Mechanical Systems and Signal Processing 25 (2011) 3125–3134

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Supervised locally linear embedding projection (SLLEP) for machinery fault diagnosis

Benwei Li, Yun Zhang*

Department of Airborne Vehicle Engineering, Naval Aeronautical and Astronautical University, Yantai 264001, China

ARTICLE INFO

Article history:
Received 20 December 2010
Received in revised form
28 March 2011
Accepted 5 May 2011
Available online 18 May 2011

Keywords:
Machinery fault diagnosis
Vibration signal
Manifold learning
Supervised locally linear embedding

ABSTRACT

Following the intuition that the measured signal samples usually distribute on or near the nonlinear low-dimensional manifolds embedded in the high-dimensional signal space, this paper proposes a new machinery fault diagnosis approach based on supervised locally linear embedding projection (SLLEP). The approach first performs the recently proposed manifold learning algorithm supervised locally linear embedding (SLLE) on the high-dimensional fault signal samples to learn the intrinsic embedded multiple manifold features corresponding to different fault modes, and map them into a low-dimensional embedded space to achieve fault feature extraction. For dealing with the new fault sample, the approach then applies local linear regression to find the projection that best approximates the implicit mapping from high-dimensional samples to the embedding. Finally fault classification is carried out in the embedded manifold space. The ball bearing data and rotor bed data are both used to validate the proposed approach. The results show that the proposed approach obviously improves the fault classification performance and outperform the other traditional approaches.

1. Introduction

Machinery fault diagnosis is essentially a kind of pattern recognition problem, in the process of which the feature extraction is fundamental and the recognition method is at the core. Sensors provide a large number of high-dimensional measurements, i.e. samples that contain much useful information for fault identification. For a measured vibration signal sample of length D, it can be treated as a data point in the high-dimensional signal space $\mathbb{R}^D$. Different fault classes usually have different data distributions in the signal space, according to which the fault recognition can be achieved. However, the ideal decision boundary between different fault classes in such signal space is often highly nonlinear. A classifier should therefore have many degrees of freedom, and consequently a large number of parameters. So applying a classifier directly on such high-dimensional data for fault recognition is quite complicated because too many parameters have to be estimated using a limited number of samples. This is the well-known "curse of dimensionality". A widely used way to attempt to resolve the problem is feature extraction, which aims to project the original high-dimensional data into a lower dimensional feature space that reflects the inherent structure of the original data and holds the useful information as much as possible. At last a suitable classifier is adopted to perform fault classification in low-dimensional space.

For fault feature extraction, principal component analysis (PCA) [1] and linear discriminate analysis (LDA) [2] are the classical dimensionality reduction methods. However, since they are linear, their performances degenerate for nonlinear data where the underlying low-dimensional structure has nonlinear manifolds (such as nonlinear curves and surfaces)

* Corresponding author. Tel./fax: +865356635573.
E-mail address: hjhy_zy@126.com (Y. Zhang).

0888-3270/$ - see front matter Crown Copyright © 2011 Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.ymssp.2011.05.001

rather than linear subspaces. Furthermore, PCA is completely unsupervised for not taking the class information of the input data into account, which may probably discard much useful information and weaken the recognition accuracy, especially when the number of sample points is very large [3]. LDA can produce an optimally discriminant projection by considering the labels of the input data. However there still exist some drawbacks in LDA. When the dimension of data is too high and training data is inadequate, LDA may suffer from "curse of dimensionality" so that it cannot find the best projection directions for classification. Moreover, the discriminating power of LDA is also limited in that the dimension of its reduced space cannot be larger than the class number minus one [3].

Some nonlinear dimensionality reduction methods have been proposed, but many of them, such as autoassociative neural network [4] and principal curves [5], involve a difficult nonlinear optimization problem. Since three papers of manifold learning are published in Science (2000) [6–8], the geometrically motivated manifold learning has drawn much attention and proven to be a powerful nonlinear dimensionality reduction tool. The goal of manifold learning is to map the high-dimensional data into an embedded space by preserving the local neighborhood information and can discover the intrinsic manifold structure of nonlinear high-dimensional data effectively. The locally linear embedding (LLE) algorithm is one of such representative manifold learning methods, which unravels the underlying data manifold in its reduced space based on the assumption that each data point and its neighbors lie on a locally linear patch of the manifold. LLE has been applied successfully in computer vision [7,8] and image processing [9,10]. However, some limitations are exposed when LLE is applied to pattern recognition. One limitation lies in that the classical LLE neglects the class information, which will impair the recognition accuracy. The supervised LLE (abbreviated as SLLE) algorithm is proposed by Ridder et al. [11] as the supervised expansion of manifold learning, which uses class label information to improve the performance of classification. Another limitation is the out-of-sample problem. Unlike PCA and LDA, LLE does not provide an explicit mapping from high-dimensional space to low-dimensional space, whereas fault diagnosis requires the mapping to conduct dimensionality reduction on newly measured data samples. Saul and Roweis [12] proposed a parametric model to process the new data on some prior information of the probability distribution. However, the learning of parameters tends to be plagued by very poor local minima [12].

In this paper, aimed at the difficulty of high-dimensional nonlinear fault data, a new machinery fault diagnosis approach based on supervised locally linear embedding projection (SLLEP) is proposed. The approach first performs the SLLE algorithm on the data of high-dimensional signal space to learn the intrinsic embedded multiple manifold features corresponding to different fault modes, and map them into an embedded space from the high-dimensional space to achieve fault feature extraction. Then the approach applies local linear regression to find the projection that best approximates the implicit mapping from high-dimensional data to the embedding; so the embedded coordinate of the new fault signal samples can be obtained by learned projection. Finally fault classification is carried out in the embedded manifold space. The vibration signals acquired from roller bearings are employed to validate the proposed approach. The results show that the proposed SLLEP-based approach obviously improves the fault classification performance and outperform the other traditional approaches.

