

Research Article

Multisensor Fused Fault Diagnosis for Rotation Machinery Based on Supervised Second-Order Tensor Locality Preserving Projection and Weighted k -Nearest Neighbor Classifier under Assembled Matrix Distance Metric

Fen Wei,¹ Gang Wang,¹ Bingyin Ren,¹ Jianghua Ge,² and Yaping Wang²

¹*School of Mechatronics Engineering, Harbin Institute of Technology, Harbin 150001, China*

²*Department of Mechanical and Dynamic Engineering, Harbin University of Science and Technology, Harbin 150080, China*

Correspondence should be addressed to Bingyin Ren; renby@hit.edu.cn

Received 15 June 2016; Revised 21 October 2016; Accepted 24 October 2016

Academic Editor: Fiorenzo A. Fazzolari

Copyright © 2016 Fen Wei 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.

In order to sufficiently capture the useful fault-related information available in the multiple vibration sensors used in rotation machinery, while concurrently avoiding the introduction of the limitation of dimensionality, a new fault diagnosis method for rotation machinery based on supervised second-order tensor locality preserving projection (SSTLPP) and weighted k -nearest neighbor classifier (WKNNC) with an assembled matrix distance metric (AMDM) is presented. Second-order tensor representation of multisensor fused conditional features is employed to replace the prevailing vector description of features from a single sensor. Then, an SSTLPP algorithm under AMDM (SSTLPP-AMDM) is presented to realize dimensional reduction of original high-dimensional feature tensor. Compared with classical second-order tensor locality preserving projection (STLPP), the SSTLPP-AMDM algorithm not only considers both local neighbor information and class label information but also replaces the existing Frobenius distance measure with AMDM for construction of the similarity weighting matrix. Finally, the obtained low-dimensional feature tensor is input into WKNNC with AMDM to implement the fault diagnosis of the rotation machinery. A fault diagnosis experiment is performed for a gearbox which demonstrates that the second-order tensor formed multisensor fused fault data has good results for multisensor fusion fault diagnosis and the formulated fault diagnosis method can effectively improve diagnostic accuracy.

1. Introduction

As one of the most common mechanical equipment classes, rotation machinery occupies an important role in industrial applications such as manufacturing, metallurgy, energy, and transportation. Due to tough working environments, similar materials, and structural properties, rotation machinery can be subject to malfunctions or failures. This can significantly decrease machinery service performance including manufacturing quality and operation safety and cause machinery to break down, which may lead to serious catastrophes [1]. Accordingly, research into fault diagnosis of rotation machinery has attracted considerable attention by researchers in related domains in recent years. The vibration signals collected from velocity or accelerator sensors located in machinery

housing are generally regarded as the foundation of fault diagnostic procedures. However, most existing studies on fault diagnosis of rotation machinery have empirically or experimentally focused on analyzing single sensor signals [2–4], and the remaining studies have performed multisensor fused fault diagnosis through complex fusion algorithms such as blind source separation (BSS) [5] and D-S evidence theory. The single-sensor-based fault diagnosis methods belonging to the former category of studies generally lead to loss of valuable information available from multiple sensors, and the multisensor fused diagnosis methods appearing in the latter category of studies are tend to cause a high computational load. To tackle these issues, this paper presents a second-order tensor representation of fault samples including fault

feature dimensions and sensor locations dimensions, which is used in an efficient multisensor fused fault diagnosis framework.

Large volumes of feature parameters generated by time-domain, frequency-domain, and time-frequency-domain analysis of vibration signals are commonly integrated into a high-dimensional data set to obtain accurate fault diagnostic results [6]. This high-dimensional feature set can provide more valuable information, but it also increases the computational load and may even trigger dimensionality issues. One approach to address this problem is to apply dimension reduction technology. Compared with classical linear dimensionality reduction methods such as principle component analysis (PCA) [7], linear discriminate analysis (LDA) [8], and multidimensional scaling (MDS) [9], a new technology for discovering intrinsic low-dimensional structure of nonlinear distributed data hidden in high-dimensional space has emerged which is known as manifold learning and has become a current research focus. Representative manifold learning methods include isometric mapping (ISOMAP) [10], locality linear embedding (LLE) [11], Laplacian eigenmaps (LE) [12], and local tangent space alignment (LTSA) [13]. The effectiveness of these basic manifold learning algorithms and their variants for fault diagnosis of rotation machinery has been validated frequently by a large number of studies. For instance, Li et al. [14] proposed a fault diagnosis method using dimension reduction with linear local tangent space alignment (LLTSA). Ding et al. [15] developed a fusion feature extraction method based on locality preserving projection (LPP) for rolling element bearing fault classification. Additionally, an envelope manifold demodulation method was investigated for planetary gear fault detection in [16]. It should be observed that the input sample for these methods is generally represented by a vector with a high-dimensional feature space. It is obvious that these manifold learning algorithms are not suitable when a multisensor fused faulty sample is represented as a second-order tensor, namely, a matrix. Furthermore, tensor representation based manifold learning methods have received little investigation for fault diagnosis. Fortunately, there are several second-order or higher-order tensor extended manifold algorithms, such as second-order tensor locality preserving projection (STLPP) [17], tensor neighborhood preserving embedding (TNPE) [18], a tensor version of discriminant locality linear embedding (DLLE/T) [19], and tensor PCA [20]. These algorithms have been progressively applied in the areas of two-dimensional or higher-dimensional image classification, computer vision, and pattern recognition and offer a feasible solution for tensor-represented fault diagnosis. Out of the methods mentioned above, STLPP possesses the ability to discover intrinsic local geometric and topological properties of a manifold embedded in a second-order tensor space, on the basis of inherited strengths of LPP. However, it has been found that there are several limitations of the STLPP algorithm. For instance, STLPP is an unsupervised method for dimension reduction and thus does not consider discriminant information which is useful for fault classification. Secondly, the similarity with second-order tensor formed samples in traditional STLPP has been computed using

the Frobenius distance measure [17] which is the same as the Euclidean distance of the vectorized version of matrix formed samples, so it may still cause a loss of spatial locality information. To tackle these problems, this paper introduces the concept of supervision into the framework of a traditional STLPP and employs an assembled matrix distance metric (AMDM) which has been successfully utilized in 2DPCA [21] into the construction of a similarity weighting matrix to obtain better matching between two second-order tensor formed faulty samples.

To further improve the accuracy and the efficiency of fault diagnosis, intelligent classification methods are considered as an indispensable component in the diagnostic procedure. These methods include artificial neural networks (ANN) [22], support vector machines (SVM) [23], and fuzzy-based systems [24] as well as Bayesian based classifiers [25]. Compared with these methods, the k -nearest neighbor classifier (KNNC) ranks k neighbors of testing samples from training samples and uses the class labels of similarity neighbors to classify input test samples by evaluating the similarity between samples in the feature space [26, 27]. The KNNC method has many benefits, including a lower calculation requirement, quicker speed, and higher pattern recognition accuracy [28]. Therefore, it is considered to be the simplest tool for faulty pattern recognition. There are some existing shortcomings for traditional KNNC which classifies the sample labels using unified weights, and thus the weighted k -nearest neighbor classifier (WKNNC) was developed which assigns different weights to nearest neighbors to represent the impact of each neighbor on each unknown sample. Therefore, this paper uses the WKNNC to establish the relationships between features of samples and conditional classifications. Additionally, the AMDM mentioned above is also employed for the similarity evaluation of low-dimensional second-order formed samples after SSTLPP based dimension reduction in WKNNC.

The remainder of this paper is organized as follows. The proposed supervised second-order tensor locality preserving projection based on assembled matrix distance metric (SSTLPP-AMDM) algorithm is discussed in detail in Section 2. The weighted k -nearest neighbor classifier with an assembled matrix distance metric (WKNNC-AMDM) is described in Section 3. Section 4 provides the overall framework for the proposed multisensors fused fault diagnosis. In Section 5, a fault diagnosis experiment is performed for a gearbox to validate the proposed method. Finally, the conclusions are given in Section 6.

