Research Article

Fault Diagnosis of Supervision and Homogenization Distance Based on Local Linear Embedding Algorithm

Guangbin Wang,[1] Jun Luo,[1] Yilin He,[1] and Qinyi Chen[2]

[1]*Hunan Provincial Key Laboratory of Health Maintenance for Mechanical Equipment, Hunan University of Science and Technology, Taoyuan Road No. 2, Xiangtan City, Hunan 411201, China*
[2]*ChuXiong Industrial School, Weichu Road No. 84, Chuxiong City, Yunnan 675099, China*

Correspondence should be addressed to Guangbin Wang; jxxwgb@126.com

In view of the problems of uneven distribution of reality fault samples and dimension reduction effect of locally linear embedding (LLE) algorithm which is easily affected by neighboring points, an improved local linear embedding algorithm of homogenization distance (HLLE) is developed. The method makes the overall distribution of sample points tend to be homogenization and reduces the influence of neighboring points using homogenization distance instead of the traditional Euclidean distance. It is helpful to choose effective neighboring points to construct weight matrix for dimension reduction. Because the fault recognition performance improvement of HLLE is limited and unstable, the paper further proposes a new local linear embedding algorithm of supervision and homogenization distance (SHLLE) by adding the supervised learning mechanism. On the basis of homogenization distance, supervised learning increases the category information of sample points so that the same category of sample points will be gathered and the heterogeneous category of sample points will be scattered. It effectively improves the performance of fault diagnosis and maintains stability at the same time. A comparison of the methods mentioned above was made by simulation experiment with rotor system fault diagnosis, and the results show that SHLLE algorithm has superior fault recognition performance.

1. Introduction

With the development of the modernization process, the structure of mechanical equipment has been more and more increasingly sophisticated while the degree of automation and function of realization grow increasingly stronger. In the fault diagnosis of rotating machinery, the more sophisticated the monitoring and control system, the more the numbers of sensors. What is more, the data of presenting state characteristics in real space is complex. These multiple variables make the data to present parameters of equipment running status become more complex which to describe the state of the data abstraction are the high-dimensional data. Faced with the characteristics of failure samples such as high-dimensional diversity, the traditional linear method has great limitations. However, the nonlinear manifold learning method was developed rapidly since it was first proposed in the journal of Science in 2000. From a huge amount of complicated and changeable high-dimensional observation data, the methods make data analysis and state decision-making extend from the original Euclidean space to the manifold, which is able to identify the key information accurately and dig out signal essential characteristics and internal rules, which will be analyzed and judged for fault diagnosis. Classic manifold learning methods such as isometric feature mapping (ISOMAP) [1], local linear embedding (LLE) [2], local tangent space alignment (LTSA) [3, 4], and Laplacian eigenmap (LE) [5, 6] algorithm are mainly applied to the fields of data mining, image processing, pattern recognition, and information retrieval [7–9]. In recent years, the manifold learning methods have been applied to many kinds of rotating machinery fault diagnosis. Yang et al. [10] put forward the noise reduction algorithm in nonlinear time series based on reconstruction of phase space and main manifold identification and successfully extracted the impact characteristics of gearbox fault signal from the noise; Yang et al. [11] put forward a method of incremental local tangent space alignment.

In the different states of rolling bearing classification recognition, the algorithm of incremental learning for new samples has higher recognition rate; Liang et al. [12] proposed impact fault feature adaptive extraction algorithm based on manifold learning and extracted the optimal impact fault feature. For the extraction of weak feature of early fault, Li et al. [13] proposed the early fault diagnosis and feature extraction methods of the rolling bearing based on manifold learning, which improved the fault pattern classification performance. However, with the complex diversities and particularity of fault samples, the most of manifold learning methods in the practical application are optimization methods, and part of them are characterized by just the combination of two kinds of the methods and the stability of method which needs to be improved. Therefore, in this study with rotating machinery as the research object, an improved local linear embedding algorithm of homogenization distance (HLLE) is proposed on the basis of the research of locally linear embedding (LLE) method to solve the problems that the real fault samples set are not distributed evenly and the phenomena that LLE is easily affected by the neighboring points. Moreover, in order to further improve the recognition performance rate and stability of the method, the supervised learning mechanism is added. And a new local linear embedding algorithm of supervision and homogenization distance (SHLLE) is proposed. Then the rotor system fault simulation experiment is used to analyze and validate the effectiveness of fault diagnosis.

2. Locally Linear Embedding Algorithm of Homogenization Distance

2.1. Locally Linear Embedding Algorithm. Euclidean distance is used in LLE algorithm for local linear fitting to show the overall topological structure, namely, to represent the global nonlinear structure characteristics according to the local linear fitting. LLE algorithm is a local weight matrix manifold learning algorithm with the assumption that observation data set is located in or approximately in the low-dimensional embedding manifold in the high-dimensional space. The basic idea of LLE is that the weight value of the data points in high-dimensional space can be best reconstructed and it can carry the local geometry information of manifold from high-dimensional space to low-dimensional space. LLE holds the opinion that constructed weight matrix can preserve the essence characteristics of the local neighborhood. That means the weight matrix can maintain the same geometric properties of local neighborhood of data set regardless of scaling or rotation.

