ICIC Express Letters
Part B: Applications
Volume **9**, Number **8**, August **2018**
 ISSN 2185-2766
pp. 765–771

DIMENSION REDUCTION METHOD FOR DATA-DRIVEN PHM SYSTEM BASED ON ULDA ALGORITHM

Chao Zhang, Erwei Li, Liang Liu and Yong Zhou

School of Aeronautics
Northwestern Polytechnical University
No. 127, West Youyi Road, Beilin District, Xi'an 710072, P. R. China
caec_zc@nwpu.edu.cn

Received January 2018; accepted April 2018

Abstract. *In data-driven prognostics and health management (PHM) system, the features extracted from the data always have a high dimension. However, the high dimension data has many redundant parts and it is difficult to calculate. Therefore, feature dimension reduction is needed after feature extraction. Most researches use the principal component analysis (PCA) method, the kernel principal component analysis (KPCA) method, and the orthogonal locality preserving projection (OLPP) method for dimension reduction. However, these methods could not be suitable for solving problems of classification tasks, and thus it should be improved in dimension reduction performance and fault diagnosis accuracy. In this paper, the uncorrelated linear discriminant analysis (ULDA) method, which can explicitly extract a set of statistically uncorrelated features, is used to reduce the data dimension of the data-driven PHM system. The comparison results show the ULDA dimension reduction method has the highest correct rate in these algorithms.*
Keywords: Data dimension reduction, Data-driven PHM, Uncorrelated linear discriminant analysis

1. Introduction.

Prognostics and health management (PHM) technology always uses as few sensors as possible to collect various performance information of the system, and uses intelligent algorithm (such as fuzzy logic, neural networks, data fusion, expert systems, and physical model) to detect the system health status, and can predict the future failures and take a series of condition-based maintenance measures according to the analysis results, so it has been widely used in aeronautics, astronautics, automobile, and energy industry. In general, the PHM system has been divided into three kinds: the physical-model PHM system, the data-driven PHM system and the knowledge-based PHM system. The physical-model PHM system needs to establish a physical model to describe the study object [1]. It is hard for complicated system, so it only suits simple system. The knowledge-based PHM system needs lots of knowledges to establish a knowledge base [2]. The data-driven PHM system studies the system state from history data, establishes an analytic method according to original monitoring data and predicts future behavior of system [3]. It has a higher correct rate and suits more complicated systems.

The data-driven PHM system usually consists of signal processing, fault diagnosis, fault prediction and health management, etc., and the framework is shown as Figure 1. The main works of signal processing include the signal acquisition, feature extraction and feature dimension reduction. In data processing step, the extracted features always have a high dimension and are complicated to calculate; thus, feature dimension reduction is needed.