2. Supervised locally linear embedding projection

2.1. Locally linear embedding algorithm

Let $\boldsymbol{X}=\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N\}$ be a set of N points in a high-dimensional data space $\mathbb{R}^D$. The data points are assumed to lie on or near a nonlinear manifold of intrinsic dimensionality $d < D$ (typically $d \ll D$). The goal of LLE is to find a low-dimensional embedding of $\boldsymbol{X}$ by mapping the D-dimensional data into a single global coordinate system in $\mathbb{R}^d$. Let us denote the corresponding set of N points in the embedded space $\mathbb{R}^d$ by $\boldsymbol{Y}=\{\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_N\}$.

The LLE algorithm [7] can be summarized as follows:

(A) for each data point $\boldsymbol{x}_i \in \boldsymbol{X}$, find the k-nearest neighbors of $\boldsymbol{x}_i$;
(B) compute the reconstruction weights of the neighbors that minimize the error of reconstructing $\boldsymbol{x}_i$ and
(C) compute the low-dimensional embedding $\boldsymbol{Y}$ for $\boldsymbol{X}$ that best preserves the local geometry represented by the reconstruction weights.

Step (A) is typically done using the Euclidean distance to define neighborhood, although more sophisticated criteria may also be used. Based on the k-nearest neighbors identified, step (B) seeks to find the best reconstruction weights. Optimality is achieved by minimizing the local reconstruction error for $\boldsymbol{x}_i$

$$\varepsilon_i = \left\| \boldsymbol{x}_i - \sum_{j=1}^{N} w_{ij} \boldsymbol{x}_j \right\|^2 \tag{1}$$

which is the squared distance between $\boldsymbol{x}_i$ and its reconstruction, subjected to the constraints $\sum_{j=1}^{k} w_{ij} = 1$ and $w_{ij}=0$ if $\boldsymbol{x}_i$ is not a neighbor of $\boldsymbol{x}_i$. Minimizing ε_i subjected to the constraints is a constrained least squares problem. After repeating steps (A) and (B) for all N data points in $\boldsymbol{X}$, the reconstruction weights are obtained from a weight matrix $\boldsymbol{W}=[w_{ij}]_{N\times N}$. Step (C) of the LLE algorithm is to compute the best low-dimensional embedding $\boldsymbol{Y}$ based on the weight matrix $\boldsymbol{W}$ obtained.

This corresponds to minimizing the following cost function:

$$\phi(\boldsymbol{Y}) = \sum_{i=1}^{N} \left\| \boldsymbol{y}_i - \sum_{j=1}^{N} w_{ij} \boldsymbol{y}_j \right\|^2 \tag{2}$$

subject to the constraints $\sum_{i=1}^{N} \boldsymbol{y}_i = \boldsymbol{0}$, $1/N \sum_{i=1}^{N} \boldsymbol{y}_i \boldsymbol{y}_i^T = \boldsymbol{I}$, where $\boldsymbol{0}$ is a column vector of zeros and $\boldsymbol{I}$ is an identity matrix. Based on $\boldsymbol{W}$, we can define a sparse, symmetric, and positive semidefinite matrix $\boldsymbol{M}$ as follows:

$$\boldsymbol{M} = (\boldsymbol{I} - \boldsymbol{W})^T (\boldsymbol{I} - \boldsymbol{W}) \tag{3}$$

Note that Eq. (2) can be expressed in the quadratic form, $\phi(\boldsymbol{Y}) = \sum_{i,j} M_{ij} \boldsymbol{y}_i^T \boldsymbol{y}_j$, based on $\boldsymbol{M} = [M_{ij}]_{N \times N}$. By the Rayleigh–Ritz theorem [13], minimizing Eq. (2) with respect to the $\boldsymbol{y}_i$'s in $\boldsymbol{Y}$ can be done by finding the eigenvectors with the smallest (non-zero) eigenvalues.

The embedding discovered by LLE is easiest to visualize for intrinsically two-dimensional manifolds. In Fig. 1, for example, the input to LLE consists of $N = 2000$ data points sampled from the S-curve manifold. The nonlinear dimensionality reduction result in Fig. 1(c) with the color coding shows how the LLE algorithm, using $k = 12$ neighbors per data point, successfully unraveled the underlying two-dimensional structure.

2.2. Supervised locally linear embedding algorithm

LLE has been used as an unsupervised technique not considering the prior information. If the class label information can be used, the recognition ability will be improved in pattern recognition field. Supervised LLE (SLLE) [11] is introduced to deal with data sets containing multiple (often disjoint) manifolds, corresponding to classes. One approach to do this is to add a distance between samples $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ from different classes to modify only the first step of the original LLE, while leaving the other two steps unchanged. This can be achieved by artificially increasing the pre-calculated Euclidean distance between samples from different classes, but leaving them unchanged if samples are from the same classes:

$$S' = S + \alpha \max(S)(1 - \delta(\boldsymbol{x}_i, \boldsymbol{x}_j)) \tag{4}$$

where $S = \|\boldsymbol{x}_i - \boldsymbol{x}_j\|$ is the Euclidean distance without considering the class label information, $\max(S) = \max_{i,j} \|\boldsymbol{x}_i - \boldsymbol{x}_j\|$ is the maximum distance between samples, and S' is the distance integrating with the class label information. If $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ belong to different classes, then $\delta(\boldsymbol{x}_i, \boldsymbol{x}_j) = 0$ and $\delta(\boldsymbol{x}_i, \boldsymbol{x}_j) = 1$ otherwise. In this formulation, α controls the amount to which class information should be incorporated. When $\alpha = 0$, SLLE is equivalent to the original unsupervised LLE; when $\alpha = 1$, the result is the fully supervised LLE (called 1-SLLE). Varying α between 0 and 1 gives a partially supervised LLE (called α-SLLE). In practice, 1-SLLE usually does not outperform α-SLLE because of its inferior generalization [14,15].

2.3. Local linear projection for mapping

SLLE does not provide an explicit mapping from high-dimensional space to low-dimensional embedded space. Yet the mapping is very necessary for obtaining the embedded coordinate of the new unlabeled signal sample. Thus based on assumption that the underlying manifold is locally linear, we establish local linear projection that best approximates the implicit mapping from high-dimensional data to the embedding. The algorithm is implemented as follows.