The basic steps (described as **Figure 1**) of LLE algorithm are as follows:

Input. The original feature matrix $X_{D \times n}$ is composed of n numbers of D dimensional vector.

Output. The feature vector matrix $Y_{D' \times n}$ is composed of n numbers of D' dimensional vector ($D' \ll D$).

Step 1. Set the neighbor points (k) and (D) dimensional embedding values.
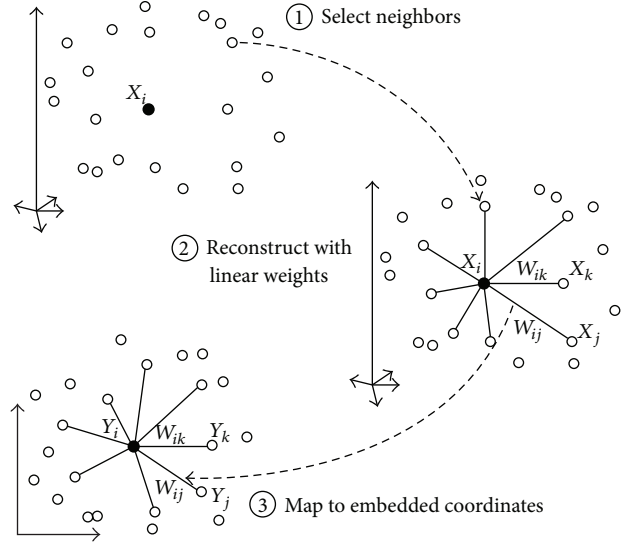

FIGURE 1: Steps of locally linear embedding algorithm.

Step 2. Calculate the Euclidean distance between any two points (x_i, x_j among $i, j \in (1, 2, 3, \ldots, n)$) in sample set $X = \{X_i, \ i = 1, 2, 3, \ldots, n\} \subset R^m$, choosing the nearest neighbor corresponding points to constitute the neighbor area based on neighbor point.

Step 3. In the neighbor area, W_{ij} represents the weights between x_i and x_j, calculating the best reconstruction weights W_{ij} of X_i to make sure that reconstruction error ($W_{ij} = \sum_{i=1} |x_i - \sum_{j=1} w_{ij} x_j|^2$) is minimum and constraint conditions is $\sum_j W_{ij} = 1$.

Step 4. Calculate D' embedded dimensional vector $Y_{D' \times n}$ by the weight matrix W_{ij} based on the minimum reconstruction error $W_{ij} = \sum_{i=1} |x_i - \sum_{j=1} w_{ij} x_j|^2$.

In LLE algorithm, nonlinear data set is smartly divided into the date representation with local linear structure to reduce dimension effectively for nonlinear date. The main parameters are neighbor points and embedding dimension. The problem of solving the least squares is transformed into the eigenvalues in calculating process; thus it reduces the amount of calculation. In general, LLE has advantages of less undetermined parameters, overall analysis optimal solution, smaller computational complexity, and direct geometric meaning.

2.2. Locally Linear Embedding Algorithm of Homogenization Distance Algorithm. It is a certain difficulty when original LLE algorithm is used to deal with complicated data samples in the engineering practice. Because the LLE method assumes that the distribution of sample points on the manifold is continuous dense and uniform sampling, using Euclidean distance directly in high dimensional space to select local neighborhood may not truly reflect the intrinsic nature of the manifold structure. When manifold shows the curls or

cascade, namely, the short distance between the two manifold surfaces, refactoring process will cause points with shorter distance from different surface to the same local neighborhood space, thus causing the distortion of the manifold structure. In addition, the numbers of neighbor points (k) have obvious effects on dimension reduction, as the local neighborhood composed of (k) nearest neighbor points on the sparse area distribution of the sample points is obviously bigger than that on the dense area, while choosing neighbor points will have difference. Therefore, an improved local linear embedding algorithm of homogenization distance (HLLE) is proposed. On the basis of the distance measurement between data in the Conformal-Isomap method and in accordance with the characteristics of practical fault samples, the following improving distance is used to replace the Euclidean distance in LLE algorithm:

$$D_{ij} = \frac{\|x_i - x_j\|}{\sqrt{M(i)M(j)}}. \tag{1}$$

Among them, $d_{ij} = \|x_i - x_j\|$ is the Euclidean distance between x_i and x_j, while $M(i)$ and $M(j)$, respectively, represent the average distance of x_i and x_j to the other points, namely,

$$\begin{aligned} M(i) &= \frac{1}{N-1}\sum_{\substack{k=1 \\ k\neq i}}^{N}\|x_i - x_k\|, \\ M(j) &= \frac{1}{N-1}\sum_{\substack{k=1 \\ k\neq j}}^{N}\|x_j - x_k\|. \end{aligned} \tag{2}$$

The purpose of homogenization distance is getting the relatively narrow distance between sample points in the relatively sparse area and the relatively increased distance in the dense area through calculating improved distance between sample points, so that the overall distribution of sample points tends to be homogeneous, which is helpful to the classification of sample set; namely, the change of distance makes categories much more classifiable and reduces dimension reduction effects of the neighbor points. Figure 2 is the effect diagram of distance measurement change.

