applies similarity measurement of image to effective evaluation of faulty diagnosis algorithm. The quantitative evaluation index of the diagnostic capability of characterization based cluster fault diagnosis model is used to provide reference for the acceptance and reliability of the diagnosis results. Finally, the effectiveness of the stability evaluation is verified by the fault data of the motor bearings.

1. Introduction

With the development of modern industrial technology and information technology, manufacturing systems in various fields such as new energy, communication, computer, and industry are becoming more and more complex. Due to the complexity of the structure and the influence of various potential factors, the system inevitably exists as the hidden trouble. Once the hidden danger is induced, the personnel and economic loss of different degree will be caused. Therefore, the method of system fault diagnosis has become the focus of researchers. There are three common fault diagnosis methods: fault diagnosis based on control model [1], diagnosis based on statistical method [2], and fault diagnosis based on Artificial Intelligence [3]. At present, a large number of studies have focused on optimizing the stability of the fault diagnosis model, usually measured by the degree of diagnosis, and the higher the stability of the model in practical application is, the more cost will be paid. Therefore, it is necessary to analyse the effect of model stability on the effectiveness of fault diagnosis. However, there is still a lack of unified system for measuring the accuracy of models. The main methods are relative deviation and residual squared sum method [4] which are error analysis method,

grey correlation theory [5] and ED train based on statistical data, Confidence interval [6], etc. The grey correlation theory can realize the diagnosis of multi data input, but it can only be applied within the range of the same characteristic parameters and cannot be used to compare with fault diagnosis with parameters of different feature ranges. Residual squared sum method is an evaluation method for regression model. It is not conducive to relative comparison between different fault diagnosis models, which is influenced by the absolute value of the dependent variable and the independent variable. The confidence interval index is based on the hypothesis that the result of the training data group's diagnosis conforms to the normal distribution. It will produce a large number of errors in the case of small data, and the upper limit of confidence interval does not converge to 1 with the increase of the accuracy. Therefore, it is not suitable for models with high accuracy in fault diagnosis.

On the basis of the above research, considering the true distribution of the fault diagnosis output, the stability evaluation method of fault diagnosis model based on Elliptic Fourier Descriptor is proposed, which apply similarity measurement to evaluation of faulty diagnosis algorithm and can provide objective evaluation without understanding the parameters of the fault diagnosis model, when using the fault

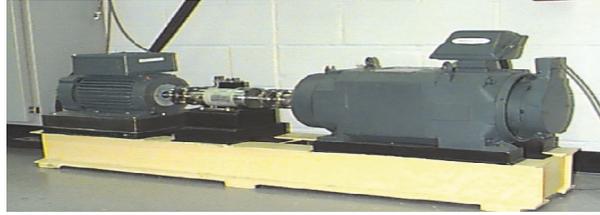


FIGURE 1: An experimental platform for fault simulation of motor bearing.

TABLE 1: Partial drive end and fan end of bearing fault data.

i	DE	FE	i	DE	FE	i	DE	FE
1	-0.0028	-0.2472	13	-0.2110	0.0119	26	0.2659	0.2786
2	-0.0963	0.1428	14	-0.0468	0.1824	27	0.0218	-0.0351
3	0.1137	0.0033	15	0.2021	-0.0754	28	-0.1639	0.1545
4	0.2573	-0.1068	16	-0.0145	-0.0536	29	0.1420	0.1999
5	-0.0583	0.1360	17	-0.1628	0.0805	30	0.2311	-0.0980
6	-0.1260	-0.0051	18	0.1092	-0.1389	31	-0.1075	0.0105
7	0.2074	-0.0625	19	0.1871	-0.0555	32	-0.1402	0.0988
8	0.1727	0.2735	20	-0.1593	0.1650	33	0.1915	-0.0555
9	-0.2199	0.1473	21	-0.1377	-0.1582
10	-0.1561	-0.0925	22	0.2505	0.0370
11	0.2240	0.1709	23	0.1075	0.3096
12	0.1137	0.0427	24	-0.2469	-0.0857	n

diagnosis model based on the feature clustering to diagnosis faulty. The application example of motor bearing diagnosis is compared and verified, which proves that the method proposed in this paper is effective.

2. System Description and Model

In order to train and verify the new technology and new theory of motor bearing fault, a motor bearing state evaluation system developed by Rockwell has obtained a series of motor performance database [7] which can be used to verify or improve the performance evaluation of motor. Some projects that have been or are making use of these databases include Winsnode state assessment technology, model-based diagnosis technology, and motor speed determination algorithm. The experimental platform is shown in Figure 1.

As shown in Figure 1, the train stand consists of a 2 hp motor (left), a torque transducer/encoder (center), a dynamometer (right), and control electronics (not shown). The train bearings support the motor shaft. Single point faults were introduced to the train bearings using electrodischarge machining with fault diameters of 7 mils, 14 mils, 21 mils, 28 mils, and 40 mils (1 mil=0.001 inches). See FAULT SPECIFICATIONS for fault depths. SKF bearings were used for the 7, 14, and 21 mils diameter faults, and NTN equivalent bearings were used for the 28 mil and 40 mil faults. Drive end and fan end bearing specifications, including bearing geometry and defect frequencies, are listed in the BEARING SPECIFICATIONS.

