

# Nonlinear machine fault detection by semi-supervised Laplacian Eigenmaps†

Quansheng Jiang\* , Qixin Zhu, Bangfu Wang and Lihua Guo

*School of Mechanical Engineering, Suzhou University of Science and Technology, Suzhou 215009, China* 

(Manuscript Received May 7, 2016; Revised March 20, 2017; Accepted April 26, 2017) --

# **Abstract**

A semi-supervised Laplacian Eigenmaps algorithm for machine fault detection is proposed. The purpose of the algorithm is to efficiently extract the manifold geometric characteristics of nonlinear vibration signal samples, and to determine fault classification of operating equipment so that the accuracy of fault detection can be improved. The data acquisition and pre-processing of the vibration signal is firstly implemented from monitoring equipment, then hybrid domain feature is obtained, and the initial sample set can be built. This is followed by implementing the semi-supervised Laplacian Eigenmaps algorithm so that the sensitive nature characteristics of manifold can be obtained from the device initial sample set. In order to establish the intelligent diagnostic model, the Least square Support vector machine (LS-SVM) is then adopted, which fault diagnosis and decisions can be achieved in the feature space of the low-dimensional manifold. The experiment results of using the IRIS data, gearbox and compressor fault data show the proposed method has more advantage when compared with the PCA and Laplacian Eigenmaps on improving the accuracy of fault detection.

*--*

*Keywords*: Semi-supervised Laplacian Eigenmaps; Fault detection; Feature extraction; Manifold learning

#### **1. Introduction**

Fault detection is a key technology to protect the safe operation of mechanical and electrical equipment, which is one of the focus areas of mechanical fault diagnosis [1]. With the increasing complexity of the structure and function of modern equipment, together with the strong non-stationary and nonlinear characteristics of equipment running state, the diagnosis information obtained is getting immense and complex, accordingly making fault detection and diagnosis process more difficult and complicated. The conventional linear feature extraction methods are such as Principal component analysis (PCA) [2], and Independent component analysis (ICA) [3]. They are not only under the assumption of global linear structure of the data, but also find the best low-dimensional projection by using different linear transformation matrix. Generally, the linear methods are difficult to obtain the classification information from the non-linear conditions.

In the need of non-linear data analysis, some non-linear feature extraction approaches are proposed. The kernel-trick has been successfully utilized to extend linear feature extraction methods for nonlinear cases, such as performing a linear method in a higher dimensional kernel feature space by a kernel function. Kernel PCA (KPCA) is the kernel extension of PCA [4]. KPCA provides superior performances for nonlinear fault classification than its linear form (PCA). However, it possesses some drawbacks that useful discriminate information is lost during the feature extraction process, and is difficult to construct the kernel function suitable for fault detection.

Wavelet transform (WT) and Empirical mode decomposition (EMD) can also be used for machine fault detection. Chen et al. [5] presented a method of wavelet transform based on inner product for rotating machinery fault diagnosis, and summarized the corresponding WT applications for offering an in-depth and comprehensive reference. Lei et al. [6] summarized applications of empirical mode decomposition in fault detection of rotating machinery, and discussed the outstanding open problems of EMD in fault diagnosis. Deng et al. [7] proposed a fault detection method for rotor-bearing system based on the Local mean decomposition (LMD) and Teager energy kurtosis (TEK). Domaneschi [8] employed finite element modeling to the material behavior dynamic simulation and detection, which achieved good identification effectiveness.

Manifold learning is a new kind of nonlinear feature extraction method. In recent years, manifold learning methods have been proposed such as isometric mapping (ISOMAP) [9], Locally linear embedding (LLE) [10], Laplacian Eigenmaps [11], Local tangent space alignment (LTSA) [12], etc. These methods are implemented to discover the intrinsic geometric structure and regularity of nonlinear high-dimensional data as the goal, namely from the observation of the phenomenon to find its essence, revealing the embedded low-dimensional

<sup>\*</sup>Corresponding author. Tel.: +86 512 68320622, Fax.: +86 512 68320622

E-mail address: qschiang@163.com

<sup>†</sup> Recommended by Associate Editor Byeng Dong Youn

<sup>©</sup> KSME & Springer 2017

smooth manifold in the high-dimensional data space.