The molecules of original Euclidean distance are constant in the homogenization distance formula (1); when sample points are distributed in sparse area, the distance between the point and other points is longer, and the average distance with other points is also longer. Therefore, the original distance divided by the larger denominator to get new distance is reduced accordingly. On the contrary, when sample points are distributed in dense area, the distance between the point and other points is shorter, and the average distance with other points is also shorter, which makes the new distance be increased accordingly. As a result, homogenization distance makes the overall distribution of sample points tend to be homogeneous. Additionally, the distribution area between two kinds of sample points also changes accordingly. The new distance in sparse area becomes shorter, while sparse area becomes downsized and dense area becomes relatively
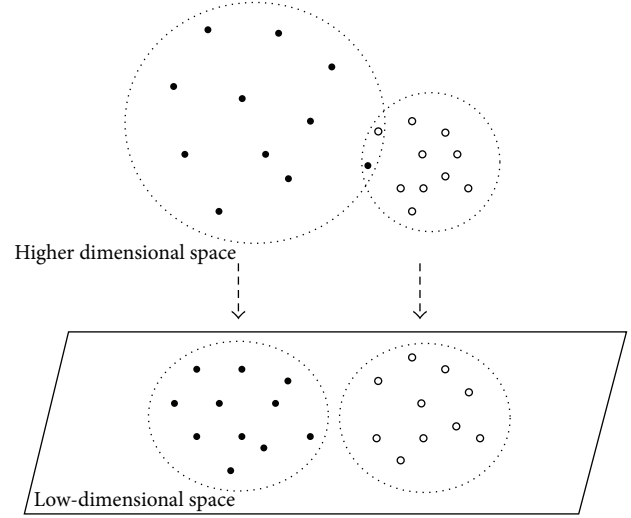

Figure 2: Effect diagram of distance measurement change.

upsized. Through distance formula, the difference of the new distance of sample points in the sparse and dense area can be released, namely, the ratio of the density.

(1) Sparse area:

$$A = \frac{d_{ij}}{D_{ij}} = \frac{\|x_i - x_j\|}{\|x_i - x_j\|/\sqrt{M(i)M(j)}} = \sqrt{M(i)M(j)}. \tag{3}$$

(2) Dense area:

$$B = \frac{d_{ij}}{D'_{ij}} = \frac{\|x_i - x_{j'}\|}{\|x_i - x_{j'}\|/\sqrt{M(i)M(j')}} = \sqrt{M(i)M(j')}. \tag{4}$$

Thereinto, x_j and $x_{j'}$ are neighbor point of x_i. If x_i is the same point, x_j and $x_{j'}$ are neighboring point in sparse area and dense regions, respectively, and the ratio of the density is as follows:

$$\frac{A}{B} = \frac{\sqrt{M(i)M(j)}}{\sqrt{M(i)M(j')}} = \frac{M(j)}{M(j')} > 1. \tag{5}$$

From formula (5), the change range ability of new distance in sparse area is bigger than that in dense area, so that sparse area can be separated from the dense one to a certain degree. In the subsequent algorithm process, the homogenization distance which chooses neighbor points according to k neighbor points can basically take the same kind of neighborhood. Thus, it can reduce the influence of the neighboring points and be more conducive to feature extraction. And gathering samples with similar features makes the samples classified effectively.

3. Locally Linear Embedding Algorithm of Supervision and Homogenization Distance

3.1. Supervised Locally Linear Embedding Algorithm. Locally linear embedding method is a higher efficient manifold method in nonlinear data dimension reduction method. But it is a nonlinear dimension reduction method without supervised learning in essence. The inadequate use of the samples category information results in a certain influence on the classification accuracy, which will not achieve an optimal effect if it is applied to areas such as classification. Therefore, de Ridder et al. [14] proposed SLLE algorithm [14]. Its main purpose is to increase the distance between classes and minimize the global reconstruction error of local data by reducing the distance within class. The key is to modify the calculation method of (k) neighbor points though adding a distance parameter to the different sample, thus increasing the category information of sample points. In SLLE algorithm, the following formula instead of Euclidean distance in LLE algorithm is used to build the neighborhood of sample points set for dimension reduction:

$$\Delta' = \Delta + \alpha \max(\Delta)\Lambda_{ij}. \tag{6}$$

From formula (6), Δ' is the distance that contains category information, Δ is the Euclidean distance without category information, and $\max(\Delta)$ is the maximum distance between classes. Λ_{ij} is the matrix whose value is 0 or 1. When two sample points belong to the same kind, the value is 0; otherwise the value is 1. α is a parameter to adjust the distance between point sets, $\alpha \in [0, 1]$, and α is an experienced parameter. When its value is 0, SLLE algorithm is equivalent to LLE algorithm.