Step 1. For the new high-dimensional sample $\boldsymbol{x}_{new}$, find its k neighbors $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ in the training set, here distance metric uses the Euclidean distance.

Step 2. Find mapping matrix $\boldsymbol{A}$ that satisfies $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X}$ in the least squares sense (an exact solution might not exist), where $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_d]^T \in \mathbb{R}^{d \times D}$, $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k]$ and $\boldsymbol{Y}[\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_k]$ is embedded coordinate corresponding to the neighbors

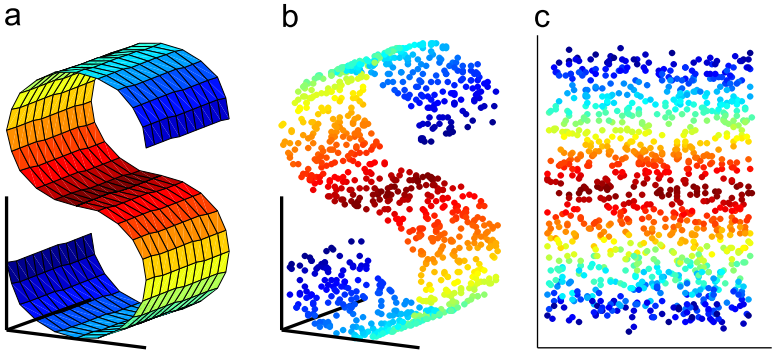

Fig. 1. (a) S-curve manifold in three-dimensional space; (b) 2000 data points sampled from the S-curve manifold; and (c) nonlinear dimensionality reduction result by LLE from three-dimensional space to two-dimensional space.

$\boldsymbol{X}$ of $\boldsymbol{x}_{new}$. Each basis vector $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_d$ is obtained by solving the linear least squares regression problem:

$$\boldsymbol{a}_j = \underset{\boldsymbol{a}}{\operatorname{argmin}} \sum_{i}^{k} (\boldsymbol{a}^T \boldsymbol{x}_i - y_i^j)^2 \tag{5}$$

where $1 \le j \le d$ and y_i^j is jth element of $\boldsymbol{y}_i$.

Step 3. The d-dimensional embedded coordinates $\boldsymbol{y}_{new}$ of $\boldsymbol{x}_{new}$ is calculates as $\boldsymbol{y}_{new} = \mathrm{A}\boldsymbol{x}_{new}$.

In step 2, if $\boldsymbol{X}^T$ has full column rank, $\boldsymbol{a}_j$ can be solved as

$$\boldsymbol{a}_j = (\boldsymbol{X}^T)^+ \boldsymbol{y}^j = (\boldsymbol{X}\boldsymbol{X}^T)^{-1}\boldsymbol{X}\boldsymbol{y}^j \tag{6}$$

where $(\boldsymbol{X}^T)^+$ is the Moore–Penrose pseudoinverse of $\boldsymbol{X}^T$ and $\boldsymbol{y}^j = [y_1^j, y_2^j, \cdots, y_N^j]^T$.

If $\boldsymbol{X}^T$ column rank is deficient (e.g. in the case where the number of neighbors k is smaller than data dimension D), problem (5) is ill posed, for which there may be infinitely many solutions. To make the learned projection generalize well to previously unseen samples, an effective way to tackle this issue is using the ridge regression [16], which minimizes a regularized least squares objective function

$$\boldsymbol{a}_j = \underset{\boldsymbol{a}}{\operatorname{argmin}} \left(\sum_{i=1}^{k} (\boldsymbol{a}^T \boldsymbol{x}_i - y_i^j)^2 + \beta \|\boldsymbol{a}\|^2 \right) \tag{7}$$

where $\beta > 0$ is the regularization parameter, and $\boldsymbol{a}_j$ can be solved as

$$\boldsymbol{a}_j = (\boldsymbol{X}\boldsymbol{X}^T + \beta \boldsymbol{I})^{-1}\boldsymbol{X}\boldsymbol{y}^j \tag{8}$$

where $\boldsymbol{I}$ denotes the $d \times d$ identity matrix.

The goal is to project each of these neighborhoods linearly into a d-dimensional space, and discard noise outside this space. Thus the value of β is determined by the noise level σ^2. Using eigenanalysis method, β can therefore be estimated by

$$\beta = \sigma^2 = \frac{1}{k-d} \sum_{i=d+1}^{k} \lambda_i \tag{9}$$

where λ_i denotes the non-zero eigenvalues of matrix $\boldsymbol{X}\boldsymbol{X}^T$ (λ_i is sorted from large to small).

3. Supervised locally linear embedding projection for fault diagnosis

Conventional fault diagnosis methods are good at solving the problems of linearity and weak nonlinearity, but not good at fault diagnosis of complexity nonlinearity mechanical system. Supervised manifold learning has the capacity of learning intrinsic geometric structure of complexity nonlinearity data. In many cases, the reality model of the samples is located in a low-dimensional manifold embedded in the high-dimensional sample space [6,7]. Suppose the length of a measured signal sample is D, we consider the signal sample as one data point in the signal space $\mathbb{R}^D$. If one fault class has m signal samples, the m signal samples can be treated as m points in $\mathbb{R}^D$. These data points usually have special distribution structure in the signal space. From the viewpoint of geometry, these data points belonging to the same fault class lie on or near a nonlinear manifold embedded in the high-dimensional signal space, and the different fault classes correspond to different manifolds [17]. So fault recognition can be considered as a multiple manifold learning problem. The supervised locally linear embedding algorithm, which considers both class label and local neighbor geometry information, can obtain the geometrical property of the multiple underlying manifolds from the fault data and has good feature extraction performance.

Thus we propose a new fault diagnosis approach based on supervised locally linear embedding projection (SLLEP) algorithm. The approach first learns the intrinsic geometric structure of fault data in the signal space to capture the nonlinear embedded manifold features, and map the high-dimensional fault data into a low-dimensional embedded space. The approach then applies local linear regression to find the projection that best approximates the implicit mapping from high-dimensional data to the embedding; so the embedded mapping of the new fault data can be obtained by the learned projection. Finally the fault classification is performed in the embedded space by comparing the distances between the embedded mapping of new fault data and the corresponding manifold.