However, these manifold learning methods do not consider the class information of the samples. Therefore, some super vised manifold learning methods have been proposed to take advantage of class label information in feature extraction. Ridder [13] presented a supervised LLE approach to solve classification problem. The method is applied as a feature extractor on a number of benchmark data sets, and is shown to be useful for high-dimensional data with a clear manifold structure. Zhang [14] proposed a supervised expansion of LTSA (named SLTSA) to deal with the high-dimensional nonlinear fault data and improve the performance of classification. Su [15] presented a fault diagnosis method using su pervised extended LTSA for dimension reduction, and applied effectively to diagnose the faults in a gearbox. Jiang [16] proposed a supervised manifold learning method for feature extraction and machinery fault diagnosis, which achieved satisfactory performance on gearbox fault diagnosis. These super vised approaches have achieved a certain effect in classification task.

More often, there are only a small labeled samples and a lot of unlabeled samples in practical applications, and it is therefore important to make more effective use of these labeled data. Semi-supervised learning mode is a good way. The goal of fault detection and classification is to realistically preserve the intrinsic structure in the samples and different classes could be clearly separated in the feature project space [17-20].

fault detection, a new fault detection approach based on semi supervised Laplacian Eigenmaps is presented, in which partial label information is used in the fault samples. The approach can efficiently extract the nonlinear manifold geometric char acteristics of signal sample, then determine fault class and improve subsequently the accuracy of the fault detection.

The rest of the paper is organized as follows. In Sec. 2, the semi-supervised Laplacian Eigenmaps algorithm is briefly described. In Sec. 3, the fault detection method by using semi supervised Laplacian Eigenmaps is presented. Application experiments and analysis of the proposed method are implemented in Sec. 4. Finally, the conclusions are drawn in Sec. 5.

## **2. Semi-supervised manifold learning algorithm**

## *2.1 Laplacian Eigenmaps*

The Laplacian Eigenmaps algorithm uses a graph embedding approach. The main principle of the algorithm is that, when the adjacent points in the high-dimensional space are projected to a low-dimensional space, the projection of image is adjacent.

The algorithmic procedure of Laplacian Eigenmaps can be summarized as follows.

Step 1. Constructing the neighbor graph. Computing the neighbor of each point via *K-*nearest neighbor (KNN) by the given neighbor parameter.

Step 2. Computing the Weighted matrix. To use heat kernel



Fig. 1. Extraction process diagram of the SSLE algorithm.

for weighting the edges, the Weighted matrix is computed.

Step 3. Eigenmaps. Computing eigenvalues and eigenvectors for the generalized eigenvector problem:

$$
Lf = \lambda Df \tag{1}
$$

In this paper, aiming at the difficulty of machine nonlinear where  $D$  is diagonal weighted matrix, its entries are column (or row, since *W* is symmetric) sums of *W*,  $D_{ii} = \sum_i W_{ii}$ .  $L = D$ . *W*, where *L* is the Laplacian matrix. *lap*  $\int_{\alpha}^{C} = \lambda Df$  (1)<br> *e D* is diagonal weighted matrix, its entries are column<br>
bw, since *W* is symmetric) sums of *W*,  $D_{ii} = \sum_{j} W_{ji}$ .  $L = D -$ <br>
here *L* is the Laplacian matrix.<br>
inimizing the objective function

Minimizing the objective function as follows:

$$
\Phi_{\text{lap}}(Y) = \sum_{i=1}^{N} \sum_{j \neq i} \frac{1}{2} \|y_i - y_j\|^2 W_{ij} = \text{trace}(Y^T L Y) \,. \tag{2}
$$

**Extraction low-dimensional**<br>
manifold features<br>
1. Extraction process diagram of the SSLE algorithm.<br>
weighting the edges, the Weighted matrix is computed.<br>
Step 3. Eigenmaps. Computing eigenvalues and eigenvec-<br>
sfor th *i* manifold features<br> *is* xtraction process diagram of the SSLE algorithm.<br> **ighting the edges, the Weighted matrix is computed.**<br> **3. Eigenmaps. Computing eigenvalues and eigenvective generalized eigenvector proble** The optimal *d*-dimensional embedding vectors *Y* are found by calculating the bottom  $d+1$  eigenvectors (corresponding to its smallest *d* + 1 eigenvalues) of the Laplacian operator *L*.