3.2. Locally Linear Embedding Algorithm of Supervision and Homogenization Distance Algorithm. Locally linear embedding algorithm of homogenization distance by the change of distance makes categories much more classifiable. But the improvement of its classification performance is limited, and it is also a kind of unsupervised algorithm. In order to promote the fault recognition rate and stability of method, supervised learning mechanism is introduced to the HLLE, and then locally linear embedding algorithm of supervision and homogenization distance algorithm (SHLLE) is proposed. On the basis of homogenization distance, to increase the category information of sample points by supervised learning, so that the same fault category are gathered and heterogeneous fault are scattered. The main steps of method are as follows.

Step 1. Set neighbor points k, embedding dimension d, and parameter values $\alpha \in [0, 1]$.

Step 2. Calculate the homogenization distance. For a given data set $X = \{x_i, i = 1, 2, \ldots, N\} \subset R^m$, calculate the Euclidean distance between sample points, and according to formula (1) receive the homogenization distance.

Step 3. Add sample information and select local neighbors. According to formula $D'_{ij} = D_{ij} + \alpha \max(D_{ij})\Lambda_{ij}$ to get D'_{ij},
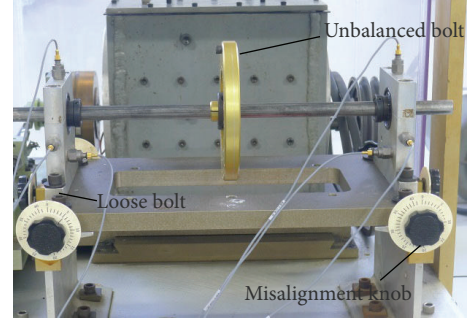

FIGURE 3: Rotor system fault simulation experiment.

look for the k $(k < N)$ nearest neighbor points of each sample point in high-dimensional space.

Step 4. Calculate and reconstruct weight matrixes. Calculate locally optimal reconstruction weights of sample points and make the reconstruction error minimized. That is, acquire the optimal solution

$$\min \quad \varepsilon(W) = \sum_{i=1}^{N} \left\| x_i - \sum_{j=1}^{N} w_{ij} x_j \right\|_2^2$$

$$\text{s.t.} \quad \sum_{j=1}^{k} w_{ij} = 1. \tag{7}$$

Among them, x_j $(j = 1, 2, \ldots, k)$ is neighbor point of x_i, and w_{ij} is the weights between x_i and x_j. When x_i and x_j are not neighbor, $w_{ij} = 0$.

Step 5. Calculate low-dimensional embedding matrix Y. From the above steps, weight matrix W is obtained which is used to get the optimal low-dimensional embedding matrix $Y = \{y_1, y_2, \ldots, y_N\}^T$, by minimizing the reconstruction error and using function $\min \Phi(Y) = \sum_{i=1}^{N} \left\| y_i - \sum_{j=1}^{N} w_{ij} y_j \right\|_2^2$. Matrix Y needs to meet $\sum_{i=1}^{N} y_i = 0$, $(1/N) \sum_{i=1}^{N} y_i y_i^T = I$. Among them I is N dimensional unit matrix. The optimization problem is transformed into the following constrained optimization problem:

$$\min \quad \Phi(Y) = \sum_{i=1}^{N} \left\| YI_i - YW_i \right\|^2$$

$$= \sum_{i=1}^{N} \left\| Y(I_i - W_i) \right\|^2 = \min \operatorname{tr} YMY^T \tag{8}$$