DOI: 10.24507/icicelb.09.08.765

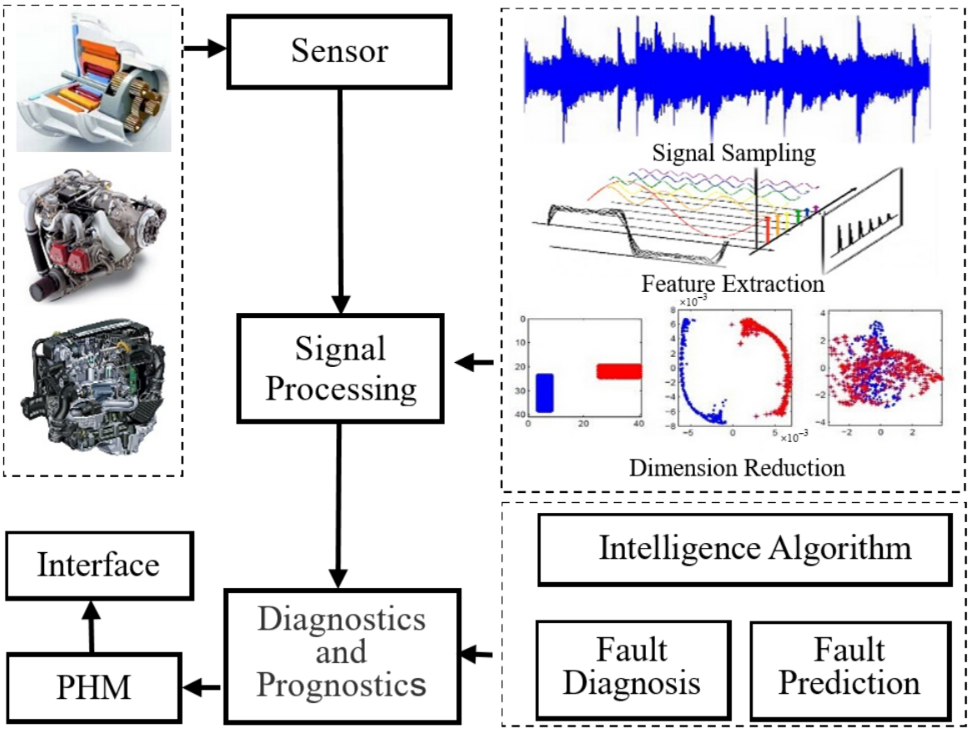


Figure 1. The framework of a data-driven PHM system

Traditional dimension reduction methods are the principal component analysis (PCA) method [4], the kernel principal component analysis (KPCA) method [5] and the orthogonal locality preserving projection (OLPP) method [6], etc. However, the fault diagnosis correct rates of these algorithms are relatively low. To improve dimension reduction performance, some adapted algorithms are proposed based on these algorithms [7-11]. For example, the kernel entropy component analysis (KECA) algorithm is proposed based on the PCA and KPCA algorithms, and it shows better performance in dimension reduction when compared with PCA algorithm and KPCA algorithm [12]. The piecewise linear dimension reduction (PLDR) technique is proposed to overcome the restriction of linearity at relatively low cost [13].

However, because these above methods may result in the loss of the optimal features in some complicated situations, and the fault diagnosis performance of above methods cannot meet the requirements of some important systems. In this paper, the uncorrelated linear discriminant analysis (ULDA) method is used for dimension reduction. The driven data of PHM are extracted from time domain, frequency domain, wavelet package or combined domain. The testing parts are used to test the fault diagnosis performance based on PCA, KPCA, ULDA and OLPP algorithms. Compared with other algorithms, the ULDA method has higher fault diagnosis correct rates and better dimension reduction results.

The remainder of this paper is organized as follows. Section 2 presents the ideas of the standard linear discriminant analysis (SLDA) method and the adapted uncorrelated linear discriminant analysis (ULDA) algorithm. Section 3 compares different algorithms in different domain features and training-testing ratios. Section 4 concludes the work.

2. **Feature Dimension.** In data processing step, the obtained data always has been analyzed by using the statistical methods, the Fourier transform method and the wavelet transform method, so the features can include time domain, frequency domain and wavelet domain. In this paper, the extracted feature parameters are divided into seven sets, that is, time domain feature parameter set, frequency domain characteristic parameter

set, wavelet domain characteristic parameter set, combined characteristic parameter set of time domain and frequency domain, combined characteristic parameter set of time domain and wavelet domain, combined characteristic parameter set of frequency domain and wavelet domain, combined characteristic parameter set of time domain, frequency domain and wavelet domain. These characteristic parameter sets all have a big dimension, so they need a simplification algorithm. The uncorrelated linear discriminant analysis (ULDA) method is used to reduce the data dimension of the driving data. The ULDA algorithm is adapted from the linear discriminant analysis (LDA) algorithm.

2.1. **Linear discriminant analysis.** Linear discriminant analysis (LDA) is a supervised learning method and the discriminant functions in the LDA method are all linear functions. LDA method finds the optimal discriminant vectors or the projection axes based on the classification criteria, such as Fisher criterion. The sample data is projected on the optimal discriminant vector. Then the data of the same kind will be concentrated while the data of different kinds will be scattered. And it makes the intra-class discretization degree and the inter-class discretization ratio reach the maximum value. Therefore, based on the Fisher criterion, the method is to calculate the optimal matrix of the following formula.

$$J(W_{opt}) = \frac{\left|W^T S_b W\right|}{\left|W^T S_w W\right|} \tag{1}$$

where W_{opt} is the optimal matrix, S_b is the inter-class discretization matrix, and S_w is the intra-class discretization matrix.