# *2.2 The Semi-supervisd Laplacian Eigenmaps (SSLE)*

The Laplacian Eigenmaps is an unsupervised learning method, and not taking the class information of dealt data into account.

The Semi-supervised Laplacian Eigenmaps algorithm (SSLE) combined the semi-supervised learning and manifold learning, with small quantities of correspondence labeled samples and plenty of unlabeled samples, to provide useful information to improve the ability of learning algorithms.

Fig. 1 is the extraction process diagram of the SSLE algorithm. The proposed semi-supervised Laplacian Eigenmaps algo-

rithm, through the following process:

(1) Input the initial sample set X, the parameters value of the low-dimensional manifolds dimension d as well as

neighborhood parameter k. Where X includes m number of label samples and u number of unlabeled samples (the total number of sample  $n = m+u$ ).

(2) For each sample point, using *k* neighbors to construct corresponding neighborhood graph and calculate the similarity matrix *S*:

$$
S_{ij} = \begin{cases} \exp(-\left\|x_i - x_j\right\|^2 / t), & x_i \in G_k(x_j) \text{ or } x_j \in G_k(x_i) \\ 0, & else \end{cases}
$$
(3)

In the Eq. (3), we use heat kernel for weighting the edges of the neighbor graph  $G(x_i)$ , where *t* denotes the width of heat kernel. <sup>21</sup> Eq. (3), we use heat kernel for weighting the edges of<br>the graph  $G(x_i)$ , where *t* denotes the width of here<br>is the use of label information to optimize the samp<br>illarity matrix, *S* is divided into:<br> $\begin{pmatrix} S_{12} \\ S_{2$ 

(3) For the use of label information to optimize the sample label similarity matrix, *S* is divided into:

$$
S = \begin{pmatrix} C & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \tag{4}
$$

where *C* is the size of  $m \times m$  matrix, and is computed by using

bel similarity matrix, *S* is divided into:

\n
$$
S = \begin{pmatrix} C & S_{12} \\ S_{21} & S_{22} \end{pmatrix}
$$
\n(4)

\nFig. 2. The pre

\nC is the size of *m* × *m* matrix, and is computed by using  
\nrel information of the labeled samples as  $\{c_1, \ldots, c_m\}$ .

\n(5)

\n
$$
C_{ji} = \begin{cases} \sqrt{m/c_j} - \sqrt{c_j/m}, & \text{if } c_j = c_i \\ -\sqrt{c_j/m}, & \text{else.} \end{cases}
$$
\n(5)

\n(6)

\nStep 1: By reduction, and the result is  $c_j$  for  $j = 1$  and  $c_j$  for  $j = 1$ .

\n(7)

\n
$$
S = \begin{cases} \sqrt{m/c_j} - \sqrt{c_j/m}, & \text{else.} \end{cases}
$$
\n(8)

\n
$$
S = \begin{cases} \sqrt{m/c_j} - \sqrt{c_j/m}, & \text{else.} \end{cases}
$$
\n(9)

\n
$$
S = \begin{cases} \sqrt{m/c_j} - \sqrt{c_j/m}, & \text{else.} \end{cases}
$$
\n(10)

\n
$$
S = \begin{cases} \sqrt{m/c_j} - \sqrt{c_j/m}, & \text{else.} \end{cases}
$$
\n(2)

\n
$$
S = \begin{cases} \sqrt{m/c_j} - \sqrt{c_j/m}, & \text{else.} \end{cases}
$$
\n(3)

\n
$$
S = \begin{cases} \sqrt{m/c_j} - \sqrt{c_j/m}, & \text{else.} \end{cases}
$$
\n(4)

\n
$$
S = \begin{cases} \sqrt{m/c_j} - \sqrt{c_j/m}, & \text{else.} \end{cases}
$$
\n(5)

\n
$$
S = \begin{cases} \sqrt{m/c_j} - \sqrt{c_j/m}, & \text{else.} \end{cases}
$$
\n(6)

\n
$$
S = \begin{cases} \sqrt{m/c_j} - \sqrt{c_j/m}, & \text{else.} \end{cases}
$$
\n(7)

\n<math display="block</p>

(4) Solving the generalized eigenvalue problem:

$$
LY = \lambda DY \tag{6}
$$

sponding eigenvectors as sample manifold features in low-

$$
Y = [U_1, U_2, ..., U_d]^T
$$
 (7)

The proposed algorithm uses a large amount of unlabeled sample data inherent in studying the geometry of the structure, and learns the label information from a small number of sam ples, to obtain the feature manifold of the whole data.