2. Supervised Second-Order Tensor Locality Preserving Projection Based on Assembled Matrix Distance Metric (SSTLPP-AMDM)

2.1. Introduction to Second-Order Tensor Locality Preserving Projection (STLPP). As the tensor extension of LPP, TLPP is essentially equivalent to finding a linear approximation of the eigenfunctions of the Laplace Beltrami operator in a tensor space. The incipient TLPP which was initially presented by He et al. [17] in 2005 is a second-order case and was reviewed

and then extended for a universal n -order version by Dai and Yeung [18] in 2008. Since the multisensor fused faulty sample studied in this paper is represented by a second-order tensor form, namely, in a matrix form, the second-order TLPP (STLPP) algorithm is the focus in the following discussion. Given N matrix formed samples $X_1, X_2, \dots, X_N \in R^{n \times m}$, the aim of STLPP is to find two transformation matrices $U \in R^{n \times n'}$ and $V \in R^{m \times m'}$ ($n', m' < n, m$) by optimizing the following formulation:

$$\begin{aligned} \arg \min \quad & \sum_{i,j} \|U^T X_i V - U^T X_j V\|_F^2 \times S_{ij} \\ \text{s.t.} \quad & \sum_i \|U^T X_i V\|_F^2 \times D_{ii} = I, \end{aligned} \quad (1)$$

where $\|\cdot\|_F$ is the Frobenius norm of the matrix; that is, $\|A\|_F = \sqrt{\sum_i \sum_j a_{ij}^2}$; S_{ij} denotes the elements of the weight matrix S of the nearest neighbor graph G , which is equal to $\exp(-\|X_i - X_j\|_F^2/t)$ when X_j is one of the k nearest neighbors of X_i or X_i is one of the k nearest neighbors of X_j ; otherwise it is equal to zero. D is a diagonal matrix; $D_{ii} = \sum_j S_{ij}$.

Using a series of mathematical derivations, the optimal values for U and V are obtained by iteratively computing the generalized eigenvectors of the following formulations:

$$\begin{aligned} (D_u - S_u) \vec{v} &= \lambda D_u \vec{v}, \\ (D_v - S_v) \vec{u} &= \lambda D_v \vec{u}, \end{aligned} \quad (2)$$

where $D_u = \sum_i D_{ii} X_i U U^T X_i^T$, $S_u = \sum_{ij} S_{ij} X_i U U^T X_j^T$, $D_v = \sum_i D_{ii} X_i V V^T X_i^T$, and $S_v = \sum_{ij} S_{ij} X_i V V^T X_j^T$.

Finally, the low-dimensional representations of the original data are obtained using $Y_i = U^T X_i V$.

2.2. Computation of a Supervised Similarity Weighting Matrix Based on AMDM. As described in the previous section, there are a certain number of limitations when using the prevailing computation method for the similarity weighting matrix S of the nearest neighbor graph G . For instance, the Frobenius

distance metric (FDM) used for the similarity evaluation between different second-order tensor formed samples is essentially the Euclidean distance of the vectorized version of the matrix formed samples, and thus it neglects spatial geometrical information of each element in matrix formed samples and thus has poor matching performance for different samples. Additionally, class label information of the training samples is not effectively used in the traditional STLPP, although this information can be helpful for subsequent accurate classification assignment. To address these issues, this paper formulates a novel supervised similarity matrix computation method that decides the similarity between matrix formed samples using an assembled matrix distance metric that takes the classification information into account.

Firstly, for any two arbitrary matrix formed samples $A = (a_{lh})_{n \times m}$ and $B = (b_{lh})_{n \times m}$, the distance between the two samples can be measured using the following assembled matrix distance metric (AMDM) [21]:

$$\text{dist}_{\text{AMDM}}(A, B) = \left(\sum_{h=1}^m \left(\sum_{l=1}^n (a_{lh} - b_{lh})^2 \right)^{(1/2)p} \right)^{1/p}, \quad (3)$$

where p denotes a variable parameter which strongly affects the representation ability of the defined distance function for subsequent classification assignment. It is obviously that the Frobenius distance metric is a special case of the AMDM with $p = 2$, and the Yang distance metric proposed by Yang et al. in [29] is another special case for $p = 1$. It has also been theoretically and experimentally verified that an assembled matrix distance metric with a lower value of p , that is, $0 < p < 1$, outperforms existing Frobenius distance and Yang distance measures in terms of the final classification accuracy. Accordingly, the value of p for the employed AMDM is set between 0 and 1, $0 < p < 1$, and its exact value is determined by repeated experiments.

Secondly, by understanding the class label information of the training samples and the AMDM based distances between samples, the proposed supervised similarity weighting matrix based on AMDM can be defined as

$$SS_{ij} = \begin{cases} 1, & X_i \text{ is among } k \text{ nearest neighbors of } X_j, \text{ vice versa, and } C_i = C_j \\ \left(\exp \left(-\frac{\text{dist}_{\text{AMDM}}^2(X_i, X_j)}{t} \right) \right)^m, & X_i \text{ is among } k \text{ nearest neighbors of } X_j, \text{ vice versa, and } C_i \neq C_j \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

where SS_{ij} denotes the element at column i and row j in the new formulated supervised similarity matrix SS , which represents the similarity degree of the matrix formed samples X_i and X_j . C_i and C_j are the class labels of samples X_i and X_j , respectively. m is the penalty coefficient which is used to characterize the reduction in the similarity degree. Since X_i is one of the k nearest neighbors of X_j or X_j is one of the

k nearest neighbors of X_i , the corresponding class labels are inconsistent, and thus the value of m should be set to $m > 1$.

The newly formulated similarity weighting matrix computation equation shown in (4) can be viewed as the combination and extension of the prevailing "0-heat kernel function" and the "0-1" binary mode, in which the former is intimately related to the manifold structure and the latter is regarded as

the direct expression of the label information. The properties and corresponding advantages of the supervised similarity weighting matrix based on AMDM can be summarized as follows. (i) A more accurate representation of the matching relationship between matrix formed samples can be achieved using AMDM rather than traditional STLPP, which uses the Frobenius distance metric. (ii) The inclusion of the penalty parameter m ($m > 1$) results in larger differences between 1 and $(\exp(-\text{dist}_{\text{AMDM}}^2/t))^m$ as the assembled matrix distance increases, which allows the interclass and intraclass similarity to be easily distinguished.

2.3. SSTLPP-AMDM Algorithm. This paper proposes a novel supervised second-order tensor locality preserving projection algorithm with the assembled matrix distance metric (SSTLPP-AMDM) that uses the improvements in both the matrix distances computation of samples in the projection space and the similarity weighting matrix computation expression. In contrast to traditional STLPP, the two transformation matrices U and V that represent both the neighborhood graph structure and the class label information are obtained by solving the following objective function:

$$\begin{aligned} \arg \min \quad & \sum_{i,j} \|U^T X_i V - U^T X_j V\|_{\text{AMDM}}^2 \times SS_{ij} \\ \text{s.t.} \quad & \sum_i \|U^T X_i V\|_{\text{AMDM}}^2 \times DD_{ii} = I. \end{aligned} \quad (5)$$

The distance between two mapped sample points $Y_i = U^T X_i V$ and $Y_j = U^T X_j V$ in the embedded tensor space is measured using the assembled matrix distance metric to achieve a better matching result. The element SS_{ij} of the supervised similarity weighting matrix which is computed by (4) is employed to represent the neighboring degree of samples X_i and X_j and considers both the local structure and class information. The diagonal matrix $DD_{ii} = \sum_j SS_{ij}$ has the ability to characterize the degree of importance of the mapped sample point Y_i in the embedded tensor space to represent the original sample point X_i .

The optimal transformation matrices U and V are solved in a similar way to traditional STLPP by applying an iterative scheme. The specific implementation process can be described as follows. Firstly, an initial matrix U is set as an identity matrix and the first iterative solution of V is then obtained by solving the generalized eigenvector problem shown in (6). Secondly, U is updated by solving the generalized eigenvector problem shown in (7). By iteratively computing the generalized eigenvectors of (6) and (7) for a predefined number of repetitions, the optimal transformation matrices U and V are obtained. Finally, the second-order low-dimensional projection Y_i of the original second-order high-dimensional sample X_i is obtained.

$$\begin{aligned} & \left(\sum_i DD_{ii} X_i U U^T X_i^T - \sum_{ij} SS_{ij} X_i U U^T X_j^T \right) \vec{v} \\ & = \lambda \left(\sum_i DD_{ii} X_i U U^T X_i^T \right) \vec{v}, \end{aligned} \quad (6)$$

$$\begin{aligned} & \left(\sum_i DD_{ii} X_i V V^T X_i^T - \sum_{ij} SS_{ij} X_i V V^T X_j^T \right) \vec{u} \\ & = \lambda \left(\sum_i DD_{ii} X_i V V^T X_i^T \right) \vec{u}. \end{aligned} \quad (7)$$

In summary, there are two main advantages to the newly proposed SSTLPP-AMDM. (1) The local structure information and the class information act cooperatively in the computation of the similarity weighting matrix, and thus the supervised similarity weighting matrix proposed in this paper outperforms other prevailing similarity weighting matrix computation methods in terms of representation of the similarity degree between samples. (2) The application of AMDM to measure the distance between both the sample points in the original second-order tensor space and the mapped sample points in the embedded second-order tensor space ensures that the measured samples have a better matching performance than the existing Frobenius distance measure. Therefore, the SSTLPP-AMDM algorithm has superior classification and dimension reduction characteristics than traditional STLPP.