In the SLLEP-based fault diagnosis approach, there are three parameters that need to set the reduced dimensionality d, the number of neighbors k and the distance parameter α. The selection of these parameters has an influence on the fault classification efficiency. If d is set too large, the mapping of SLLEP will contain much redundant information and enhance noise. If d is set too small, the data of different fault class might be mapped to overlapping each other. If k is set too large, the mapping will lose its nonlinear character and behave like traditional PCA. If k is too small, the mapping will not be able to maintain the data's topology in the low-dimensional space. The parameter α controls the amount to which class information should be incorporated. If α is set incorrectly, the mapping might not generalize well to the new data.

An effective way to estimate d can be described as follows [14]: first, perform an eigenanalysis on every local covariance matrix $\boldsymbol{Q}_i$ whose element is $\boldsymbol{Q}_i(m,n) = (\boldsymbol{x}_i - \boldsymbol{x}_m)^T(\boldsymbol{x}_i - \boldsymbol{x}_m)$. Then for each sample $\boldsymbol{x}_i$, a d_i can be found by specifying $\sum_{j=1}^{d_i} \lambda_j / \sum_{j=1}^{k} \lambda_j \ge 0.90$, where λ_j is the non-zero eigenvalue of $\boldsymbol{Q}_i$. Finally one can use majority voting over the entire data set to choose the largest d_i as the overall d. The parameters k and α can be determined by the ten-fold cross validation

scheme. First, select k with a step size of 1 from 2 to 40 and α with a step size of 0.05 from 0 to 1. Then for every pair of parameter (k,α) chosen, the ten-fold cross validation is run. That is, the original data set is randomly divided into ten equal-sized subsets. In each fold, one subset is used as testing set and the union of the remaining ones is used as training set. After ten folds, each subset has been used as testing set once. The average classification result of these ten folds is recorded. Finally choose (k,α) which provides the best classification result as the optimal parameter.

The SLLEP-based fault diagnosis approach is summarized below.

(A) Normalize each training vibration signal $\boldsymbol{x}_i=[x_i(1),x_i(2),\cdots,x_i(D)]$ to zero mean and unit variance. D is the length of training signal, and $i=1,2,\cdots,N$, where N is the number of training signals.
(B) Construct a high-dimensional signal space $\mathbb{R}^D$, where $\boldsymbol{x}_1, \boldsymbol{x}_2,\cdots,\boldsymbol{x}_N$ are treated as N data points.
(C) Choose the proper parameters such as d, k, and α. Perform the SLLE on the training data $\boldsymbol{x}_1, \boldsymbol{x}_2,\cdots,\boldsymbol{x}_N$ in high-dimensional signal space to learn the intrinsic multiple manifold features corresponding to the different fault class, and map them into a reduced d-dimensional feature space to obtain the mapped training data's features $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_N$.
(D) Train a classifier on $\boldsymbol{y}_1, \boldsymbol{y}_2,\cdots,\boldsymbol{y}_N$ in the reduced feature space to achieve fault classification. For a newly measured vibration signal $\boldsymbol{x}_{new}$, normalize it to zero mean and unit variance, and apply local linear projection as described in Section 2.3 to map it into the feature space, and then predict its fault class label using the trained classifier.

4. Experiments and analysis

4.1. Bearing fault data

The bearing vibration signals with different faults are obtained from the website of the bearing data center [18] of Case Western Reserve University. The bearing data have been validated in many researches [19,20] and has become a standard data set of the ball bearings. In this experiment, the bearing type is SKF 6205, a deep groove ball bearing. Single point fault is introduced to the test bearing inner race and outer race, respectively, using an electro-discharge machining with the fault diameter of 0.021 inches and the fault depth of 0.011 inches (1 inch=25.4 mm). Three data sets (inner race fault, outer race fault, and ball fault) are obtained from the experiment system with the sampling frequency of 12 kHz and the motor speed of 1730 rpm. Fig. 2 shows three typical vibration signals of the three corresponding fault classes.

From the data sets, up to 150 signal samples per class are selected randomly as a training set, and 100 signal samples per class are used for testing. Here "signal sample" means a section of signal consisting of a number of data points. This procedure is repeated ten times and gets ten classification results. Then the average of the ten results is recorded and shown. The length of the signal sample is set to 1030, which is equal to the number of sampling points within one period of the modulating frequency of ball faults. So each signal sample is transformed into a data point in a 1030-dimensional signal space.

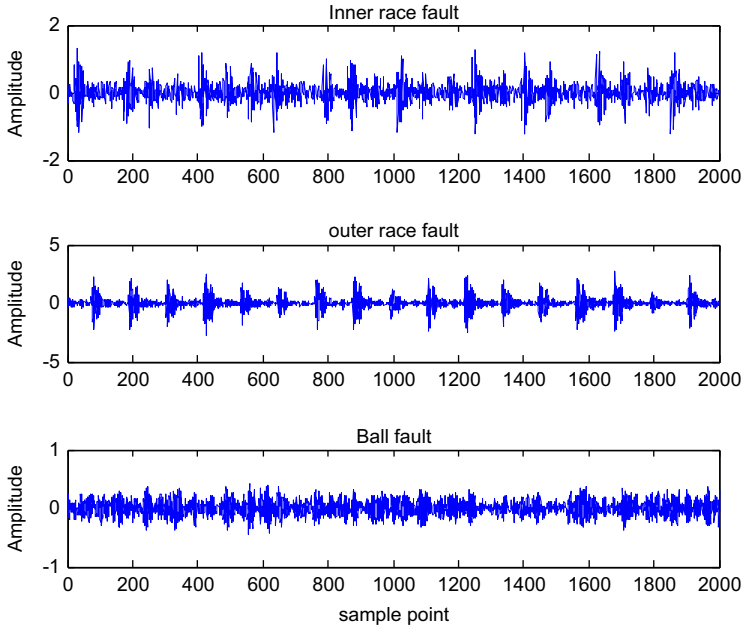

Fig. 2. Three typical vibration signals of the three corresponding fault classes (inner race fault, outer race fault, and ball fault).