Vibration data was collected using accelerometers, which were attached to the housing with magnetic bases.

Accelerometers were placed at the 12 o'clock position at both the drive end and the fan end of the motor housing. During some experiments, an accelerometer was attached to the motor supporting base plate as well. Vibration signals were collected using a 16 channel DAT recorder and were post processed in a Matlab environment. All data files are in Matlab (*.mat) format. Digital data was collected at 12,000 samples per second, and data was also collected at 48,000 samples per second for drive end bearing faults. Speed and horsepower data were collected using the torque transducer/encoder and were recorded by hand.

Outer raceway faults are stationary faults; therefore placement of the fault relative to the load zone of the bearing has a direct impact on the vibration response of the motor/bearing system. In order to quantify this effect, experiments were conducted for both fan and drive end bearings with outer raceway faults located at 3 o'clock (directly in the load zone), at 6 o'clock (orthogonal to the load zone), and at 12 o'clock.

Data files are in Matlab format. Each file contains fan and drive end vibration data as well as motor rotational speed. For all files, the following item in the variable name indicates

DE: drive end accelerometer data

FE: fan end accelerometer data

BA: base accelerometer data

time: time series data

RPM: rpm during training

Part of the data is shown in Table 1 part of the drive end bearing failure data.

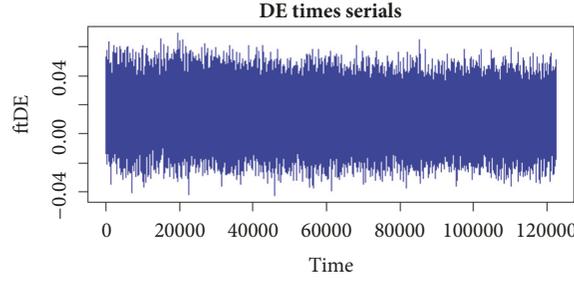


FIGURE 2: DE time series.

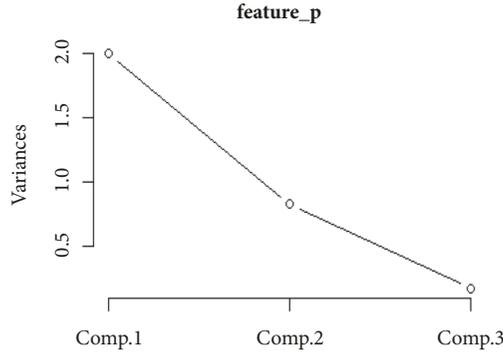


FIGURE 3: DE eigenvector principal component analysis.

3. Fault Diagnosis Model Based on Feature Clustering

Clustering analysis is a kind of unsupervised learning. It does not need to define the classes in advance or give a training sample to indicate what the data should have. Data sets can be divided into a number of different classes, and the intraclass data have very high similarity. This is very applicable where no standard information signs can be identified, such as fault diagnosis. Because some system parameters, environmental interference and noise are difficult to be confirmed accurately in the real environment, it is difficult to establish an accurate model of fault diagnosis model. Based on the data driven method, it avoids the mathematical modeling of the process and can be learned through historical data when the information of the diagnosis of the object mechanism is not clear. It can learn and model through historical data to complete the fault diagnosis. Commonly used clustering algorithms are Kmeans [8], BRICH [9], EM, DBSCN [10], CLARANS [11], etc. This paper mainly studies the stability evaluation method of fault diagnosis model (EFDSE). Therefore, Clara, Kmeans, and Dbscan are directly used for fault diagnosis of motor driven end bearing (DE) based on feature clustering. These clustering algorithms are distance-based clustering and density-based clustering, respectively. The data DE is a time series, as shown in Figure 2

In the process of fault diagnosis of motor bearings, the effect of fault feature extraction determines the final diagnosis rate. The peak to average ratio (PAR), kurtosis (KURTOSIS), and bias (SKEWNESS) of the vibration data cover the distribution features, statistical characteristics, and linear

characteristics of the vibration, which can effectively reflect the main characteristics of the vibration events. Therefore, this paper regards these three characteristics as the basis of fault diagnosis, specific calculation method, and specific calculation method (1)(2)(3):

$$PAR = E \left\{ \max \frac{(y^2)}{E(y^2)} \right\} \quad (1)$$

y: peak power

$$K = \frac{\int_{-\infty}^{+\infty} [x(t) - \bar{x}]^4 p(x) dx}{\sigma^4} \quad (2)$$

$$S = \frac{E(X - \bar{x})^3}{\sigma^3} \quad (3)$$

$x(t)$: instantaneous amplitude; \bar{x} : mean amplitude; $p(x)$: probability density; σ : standard deviation

Because of some eigenvectors of vibration may have certain correlation, the stability of clustering fault diagnosis models will be affected. Therefore, removing relevant eigenvectors is the first step to accomplish fault diagnosis. In this paper, principal component analysis (PCA) is used to extract unrelated feature vectors. The calculation results are shown in Figure 3. It can be seen that PAR and KURTOSIS can represent the feature vectors of DE, as shown in Table 2.