3. Weighted k -Nearest Neighbor Classifier with Assembled Matrix Distance Metric (WKNNC-AMDM)

As stated above, the KNNC method proposed by Cover and Hart in 1967 [28] is regarded by many as the simplest pattern classification algorithm. Due to its advantages of a lower calculation requirement, quicker speed, and higher identification accuracy, KNNC has been widely applied to various types of pattern recognition problems, especially fault diagnosis issues. The main KNNC concept is described in the following two steps.

Step 1. For a given unknown labeled sample X' , k similar samples $\{X_1, X_2, \dots, X_k\}$ in the training sample set are searched to construct a neighbor set N .

Step 2. A maximum voting rule is used on all samples in N to obtain the class that X' belongs to.

The above description shows that there are two focus points to KNNC: a similarity measurement method between samples and the establishment of a decision rule. For the first focus point, there have been many similarity measurement methods suggested by previous publications, such as the Euclidean distance, the Manhattan distance, and the cosine angle. However, these vector representations of the data-based metric indexes described above are unsuitable for similarity measurement of the matrix formed data points appearing in this paper. Thus, the AMDM is introduced for the similarity computation of samples in KNNC. It is known that AMDM outperforms common FDM in terms of the similarity presentation between matrix formed samples for classification. Additionally, since selection of neighbors is greatly impacted by the sparsity of the sample distribution,

this paper employs a novel assembled matrix distance based on density to efficiently measure the similarity between X' and its neighbor X_i , ($i = 1, 2, \dots, k$) using the following formula:

$$\begin{aligned} \text{dist}(X', X_i) \\ = \frac{\|X' - X_i\|_{\text{AMDM}}}{\sqrt{(1/n) \sum_{j=1}^n \|X' - X_j\|_{\text{AMDM}} \sqrt{(1/n) \sum_{j=1}^n \|X_i - X_j\|_{\text{AMDM}}}}. \end{aligned} \quad (8)$$

$$\omega_i = \begin{cases} \frac{\text{dist}(X', X_k) - \text{dist}(X', X_i)}{\text{dist}(X', X_k) - \text{dist}(X', X_1)} & \text{if } \text{dist}(X', X_k) \neq \text{dist}(X', X_1) \\ 1 & \text{if } \text{dist}(X', X_k) = \text{dist}(X', X_1). \end{cases} \quad (9)$$

Consequently, the class label of an unknown labeled sample X' can be determined as follows:

$$Y' = \arg \max_y \left(\sum_{X \in N} \omega_i \times \delta(y = y_i) \right), \quad (10)$$

where y_i denotes the class label of X_i in N and $\delta(y = y_i)$ is the Di carat function which has the functional value equal to 1 when $y = y_i$, and otherwise it is equal to zero.

Additionally, the selection of k is an issue that requires attention in the WKNNC algorithm. In this paper the value of k is set to $k = 2 \times c + 1$, since the classification precision is only just assured when the number of samples in N equals $c + 1$, where the number of classes in training set is c [30].

4. Overall Framework of the Proposed Fault Diagnostic Method

Based on the preparations above, this paper proposes a novel multisensor fused fault diagnosis method based on SSTLPP-AMDM and WKNNC-AMDM for rotation machinery. The flow chart for the proposed method is shown in Figure 1. There are three main steps to the diagnostic procedure, which will be discussed in detail in this section.

Firstly, through prevalent multidomain signal analysis and truncated sampling, a multisensor fused faulty sample set with an $n \times m \times N$ -dimensional third-order tensor representation is constructed and then decomposed into an $n \times m \times N_1$ -dimensional training sample set and an $n \times m \times N_2$ testing sample set, where n is the number of vibration sensors located in the equipment being diagnosed, m is the number of features originating from the vibration signal of a single sensor, and N is the number of samples and $N = N_1 + N_2$.

The second step is compression of the high-dimensional $n \times m \times N_1$ tensor into a relatively low-dimensional $n' \times m' \times N_1$ tensor using SSTLPP-AMDM. By constructing a supervised similarity weighting matrix based on AMDM, the minimization problem is formulated for the weighted sum of the assembled matrix distances between samples in the embedded tensor space, in order to find the optimal transformation matrices U and V .

Unlike the classical KNNC voting strategy that uses unified weights for neighbors, in this paper, a weighted voting strategy is used to form the weighted k -nearest neighbor classifier (WKNNC), which assigns different weights to each sample X_i in N , reflecting the influence each neighbor has on an unknown sample X' . A new neighbor set $NN = (X_1, X_2, \dots, X_k)$ is generally reconstructed in ascending order of distance; that is, $\text{dist}(X', X_1) \leq \text{dist}(X', X_k)$, and thus the voting weight ω_i of sample X_i is computed using the following equation:

Finally, the low-dimensional projection of the testing sample set and the low-dimensional projection of the training sample set obtained in the previous step are input into WKNNC for fault diagnosis.

5. Experimental Results and Analysis

5.1. Experimental Setup. The validity of the newly proposed method will now be demonstrated using a fault diagnosis experiment of a single-stage gearbox. As shown in Figure 2(a), this paper employed a rotation machinery fault diagnosis experiment platform system of type QPZZ-II, which was converted into a gearbox fault test bench while the timing belt pulley at the side of gearbox was connected to the motor shaft. A diagram of the gearbox fault experiment system is displayed in Figure 2(b). Seven displacement sensors and accelerometers which were reinstalled to the input shaft and the end housings of the four bearings were employed to collect faulty vibration signals. Specific location information is shown in Table 1.

During the experiment, the sampling frequency was 5120 Hz, there were 53248 sampling points, the rotation speed of the drive motor was 880 rev/min, and the load was 0.2 A. There were six types of conditions used in the gearbox fault simulation experiment: (1) normal (Norm), (2) corrosion of the gearwheel (C_G), (3) broken teeth in the gearwheel (B_G), (4) wear of the pinion (W_P), (5) broken teeth in the gearwheel coupled with wear of the pinion (B_G_C_W_P), and (6) corrosion of the gearwheel coupled with wear of the pinion (C_G_C_W_P). Figure 3 shows the time-domain waveforms of the faulty samples originating from the seven different sensors under each condition, and the time-domain waveforms of the faulty samples originating from a single sensor under the six conditions are displayed in Figure 4. It can be observed from these graphs that the sensors reinstalled to different equipment positions have the distinct ability to characterize changes in the machinery condition, and thus it is a feasible method of fusing faulty information from multiple sensors for accurate fault diagnosis.

50 samples under each condition from a single sensor were subsequently selected, and 30 of these samples were used