$$\text{s.t.} \quad YY^T = I.$$

Among them, $M = (I - W^T)(I - W)$ is solved to $MY^T = \lambda Y^T$ by Lagrange multiplier method. Select the eigenvectors which are corresponded with d minimum nonzero eigenvalues of M that is the requested low-dimensional coordinate matrix Y. As usual, the minimum eigenvalue is almost zero. Taking
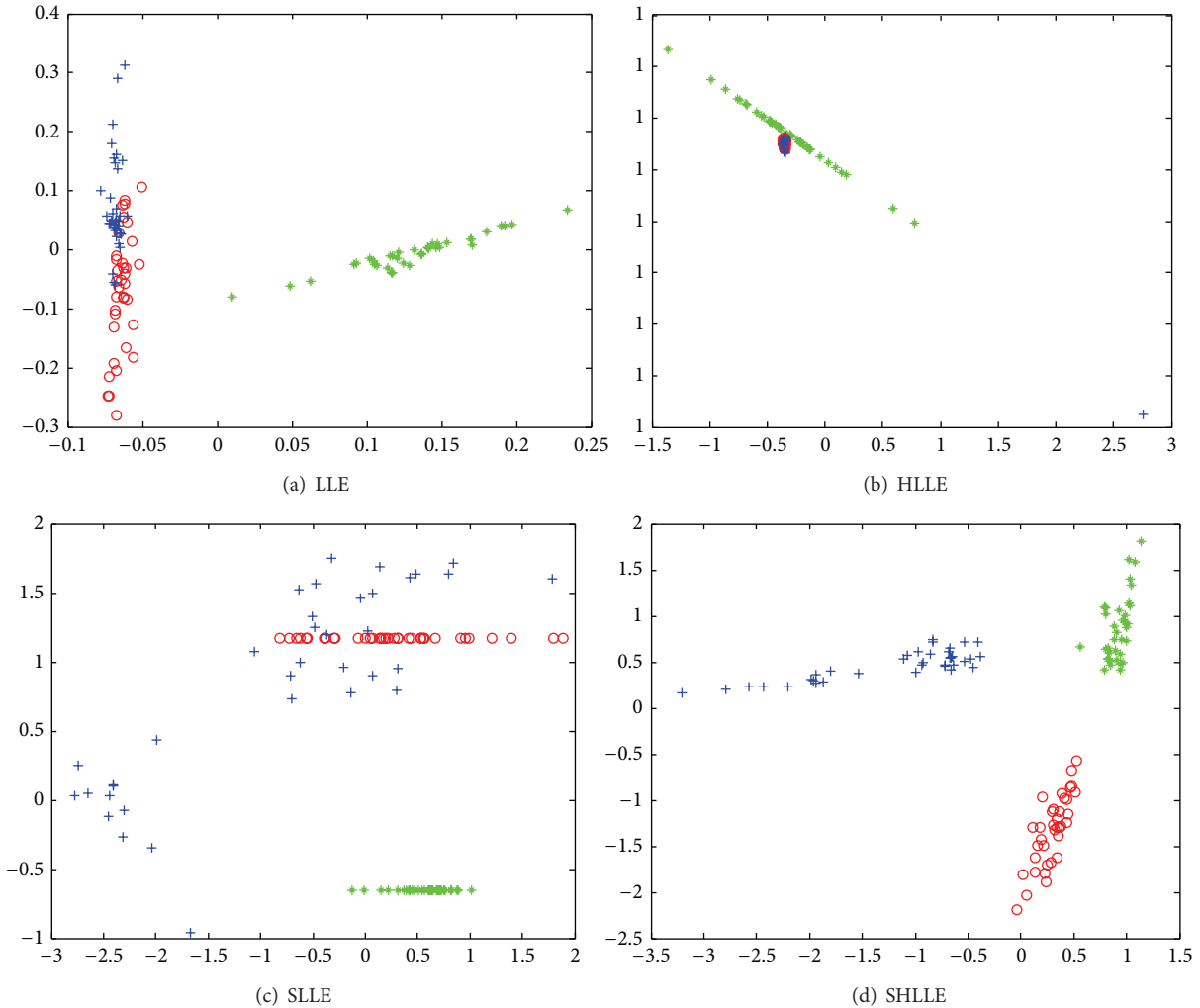
(a) LLE

(b) HLLE

(c) SLLE

(d) SHLLE

FIGURE 4: Fault recognition effect of data 1.

the eigenvectors $(u_1, u_2, \ldots, u_d)$ corresponded with the eigenvalues between $2 \sim (d+1)$, so the optimal embedding results are obtained.

4. Simulation Experiment of Rotor System Fault

Comprehensive fault simulation test bed of Spectra Quest Company (USA) is used as experimental platform. Specifically, fault simulation experiment system is composed of Spectra Quest integrated fault simulation test bench and PULSE data acquisition system. As shown in Figure 3, the rotor disk contains two laps of a total of 36 threaded holes, in which the bolt can be randomly installed and can be used to simulate the rotor unbalance fault, and the quality of bolt is 5.596 g in the experiment. Adjusting two same knobs of the plinth to control both ends in the relative position of the axial and radial bearing can simulate rotor misalignment fault, 0.025 mm/scale. In the experiment, the left clockwise rotation for 5 scales is 0.125 mm, and the right end clockwise rotation for 10 scales is 0.25 mm. For the pedestal loose fault

of the experiment, it can be simulated to loosen the left end part of the bolt.

Through the simulation of the fault of the rotor system which is normal, unbalanced, misalignment, and loose in the pedestal at the running speed of 10 Hz, 20 Hz, and 30 Hz and using a total of six acceleration sensors at two bearing seats on three directions to carry out the vibration signal acquisition, a total of 144 groups of data signal were obtained.