In order to effectively verify the reliability of the proposed evaluation method, this paper adopts DB (Davies-Bouldin) [12], Dunn Validity Index (DVI) [13], and Silhouette coefficient (SC) [14]. More than ten indexes of clustering evaluation are used to pre-evaluate the optimal number of

TABLE 2: DE eigenvector table.

i	PAR	KURT	i	PAR	KURT	i	PAR	KURT
1	3.2648	-0.3647	10	3.1979	-0.1333	19	3.4168	1.0757
2	3.0934	-0.3537	11	3.1780	-0.2197	20	3.7441	1.3899
3	3.0752	-0.4350	12	3.2202	-0.2486	21	3.7571	1.5904
4	3.0823	-0.4159	13	3.8729	1.5753	22	3.7994	1.5775
5	3.2896	0.0833	14	3.7868	1.7356	23	3.7263	1.3241
6	3.3391	0.2690	15	3.6880	1.5991	24	3.5290	0.6109
7	3.3191	0.1479	16	3.6230	1.4547
8	3.2075	0.2307	17	4.1758	1.8610
9	3.5993	-0.1001	18	3.5839	1.7459	n

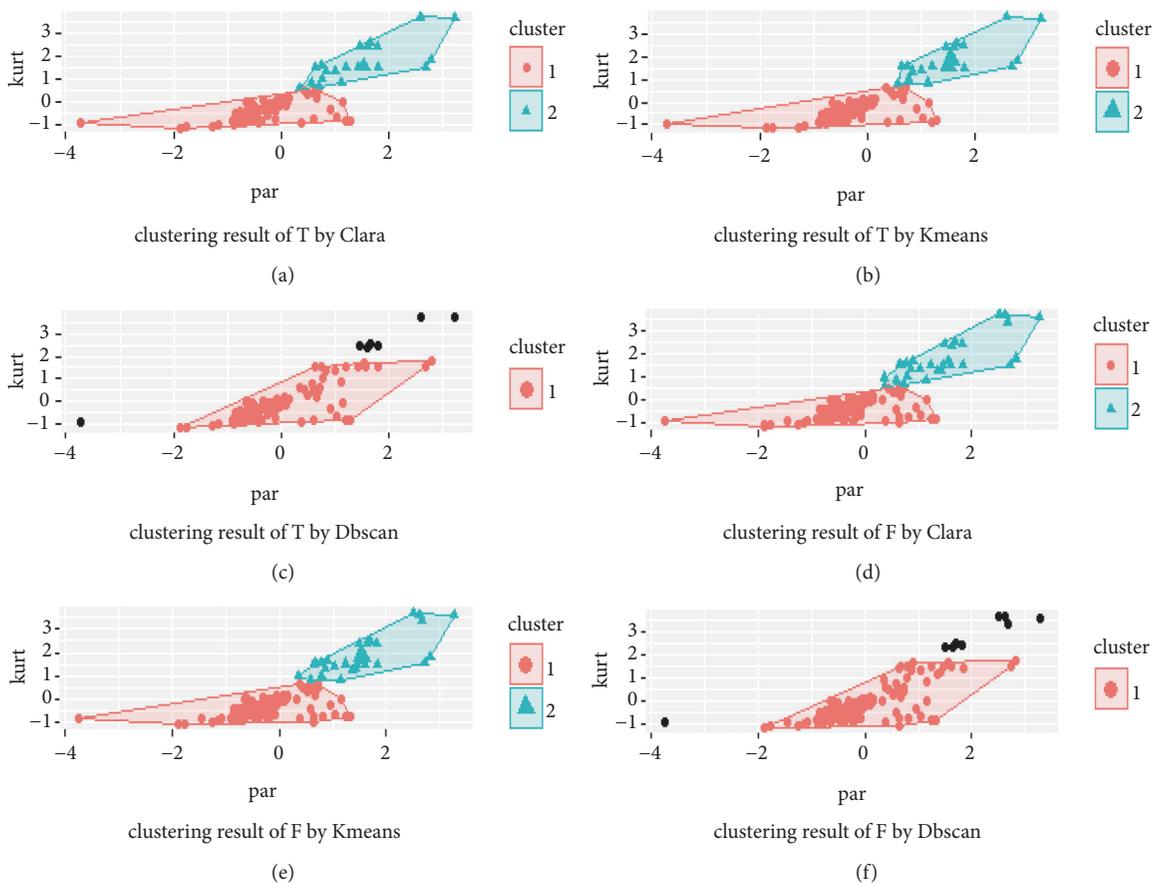


FIGURE 4: (a) Results of Clara clustering of T sets. (b) Results of Kmeans clustering of T sets. (c) Results of Dbscan clustering of T sets. (d) Results of Clara clustering of F sets. (e) Results of Kmeans clustering of F sets. (f) Results of Dbscan clustering of F sets.