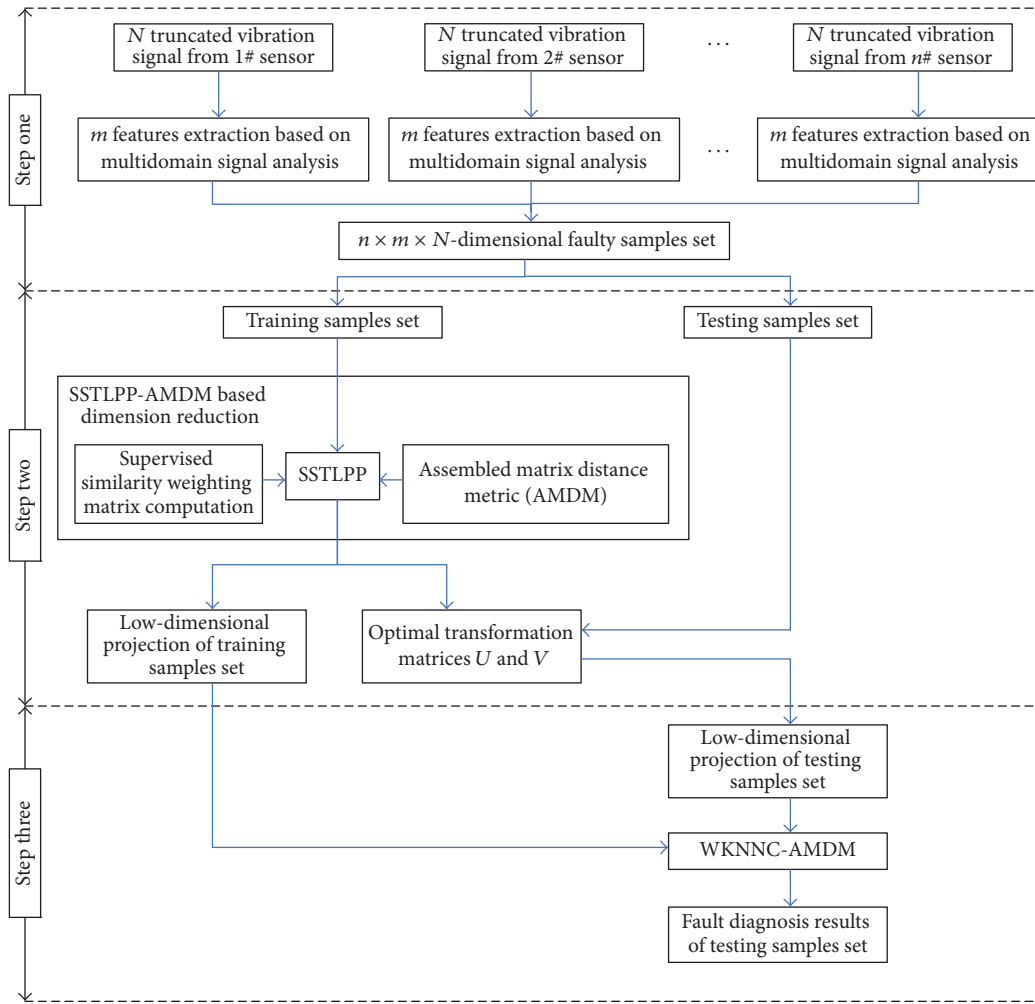


FIGURE 1: Implementation process of the proposed fault diagnosis method based on SSTLPP and WKNNC.

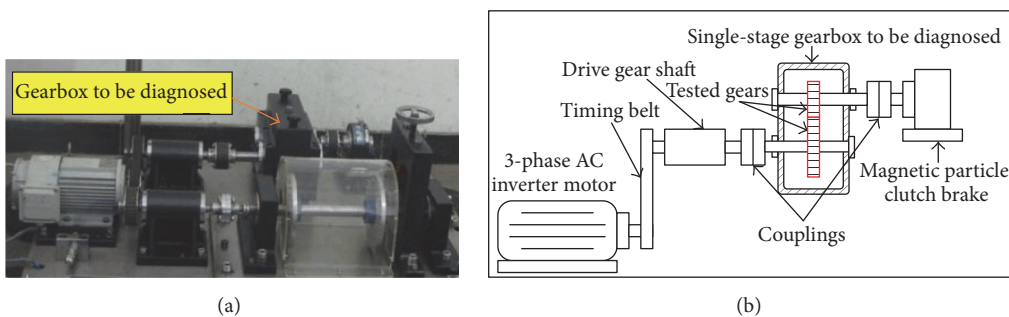


FIGURE 2: Gearbox fault simulation test setup: (a) overview and (b) diagram.

to train the fault diagnosis model, with the remaining samples used for the testing purposes. The length of each sample was 1024. Furthermore, five time-domain feature parameters and five frequency-domain parameters were calculated to construct a feature set: root mean square, skewness, kurtosis, impulse factor, peak factor, mean frequency, frequency

center, root mean square frequency, standard deviation frequency, and kurtosis frequency, as commonly defined in the previous literature [31, 32]. Accordingly, a $7 \times 10 \times 300$ -dimensional tensor formed sample set labeled with the corresponding classes was modeled to act as the input for the entire fault diagnosis experiment, which was composed of

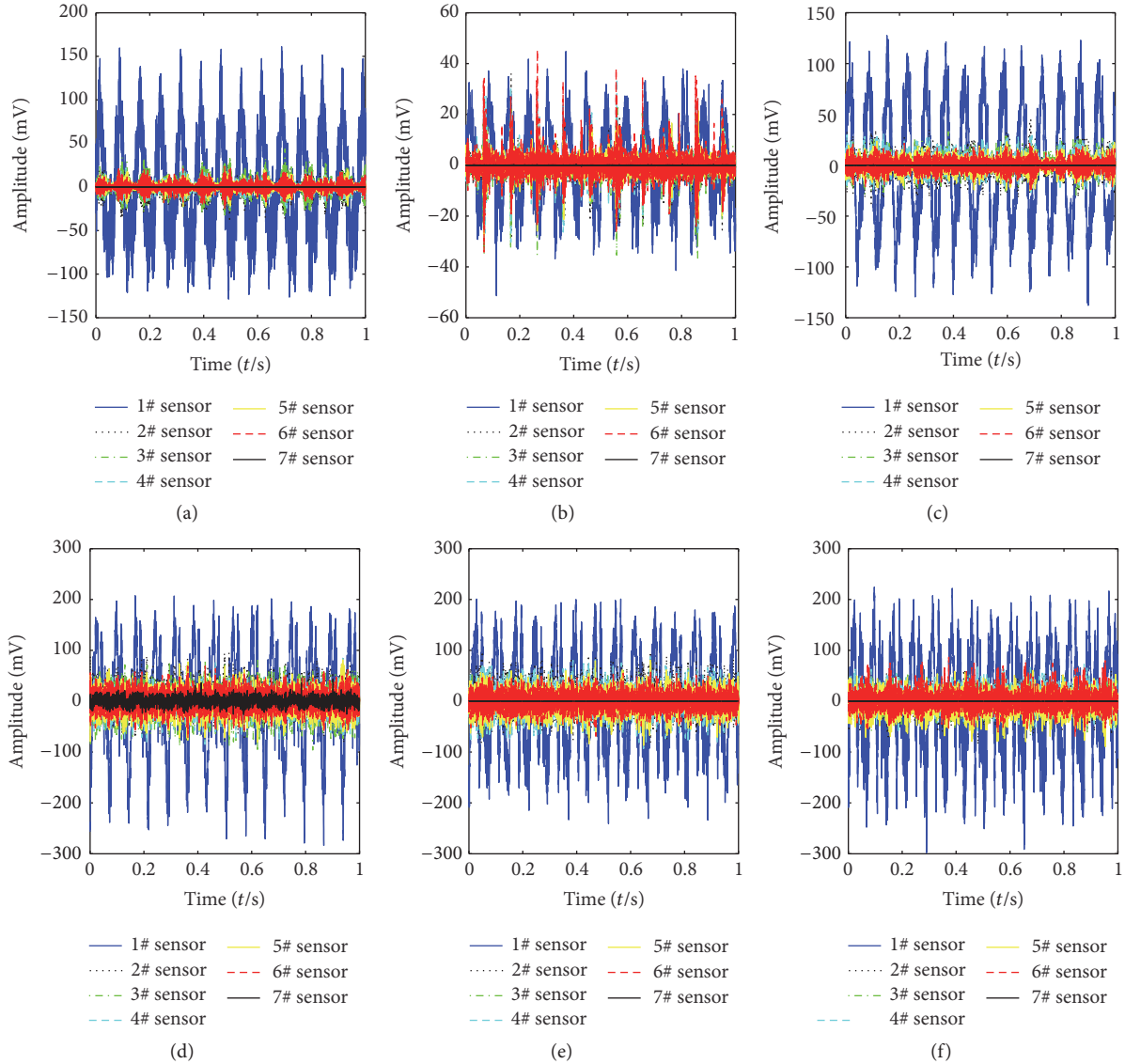


FIGURE 3: Test signals originating from the seven different sensors under the following conditions: (a) Norm, (b) C_G, (c) B_G, (d) W_P, (e) B_G_C_W_P, and (f) C_G_C_W_P.