Using the semi-supervised Laplacian Eigenmaps algorithm, we can fulfill feature extraction of the equipment samples manifolds, and obtain the failure sensitive nature of low dimensional manifold characteristics.

## **3. The nonlinear fault detection approach by SSLE**

A nonlinear fault detection method based on the semi supervised Laplacian Eigenmaps algorithm is proposed in the paper. The process diagram of the method is as shown in Fig. 2. The approach can efficiently extract the manifold geometric characteristics of nonlinear vibration signal samples, and determine fault type of equipment for improving the accu-



Fig. 2. The process diagram of the fault detection method.

racy of fault detection.

The implementation steps of the proposed method are listed as following:

*LY* By the eigenbor graph  $G(x)$ , where *l* denotes the width of heat<br>
Lement.<br>
Lement,<br>
Lementary matrix, *S* is divided into:<br>
Lementary matrix, *S* is divided into:<br>  $S = \begin{pmatrix} S_{S_2} \\ S_{S_3} \end{pmatrix}$  (4) <br>  $\begin{pmatrix} \text{Dagness and decisions}$ Where *C* is the size of  $m \times m$  matrix, and is computed by using<br>
label information of the labeled samples as { $c_1,..., c_m$ }.<br>  $\left\{\sqrt{m/c_1} - \sqrt{c_1/m}$ , if  $c_j = c_j$ <br>  $-\sqrt{c_j/m}$ , if  $c_j = c_j$ <br>  $\sqrt{m/m}$ ,  $\cos \theta$ <br>  $\cos \theta$  as following: Step 1: By implementing data acquisition and wavelet noise reduction pre-processing to the vibration signal of monitoring mechanical and electrical equipment, then the hybrid domain feature can be obtained. Here the hybrid domain feature in cludes the time domain and frequency domain information features of the samples, such as RMS, kurtosis, absolute mean, amplitude spectrum, etc. Thus the dealt samples are acquired and the initial feature space can be constituted.

Step 2: To obtain the sensitive nature of manifold characteristics through using the proposed semi-supervised Laplacian Eigenmaps algorithm, the manifolds feature extraction of the equipment samples can be implemented;

Step 3: To build the intelligent diagnostic model, the Least square Support vector machine (LS-SVM) classifier is adopted, and then fault diagnosis and decisions can be achieved in the feature space of the low-dimensional manifold.

In contrast to the traditional detection methods, the new approach can treat high-dimensional nonlinear data and distin guish failure types from the obtained manifold feature, which avoiding "Curse of dimensionality". Furthermore, the fault detection process is quiet brief and the execution speed is fast.

# **4. Experiments and results**

To verify the validity of the proposed method, we utilize three datasets to implement the experiments as follows: IRIS, gearbox data and compressor failure data.

#### *4.1 IRIS data*

IRIS data is a benchmark database in the pattern recognition

Dataset	<b>Numbers</b>			Dimen-	Low-	Neighb
	Training	Testing	Classes (C)	sions (D)	dimen- sion (d)	<b>ors</b> (k)
<b>IRIS</b> Data	120	30	3		$\overline{c}$	20
Gearbox data	48	12	3		$\overline{c}$	
Compressor data	168	42	3	512	$\overline{c}$	9

Table 1. The data structure of the experiment and algorithm parameters.

from UCI machine learning repository [21]. The dataset contains 150 instances of iris flowers collected in Hawaii. The flowers include 3 classes of Iris Setosa, Iris Versicolour and Iris Virginica, based on 4 measures of sepal's width and length, 2 and petal's width and length. The two types in IRIS are nonlinear and difficult to separable from each other.

