Journal of Southeast University (English Edition) Vol. 23, No. 1, pp. 51 – 56 Mar. 2007 ISSN 1003—7985

Nonlinear online process monitoring and fault diagnosis of condenser based on kernel PCA plus FDA

Zhang Xi Yan Weiwu Zhao Xu Shao Huihe

(Department of Automation, Shanghai Jiaotong University, Shanghai 200040, China)

Abstract: A novel online process monitoring and fault diagnosis method of condenser based on kernel principle component analysis (KPCA) and Fisher discriminant analysis (FDA) is presented. The basic idea of this method is: First map data from the original space into high-dimensional feature space via nonlinear kernel function and then extract optimal feature vector and discriminant vector in feature space and calculate the Euclidean distance between feature vectors to perform process monitoring. Similar degree between the present discriminant vector and optimal discriminant vector of fault in historical dataset is used for diagnosis. The proposed method can effectively capture the nonlinear relationship among process variables. Simulating results of the turbo generator's fault data set prove that the proposed method is effective.

Key words: nonlinear; kernel PCA; FDA; process monitoring; fault diagnosis; condenser

The turbine generator involves a very complex process and consists of many subsystems. Among them, the condenser is considered as one of the important parts. The condenser fault will damage the generator and lead to the breakdown of the power supply. The economic losses are usually beyond our visualization. Therefore, it is of great importance to recognize the incipient fault of the condenser as early as possible[1]. In the past decade, much research has been done in this field[1–7].

But the diagnosis of a condenser requires monitoring many variables which are highly correlated and the traditional methods often show poor results when applied to the turbine generator process. In order to overcome these shortcomings, a statistical method was considered to do the process monitoring and fault diagnosis of the condenser. Based on principle component analysis (PCA), partial least squares (PLS) and canonical correlation analysis (CVA), the statistical method can project the data from high-dimensional feature space into a low-dimensional space and identify important features of the data. These methods have been used and extended in various applications[8–10]. However, some complicated cases such as the generator process, PCA, PLS and CVA perform poorly due to their linear characteristics. To address this problem, a new combination technique called kernel PCA plus FDA has been developed[11–12].

The basic idea of kernel PCA plus FDA is to nonlinearly map the data into high-dimensional feature space, in which the data has a linear structure, then performs FDA in the feature space. It uses "kernel tricks" to solve the computation of independent projection directions in high-dimensional feature space and ultimately convert the problem of performing FDA in feature space into a problem of implementing FDA in the kernel principal component analysis (KPCA) transformed space[13–14]. Its essence is equivalent to kernel Fisher discriminant analysis (KFDA), but the algorithm is more transparent.

In this paper, a novel nonlinear combination process monitoring and fault diagnosis method of condenser based on kernel PCA plus FDA is proposed. If a fault occurs, the similar degree between the present discriminant vector and the optimal vector of fault in a historical dataset is used for diagnosis. The simulation results show that the proposed method is effective.

1 Principles of KPCA plus FDA

1.1 Fundamentals

Suppose that a set of M training samples $x_1, x_2, \ldots, x_M$ take values in an n-dimensional space. l_i is the number of training samples of class i and satisfies $\sum_{i=1}^{c} l_i = M$, c is the number of classes. Let the input space X be mapped into the feature space F:

$$\Phi: X \rightarrow F \quad x \rightarrow \Phi(x) \tag{1}$$

The idea of kernel FDA is to solve the problem of FDA in the feature space F, thereby yielding a set of nonlinear discriminant vectors in input space. This can be achieved by maximizing the following Fisher criterion:

$$J^{\Phi}(\boldsymbol{\varphi}) = \frac{\boldsymbol{\varphi}^{\mathrm{T}} \boldsymbol{S}_{\mathrm{b}}^{\Phi} \boldsymbol{\varphi}}{\boldsymbol{\varphi}^{\mathrm{T}} \boldsymbol{S}_{\mathrm{t}}^{\Phi} \boldsymbol{\varphi}} \qquad \boldsymbol{\varphi} \neq 0 \tag{2}$$

Received 2006-10-18.

Foundation item: The National Natural Science Foundation of China (No. 60504033).