DE eigenvectors, and motor bearing faults are classified into two categories. That is to say, Clara, Kmeans, and Dbscan will be used to divide the faults into two categories by means of Euclidean distance, respectively. In order to verify the effectiveness of the proposed method, the training sets T is randomly extracted from the feature sets F according to the proportion of 3:1, and then the training set T and the feature set F are divided into two categories, respectively. The results of the classification are as in Figure 4.

The first class data corresponding to the training set T and the feature set F are extracted respectively. See Figure 5.

In order to evaluate the stability of the fault diagnosis model based on feature clustering, the 2D points need to be mapped to a graphic based on a certain rule. This paper will use Dirichlet tessellation to accomplish points mapped. The Dirichlet (Delaunay) mosaic, also known as Voronoi Diagram or Thiessen Polygon, is a structure of computational geometry, which can be used for qualitative analysis, statistical analysis, and adjacent analysis [15]. In this paper, the Euclidean distance between any two points in the first category is computed. Any point is seen as the vertex of a triangle will be connected to two nearest points of Euclidean

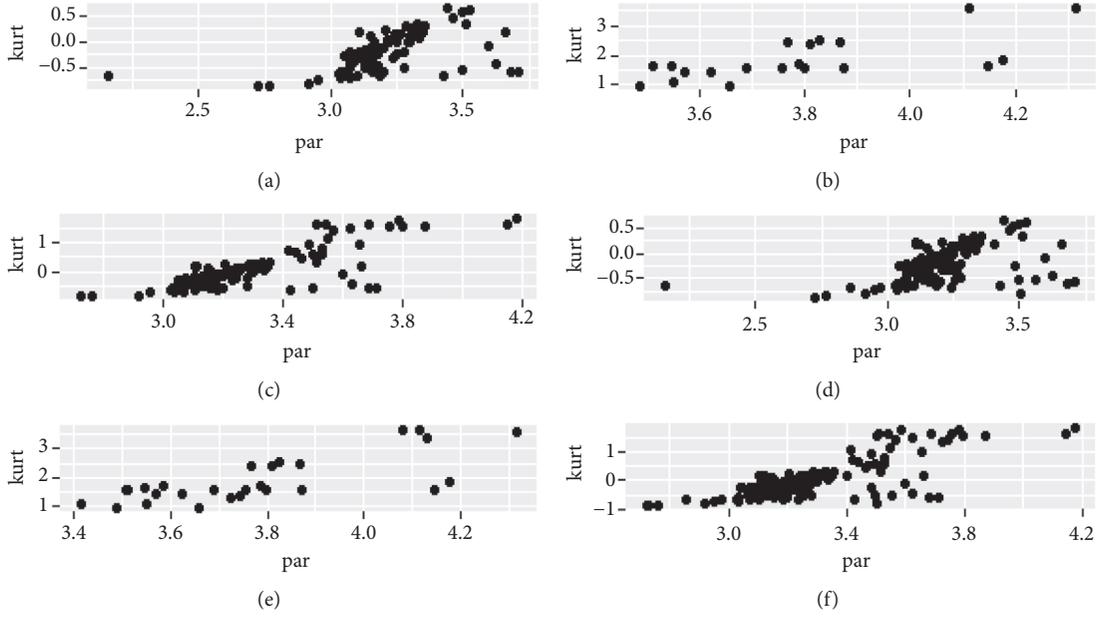


FIGURE 5: (a) The class I of Clara clustering of T sets. (b) The class I of Kmeans clustering of T sets. (c) The class I of Dbscan clustering of T sets. (d) The class I of Clara clustering of F sets. (e) The class I of Kmeans clustering of F sets. (f) The class I of Dbscan clustering of F sets.

distance, and Delaunay triangulation can be obtained by N iterations. All triangle of the common point is recorded, the circle center of the triangle is found, and the clockwise connection of the center of the center is the corresponding Thiessen Polygon. The time complexity of the triangle is the complexity of the polygon. The algorithm flow is as follows.

Hypothesis: $(X, Y) = \{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)\}$. It represents the point set P consisting of N nonrepeating points on the plane, and the specific steps of constructing the Delaunay triangulation of point set P is as follows.

The N points are mainly based on X coordinates and are sorted by Y coordinates.

Step 1. The N points are sorted by mainly based on X coordinates.

Step 2. Structural process is as follows:

- (i) If $N = 2$, return
- (ii) If $N = 3$, three points are connected to construct a triangulation net and return
- (iii) The N points are divided into subsets P_l and P_r on the basis of evenly principle or nearest neighbor principle
- (iv) Construct triangular net $DT(P_l)$ of P_l
- (v) Construct triangular net $DT(P_r)$ of P_r
- (vi) $DT(P_l)$ merge with $DT(P_r)$ and put back

Step 3. Merge process

- (i) For given $DT(P_l)$ and $DT(P_r)$, calculate convex hull of P_l and P_r
- (ii) Obtain the top tangent UCT and the bottom tangent BCT

- (iii) Start from BCT , according to left endpoint, right endpoint, and their adjacent points to complete $DT(P_l)$ and merge with $DT(P_r)$ until the UCT is encountered.