In a first experiment, SLLEP is compared to LDA as a supervised feature extractor. The training set consists of 150 samples per class. For SLLEP, the parameter k is set to 12. Fig. 3(a)–(c) shows the mapping of training data onto two-dimensions using LDA and SLLEP for various values of α. For $\alpha=0.3$, SLLEP succeeds in mapping all points from the same class into one point in the reduced space. Mathematically, it means that the best solution for the minimization problem (2) is found, i.e., $\phi(\boldsymbol{Y})=0$. Fig. 3(d)–(f) shows the mapped test data and classification results with the simple minimum-distance classifier. The performance of minimum-distance classifier trained on SLLEP-mapped data is very good (the average classification accuracy reaches 99.33% when $\alpha=0.1$, whereas for LDA-mapped data the accuracy is 94.06%). This indicates that SLLEP can extract nonlinear manifolds embedded in the fault data in a supervised way and can achieve better classification performance than LDA. Moreover, the performance of LDA becomes very poor as the training sample size is decreased. The average classification accuracy of LDA sharply decreases to 65.33% and 31.66% when the training sample sizes are 140 and 100, respectively. It means LDA suffers from "curse of dimensionality" and get highly biased estimates because the dimension of the sample is too high and training sample is inadequate. However, the SLLEP still provides a good performance (the average classification accuracies are 99.16% and 98.33% as the training sample sizes are 140 and 100, respectively).

In a second experiment, SLLEP is compared to the unsupervised feature extraction technique PCA. The 1-nearest neighbor classifier is used for fault classification in the reduced space. The training sample size is 100 per class. Fig. 4 shows the performances on the test set as a function of reduced dimension d (for SLLEP, the optimal parameters k and α are set to 12 and 0.1, respectively). It can be seen that SLLEP performs much better than PCA when mapping down to a very small number of dimensions. This is because for small d, PCA will lose much information and introduce overlap in the mapping where there is none in the original data, whereas SLLEP performs well due to its nonlinear nature and supervised ability. SLLEP reaches its best classification accuracy 99.67% at the small dimension ($d=6$). However, PCA gives the best classification accuracy of 97.64% at high dimension ($d=40$). It indicates that comparing to PCA, SLLEP can provide a much lower reduced space where the time spent on training classifier can be significantly shortened.

Table 1 shows classification results with 1-nearest neighbor classifier as a function of training sample size (the optimal parameters k,d and α are set as described in Section 3). The performances of SLLEP and PCA consistently improve with more training samples. This further corroborates the benefit of dimensionality reduction for fault classification. SLLEP provides obviously the better performances than PCA with each size of training samples. Especially, as the training sample size is 20 per class, the classification accuracy of PCA (76.33%) is much lower than that of SLLEP (85.66%). This means SLLEP can still perform well in case of a small number of samples by choosing proper parameter that controls the nonlinearity of mapping.

In the third experiment, the performance of a number of classifiers trained on the SLLEP-mapped data is compared to that of these classifiers trained on the original data (150 per class). Three classifiers are used: the minimum-distance

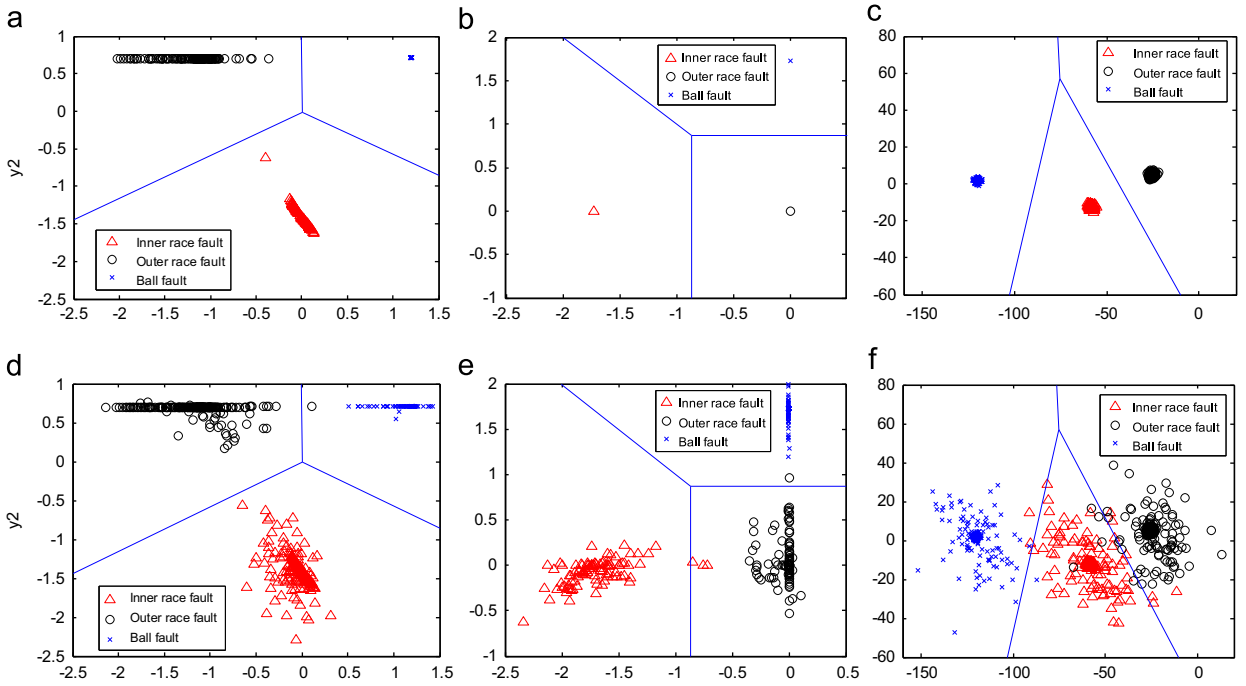

Fig. 3. (a) Mapping of training roller bearing data onto two-dimensions using SLLEP for $\alpha=0.1$, with a trained minimum-distance classifier (decision boundary is drawn). (b) Mapping of the same training data using SLLEP for $\alpha=0.3$, with trained minimum-distance classifier. (c) Mapping of the same data using LDA, with trained minimum-distance classifier. (d)–(f) Mapped test data and the average classification accuracy for the classifier ((d) average accuracy=99.33%, (e) average accuracy=98.67%, (f) average accuracy=94.06%).