TABLE 1: Specific information of seven sensors.

| Number of sensor | Type of sensor | Location | Direction |
|------------------|----------------|--|-----------|
| 1# | Displacement | Input shaft | X |
| 2# | Displacement | Input shaft | Y |
| 3# | Accelerometer | Bearing housing of input shaft in the side of motor | Y |
| 4# | Accelerometer | Bearing housing of output shaft in the side of motor | Y |
| 5# | Accelerometer | Bearing housing of input shaft in the side of load | Y |
| 6# | Accelerometer | Bearing housing of output shaft in the side of load | Y |
| 7# | Accelerometer | Bearing housing of output shaft in the side of load | X |

a $7 \times 10 \times 180$ -dimensional tensor formed sample set for the training sample set and the remainder as the testing set.

5.2. Performance Analysis and Comparison of Different Dimension Reduction Algorithms. This subsection validates the effectiveness of the proposed SSTLPP-AMDM algorithm for dimension reduction in fault diagnosis of rotation machinery, as well as its superiority to the traditional STLPP method. Using the calculation procedure described in Section 2.2 with the training sample set constructed above as the input, SSTLPP-AMDM based dimension reduction is implemented to obtain the explicit transformation matrices U and V , as well as the low-dimensional second-order tensor formed projection Y . In this experiment, the neighbor parameter k was set to 13, the similarity penalty

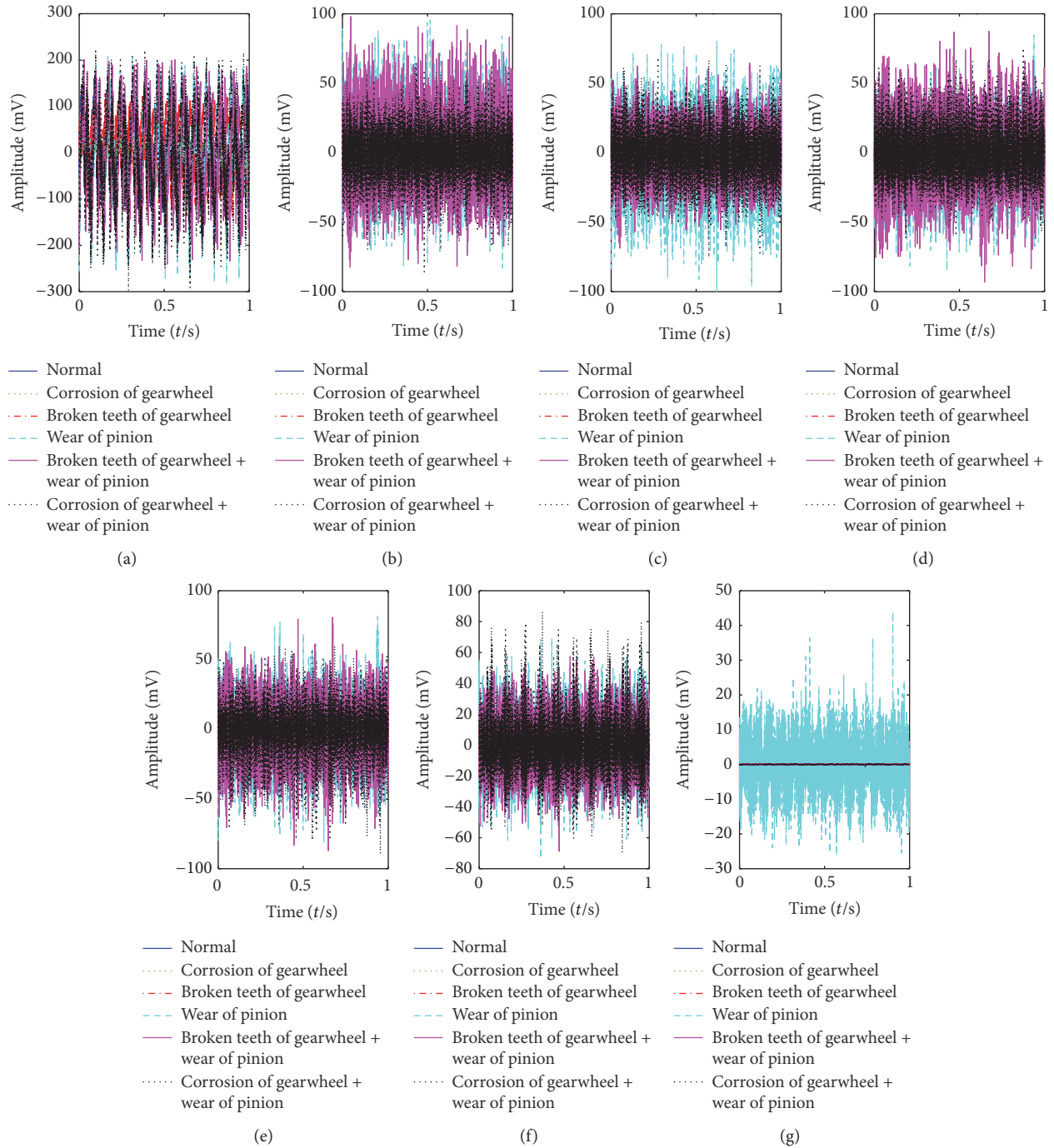


FIGURE 4: Test signals originating from each single sensor under six conditions: (a) 1# sensor, (b) 2# sensor, (c) 3# sensor, (d) 4# sensor, (e) 5# sensor, (f) 6# sensor, and (g) 7# sensor.

coefficient m was 2, and the value of p in the employed AMDM was 0.35. The first three-dimensional diagram of the vector-represented dimension reduction result Y is shown in Figure 5(a). For comparison purposes, the traditional STLPP-based dimension reduction result for the same input is also displayed in Figure 5(b). It can be observed from the scatter plots that, after dimension reduction, the SSSLPP-AMDM based samples have better separation with a distinct

clustering distribution between classes. In contrast, after dimension reduction, the traditional STLPP-based samples show inferior plots with some overlapping samples, which indicates that the proposed SSSLPP-AMDM dimension reduction algorithm is superior to the traditional STLPP in terms of the clustering performance of low-dimensional projection of the original high-dimensional second-order tensor formed samples.

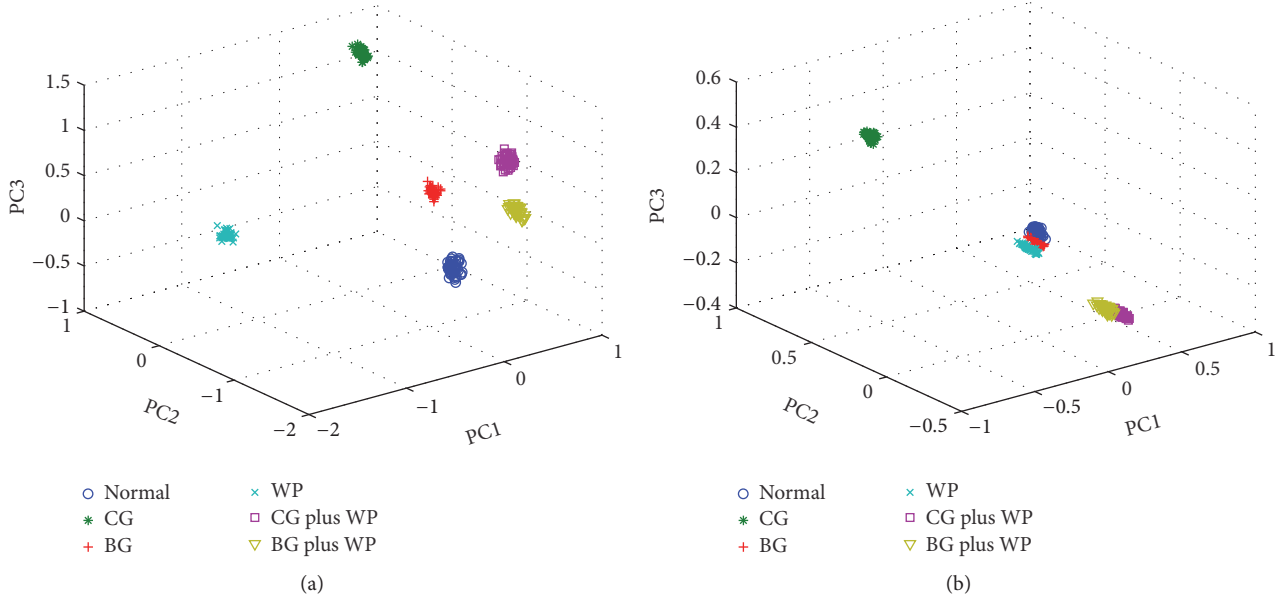


FIGURE 5: Scatter plots of the vector-formed dimension reduction result based on different algorithms for the training sample set: (a) SSTLPP-AMDM and (b) STLPP.