The Voronoi diagram of the class I of T sets and F sets is shown in Figure 6.

4. EFDSE Algorithm

4.1. Edge Detection. Canny edge detection operator has obvious advantages compared with Roberts Cross operator, Prewitt operator, Sobel operator, and Kirsch operator. So in this article, edge detection is performed by Canny operator to identify its contour boundaries.

First, in order to smooth the image to reduce the obvious noise influence on the edge detector, the image adopted is the Gauss filter to check the image by Gauss filter whose size is 2×2 , as follows:

$$H_{ij} = \frac{1}{2\pi} \exp\left(-\frac{(i - (k+1))^2 + (j - (k+1))^2}{2\sigma^2}\right) \quad (4)$$

If a window A whose size is 3×3 in the image is, pixel e will be filtered. Then, after Gauss filtering, the brightness value of the pixel e is

$$e = H * A = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} * \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad (5)$$

$$= \text{sum} \left(\begin{bmatrix} a \times h_{11} & b \times h_{12} & c \times h_{13} \\ d \times h_{21} & e \times h_{22} & f \times h_{23} \\ g \times h_{31} & h \times h_{32} & i \times h_{33} \end{bmatrix} \right)$$

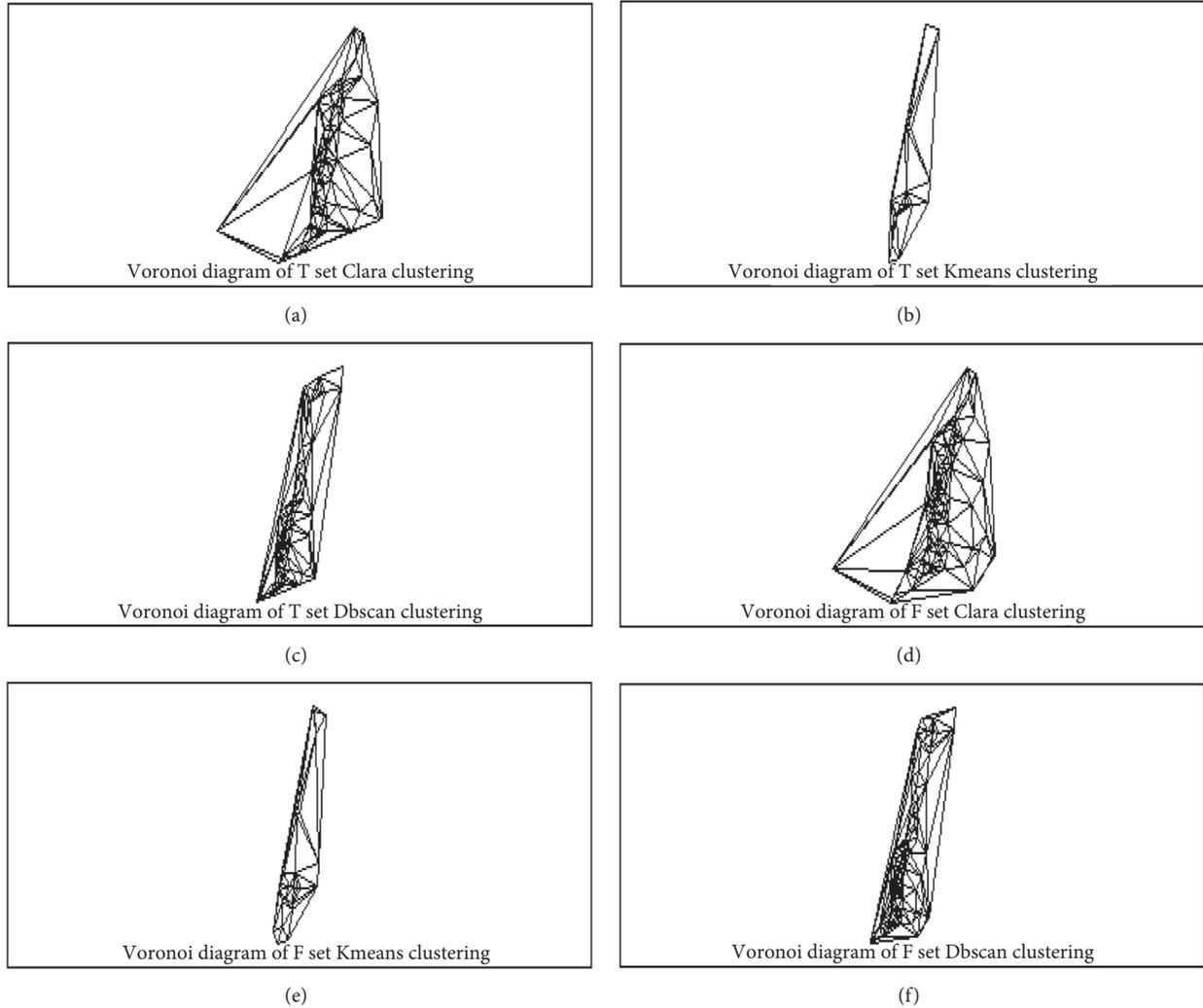


FIGURE 6: (a) Voronoi diagram of T set clara clustering. (b) Voronoi diagram of T set kmeans clustering. (c) Voronoi diagram of T set dbSCAN clustering. (d) Voronoi diagram of F set clara clustering; (e) Voronoi diagram of F set kmeans clustering. (f) Voronoi diagram of F set DbSCAN clustering.

where $*$ is a convolution symbol; *sum*: the sum of all elements in the representation matrix. The Canny algorithm uses four operators to detect the horizontal, vertical, and diagonal edges of the image. The operator of the edge detection returns the first-order value θ of the horizontal and vertical directions G_y ; thus the gradient intensity G and θ which is direction gradient of the pixels can be determined.

$$G = \sqrt{G_x^2 + G_y^2} \quad (6)$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right)$$

The gradient strength of the current pixel is compared with the two pixels along the positive and negative gradient direction. If the gradient intensity of the current pixel is maximum compared to the other two pixels, the pixel is retained as the edge point; otherwise the pixel is suppressed, which is called the maximum value suppression. After exerting