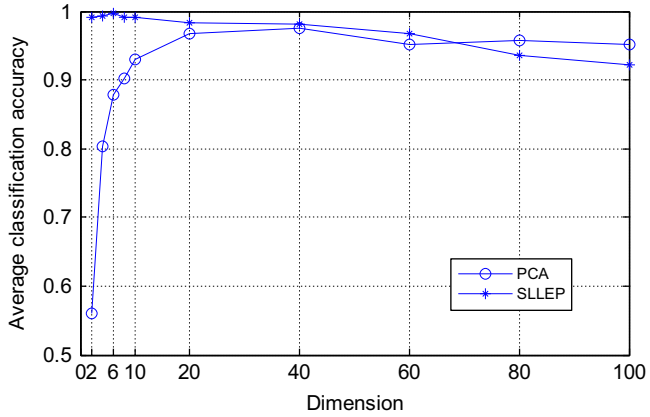

Fig. 4. Average classification accuracy of SLLEP and PCA on the test set with 1-nearest neighbor classifier as a function of reduced dimension d.

Table 1
Bearing fault data: average classification accuracy (%) with 1-nearest neighbor classifier as a function of training sample size per class.

Training sample size	20	40	80	100	150
PCA	76.33 ($d=40$)	86.83 ($d=40$)	96.83 ($d=40$)	97.40 ($d=40$)	97.64 ($d=40$)
SLLEP ($\alpha=0.1$)	85.66 ($k=6$, $d=4$)	93.86 ($k=8$, $d=4$)	97.67 ($k=12$, $d=6$)	99.67 ($k=12$, $d=6$)	99.67 ($k=12$, $d=6$)

Table 2
Bearing fault data: average classification accuracy (%) for a number of classifiers trained on the original data and SLLEP-mapped data.

Classifier	Minimum-distance	K-NN	SVM (RBF, σ)
Original data ($d=1030$)	28.66	96.64 ($K=4$)	97.33 ($\sigma=12$)
SLLEP-mapped data ($k=12$, $d=6$, $\alpha=0.1$)	99.42	99.67 ($K=1$)	99.67 ($\sigma=6$)

classifier, support vector machine (SVM) classifier with RBF kernel and K-nearest neighbors (K-NN) classifier with K optimized by the leave-one-out procedure. The experiment results are listed in Table 2. Clearly, SLLEP is suitable as a feature extraction step prior to classification. All the classifiers perform better on the SLLEP-mapped data than on the original data. Especially, the minimum-distance classifier works very poorly on the original data (the classification accuracy is 28.66%), whereas the SLLEP followed by the minimum-distance classifier performs remarkably well (the classification accuracy reaches 99.42%, which is even higher than that of the SVM classifier on the original data). This is because SLLEP maps the data nonlinearly such that this simple linear classifier can do well.

Furthermore, in order to analyze the influence of parameter k on the fault classification performance, the SLLEP is performed for various values of k. The optimal α is set to 0.1 and d is set as described in Section 3. The training sample size is 150 and the 1-nearest neighbor classifier is used. The experiment results are shown in Fig. 5. It can see that the classification performance is not ideal (the accuracy is 81.33%) as $k=2$. However, when k changes from 4 to 20, the performance becomes good (the accuracy is more than 96%) and the change in the fault classification accuracy is about 3%. Thus it is thought that if the parameter k is not set too small, there is no significant influence of k on the fault classification performance in the experiment.

4.2. Rotor bed fault data

Experiments have also been conducted on a rotor fault test-bed shown in Fig. 6. It consists of AC motor, rotor, bearing, shaft, and coupling. An accelerometer is attached to the bearing housing to collect the vibration signals. Three types of fault conditions are simulated as follows: rotor rub-impact, bearing oil-film whirl and rotor misalignment fault. The sampling frequency is 10 kHz and the motor speed is 2400 rpm.

From the data sets, up to 120 signal samples per class are selected randomly as a training set, and 60 signal samples per class are used for testing. The length of signal sample is set to 250, which is equal to the number of sampling points within one period of the shaft rotating frequency. So each signal sample is transformed into a data point in a 250-dimensional signal space.

We carry out feature extraction to the signal samples set with the three methods, PCA, LDA, and SLLEP, and then capture its embedded manifold to implement fault classification. The minimum-distance classifier is used for classification

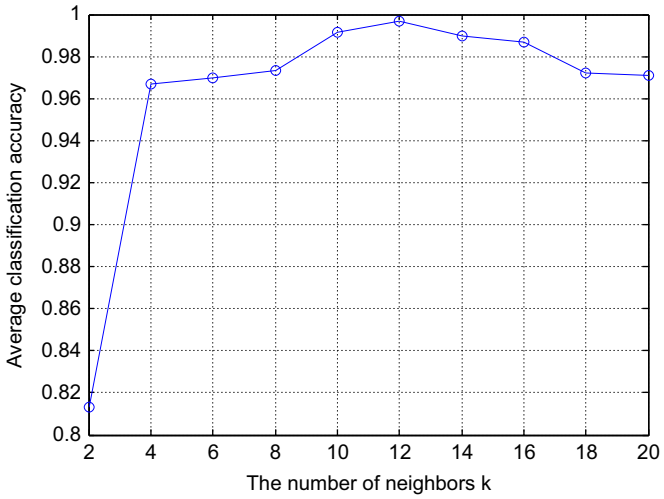

Fig. 5. Average classification accuracy as a function of parameter k.