In the experiment, IRIS is randomly divided into training set and testing set, the label number of each dataset is 20 %.<br>The data structure and algorithm parameters in the experiment are shown in Table 1. s<br>f. sock detecet is 20

In the six parameters of Table 1, the first four parameters are the intrinsic properties of proceeding datasets and no selection required. Among them, the first two parameters are the number of training and testing samples, the third parameter is 8 the number of sample categories, and the fourth parameter is the dimensions of samples. The last two parameters (Low- 6 dimension and neighbors) are the essential parameter of the proposed algorithm and need to choose. The value of low- 4 dimension is not well estimated at present, we generally choose it based on experience with the dataset classes C subtract 1. The neighbors in the manuscript have been optimized by the leave-one-out procedure on the training set.

In order to compare the performance of the proposed method, we also use PCA and Laplacian Eigenmaps to execute feature extraction. The two-dimensional feature distribution of IRIS data using the three methods are indicated in Fig. 3.

After implementing feature extraction on IRIS Data with the three methods, we adopt the Least square Support vector machine (LS-SVM) classifier for final recognition. The classification accuracy rates of the three algorithms are indicated in 8 Table 2.

From Figs.  $3(a)$  and (b) we can see, there are two classes points overlapped. This means that the effects of pattern clas-4 sification are not satisfactory when using PCA and Laplacian Eigenmaps. In Table 2, their classification accuracy rates with 2 LS-SVM classifier are 86.67 % and 90.4 %, respectively. Compared with them, the effect of pattern classification with 5 0 the SSLE is satisfactory, because each class of the testing data 4 can be classified obviously from Fig. 3(c). The classification accuracy rate of the SSLE with LS-SVM classifier is 93.33 %. 4 2 Therefore, the proposed SSLE algorithm has shown better identification of performance than PCA and Laplacian Eigenmaps when executing feature extraction to IRIS data.

Table 2. The classification accuracy rates of the methods to the IRIS using LS-SVM classifier.

Datasets	<b>PCA</b>		SSLE	
IRIS data	86.67%	90.4%	93.33 %	



Fig. 3. Feature extraction to IRIS data.

# *4.2 Gearbox data*

In the machinery fault detection, the gearbox fault diagnosis is quite representative. The gearbox samples in this experiment are collected from an on-line monitoring system of run ning gearbox in a transmission test bed [22]. The testing object is a group of low-carbon steel helical gear with gear ratio  $Z1 / Z2 = 41/37$ , and the modulus is 5 mm. The sampling frequency is 10 kHz. The sample consists of three states, such as gear surface pitting, tooth wear and normal gear. The number of the gearbox data is 60, and the feature dimensions include 1 the standard deviation, kurtosis, RMS, absolute mean, peak factor, pulse factor and margin coefficient.

The two-dimensional feature distribution of the gearbox data with PCA, Laplacian Eigenmaps and SSLE are indicated in Fig. 4.

For final fault pattern recognition, we use the LS-SVM classifier to the two-dimensional manifold feature. The classification accuracy rates of the three algorithms are indicated in Table 3.

As shown in Fig.  $4(a)$ , there are two classes of overlapping (Classes C1 and C2), so the classification effect of PCA to gearbox data is not satisfied. There are some points of differ ent classes overlapped in Fig. 4(b), the classification perform ance of Laplacian Eigenmaps is also unsatisfactory. In Table 3, the classification accuracy rates of PCA and Laplacian Eigenmaps are respectively 91.67 % and 93.33 %. In compari-1 son, the classification effect of the SSLE to gearbox data is better than PCA and Laplacian Eigenmaps from Fig. 4(c) and Table 3. The classification accuracy rate of SSLE with LS- 0 SVM classifier is 96.67 %.

#### *4.3 Compressor fault data*

The compressor fault sample is collected from an operating 1 high-speed air compressor unit in a power plant. The compressor unit has three shafts, the motor drives the middle gear shaft, the front high-speed shaft and the rear high-speed shaft through the coupling. Both ends of the high-speed shafts are equipped with cantilever impeller, the two high-speed shafts are flexible shaft, speed of 15240 r / min and 23400 r / min. There are 210 points and 512 dimensions in the compressor data by the amplitudes after FFT transform to the collected 1 vibration signal. The fault data includes three failure types as follows: mechanical looseness (Class C1), oil film whirl (Class C2), and rotor imbalance fault (Class C3). 0

The SSLE, PCA and Laplacian Eigenmaps are also executed to the compressor data. The algorithm parameters in the experiment are listed in Table 1.