maximum suppression, there are still some edge pixels caused by the change of noise and color. In order to solve these stray responses, the selection of high and low threshold is established. If edge pixels are higher than the high threshold, the edge pixels are marked as strong edge pixels; if the gradient value of the edge pixels is less than the high threshold is larger than the low threshold, the edge pixels are recorded as the weak edge pixels, but if the 8 neighborhood pixels of the weak edge pixels have one strong edge pixel, they can be retained as edge points; if the edge pixels are pixels, the edge pixels can be retained as edge pixels. The gradient is less than the low threshold, and it is suppressed. Figure 6 is detection by Canny operator. Their Edge contour IS detected (see Figure 7).

4.2. Fourier Descriptor. Shape is one of the most important visual features of a target. The existing shape representation methods can be divided into two categories: shape representation based on region feature and shape representation based

$$\begin{aligned}
&= \frac{a_0}{2} + \sum_{k=1}^{\infty} \left[\frac{a_k - jb_k}{2} e^{jkwt} + j \frac{a_k + jb_k}{2} e^{-jkwt} \right] \\
&= c_0 + \sum_{k=1}^{\infty} [c_k e^{jkwt} + c_{-k} e^{-jkwt}] = \sum_{k=1}^{\infty} [c'_k e^{jkwt}]
\end{aligned} \tag{8}$$

If we define $c_k = c_{k1} - jc_{k2}$, $c_{-k} = c_{k1} + jc_{k2}$
Then, (8) can be derived by

$$\begin{aligned}
c(t) &= c_0 + 2 \sum_{k=1}^{\infty} \left[c_{k1} \left(\frac{e^{jkwt} + e^{-jkwt}}{2} \right) \right. \\
&\quad \left. - jc_{k2} \left(\frac{-e^{jkwt} + e^{-jkwt}}{2} \right) \right] = c_0 \\
&\quad + 2 \sum_{k=1}^{\infty} [c_{k1} \cos(kwt) + c_{k2} \sin(kwt)] = c_0 \\
&\quad + \sum_{k=1}^{\infty} [a_k \cos(kwt) + b_k \sin(kwt)]
\end{aligned} \tag{9}$$

where a_k and b_k are said to Fourier Descriptor

Then, c_k and c_{-k} can be derived by (8) and (9)

$$\begin{aligned}
c_k &= \frac{a_k - jb_k}{2}, \\
c_{-k} &= \frac{a_k + jb_k}{2}
\end{aligned} \tag{10}$$

The coefficients in (9) can be obtained by considering the orthogonal property. Thus, one way to compute values for the descriptors is

$$\begin{aligned}
a_k &= \frac{2}{T} \int_0^T c(t) \cos(kwt) dt, \\
b_k &= \frac{2}{T} \int_0^T c(t) \sin(kwt) dt
\end{aligned} \tag{11}$$

4.3. Fourier Description of the Edge Features of Fault Classification. Determining a starting point (x_0, y_0) of the target boundary and moving along the counter clockwise direction at a certain speed, the boundary of the boundary point coordinates can be used to describe the boundary. The cluster boundary curve of the first class data set is defined as

$$s(t) = x(t) + jy(t), \quad t = 0, 1, \dots, N-1 \tag{12}$$

t is the unit arc length along with boundary circle. In order to describe the outline of the image, the selected starting point needs to circle along the boundary curve. So, $s(t)$ is a periodic function of a period in which periodic is 2π . In order to obtain the Elliptic Fourier Descriptor of the boundary curve, Fourier series expansion is first carried out, and it can be expanded by 1D Fourier series.