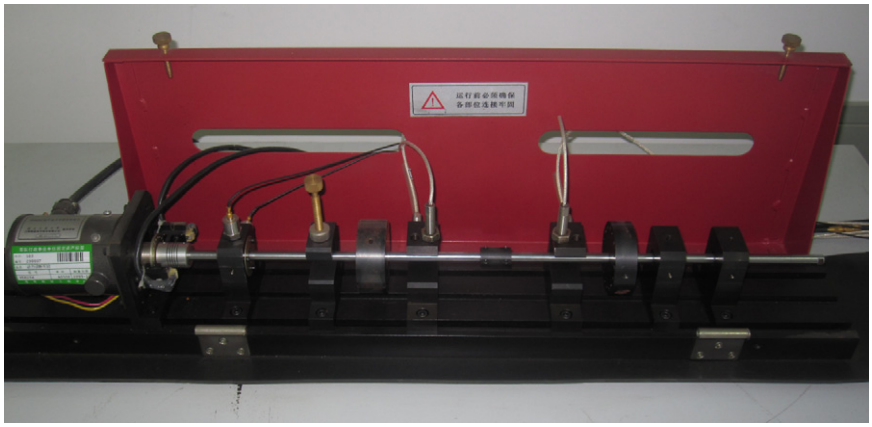

Fig. 6. Rotor test-bed.

Table 3
Rotor bed fault data: average classification accuracy (%) with minimum-distance classifier as a function of training sample size per class.

Training sample size	40	60	80	100	120
PCA	84.56 (d=28)	88.22 (d=28)	90.45 (d=32)	93.12 (d=32)	93.83 (d=36)
LDA	81.24 (d=2)	86.32 (d=2)	91.72 (d=2)	92.46 (d=2)	92.78 (d=2)
SLLEP (α=0.2)	90.57 (k=6, d=3)	91.66 (k=8, d=3)	93.88 (k=10, d=4)	95.00 (k=10, d=4)	96.11 (k=10, d=4)

in the reduced space. Table 3 shows classification results of the three method with different training sample size (for SLLEP, the optimal parameters k, d and α are set as described in Section 3). Fig. 7(a) and (b) shows the mapping of training data onto two-dimensions using SLLEP and LDA when the training sample size is 120 per class. Fig. 7(c) and (d) shows the mapped test data and classification results. It can be seen that SLLEP obviously performs better than PCA and LDA with each size of training samples. Especially, when the training sample size is small (40 per class), the classification accuracies of PCA (84.56%) and LDA (81.24%) are sharply decreased, whereas that of SLLEP is still high (90.57%). Consequently, SLLEP shows better classification performance than PCA and LDA, especially in case of small sample size.

The performance of minimum-distance SVM and K-NN classifiers trained on the SLLEP-mapped data is also compared to that of these classifiers trained on the original data (120 per class). From Table 4, it can be seen that all the classifiers perform better on the SLLEP-mapped data than on the original data. Especially, the classification accuracy (96.11%) of the SLLEP followed by the minimum-distance classifier is much higher than the pure minimum-distance classifier on the original data (70.32%). It demonstrates that SLLEP is a suitable feature extraction method, which when coupled with simple classifiers can yield very promising fault classification results. Moreover, the SVM training on SLLEP-mapped data is