Biographies: Zhang Xi (1974—), male, graduate; Shao Huihe (corresponding author), male, professor, HHshao@sjtu.edu.cn.

where S_b^Φ and S_t^Φ are the between-classes and total scatter matrices defined in feature space F:

$$S_b^\Phi = \frac{1}{M}\sum_{i=1}^{c} l_i(u_i^\Phi - u_0^\Phi)(u_i^\Phi - u_0^\Phi)^T$$

$$S_t^\Phi = \frac{1}{M}\sum_{i=1}^{M} (\Phi(x_i) - u_0^\Phi)(\Phi(x_i) - u_0^\Phi)^T \quad (3)$$

$$S_t^\Phi = \frac{1}{M}\sum_{i=1}^{M} (\Phi(x_i) - u_0^\Phi)(\Phi(x_i) - u_0^\Phi)^T \quad (4)$$

where u_i^Φ is the mean vector of the mapped training samples of class i and u_0^Φ is the mean vector across all mapped training samples[11-12].

1.2 Kernel Fisher optimal discriminant vectors

The optimal discriminant vectors with respect to the Fisher criterion are actually the eigenvectors of the generalized equation $S_b^\Phi \varphi = \lambda S_t^\Phi \varphi$. Since any of its eigenvectors can be expressed by a linear combination of the observations in feature space, we have

$$\varphi = \sum_{j=1}^{M} \alpha_j \Phi(x_j) = Q\alpha \quad (5)$$

where $Q = [\Phi(x_1), \ldots, \Phi(x_M)]$ and $\alpha = \{a_1, \ldots, a_M\}^T$.

Substituting Eq. (5) into Eq. (2), the Fisher criterion is converted to[15]

$$J^K(\alpha) = \frac{\alpha^T(KWK)\alpha}{\alpha^T(KK)\alpha} \quad (6)$$

where the matrix K is defined as

$$K = \tilde{K} - 1_M\tilde{K} - \tilde{K}1_M + 1_M\tilde{K}1_M \quad (7)$$

where, $1_M = \frac{1}{M}\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix}_{M\times M}$, $\tilde{K} = Q^TQ$ is an $M \times M$ matrix, and its elements are determined by

$$\tilde{K}_{ij} = \Phi(x_i)^T\Phi(x_j) = (\Phi(x_i)\cdot\Phi(x_j)) = k(x_i, y_j) \quad (8)$$

where $k(x, y)$ is the kernel function corresponding to a given nonlinear mapping Φ, and $W = \mathrm{diag}(W_1, \ldots, W_c)$, W_j is an $l_j \times l_j$ matrix with terms all equal to $1/l_j$. Thereby, W is an $M \times M$ block diagonal matrix.

Now, let us consider the QR decomposition of matrix K. Suppose that $\gamma_1, \gamma_2, \ldots, \gamma_m$ are K's orthonormal eigenvectors corresponding to m (m is the rank of K) nonzero eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_m$. Then, K can be expressed by $K = P\Lambda P^T$, where $P = (\gamma_1, \gamma_2, \ldots, \gamma_m)$ and $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots \lambda_m)$.

Obviously, $P^TP = I$, where I is the identity matrix. Substituting $K = P\Lambda P^T$ into Eq. (6) and let

$$\beta = \Lambda^{\frac{1}{2}}P^T\alpha \quad (9)$$

we have

$$J^K(\alpha) = \frac{(\Lambda^{\frac{1}{2}}P^T\alpha)^T(\Lambda^{\frac{1}{2}}P^TWP\Lambda^{\frac{1}{2}})(\Lambda^{\frac{1}{2}}P^T\alpha)^T}{(\Lambda^{\frac{1}{2}}P^T\alpha)^T\Lambda(\Lambda^{\frac{1}{2}}P^T\alpha)} \quad (10)$$

Then, Eq. (10) becomes

$$J(\beta) = \frac{\beta^TS_b\beta}{\beta^TS_t\beta} \quad (11)$$

where

$$S_b = \Lambda^{\frac{1}{2}}P^TWP^T\Lambda^{\frac{1}{2}}, \quad S_t = \Lambda \quad (12)$$

It is easy to know that S_t is positive definite and S_b is semi-positive definite. Thus, Eq. (11) is a standard generalized Rayleigh quotient. By maximizing this Rayleigh quotient, we can obtain a set of optimal solutions $\beta_1, \beta_2, \ldots, \beta_d$, which are actually the eigenvectors of $S_t^{-1}S_b$ corresponding to d ($d \leqslant c-1$) largest eigenvalues.