For further confirmation of the superiority of the proposed second-order tensor formed faulty samples originating from multisensor fusion over the vector-represented multisensor fused samples and the prevailing vector-formed faulty samples that originated from merely a single sensor, two further groups of experiments were designed. These experiments are the LPP-based dimension reduction of vector expressed multisensor fused samples (LPP-VM) and the LPP-based dimension reduction of faulty samples from any single sensor (LPP-VS) and their purpose is to compare with SSTLPP-AMDM which is input by the proposed second-order tensor-represented faulty samples in terms of the dimension reduction effect shown as Figure 5(a). The dimension reduction results of these two experiments are displayed in Figures 6(a) and 6(b), respectively. In contrast to Figure 5(a), these results demonstrate that the proposed second-order tensor-represented multisensor fused samples combined with SSTLPP-AMDM achieve the best clustering performance of the three experiments, that is, LPP-VM, LPP-VS, and the proposed SSTLPP-AMDM. Beyond that, by comparing Figure 6(a) with Figure 6(b), it can be intuitively seen that the first diagram shows better sample clustering results than the second diagram, not only in terms of between-class decentralization but also in terms of within-class aggregation. This indirectly demonstrates the benefit of using multisensor data fusion to increase the integrity of the fault information. In order to ensure the precision of these experimental verification conclusions, two other sets of comparison experiments were further implemented. The first was a quantitative analysis of the dimension reduction results of ten types of approaches which respectively adopted SSTLPP-AMDM, STLPP, and LPP as well as three different inputs including the second-order tensor formed multisensor fused data, the vector-represented multisensor fused data,

and the vectored samples data from a single sensor. The detailed comparison results are provided in Table 2 and Figure 7. The other group of experiments compared the classification accuracies of these ten types of fault feature dimension reduction results using three popular intelligent fault classifiers: a support vector machine (SVM), a multilayer perception (MLP) neural network, and a support vector data description (SVDD). The specific experimental description is discussed in the following paragraphs and the comparison results are shown in Table 3.

The results shown in the scatter distribution diagrams in Figure 6 were used for the qualitative analysis to assess the characteristics of the dimension reduction results based on SSTLPP-AMDM, STLPP, and LPP combined with different inputs. Three commonly used clustering performance measure indicators were used to quantitatively evaluate the ability of the dimension reduction algorithms to be used for subsequent fault classification: the within-class scatter S_w , the between-class scatter S_b , and the synthesized within-class-between-class scatter SS . The mathematical equations for these three indicators can be written as follows:

$$\begin{aligned}
 S_w &= \frac{1}{C} \sum_{i=1}^C \left(\frac{1}{N_i} \sum_{j=1}^{N_i} \|f_j^i - \mu_f^i\|^2 \right) \\
 S_b &= \frac{1}{C} \sum_{i=1}^C \|\mu_f^i - \bar{\mu}_f\|^2 \\
 SS &= \frac{S_b}{S_w},
 \end{aligned} \tag{11}$$

where C is the number of conditional classes, $N = \sum_{i=1}^C N_i$ is the total number of samples where N_i is the number of samples belonging to the i th class, f_j^i is the feature value of

TABLE 3: Comparison of fault classification results based on three classifiers and ten types of reduced feature sets.

| Classifier | Classification accuracy (%) | | | | | | | | | |
|--------------------|-----------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | F1 | F2 | F3 | F4 | F5 | F6 | F7 | F8 | F9 | F10 |
| SVM | 99.8 | 97.5 | 96.0 | 88.67 | 89.17 | 92.22 | 84.5 | 86.8 | 87.5 | 83.33 |
| MLP neural network | 83.33 | 66.67 | 66.67 | 65.0 | 64.17 | 63.33 | 65.0 | 63.5 | 63.33 | 65.0 |
| SVDD | 98.5 | 97.22 | 94.33 | 90.33 | 94.17 | 67.5 | 83.33 | 76.67 | 82.22 | 82.5 |

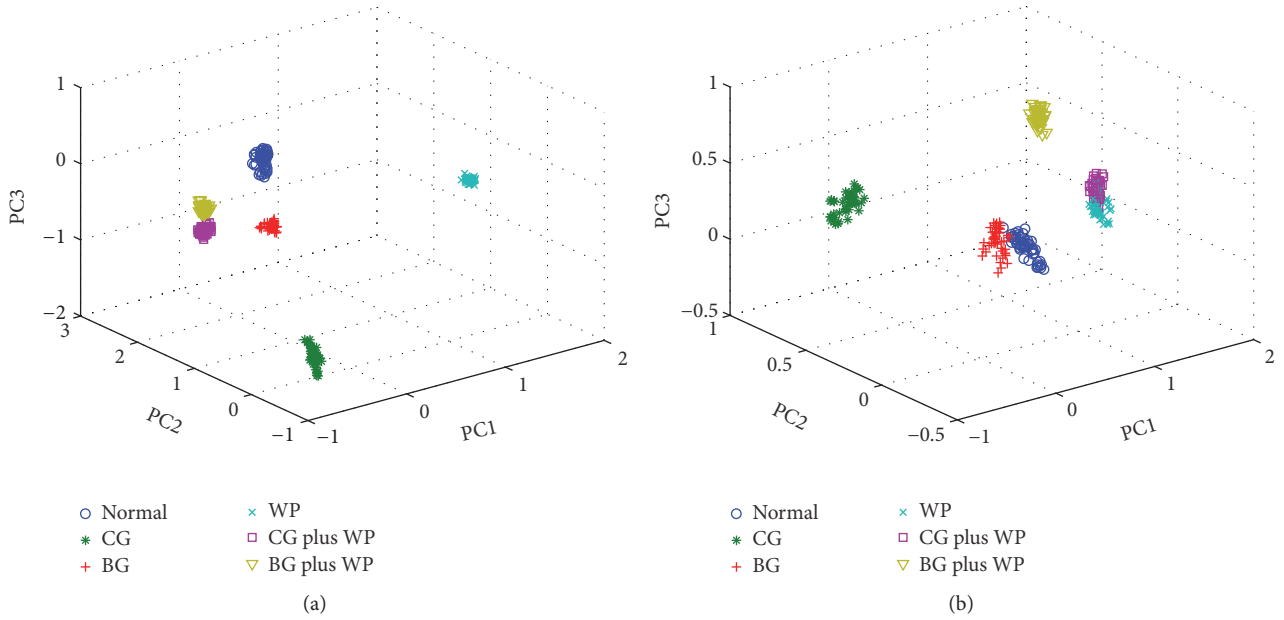


FIGURE 6: Scatter plots of dimension reduction results based on LPP with different input data: (a) LPP-VM and (b) LPP-VS.