5. Fault Diagnosis Based on SHLLE

Merging and reconstructing raw signal data, through analysis and comparison, a set of data 1 is shown as vector index of fault feature which could represent normal, misalignment, and unbalanced fault. Another group of data 2 is shown as vector index of fault feature which represent loose, misalignment, and unbalanced fault. Then two groups of the data space are filtered and extract 8 time domain parameters to form the original featured space which can be rolled into two $48 * 108$ data matrices.
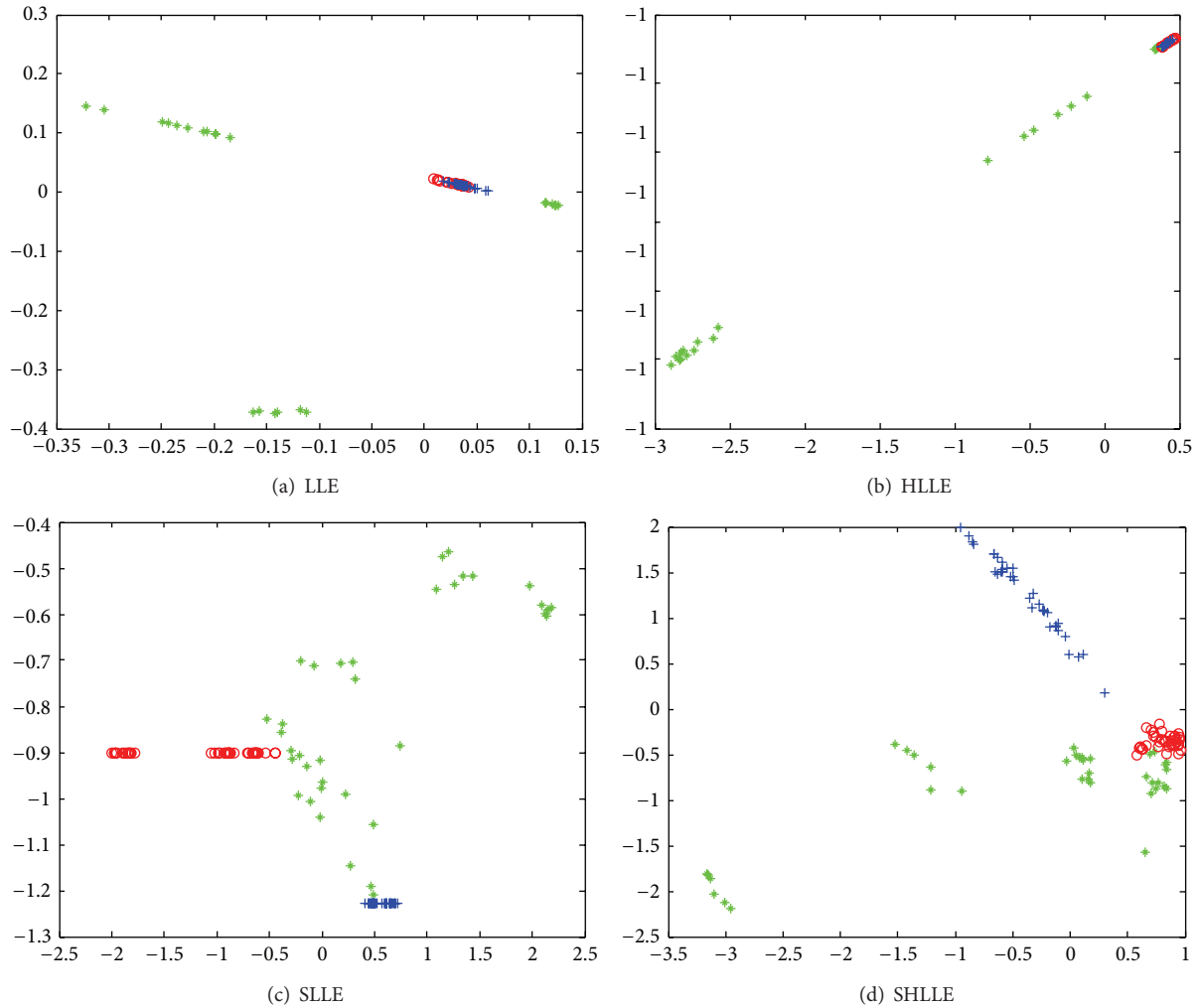
(a) LLE

(b) HLLE

(c) SLLE

(d) SHLLE

FIGURE 5: Fault recognition effect of data 2.

The analysis and comparison of classification and results of fault diagnosis between LLE, HLLE, SLLE, and SHLLE are as follows.

5.1. Fault Identify and Comparison. As shown in Figures 4 and 5, (a), (b), (c), and (d) are two-dimensional classification map of LLE, HLLE, SLLE, and SHLLE algorithm for original feature datasets, while horizontal and vertical axes represent the component 1 and component 2 of the main characteristics, respectively.