From Eq. (9), we know that for a given β, there exists at least one α satisfying $\alpha = P\Lambda^{-1/2}\beta$. Thus, after determining $\beta_1, \beta_2, \ldots, \beta_d$, we can obtain a set of optimal solutions $\alpha_j = P\Lambda^{-1/2}\beta_j$ ($j = 1, 2, \ldots, d$) with respect to the criterion in Eq. (6). Thereby, the optimal discriminant vectors with respect to the Fisher criterion in Eq. (2) in feature space are

$$\varphi_j = Q\alpha_j = QP\Lambda^{-\frac{1}{2}}\beta_j \qquad j = 1, 2, \ldots, d \quad (13)$$

1.3 Essence of KFD transformation: KPCA plus FDA

Given a sample x and its mapped image $\Phi(x)$, we can obtain the discriminant feature vector z by the following KFD transformation

$$z = \psi^T\Phi(x) \quad (14)$$

where $\psi = \{\varphi_1, \varphi_2, \ldots, \varphi_d\} = (QP\Lambda^{-1/2})\{\beta_1, \beta_2, \ldots, \beta_d\}$.

The above transformation can be divided into two items

$$y = (QP\Lambda^{-\frac{1}{2}})^T\Phi(x) \quad (15)$$

$$z = G^Ty, \quad G = (\beta_1, \beta_2, \ldots, \beta_d) \quad (16)$$

Let us consider the transformation in Eq. (15) first. Since $Q = [\Phi(x_1), \ldots, \Phi(x_M)]$, $P = \{\gamma_1, \gamma_2, \ldots, \gamma_m\}$ and $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)$, Eq. (15) can be rewritten as

$$y = \left(\frac{\gamma_1}{\sqrt{\lambda_1}}, \ldots, \frac{\gamma_m}{\sqrt{\lambda_m}}\right)^T (\Phi(x_1), \ldots, \Phi(x_m))^T\Phi(x) = \left(\frac{\gamma_1}{\sqrt{\lambda_1}}, \ldots, \frac{\gamma_m}{\sqrt{\lambda_m}}\right)^T [k(x_1, x), \ldots, k(x_M, x)] \quad (17)$$

Since $\gamma_1, \gamma_2, \ldots, \gamma_m$ are K's orthonormal eigenvectors and $\lambda_1, \lambda_2, \ldots, \lambda_m$ are the associated nonzero eigenvalues, the transformation in Eq. (17) is exactly the KPCA transformation[16], which transforms feature space F into Euclidean space R^m.

Let us view the issues in the KPCA-transformed space R^m. Looking back at Eq. (11) and considering the matrices S_b and S_t within the function $J(\beta)$, it is easy to verify that they are actually the between-class and total scatter matrices in the KPCA-transformed space R^m. Since the expression of S_b in Eq. (12) is not so intuitive, we can construct it directly in R^m based on

KPCA-transformed features

$$S_{\mathrm{b}} = \frac{1}{M}\sum_{i=1}^{c} l_i(\boldsymbol{u}_i - \boldsymbol{u}_0)(\boldsymbol{u}_i - \boldsymbol{u}_0)^{\mathrm{T}} \tag{18}$$

where l_i is the number of training samples in class i, $\boldsymbol{u}_i$ is the mean vector of the training samples in class i, and $\boldsymbol{u}_0$ is the mean vector across all training samples.

Since $\boldsymbol{S}_{\mathrm{b}}$ and $\boldsymbol{S}_{\mathrm{t}}$ are the between-class and total scatter matrices in the KPCA-transformed space R^m, the function $J(\boldsymbol{\beta})$ is actually the Fisher criterion in such space and its stationary points $\beta_1, \beta_2, \ldots, \beta_d$ are the associated Fisher optimal discriminant vectors. Correspondingly, the transformation in Eq. (16) is essentially the Fisher linear discriminant transformation in the KPCA-transformed space.