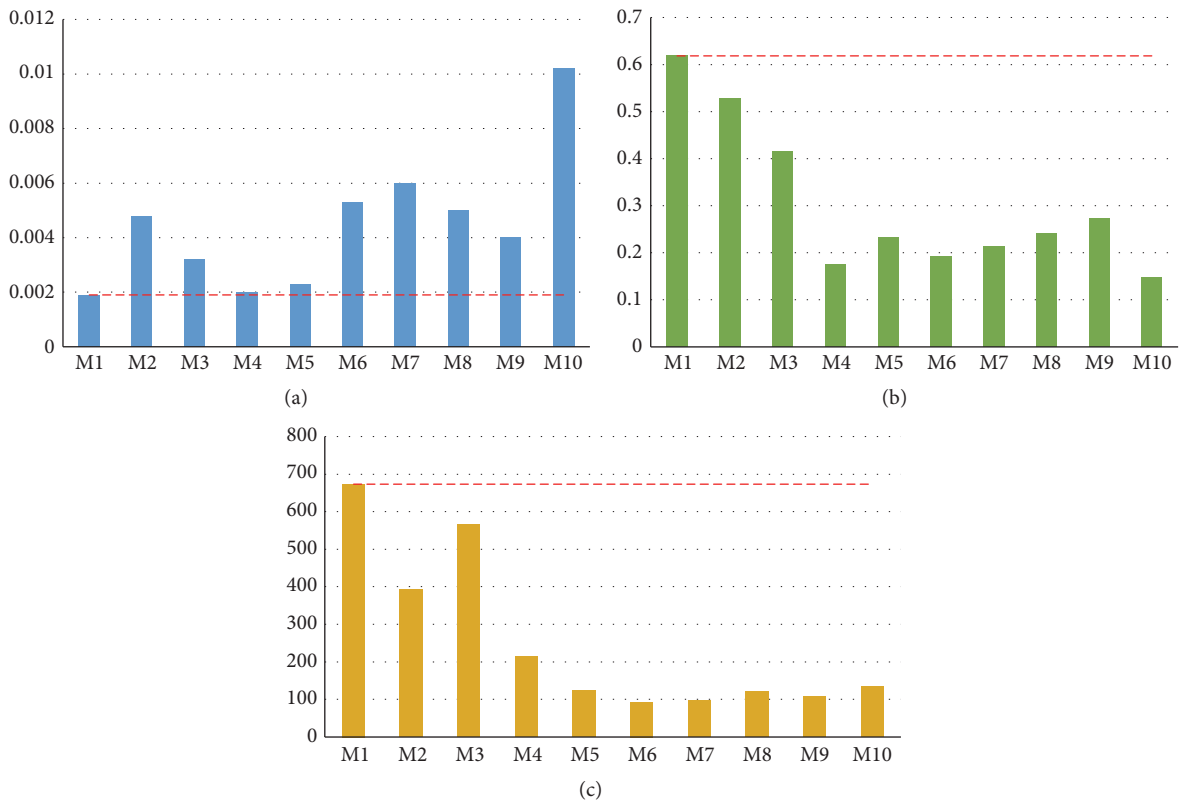


FIGURE 7: Comparison of scatter parameter values based on different methods: (a) S_w , (b) S_b , and (c) SS.

the j th sample in the i th class, μ_f^i is the mean feature value of the i th class, and $\bar{\mu}_f = 1/C \sum_{i=1}^C \mu_f^i$ is the total mean feature value of all classes. It should be noted that the clustering performance of each feature is proportional to the values of the between-class scatter S_b and the synthesized within-class-between-class scatter SS and yet is inversely proportional to the value of the within-class scatter S_w .

The previously-mentioned training sample set is used, which contains six-class faulty condition data for the gearbox as the input. Ten groups of experiments are performed to calculate the corresponding scatter parameters of the first three-dimensional features of the vectored dimension reduction results: (1) SSTLPP-AMDM based dimension reduction for second-order tensor formed multisensor fused data (SSTLPP-AMDM for STMD), (2) traditional STLPP-based dimension reduction for second-order tensor formed multisensor fused data (STLPP for STMD), (3) LPP-based dimension reduction for vector-represented multisensor fused data (LPP for VMD), and (4)–(10) LPP-based dimension reduction of vectored sample data from seven different positional sensors (LPP for VSD1~7). The scatter computation results for each dimensional feature based on each of the different ten methods are shown in Table 2 and the corresponding average scatter parameter values are displayed in Figure 7. It can be seen that the SSTLPP-AMDM based dimension reduction results of tensor formed multisensor fused samples have the smallest within-class scatter S_w , the largest between-class scatter S_b , and thus the largest synthesized scatter SS . The traditional STLPP-based dimension reduction for tensor formed multisensor fused data and the LPP-based dimension reduction for vector-represented multisensor fused data obtain larger S_w values, smaller S_b values, and smaller SS values than SSTLPP-AMDM for STMD but achieve smaller S_w values, larger S_b values, and larger SS values than the other seven types of LPP-based dimension reduction approaches for vectored sample data originating from a single sensor. These comparison and analysis results indicate that SSTLPP-AMDM for STMD is much more effective than any of the other nine dimension reduction methods for different types of samples data in terms of the clustering performance of the dimension reduction results.

As mentioned earlier, in order to acquire direct evidence of the superiority of the proposed SSTLPP-AMDM algorithm as well as the multisensor data fusion, three frequently-used intelligent classifiers (SVM, MLP neural network, and SVDD), respectively, acted on the first three-dimensional features of the vectored dimension reduction results of the ten methods (M1~M10), which are marked as F1~F10. Each experiment is carried out ten times. For the SVM classifier, this paper employs a radial basis kernel function and the value of the kernel parameter is 1. For the MLP neural network, the commonly used three layers structure is employed: input layer, hidden layer, and output layer, and the numbers of nodes in the input and output layers are set to 3 and 6. These values depend on the number of input features and output classes. The geometric pyramid rule determines that the number of hidden layer nodes is 5. The Gaussian kernel function is used for the SVDD model, and the corresponding

kernel parameter is set to 3. The classification results of the three models which are applied to each of the ten types of feature sets originating from the previous experiment are listed in Table 3.

As shown in Table 3, the reduced feature set of the tensor formed multisensor fused fault data based on the proposed SSTLPP-AMDM dimension reduction algorithm (F1) achieves higher classification accuracy than the other nine types of reduced feature sets (F2~F10) for all three classifiers. The reduced feature sets of the tensor formed multisensor fused fault data based on the traditional STLPP algorithm (F2) and that of the vector-represented multisensor fused data based on LPP (F3) achieve the second and third highest classification accuracy. The seven types of reduced feature sets of vectored fault data originating from a single sensor (F4~F10) have the lowest level of classification accuracy. These results further confirm the effectiveness of the proposed SSTLPP-AMDM combined with the formulated tensor-represented multisensor data fusion to increase the amount of useful information in the feature set and facilitate the subsequent classification task.

5.3. Overall Performance Validation of the Proposed Fault Diagnosis Approach. The following experiments and analysis were also employed to verify the superiority of the proposed WKNNC-AMDM method, as well as the overall fault diagnosis approach proposed by this paper. Using the implementation procedure for the proposed fault diagnosis method shown in Figure 1, a final fault diagnostic result is achieved by inputting the low-dimensional tensor formed testing sample set after dimension reduction with SSTLPP-AMDM into WKNNC-AMDM. Furthermore, the classification performance of WKNNC-AMDM combined with the low-dimensional second-order tensor formed multisensor fused sample data after dimension reduction is compared with the WKNNC-FDM and the KNNC-FDM, which both have the same input data. For all three classifiers, the neighborhood size was set to 13. Each experiment was performed ten times and the classification results of the three classifiers for the fault sample data of a gearbox including the cumulative number of false classification samples (Cum. number of FCS), the distribution of false classification samples within the six different faulty classes (number of FCS within Classes 1~6), and the total testing accuracy are listed in detail in Table 4.

It can be seen from Table 4 that although the same input data is used for the three classifiers, namely, the low-dimensional tensor formed testing sample set after dimension reduction with SSTLPP-AMDM, the proposed WKNNC under AMDM has a classification accuracy of 100%, which is higher than the 89.17% accuracy achieved using WKNNC under traditional FDM and the 70.83% accuracy achieved using the classical KNNC with FDM. These results indicate that the WKNNC-AMDM method has superior classification performance to WKNNC-FDM and KNNC-FDM, due to the addition of the assembled matrix distance metric for the similarity representation of the second-order tensor formed samples and the weighted voting strategy for the nearest neighbor classifier. This experiment also effectively

TABLE 4: Fault diagnosis results of three different classifiers with the reduced low-dimensional tensor formed multisensor fused samples.

| Classifier | Cum. number of FCS | Norm | C_G | B_G | W_P | B_G_C_W_P | C_G_C_W_P | Total testing accuracy (%) |
|------------|--------------------|------------------------------|------------------------------|------------------------------|------------------------------|------------------------------|------------------------------|----------------------------|
| | | Number of FCS within Class 1 | Number of FCS within Class 2 | Number of FCS within Class 3 | Number of FCS within Class 4 | Number of FCS within Class 5 | Number of FCS within Class 6 | |
| KNNC-FDM | 35 | 16 | 15 | 0 | 3 | 0 | 1 | 70.83 |
| WKNNC-FDM | 13 | 0 | 13 | 0 | 0 | 0 | 0 | 89.17 |
| WKNNC-AMDM | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 |

demonstrated the performance of the overall proposed total fault diagnosis framework, which comprehensively includes the SSSLPP-AMDM based dimension reduction and the WKNNC-AMDM.

6. Conclusions

This paper has presented a novel multisensor fused fault diagnosis approach for rotation machinery based on SSSLPP-AMDM and WKNNC-AMDM. Based on significant experimental analysis and comparisons that were performed, the main conclusions can be summarized as follows.