Define the detection matrix as $X = (x_{ij}) \in R^{n\times p}$, where n and p represent the dimensions of sample parameters and variable parameters. Define the $c_i^T \in R^{l\times p}$ as the i-th sample mean vector and $c^T \in R^{l\times p}$ represents the mean vector of all samples. Define matrix:

$$H_w = \frac{1}{\sqrt{n}}\begin{bmatrix} \left(X_1 - ec_1^T\right) \\ \vdots \\ \left(X_k - ec_1^T\right) \end{bmatrix},\ H_b = \frac{1}{\sqrt{n}}\begin{bmatrix} \sqrt{n_1}\left(X_1 - ec_1^T\right) \\ \vdots \\ \sqrt{n_k}\left(X_k - ec_1^T\right) \end{bmatrix},\ H_t = \frac{1}{\sqrt{n}}\left[\, X - ec^T \,\right] \tag{2}$$

where the X_i represents the i-th sample matrix, n represents the number of the i-th sample, and e represents an all-one matrix. Then discretization matrixes are:

$$S_w = H_w^T H_w, \quad S_b = H_b^T H_b, \quad S_t = H_t^T H_t \tag{3}$$

where S_t represents the total discretization matrix of detection sample.

Calculate the optimal matrix by Lagrange multiplier method, define $W^T S_w W = a \neq 0$, and then the optimal function becomes:

$$L(W, \lambda) = W^T S_b W - \lambda\left(W^T S_w W - a\right) \tag{4}$$

The partial derivative of W is:

$$\frac{\partial L(W, \lambda)}{\partial W} = S_b W - \lambda S_w \tag{5}$$

Set $\frac{\partial L(W,\lambda)}{\partial W} = 0$, then $S_b W = \lambda S_w$, and $W = S_w^{-1} S_b$. The optimal matrix is the eigenvector corresponding to the former d maximum eigenvalue. However, if the data dimension is very big, the inter-class discretization matrix is singular. So the LDA method cannot be used to reduce the dimension of the high dimension data.

2.2. **Uncorrelated linear discriminant analysis.** The uncorrelated linear discriminant analysis (ULDA) method is improved by the LDA method. The feature vector of ULDA method is irrelevant, if the former r vectors are $\{g_1, g_2, \ldots, g_r\}$, then the $r + 1$ vector g_{r+1} obeys the rule:

$$g_{r+1}^T S_t g_k = 0 \quad (k = 1, 2, \ldots, r) \tag{6}$$

The ULDA method overcomes the shortage of LDA method, and it fits high dimension situations. The characteristic matrix transformed by the orthogonal discriminant vector of the detection sample data is linearly independent and does not have any redundant features. The conversion matrix meets:

$$G_{ULDA} = \arg\left(\max\left(trace\left(G^T S_t G\right)^{-1}\right) G^T S_b G\right) \quad \left(G \in R^{m\times l},\ G^T S_t G = I_l\right) \tag{7}$$

The steps of ULDA method are as follows.

Step1: Calculate the column space U of matrix S_t: do the singular value decomposition of H_t, make $H_t = U_1 \Lambda_t V^T$, where $U_1 \in R^{m\times r}$, $\Lambda_t \in R^{r\times r}$ are nonzero diagonal matrixes arranged in descending order, $V \in R^{r\times r}$ is an orthogonal matrix, and $r = rank(H_t)$.

Step2: Whitening processing. Set $W = U_1\Lambda_t^{-1}$, and then the discrete matrix $\tilde{S}_t$ and $\tilde{S}_b$ are:

$$\tilde{S}_t = W^T S_t W = W^T H_t H_t^T W = \Lambda_t^{-1} U_1^T U_1 \Lambda_t U_1^T U_1 \Lambda_t^{-1} = I_r \tag{8}$$

$$\tilde{S}_b = W^T H_b H_b^T W \tag{9}$$

Step3: Do the singular value decomposition of the matrix $W^T H_b$. Set $W^T H_b = M\Lambda_b V^T$, and the corresponding vectors of the largest l non-zero singular values in M are selected to make the matrix M_1 by column.

The dimension reduction steps of ULDA method are as Figure 2, where i represents the number of feature parameters, k is the number of class states, N is total number of feature parameters, M represents the total number of class states, a is the sample data minus the sample mean, and b represents the number of samples when a is under the secondary root.