So, the essence of KFD has been revealed and it is equal to KPCA plus FDA. KPCA is first employed to reduce the dimension of feature space to m, and then FDA is used for further feature extraction in the KPCA-transformed space R^m[11-13].

2 Online Monitoring and Fault Diagnosis Strategy using KPCA plus FDA

Process monitoring methods based on PCA monitor variables in the principle component and residual space using Hotelling's T^2 and Q statistics. The essence of the T^2 statistic is a kind of weighted statistical distance[17]. The process monitoring method based on kernel PCA plus FDA also uses distance as a statistic and it compares the Euclidean distance of the optimal kernel Fisher feature vector between present data and reference data to perform process monitoring.

Fault diagnosis was performed using contribution plots in traditional statistical methods. When a fault occurs, through the contribution rate of variables, we can find the root cause of the fault. However, in high-dimensional feature space, it is difficult or even impossible to find an inverse mapping to the original space to calculate the contribution rate of variables to the fault. So, the monitoring method based on kernel PCA plus FDA performs fault diagnosis not using contribution plots but by using pattern matching technology. As we know, in KPCA plus FDA analysis, the optimal kernel Fisher discriminant vectors extracted from different faults are different. Through calculating the similar degree between the present kernel feature vector and the optimal kernel vectors of the faults, we can decide that the one in the historical dataset, which is mostly similar to the present optimal feature vector, can be recognized as the present fault. In practical application, we set a diagnosis limit τ. If the similar degree between the present discriminant vector and the optimal vector of each fault is smaller than τ, a new fault can be recognized. We can discriminate the fault according to our experience and add it to the historical fault data set.

2.1 Definition of confidence bounds

Once a model has been developed that reflects the normal operation region, it is necessary to detect any departure of the process from its standard behavior. That is, we must calculate the confident limit value to determine whether the process is in control or not. In PCA or PLS monitoring, Hotelling's T^2 analysis and Q charts are effective tools for extracting the critical features of the data. These analyses are based on the assumption that the probability density functions of the latent variables follow a multivariate Gaussian distribution. However, contrary to this assumption, Martin and Morris[18] reported that the latent variables in many industrial processes rarely have a multivariate Gaussian distribution through tests for multivariate normality on the scores. An alternative approach to defining the nominal operating regions is to use data-driven techniques such as non-parametric empirical density estimates using kernel extraction [18-19].

The latent variables in the KPCA plus FDA monitoring method also do not follow a Gaussian distribution. So, we can use kernel density estimation to calculate the confidence limits of the D statistic.

A univariate kernel estimator with kernel K is defined by

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \tag{19}$$

where x is the data point under consideration, x_i is an observation value from the dataset, h is the window width, n is the number of observations, and K is the kernel function.

The control limit used in FDA monitoring charts in the KPCA-transformed feature space can be obtained using kernel density estimation as follows. First, the D values from normal operating data are required. Then, the univariate kernel density estimator is used to estimate the density function of the normal D values. The point, occupying 99% area of density function, can be obtained and becomes the control limit of normal operating data[14, 18-19]. We denoted it as D^*.

2.2 Outline of online process monitoring and fault diagnosis using KPCA plus FDA

Step 1 Select the appropriate nonlinear kernel function and map the reference dataset $\boldsymbol{X}_{\mathrm{ref}}$ and new sample dataset $\boldsymbol{X}_{\mathrm{new}}$ from original space into high-dimensional feature space. Acquire the kernel reference dataset $\boldsymbol{\xi}_{x_{\mathrm{ref}}}$ and the kernel sample dataset $\boldsymbol{\xi}_{x_{\mathrm{new}}}$.

Step 2 Construct the centralized inner product matrix $\boldsymbol{K}$ using Eq. (7) and calculate its orthonormal eigenvectors $\gamma_1, \gamma_2, \ldots, \gamma_m$ corresponding to m largest nonzero eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$. Carry out KPCA transformation using Eq. (17).