- (1) In contrast with traditional STLPP, the proposed SSSLPP-AMDM algorithm can obtain better dimension reduction effects for the original high-dimensional second-order tensor-represented samples. This was achieved by the addition of the class label information and improvement of the similarity evaluation method for matrix formed samples by AMDM. Furthermore, it was also verified that SSSLPP-AMDM based dimension reduction of multisensor fused second-order tensor formed samples is superior to LPP-based dimension reduction of multisensor fused vector-formed samples and LPP-based dimension reduction of vector-formed samples from a single sensor in terms of the clustering performance of samples of different classes after reduction.
- (2) The proposed WKNNC-AMDM can obtain higher classification accuracy than WKNNC-FDM and KNNC-FDM due to the introduction of weighted voting strategy and assembled matrix distance metric for similarity representation of second-order tensor formed samples.
- (3) Using the advantages of second-order tensor formed multisensor fused faulty sample representation, SSSLPP-AMDM for efficient dimension reduction, and WKNNC-AMDM for rapid fault classification, the proposed fault diagnosis approach achieves higher classification accuracy for rotation machinery than the other homogenous methods.

In summary, the proposed fault diagnosis approach has the following strengths: more adequate fault information, lower calculation complexity, and higher fault recognition accuracy. Therefore, it is extremely suited to engineering applications for fault diagnosis of rotation machinery.

Competing Interests

The authors declare that they have no competing interests.

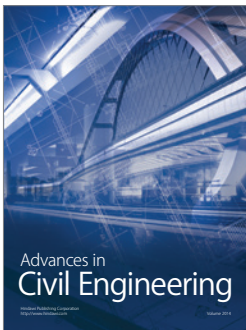
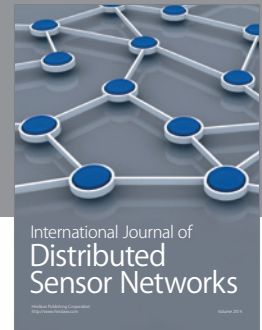
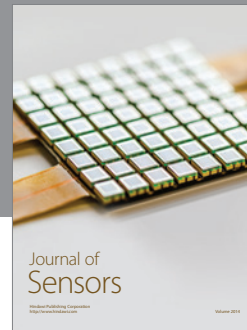
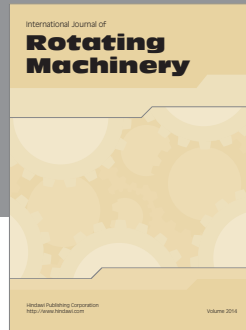
Acknowledgments

This research is supported by National Natural Science Foundation of China (no. 51575143).

References

- [1] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 108–126, 2013.
- [2] Z. Feng, M. Liang, and F. Chu, "Recent advances in time-frequency analysis methods for machinery fault diagnosis: a review with application examples," *Mechanical Systems and Signal Processing*, vol. 38, no. 1, pp. 165–205, 2013.
- [3] Y. Lei, J. Lin, M. J. Zuo, and Z. He, "Condition monitoring and fault diagnosis of planetary gearboxes: a review," *Measurement*, vol. 48, no. 1, pp. 292–305, 2014.
- [4] Y. Wang, J. Xiang, R. Markert, and M. Liang, "Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: a review with applications," *Mechanical Systems and Signal Processing*, vol. 66-67, pp. 679–698, 2016.
- [5] Z. Li, X. Yan, Z. Tian, C. Yuan, Z. Peng, and L. Li, "Blind vibration component separation and nonlinear feature extraction applied to the nonstationary vibration signals for the gearbox multi-fault diagnosis," *Measurement*, vol. 46, no. 1, pp. 259–271, 2013.
- [6] J. Zhang, W. Ma, J. Lin, L. Ma, and X. Jia, "Fault diagnosis approach for rotating machinery based on dynamic model and computational intelligence," *Measurement*, vol. 59, pp. 73–87, 2015.
- [7] Z. Li, X. Yan, C. Yuan, Z. Peng, and L. Li, "Virtual prototype and experimental research on gear multi-fault diagnosis using wavelet-autoregressive model and principal component analysis method," *Mechanical Systems and Signal Processing*, vol. 25, no. 7, pp. 2589–2607, 2011.
- [8] S. W. Ji and J. P. Ye, "Generalized linear discriminant analysis: a unified framework and efficient model selection," *IEEE Transactions on Neural Networks*, vol. 19, no. 10, pp. 1768–1782, 2008.
- [9] T. F. Cox and M. A. Cox, *Multi-Dimensional Scaling*, Chapman & Hall, London, UK, 1994.
- [10] 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.

- [11] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
- [12] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation*, vol. 15, no. 6, pp. 1373–1396, 2003.
- [13] Z. Y. Zhang and H. Y. Zha, "Principal manifolds and nonlinear dimensionality reduction via tangent space alignment," *SIAM Journal on Scientific Computing*, vol. 26, no. 1, pp. 313–338, 2004.
- [14] F. Li, B. Tang, and R. Yang, "Rotating machine fault diagnosis using dimension reduction with linear local tangent space alignment," *Measurement: Journal of the International Measurement Confederation*, vol. 46, no. 8, pp. 2525–2539, 2013.
- [15] X. Ding, Q. He, and N. Luo, "A fusion feature and its improvement based on locality preserving projections for rolling element bearing fault classification," *Journal of Sound and Vibration*, vol. 335, pp. 367–383, 2015.
- [16] W. Wen, R. X. Gao, and W. Cheng, "Planetary gearbox fault diagnosis using envelope manifold demodulation," *Shock and Vibration*, vol. 2016, Article ID 3952325, 13 pages, 2016.
- [17] X. He, D. Cai, and P. Niyogi, "Tensor subspace analysis," in *Advances in Neural Information Processing Systems*, pp. 499–506, 2005.
- [18] G. Dai and D.-Y. Yeung, "Tensor embedding methods," in *Proceedings of the Neural Conference on Artificial Intelligence*, pp. 330–335, July 2006.
- [19] X. Li, S. Lin, S. Yan, and D. Xu, "Discriminant locally linear embedding with high-order tensor data," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 38, no. 2, pp. 342–352, 2008.
- [20] K. Lee and H. Park, "Probabilistic learning of similarity measures for tensor PCA," *Pattern Recognition Letters*, vol. 33, no. 10, pp. 1364–1372, 2012.
- [21] W. Zuo, D. Zhang, and K. Wang, "An assembled matrix distance metric for 2DPCA-based image recognition," *Pattern Recognition Letters*, vol. 27, no. 3, pp. 210–216, 2006.
- [22] N. Saravanan and K. I. Ramachandran, "Incipient gear box fault diagnosis using discrete wavelet transform (DWT) for feature extraction and classification using artificial neural network (ANN)," *Expert Systems with Applications*, vol. 37, no. 6, pp. 4168–4181, 2010.
- [23] F. Chen, B. Tang, T. Song, and L. Li, "Multi-fault diagnosis study on roller bearing based on multi-kernel support vector machine with chaotic particle swarm optimization," *Measurement*, vol. 47, no. 1, pp. 576–590, 2014.
- [24] N. Saravanan, S. Cholairajan, and K. I. Ramachandran, "Vibration-based fault diagnosis of spur bevel gear box using fuzzy technique," *Expert Systems with Applications*, vol. 36, no. 2, pp. 3119–3135, 2009.
- [25] 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.
- [26] R. Stoklasa, T. Majtner, and D. Svoboda, "Efficient k -NN based HEp-2 cells classifier," *Pattern Recognition*, vol. 47, no. 7, pp. 2409–2418, 2014.
- [27] F. Li, J. Wang, B. Tang, and D. Tian, "Life grade recognition method based on supervised uncorrelated orthogonal locality preserving projection and K -nearest neighbor classifier," *Neurocomputing*, vol. 138, pp. 271–282, 2014.
- [28] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.
- [29] J. Yang, D. Zhang, A. F. Frangi, and J.-Y. Yang, "Two-dimensional PCA: a new approach to appearance-based face representation and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131–137, 2004.
- [30] G. Gates, "The reduced nearest neighbor rule," *IEEE Transactions on Information Theory*, vol. 18, no. 3, pp. 431–433, 1972.
- [31] B. Samanta, "Gear fault detection using artificial neural networks and support vector machines with genetic algorithms," *Mechanical Systems and Signal Processing*, vol. 18, no. 3, pp. 625–644, 2004.
- [32] Y. Lei, Z. He, Y. Zi, and X. Chen, "New clustering algorithm-based fault diagnosis using compensation distance evaluation technique," *Mechanical Systems and Signal Processing*, vol. 22, no. 2, pp. 419–435, 2008.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