$$s_k = \frac{1}{T} \int_0^T s(t) e^{-jkw t} dt \tag{13}$$

Then an expression of ellipse coefficient can be computed by (12)

$$s_k = s_{xk} + js_{yk} \tag{14}$$

Then

$$\begin{aligned}
s_{xk} &= \frac{1}{T} \int_0^T x(t) e^{-jkwt} dt, \\
s_{yk} &= \frac{1}{T} \int_0^T y(t) e^{-jkwt} dt
\end{aligned} \tag{15}$$

According to the relationship between trigonometric function and exponential function, there are

$$\begin{aligned}
s_{xk} &= \frac{a_{xk} - jb_{xk}}{2}, \\
s_{yk} &= \frac{a_{yk} - jb_{yk}}{2}, \\
s_{x-k} &= \frac{a_{xk} + jb_{xk}}{2}, \\
s_{y-k} &= \frac{a_{yk} + jb_{yk}}{2}
\end{aligned} \tag{16}$$

Then

$$\begin{aligned}
a_{xk} &= \frac{2}{m} \sum_i^m x_i \cos(kwit), \\
b_{xk} &= \frac{2}{m} \sum_i^m x_i \sin(kwit), \\
a_{yk} &= \frac{2}{m} \sum_i^m y_i \cos(kwit), \\
b_{yk} &= \frac{2}{m} \sum_i^m y_i \sin(kwit)
\end{aligned} \tag{17}$$

m is the number of sampling points in the contour curve, it is usually the half of the number of pixels in the contour curve, and x_i and y_i is the value at the sample point $x(t)$ and $y(t)$ when they lie in i . According to (16) and (17), s_k can be regarded as the sum of complex numbers. That is,

$$\begin{aligned}
s_k &= A_k - jB_k, \\
s_{-k} &= A_k + jB_k
\end{aligned} \tag{18}$$

Here

$$\begin{aligned}
A_k &= \frac{a_{xk} + ja_{yk}}{2}, \\
B_k &= \frac{b_{xk} + jb_{yk}}{2}
\end{aligned} \tag{19}$$

Equation (13) can be expressed:

$$\begin{aligned}
c(t) &= c_0 \\
&\quad + \sum_1^{\infty} (A_k - jB_k) e^{jkwt} + \sum_{-\infty}^{-1} (A_k + jB_k) e^{jkwt}
\end{aligned} \tag{20}$$

TABLE 3: Elliptical Fourier Description on the edge of fault classification.

Training set T	0.0039	0.0079	0.0157	0.0196	0.0235	0.0275	0.0314	0.0353
Feature set F	0.0034	0.068	0.0101	0.0135	0.0169	0.0203	0.0237	0.0271

TABLE 4: The index of stability Clara Kmeans Dbscan.

Index	Clara	kmeans	Dbscan
within.cluster.ss	26.26355	25.70769	49.13795
avg.silwidth	0.719723	0.734621	0.706651
EFDSE	0.8162398	0.993007	0.6741096

Then, (21) is called Elliptical Fourier Description

$$EFD = \frac{|A_k|}{|A_1|} + \frac{|B_k|}{|B_1|} \quad (21)$$

According to (21), the Ellipse Fourier Descriptor of the fault classification contour curve is to be obtained and is normalized, shown in Table 3.

4.4. Stability Evaluation of Fault Diagnosis Model Based on Elliptical Fourier Descriptor. Assume that the class I contour edge descriptor of the training set T and the feature set F is *EFD1* and *EFD2*, respectively. This paper shows that if the stability of the fault diagnosis model based on the feature clustering is good, and the cohesiveness of the class center is stronger. That is to say, when adding or removing the same characteristic of data to a certain class, changed degree of boundary shape of the cluster is very small and vice versa. Therefore, we define the similarity of the contour shapes of the two fault classification results defined as the stability evaluation criteria of the fault diagnosis model, as shown in

$$sim = \frac{|\text{cov}(EFD1, EFD2)|}{\sigma_{EFD1} * \sigma_{EFD2}} \in [0, 1] \quad (22)$$

$\text{cov}()$ represents the covariance of the two descriptors; σ represents the standard deviation of the descriptor vectors. The range of *sim* is $[0, 1]$. Value of *sim* is close to 1; then stability of fault diagnostic is better.

5. Experimental Results and Discussion

The clustering results are usually verified by two kinds of techniques: one is the intracluster distance such as the within.cluster.ss calculation that is the square of each internal distance. The more similar the characteristics of the data in cluster, the better the clustering effect. The other is the distance between clusters such as the average contour coefficient calculated by avg.silwidth. The larger the value, the larger the difference of the data feature of different classes and the better the data area diversity of the clustering algorithm. The within.cluster.ss index and avg.silwidth index of the three clustering results are compared with the EFDSE index in this paper (show in Table 4). As seen from the table, EFDSE indicates that diagnosis effect of Kmeans is the best, Clara is the second, and the Dbscan is the worst. It is consistent with

the conclusion of avg.silwidth and within.cluster.ss, which proves that proposed EFDSE method in this paper is effective for the stability evaluation of the fault diagnosis model based on the feature clustering.