In Figure 4, in data 1, "∗" indicates normal state, "○" indicates misalignment fault, and "+" indicates unbalanced fault. To know, (1) when using LLE and HLLE methods, normal state can be better separated, and overlapping phenomenon exists about unbalanced and misalignment faults. (2) When using SLLE method and setting parameter $\alpha = 0.5$, the separated data set is gathered into columnar shape. Normal data set is identified, but misalignment and unbalanced fault are partially overlapped. (3) As for the SHLLE method, setting parameter $\alpha = 0.5$, normal, misalignment, and unbalanced

fault can be significantly classified and aggregated. It has better recognition effect compared with LLE and SLLE.

In Figure 5, "∗" represents loose fault; "○" indicates misalignment fault; "+" indicates unbalanced fault. To know, (1) when using LLE and HLLE methods, loose fault can be separated but not gathered, because of its impact that it is difficult to separate misalignment fault from unbalanced fault. (2) When using SLLE method and setting parameter $\alpha = 0.5$, misalignment and unbalanced fault can be better separated and gathered without the effect of loose fault, but loose fault still cannot be gathered. (3) When using SHLLE method and also setting parameter $\alpha = 0.5$, normal, misalignment, and unbalanced fault can be effectively classified, and fault recognition effect is relatively better. In short, there is great superiority when using SHLLE to handle data sets of original features which is more effective on classification and identification of faults.

5.2. Comparison of Fault Recognition Rate. Two-dimensional map of data set according to various algorithms cannot fully reflect effect of fault recognition, because some methods also
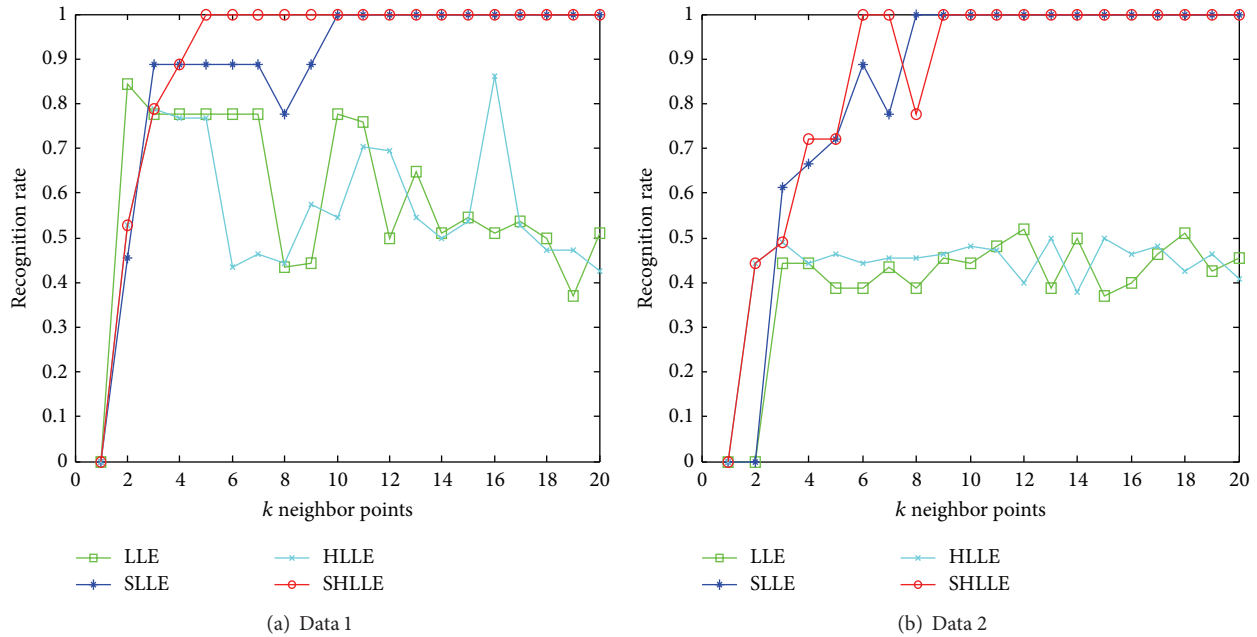
FIGURE 6: Recognition rate based on LLE, HLLE, SLLE, and SHLLE.

can be used to classify and identify the fault when the number of neighborhood is small. Therefore, a comparative analysis of the recognition rate changing with the neighbor points was shown below.

Figure 6(a) is the changeable map for normal, misalignment, and unbalanced data sets (data 1) fault diagnosis accuracy rate with neighboring points of change in k based on LLE, HLLE, SLLE, and SHLLE algorithms. Taking the embedding dimension $d = 2$, $\alpha = 0.5$. "Box" Line, "∗" Line, "x" Line, and "○" Line represent a recognition rate curve of LLE, SLLE, HLLE, and SHLLE, respectively. It can be obtained according to the diagram and the data: about 80% recognition rate of LLE and HLLE is achieved with smaller neighbor points, which shows an unstable state of fluctuation obviously with the increase of k; recognition rate of SLLE algorithm shows a higher trend which reaches 100% and remains stable when k values 10, which is suitable for the case where neighbor number is more than 10; SHLLE algorithm has great advantages of recognition rate of normal, misalignment, and unbalanced faults which reaches 100% and keeps stable when k values 5.