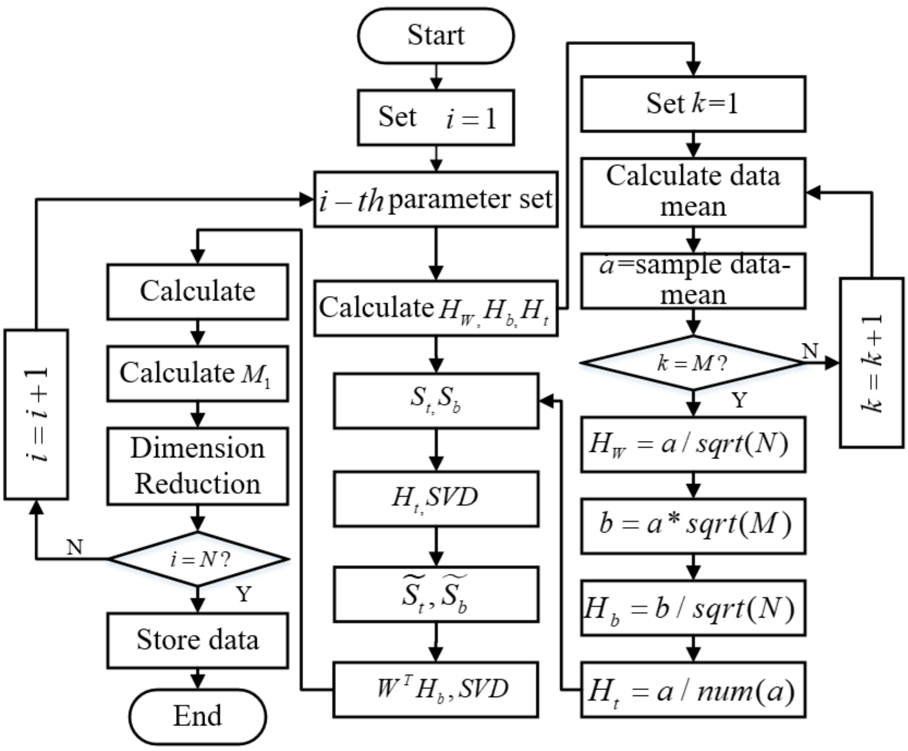
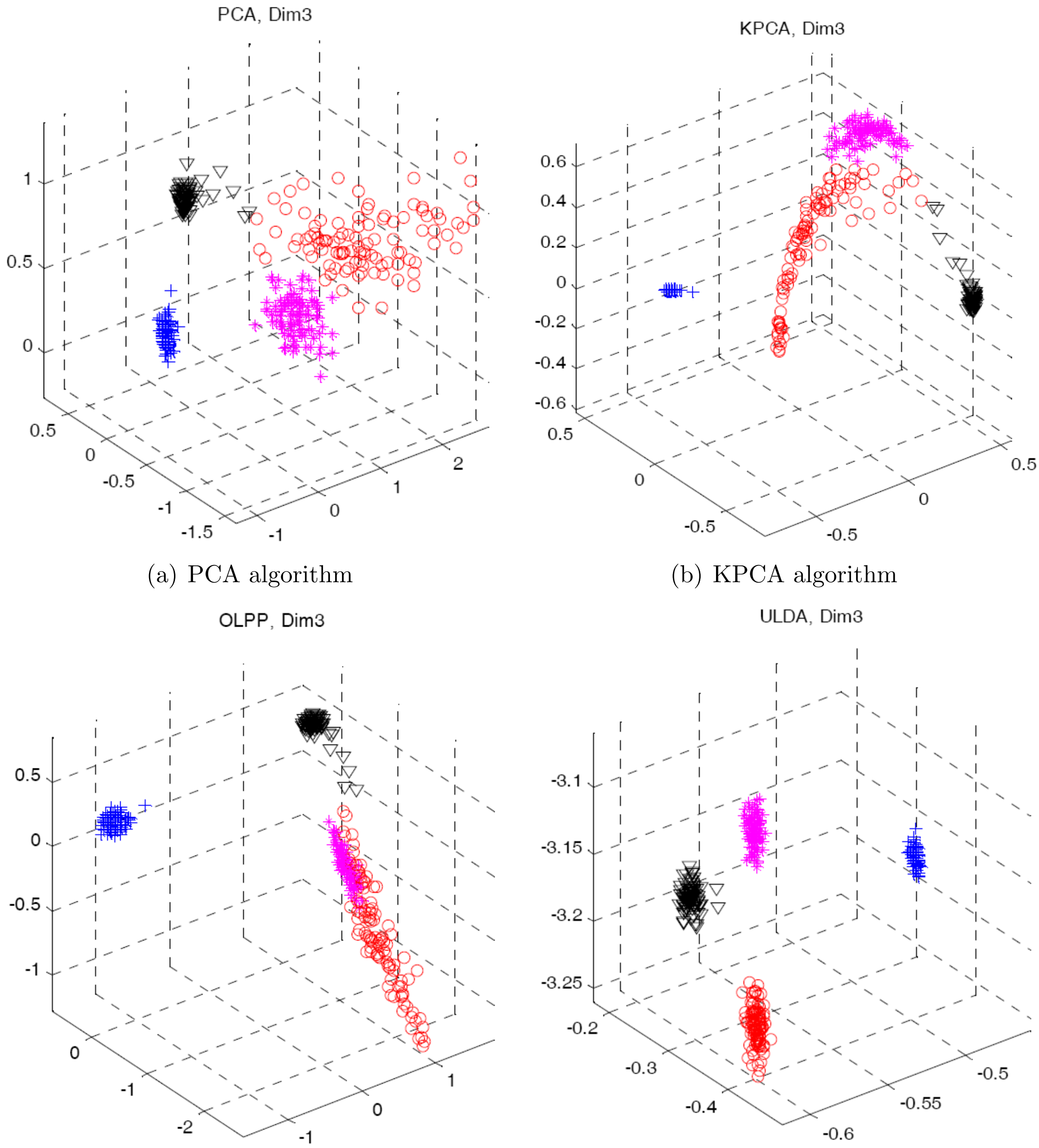