6. Conclusion

EFDSE that map the fault classification results in 2D graphic, using graphic edge detection technology. By extracting the feature vectors of the contour curves of fault classification results, the contour shape similarity is calculated to evaluate the effect of fault diagnosis. It is a new method of stability evaluation based on feature clustering fault diagnosis model. It applies similarity measurement of image to valuation of faulty diagnosis algorithm. In the case of unknown data samples and data methods, the stability of the model fault diagnosis effect is evaluated only by the visual contour feature vectors of the fault classification results. From the experimental method and principle, the evaluation is applicable to the stability evaluation of fault diagnosis models based on feature clustering. But used clustering algorithm should be distance-based clustering and density-based clustering. Making EFDSE is fit for more and more faulty diagnosis method is our work direction in future.

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.

Acknowledgments

The authors thank all challenge participants for their invaluable contribution. This paper is supported by National key research and development plan (2016YFC0701309) and by the National Natural Science Foundation of China (61627816).

References

- [1] W. Wang, L. Cui, and D. Chen, "Multi-scale morphology analysis of acoustic emission signal and quantitative diagnosis

- for bearing fault,” *Acta Mechanica Sinica*, vol. 32, no. 2, pp. 265–272, 2016.
- [2] X. Chen, S. Wang, B. Qiao, and Q. Chen, “Basic research on machinery fault diagnostics: Past, present, and future trends,” *Frontiers of Mechanical Engineering*, vol. 13, no. 2, pp. 264–291, 2018.
- [3] M. Witczak, M. Mrugalski, and J. Korbicz, “Towards Robust Neural-Network-Based Sensor and Actuator Fault Diagnosis: Application to a Tunnel Furnace,” *Neural Processing Letters*, vol. 42, no. 1, pp. 71–87, 2015.
- [4] Du. Kun, *water supply network hydraulic model checking and leakage location research*, Chongqing University, 2014.
- [5] K. S. Prasad, S. R. Chalamalasetti, and N. R. Damera, “Application of grey relational analysis for optimizing weld bead geometry parameters of pulsed current micro plasma arc welded inconel 625 sheets,” *The International Journal of Advanced Manufacturing Technology*, vol. 78, no. 1-4, pp. 625–632, 2015.
- [6] Y. Xiyun, L. Yuqi, and L. Jianlin, “method of calculating confidence interval of energy storage photovoltaic power station based on four division method,” *Journal of electrical technology*, vol. 15, pp. 136–144, 2017.
- [7] *Fault data of motor bearing*, 2009, <http://csegroups.case.edu/bearingdatacenter/pages/download-data-file>.
- [8] S. H. Steinberg, “United States of America,” p. 765, 1954.
- [9] T. Zhang, R. Ramakrishnan, and M. Livny, “BIRCH: a new data clustering algorithm and its applications,” *Data Mining and Knowledge Discovery*, vol. 1, no. 2, pp. 141–182, 1997.
- [10] T. Marwala, *Condition Monitoring Using Computational Intelligence Methods*, Springer London, London, 2012.
- [11] A. Hinneburg and D. A. Keim, “A General Approach to Clustering in Large Databases with Noise,” *Knowledge and Information Systems*, vol. 5, no. 4, pp. 387–415, 2003.
- [12] D. L. Davies and D. W. Bouldin, “A cluster separation measure,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 2, pp. 224–227, 1978.
- [13] J. C. Dunn, “A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters,” *Journal of Cybernetics*, vol. 3, no. 3, pp. 32–57, 1973.
- [14] A. F. Mcdaid, D. Greene, and N. Hurley, *Normalized Mutual Information to evaluate overlapping community finding algorithms*, Computer Science, 2011.
- [15] J. E. Goodman, *Handbook of discrete and computational geometry*, J. O’Rourke, Ed., CRC Press Series on Discrete Mathematics and its Applications, Boca Raton, FL, USA, 1997.
- [16] M. Nikolic, E. Tuba, and M. Tuba, “Edge detection in medical ultrasound images using adjusted Canny edge detection algorithm,” in *Proceedings of the 24th Telecommunications Forum, TELFOR 2016*, Serbia, November 2016.
- [17] H. A. A. Alasadi and M. A. Abed, *Object Recognition Using Artificial Fish Swarm Algorithm on Fourier Descriptors*, Social Science Electronic Publishing, 2017.
- [18] D. Zhang and G. Lu, “Enhanced generic Fourier descriptors for object-based image retrieval,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2011.
- [19] S. S. Li, Y. D. Huang, and J. W. Yang, “Affine Invariant Ring Fourier Descriptors,” *International Conference on Wavelet Analysis and Pattern Recognition*, 2013.
- [20] B. Belkhaoui, A. Toumi, and A. Khalfallah, “Fusion Fourier descriptors from the EM,” *International Journal of Computer & Information Technology*, 2014.