The two-dimensional feature distribution of the compressor data using the three methods are shown in Fig. 5.

We use the LS-SVM classifier to the compressor data after feature extraction for final recognition. The classification accuracy rates of the three algorithms are indicated in Table 4. -1

As shown in Fig. 5(a), there are three classes of overlapping (Class C1, C2 and C3), the classification effect of PCA to

Table 3. The classification accuracy rates of the methods to the gear data using LS-SVM classifier.

Datasets	<b>PCA</b>	Laplacian Eigenmaps	SSLE	
Gearbox data	91.67%	93.33 %	96.67%	



Fig. 4. Feature extraction to gearbox data.

(c) Using SSLE

Table 4. The classification accuracy rates of the methods to the com press data using LS-SVM classifier.



Fig. 5. Feature extraction to compressor data.

compressor data is poor. From Fig. 5(b) and Table 4, Lapla cian Eigenmaps has better classification effect than PCA. To compressor data, the classification accuracy rates of the two methods with LS-SVM classifier are respectively 93.33 % and 59.92 %. In comparison, the classification effect of the SSLE is more prominent than PCA and Laplacian Eigenmaps from Fig. 5(c) and Table 4. Its classification accuracy rate with LS- SVM classifier is 97.32 %.

From the experimental results above, we can draw a con clusion that, the proposed SSLE has better classification performance than Laplacian Eigenmaps and PCA. Therefore, the SSLE can efficiently extract the nonlinear manifold geometric characteristics of signal sample, determine fault class and improve the accuracy of fault detection. The new approach is proved to be an effective method for machine fault detection.

## **5. Conclusions**

This paper proposes a novel method of nonlinear fault detection based on Semi-supervised Laplacian Eigenmaps algorithm (SSLE). The approach takes full advantage of labeled class as well as unlabeled class information, and preserves the whole intrinsic geometry amongst the samples in the feature extraction process. The Least square Support vector machine (LSSVM) classifier is then adopted to establish the intelligent diagnostic model. As shown by testing the IRIS data, gearbox data and compressor fault data, the presented method has satisfied classification performance when treating the task of fault detection and classification. The experimental results show that the presented method is more effective than PCA and Laplacian Eigenmaps on improving the performance of fault detection.

In the future work, we plan to investigate a parameter optimization method of the algorithm of SSLE, and then find the optimal neighborhood parameter and the low-dimensional manifolds dimension to further improve the efficiency of the proposed algorithm.

#### **Acknowledgment**

This work is supported by National Natural Science Foun dation of China (No.51005025, 51375323), Natural Science Foundation of Jiangsu Province (No. BK20151199), and Scientific Fund of Suzhou University of Science and Technology (No. XKZ201408).

# **References**

- [1] G. F. Wang, X. L. Feng and C. Liu, Bearing fault classification based on conditional random field, *Shock and Vibration*, 20 (2013) 591-600.
- [2] W. Sun, J. Chen and J. Li, Decision tree and PCA-based fault diagnosis of rotating machinery, *Mechanical Systems and Signal Processing*, 21 (3) (2007) 1300-1317.
- [3] A. Widodo, B. S. Yang and T. Han, Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors, *Expert Sys-*

*tems with Applications*, 32 (2) (2007) 299-312.