FIGURE 2. The detailed steps of ULDA method

3. **Feature Dimension Algorithm Comparison.**

To verify the superiority of the ULDA method in dimension reduction, the experiments are compared with PCA, KPCA and OLPP algorithms. The fault diagnosis results of different algorithms are as Figure 3, where cross point set represents the normal state, circle point set represents the inner



(a) PCA algorithm (b) KPCA algorithm

(c) OLPP algorithm (d) ULDA algorithm

FIGURE 3. Feature reduction dimension scatter plot

ring fault state, triangle point set represents the rolling failure state, and star point set represents the outer ring fault state.

Studying the features of time domain, frequency domain, wavelet domain feature data sets, the ratio of the training data and the detection data is 3:7, 4:6, 5:5 and 6:4. The fault diagnosis results are shown in Tables 1-4, where TD is time domain, FD is frequency domain and WD is wavelet domain.

From the tables, the ULDA algorithm has higher diagnosis results than other algorithms in different ratios. The ratio of 5:5 has higher correct rates than other ratios. The highest correct rate is 99.80±0.16%, when the driven data is the combined features from time domain, frequency domain and wavelet package.

4. **Conclusions.** In this paper, a feature dimension reduction method for data-driven PHM system based on ULDA algorithm is proposed. The driven data are features extracted in time domain, frequency domain and wavelet domain. The ULDA algorithm

Table 1. The TD and FD diagnosis results (%)

Domain	Train-test ratio	PCA	KPCA	ULDA	OLPP
Time domain	3:7	91.36±4.46	62.30±28.23	**97.82±0.79**	92.91±1.56
	4:6	90.29±6.63	77.83±15.83	**98.15±0.65**	92.98±1.13
	5:5	93.48±2.14	81.08±19.60	**98.40±0.43**	93.95±1.61
	6:4	90.81±4.96	83.09±15.35	**98.84±0.43**	93.90±1.66
Frequency domain	3:7	95.21±1.64	80.55±13.33	**99.11±0.32**	93.89±1.09
	4:6	94.98±1.07	79.40±15.59	**99.14±0.58**	95.35±0.69
	5:5	95.93±0.71	84.95±9.00	**99.55±0.25**	94.00±2.05
	6:4	95.47±0.86	76.56±18.81	**99.53±0.31**	94.41±1.09

Table 2. The WD and TDFD diagnosis results (%)

Domain	Train-test ratio	PCA	KPCA	ULDA	OLPP
Wavelet domain	3:7	90.98±0.85	67.94±15.23	**90.77±1.20**	88.09±3.63
	4:6	90.27±1.27	67.83±16.64	**90.71±0.98**	89.17±1.75
	5:5	91.18±0.83	71.60±10.95	**91.38±0.85**	90.83±0.61
	6:4	90.94±0.88	78.91±7.93	**91.06±1.06**	91.53±1.10
Time domain & Frequency domain	3:7	97.96±0.69	93.52±6.30	**99.60±0.16**	94.41±1.90
	4:6	97.73±0.73	92.50±8.13	**90.69±1.02**	95.54±0.88
	5:5	97.35±0.91	87.63±14.53	**99.71±0.05**	95.65±1.23
	6:4	97.28±0.72	98.66±0.51	**99.37±0.10**	95.69±1.08

Table 3. The TD-WP and FD-WD diagnosis results (%)