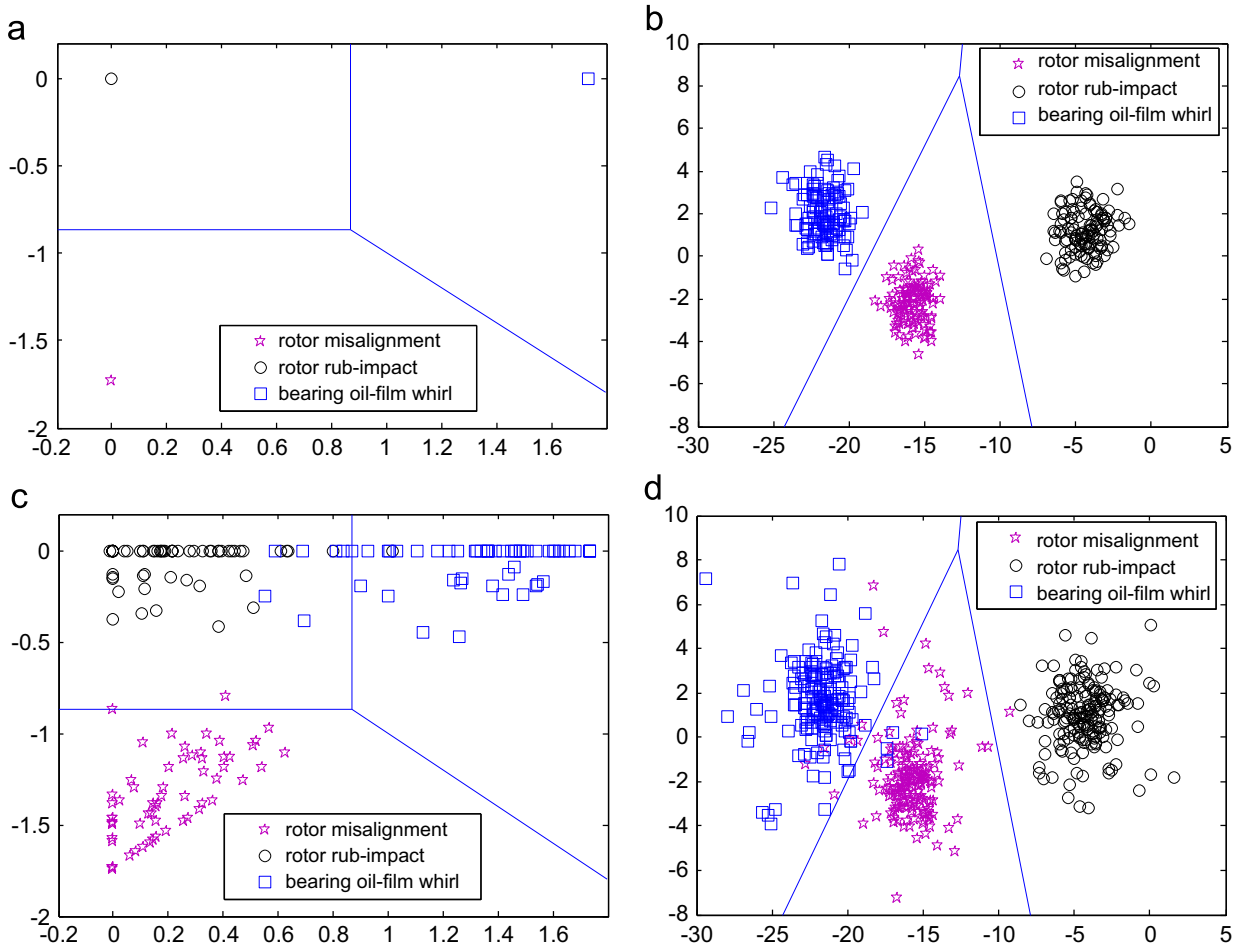

Fig. 7. (a) Mapping of training rotor bed data onto two-dimensions using SLLEP, with a trained minimum-distance classifier ($\alpha=0.2$, $k=10$, $d=2$, and the training sample size is 120 per class). (b) Mapping of the same data using LDA, with trained minimum-distance classifier. (c)–(d) Mapped test data and the average classification accuracy for the classifier ((c) average accuracy=95.56%, (d) average accuracy=92.78%).

Table 4
Rotor bed fault data: average classification accuracy (%) for a number of classifiers trained on the original data and SLLEP-mapped data.

Classifier	Minimum-distance	K-NN	SVM (RBF, σ)
Original data ($d=250$)	70.32	92.68 ($K=6$)	93.32 ($\sigma=30$)
LLEP-mapped data ($k=10$, $d=4$, $\alpha=0.2$)	96.11	96.11 ($K=2$)	96.84 ($\sigma=12$)

very fast since the dimensionality of embedded space is much lower than that of the original data space, and immediately gives us support vectors, 18 support vectors are found, whereas the number is 124 when SVM is trained on the original data.

5. Conclusions

In this paper, a new machinery fault diagnosis approach based on supervised locally linear embedding projection (SLLEP) is proposed. The approach first performs the recently proposed manifold learning algorithm supervised locally linear embedding (SLLE) on the high-dimensional fault signal samples to learn the nonlinear embedded manifold features, and map them into a low-dimensional embedded space to achieve fault feature extraction. For dealing with the new fault sample, the approach then applies local linear regression to find the projection that best approximates the implicit mapping from high-dimensional samples to their embedding. Finally fault classification is carried out in the embedded space. The ball bearing data and rotor bed data both are used to validate the proposed approach. The results show that the proposed approach obviously improves the fault classification performance and outperforms the other traditional approaches such as LDA and PCA.

The approach still needs a number of parameters to be set, which influence the fault classification performance. Thus how to choose the number of neighbors k and the distance parameter α in a more well-founded way is worthy of further study. It would also be interesting to compare the proposed approach with other combinations of nonlinear mapping methods and classifiers on more machinery fault data sets.

Acknowledgment

This work is supported by military scientific research foundation. Thanks are also due to the Case Western Reserve University, Rockwell Science Office of Naval Research and CVX for the bearing fault data sets used in this research. Finally, the authors appreciate very much the valuable comments from the anonymous reviewers who helped to improve this paper.

References

[1] I.T. Jolliffe, Principal Component Analysis, Springer, 1986.
[2] K. Fukunaga, Introduction to Statistical Pattern Recognition, Academic Press, 1990.
[3] A.M. Martinez, A.C. Kak, PCA versus LDA, IEEE Trans. Pattern Anal. Mach. Intell. 23 (2) (2001) 228–233.
[4] M.A. Kramer, Nonlinear principal component analysis using autoassociative neural networks, AIChE J. 37 (2) (1991) 233–243.
[5] D. Dong, T.J. McAvoy, Nonlinear principal component analysis based on principal curves and neural networks, Comput. Chem. Eng. 20 (1) (1996) 65–78.
[6] H.S. Seung, D.L. Daniel, The manifold ways of perception, Science 290 (5500) (2000) 2268–2269.
[7] S.T. Roweis, L.K. Saul, Nonlinear dimensionality reduction by locally linear embedding, Science 290 (2000) 2323–2326.
[8] J.B. Tenenbaum, V.D. Silva, J.C. Langford, A global geometric framework for nonlinear dimensionality reduction, Science 290 (2000) 2319–2323.
[9] K. Samuel, D.L. Martin, Face detection in gray scale images using locally linear embeddings, Comput. Vision Image Understanding (2006) 1–20.
[10] A. Hadid, O. Kouropteva, M. Pietikainen, Unsupervised learning using locally linear embedding: experiments with face pose analysis, in: Proceedings of the 16th International Conference on Pattern Recognition, 2002, pp. 111–114.
[11] D.D. Ridder, O. Kouropteva, O. Okun, Supervised locally linear embedding, in: Proceedings of the Joint International Conference, ICANN/ICONIP 2003. Lecture Notes in Computer Science, Springer, Heidelberg, 2003, pp. 333–341.
[12] L.K. Saul, S.T. Roweis, Think globally, fit locally: unsupervised learning of low dimensional manifolds, J. Mach. Learn. Res. 4 (2) (2004) 119–155.
[13] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, UK, 1990.
[14] D.D. Ridder, R.P.W. Duin, Locally linear embedding for classification, Technical Report PH-2002-01, Delft University of Technology, Delft, The Netherlands, 2002.
[15] O. Kayo, Locally linear embedding algorithm extensions and applications, Ph.D. Dissertation, Faculty of Technology, University of Oulu, 2006.
[16] A.N. Tikhonov, Solutions of Ill-Posed Problems, Wiley, New York, 1997.
[17] Quansheng Jiang, Minping Jia, Jianzhong Hua, Feiyun Xu, Machinery fault diagnosis using supervised manifold learning, Mech. Syst. Signal Process. 23 (2009) 2301–2311.
[18] K.A. Loparo, Bearings vibration data set, Case Western Reserve University, 〈http://www.eecs.case.edu/ laboratory/beaing/〉.
[19] L. Zhang, J. Xu, J. Yang, et al., Multiscale morphology analysis and its application to fault diagnosis, Mech. Syst. Signal Process. 22 (2008) 597–610.
[20] H. Yang, J. Mathew, L. Ma, Fault diagnosis of rolling element bearings using basic pursuit, Mech. Syst. Signal Process. 19 (2005) 341–356.