- [4] K. Shi, S. Liu and H. Zhang, Kernel local linear discriminate method for dimensionality reduction and its application in machinery fault diagnosis, *Shock & Vibration*, 20 (2014) 1- 11.
- [5] J. Chen et al., Wavelet transform based on inner product in fault diagnosis of rotating machinery: A review, *Mechanical Systems and Signal Processing*, 38 (70-71) (2016) 1-35.
- [6] Y. G. Lei et al., A review on empirical mode decomposition in fault diagnosis of rotating machinery, *Mechanical Systems and Signal Processing*, 35 (1-2) (2013) 108-126.
- [7] L. Deng and R. Zhao, Fault feature extraction of a rotor system based on local mean decomposition and Teager energy kurtosis, *Journal of Mechanical Science and Technology*, 28 (4) (2014) 1161-1169.
- [8] M. Domaneschi, Experimental and numerical study of standard impact tests on polypropylene pipes with brittle behaviour, *Proc. IMechE Part B: Journal of Engineering Manufacture*, 226 (2012) 2035-2046.
- [9] J. Tenenbaum, D. D. Silva and J. Langford, A global geometric framework for nonlinear dimensionality reduction, *Science*, 290 (5500) (2000) 2319-2323.
- [10] S. Roweis and L. Saul, Nonlinear dimensionality reduction by locally linear embedding, *Science*, 290 (5500) (2000) 2323-2326.
- [11] M. Belkin and P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation, *Neural Computation*, 15 (6) (2003) 1373-1396.
- [12] Z. Y. Zhang and H. Y. Zha, Principal manifolds and nonlinear dimensionality reduction via tangent space alignment, *SIAM Journal of Scientific Computing*, 26 (1) (2003) 313-338.
- [13] D. D. Ridder, O. Kouropteva and O. Okun, Supervised locally linear embedding, *Lecture Notes in Computer Sci ence*, Springer, Heidelberg (2003) 333-341.
- [14] Y. Zhang, B. Li and W. Wang, Supervised locally tangent space alignment for machine fault diagnosis, *Journal of Mechanical Science and Technology*, 28 (8) (2014) 2971-2977.
- [15] Z. Su, B. Tang and L. Deng, Fault diagnosis method using supervised extended local tangent space alignment for dimension reduction, *Measurement*, 62 (2015) 1-14.
- [16] Q. Jiang et al., Machinery fault diagnosis using supervised manifold learning, *Mechanical Systems and Signal Processing*, 23 (7) (2009) 2301-2311.
- [17] B. S. Yang, T. Han and W. W. Hwang, Application of multi-class support vector machines for fault diagnosis of rotating machinery, *Journal of Mechanical Science and Technology*, 19 (3) (2005) 845-858.
- [18] J. Wei and H. Peng, Neighbourhood preserving based semisupervised dimensionality reduction, *Electronics Letters*, 44 (20) (2008) 1190-1191.
- [19] M. Fan, H. Qiao and B. Zhang, Intrinsic dimension estimation of manifolds by incising balls, *Pattern Recognition*, 42 (5) (2009) 780-787.
- [20] M. Li et al., Multiple manifolds analysis and its application to fault diagnosis, *Mechanical Systems and Signal Processing*, 23 (8) (2009) 2500-2509.
- [21] C. L. Blake and C. J. Merz, *UCI Repository of Machine Learning Databases* (1998).
- [22] G. L. Liao, Research on technology of Mechanical condition monitoring and fault diagnosis based on unsupervised learning, *Ph.D. Thesis*, School of Mechanical Engineering, Huazhong University of Science and Technology, Wuhan, China (2003).
- [23] F. Zhang, Y. Liu, C. Chen, Y.-F. Li and H.-Z. Huang, Fault diagnosis of rotating machinery based on kernel density estimation and Kullback-Leibler divergence, *Journal of Mechanical Science and Technology*, 28 (11) (2014) 4441-4454.



**Quansheng Jiang** received his Ph.D. degree in mechanical engineering from Southeast University, China, in 2009. Currently he is an Associate Professor at School of Mechanical Engineering, Suzhou University of Science and Technology, China. His research interests include signal processing and fault de-

tection for mechanical systems.



**Qixin Zhu** received his Ph.D. degree in Control Theory and Control Engineering from Nanjing University of Aeronautics and Astronautics in 2003. Currently he is a Full Professor at School of Mechanical Engineering, Suzhou University of Science and Technology, Suzhou, China. His research interests include

intelligent control, robot and the applications of control theory in engineering.





**Bangfu Wang** received his M.S. degree in mechanical engineering from Jiangsu University, China, in 2005. Currently he is a Lecturer at School of Mechanical Engineering, Suzhou University of Science and Technology, China. His research interests are Mechatronics and fault detection.

**Lihua Guo** received her Ph.D. degree in mechanical engineering from Southeast University, China, in 2011. Currently she is a Lecturer at School of Mechanical Engineering, Suzhou University of Science and Technology, China. Her research interests are in optimization algorithm for mechanical engineering.