Figure 6(b) is the changeable map for loose fault, misalignment, and unbalanced data sets (data 2) fault diagnosis accuracy rate with neighboring points of change in k based on LLE, HLLE, SLLE, and SHLLE algorithms. Take the embedding dimension $d = 2$, $\alpha = 0.5$. "Box" Line, "∗" Line, "x" Line, and "○" Line represent a recognition rate curve of LLE, SLLE, HLLE, and SHLLE, respectively. To know, loose fault still has some bad impacts on the effect of LLE and HLLE algorithm, and even the purpose of fault diagnosis cannot be achieved; recognition rate of SLLE algorithm increases in general which shows unstable fluctuation when k is smaller. Recognition rate of SLLE algorithm reaches 100% and keeps

stable when k is equal to 8, which is still suitable for the case where the neighbor number is larger.

SHLLE algorithm still has a soft spot for recognition rate of loose, misalignment, and unbalanced faults which reaches 100% firstly and keeps stable after identifying one case of fluctuation when k value is euqal to 8.

In conclusion, fault diagnosis based on SHLLE algorithm has superior performance compared to the other LLE algorithm. SLLE algorithm is suitable for fault identification of which the neighborhood is slightly larger, and it is relatively stable. However, SHLLE algorithm has optimal performance, of which fault identification is more stable than others.

6. Conclusion

The paper researches on rotating machinery. In order to get better recognition effect, LLE of homogenization distance (HLLE) and LLE of supervision and homogenization distance (SHLLE) are proposed. Proving the validity of the fault diagnosis by simulating rotor system failure experiment, the following conclusions are reached.

(1) In two-dimensional map of each algorithm for two types of data set, overlapping phenomenon exists between unbalanced fault and misalignment fault when using LLE, HLLE, and SLLE methods. However, SHLLE has a strong advantage and is more effective in fault classification.

(2) In the map with the changes of neighbor points of each algorithm for two types of data set, SLLE algorithm is suitable for the fault diagnosis when the number of neighbor points is slightly larger. However, the fault diagnosis of SHLLE algorithm has superior

performance compared to the other LLE algorithm, even fault identification is more stable.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[2] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[3] Z. Y. Zhang and H. Y. Zha, "Nonlinear dimension reduction via local tangent space alignment," in *Proceedings of the 4th International Conference on Intelligent Data Engineering and Automated Learning (IDEAL '03)*, Hong Kong, March 2003.

[4] Z. Y. Zhang and H. Y. Zha, "Principal manifolds and nonlinear dimension reduction via local tangent space alignment," *SIAM Journal on Scientific Computing*, vol. 26, no. 1, pp. 313–338, 2004.

[5] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation*, vol. 15, no. 6, pp. 1373–1396, 2003.

[6] M. Belkin and P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," in *Advances in Neural Information Processing Systems*, MIT Press, Whistler, Canada, 2002.

[7] X. F. He, S. C. Yan, Y. X. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using laplacianfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 328–340, 2005.

[8] J. Nilsson, "Nonlinear dimensionality reduction of gene expression data," Printed in by KFS, Lund, Sweden, 2006.

[9] H. Chang and D.-Y. Yeung, "Locally linear metric adaptation with application to semi-supervised clustering and image retrieval," *Pattern Recognition*, vol. 39, no. 7, pp. 1253–1264, 2006.

[10] J. Yang, J. Xu, D. Yang, and M. Li, "Noise reduction method for nonlinear time series based on principal manifold learning and its application to fault diagnosis," *Chinese Journal of Mechanical Engineering*, vol. 42, no. 8, pp. 154–158, 2006.

[11] Q. Yang, G. Chen, X. Tong, and Q. He, "Application of incremental local tangent space alignment algorithm to rolling bearings fault diagnosis," *Journal of Mechanical Engineering*, vol. 48, no. 5, pp. 81–86, 2012.

[12] L. Liang, G. Xu, M. Li, Y. Zhang, and X. Liang, "Nonlinear manifold learning method of mechanical impact faults extraction," *Journal of Xi'an Jiaotong University*, vol. 43, no. 11, 2009.

[13] M. Li, S. Wang, and L. Liang, "Feature extraction for incipient fault diagnosis of rolling bearings based on nonlinear manifold learning," *Journal of Xi'an Jiaotong University*, vol. 44, no. 5, pp. 45–49, 2010.

[14] D. de Ridder, O. Kouropteva, O. Okun et al., "Supervised locally linear embedding," in *Artificial Neural Networks and Neural Information Processing—ICANN/ICONIP-2003*, vol. 2714 of *Lecture Notes in Computer Science*, pp. 333–341, Springer, Berlin, Germany, 2003.