Domain	Train-test ratio	PCA	KPCA	ULDA	OLPP
Time domain & Wavelet domain	3:7	97.05±0.93	93.48±4.68	**99.36±0.25**	91.98±2.52
	4:6	97.79±0.50	77.08±11.08	**98.96±0.50**	94.02±2.22
	5:5	97.68±0.54	83.68±10.36	**99.52±0.23**	95.30±1.51
	6:4	98.38±0.56	90.00±9.38	**99.75±0.20**	96.19±1.76
Frequency domain & Wavelet domain	3:7	88.39±1.93	72.29±16.94	**99.43±0.06**	95.07±0.91
	4:6	88.63±1.08	55.42±19.63	**99.50±0.21**	96.44±1.61
	5:5	87.33±2.00	70.48±18.33	**99.70±0.10**	93.80±2.00
	6:4	88.66±1.45	55.91±32.72	**99.23±0.13**	93.78±1.66

Table 4. The TD-FD-WD diagnosis results (%)

Domain	Method	Ratio	Correct rate	Method	Ratio	Correct rate
Time domain & Frequency domain & Wavelet domain	PCA	3:7	98.73±0.74	ULDA	3:7	**99.60±0.23**
		4:6	98.44±0.79		4:6	**99.73±0.14**
		5:5	98.23±0.73		5:5	**99.80±0.16**
		6:4	98.81±0.63		6:4	**99.75±0.23**
	KPCA	3:7	97.66±0.77	OLPP	3:7	94.73±2.05
		4:6	94.23±6.97		4:6	92.65±3.46
		5:5	97.28±1.66		5:5	95.05±1.25
		6:4	99.00±0.39		6:4	96.13±1.11

is compared with PCA, KPCA and OLPP algorithms in different domains and training-testing ratios. The results show the highest correct rate occurs when the training-testing ratio is 5:5 and the ULDA algorithm is used.

Acknowledgment. This work is partially supported by the Basic Research Foundation of NWPU under Grant 3102016ZY002. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] C. Zhang, G. Fu, H. Gu et al., Simulation of temperature effects on GaAs MESFET based on physical model, *IEEE Prognostics and System Health Management Conference*, 2012.
[2] X. Luo, J. Zhang, F. Ye et al., Power series representation model of text knowledge based on human concept learning, *IEEE Trans. System, Man, and Cybernetics*, vol.44, no.1, pp.86-102, 2014.
[3] Y. Lu and Z. Gao, Data-driven model reduction and fault diagnosis for an aero gas turbine engine, *Industrial Electronics and Applications Conference*, pp.1936-1941, 2014.
[4] H. Yu and M. Bennamoun, 1D-PCA, 2D-PCA to nD-PCA, *The 18th International Conference on Pattern Recognition*, pp.181-184, 2006.
[5] Y. Meng, Research on KPCA fault diagnosis method based on multi-domain features, *International Conference on Consumer Electronics, Communications and Networks (CECNet)*, pp.642-645, 2011.
[6] H. Zhao and S. Sun, Optimal locality preserving projection, *Proc. of the International Conference on Image*, 2010.
[7] D. Zhan, H. Zhang, W. Hao et al., An improved LSI dimension reduction algorithm based on vector space modal, *International Conference on Software Engineering and Service Science*, pp.173-176, 2012.
[8] D. Wang, H. Shen and Y. Truong, Efficient dimension reduction for high-dimensional matrix-valued data, *Neurocomputing*, vol.190, pp.25-34, 2016.
[9] D. Heider, C. Bartenhagen, J. N. Dybowski et al., Unsupervised dimension reduction methods for protein sequence classification, *Data Analysis, Machine Learning and Knowledge Discovery*, pp.295-302, 2014.
[10] S. Sengupta and A. K. Das, Dimension reduction using clustering algorithm and rough set theory, *International Conference on Swarm, Evolutionary, and Memetic Computing*, pp.705-712, 2012.
[11] K. Luebke and C. Weihs, Linear dimension reduction in classification: Adaptive procedure for optimum results, *Advances in Data Analysis and Classification*, vol.5, no.3, pp.201-213, 2011.
[12] Y. D. Hu, J. C. Pan and X. Tan, High-dimensional data dimension reduction based on KECA, *Applied Mechanics & Materials*, vols.303-306, pp.1101-1104, 2013.
[13] B. Shen and J. P. Allebach, Piecewise linear dimension reduction for nonnegative data, *Proc. of SPIE – The International Society for Optical Engineering*, 2015.