Step 3 Construct the between-class scatter matrix $\boldsymbol{S}_{\mathrm{b}}$ using Eq. (18) and the total scatter matrix $\boldsymbol{S}_{\mathrm{t}} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)$. Calculate $\boldsymbol{S}_{\mathrm{t}}^{-1}\boldsymbol{S}_{\mathrm{b}}$'s eigenvector $\boldsymbol{\beta}_{\max}$ corresponding to the largest eigenvalue.

Step 4 Acquire the optimal kernel Fisher discriminant vector $\boldsymbol{\varphi}_{\mathrm{opt}}$ using $\boldsymbol{\varphi}_{\mathrm{opt}} = \boldsymbol{QP\Lambda}^{-1/2}\boldsymbol{\beta}_{\max}$.

Step 5 Project the kernel reference dataset $\boldsymbol{\xi}_{x_{\mathrm{ref}}}$ and the kernel sample dataset $\boldsymbol{\xi}_{x_{\mathrm{new}}}$ to the optimal kernel discriminant vector $\boldsymbol{\varphi}_{\mathrm{opt}}$ and acquire the kernel optimal feature vectors $\boldsymbol{T}_{\mathrm{ref}}$ and $\boldsymbol{T}_{\mathrm{new}}$.

Step 6 Calculate the Euclidean distance between the two kernel fisher feature vectors via $D = \|\boldsymbol{T}_{\mathrm{ref}} - \boldsymbol{T}_{\mathrm{new}}\|^2$.

Step 7 If the statistic D is larger than the control limit D^*, the fault may occur and calculate the similar coefficient S to diagnose the type of faults.

$$S = \frac{\boldsymbol{\varphi}_{\mathrm{opt}}\boldsymbol{\varphi}_i^{\mathrm{T}}}{\|\boldsymbol{\varphi}_{\mathrm{opt}}\| \cdot \|\boldsymbol{\varphi}_i\|} \tag{20}$$

where $\boldsymbol{\varphi}_i$ is the optimal kernel Fisher discriminant vector of the present data set, and $\boldsymbol{\varphi}_{\mathrm{opt}}$ is the optimal kernel Fisher discriminant vector of the fault data set (assuming that there are n kinds of faults in the fault data set). From Eq. (20), we know that the similar coefficient S is the cosine value of the two optimal kernel fisher vectors' angle.

3 Application Studies

As indicated in Ref. [2], there are 21 kinds of typical faults in a condenser. The variables which are related to the faults are of 33 kinds. A complete list of the variables can be found in Ref. [2]. We select all the 33 kinds of variables as monitoring variables and nine kinds of typical faults for simulation. The simulation faults are listed in Tab. 1. The data come from the historical data set of some 300 MW turbine generators. We collect 600 samples including fault 2 as test data. The fault is introduced at sample 300 and is followed to the end of the process.

Tab. 1 Typical faults of condenser

Fault ID	Fault description
1	Recycle pump fault
2	Water filling in condenser
3	Cooling pump fault
4	Walling up of cooling pipe in condenser
5	Stream pump fault
6	Cooling tower fault
7	Vacuum pump fault
8	Smudginess of cooling pipe in condenser
9	Heater pipe bursting

We select a group of normal data and three groups of fault data (fault 1, fault 2 and fault 3) from the historical data set and extract the first and the second optimal discriminant vector using Fisher discriminant analysis. Then we project the data to the optimal discriminant vector and obtain a scatter plot of the first and the second feature vector in the original space. As shown in Fig. 1, only fault 2 can be differentiated clearly from the normal data and the other faults, but it cannot differentiate fault 1 and fault 3 from the normal data. This is because there exist complicated nonlinear relationships among variables. FDA is a linear method intrinsically, so it shows poor performance in dealing with data with strong nonlinearity. The scatter plot of the first kernel feature vector and the second one via FDA in the KPCA transformed space is shown in Fig. 2. It can be seen from the figure that after projecting data to the high-dimensional feature space through selecting the appropriate kernel function, the KPCA plus FDA method can easily differentiate among different data. We used the RBF function as the selected kernel function.

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(-\frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2}{c}\right) \tag{21}$$

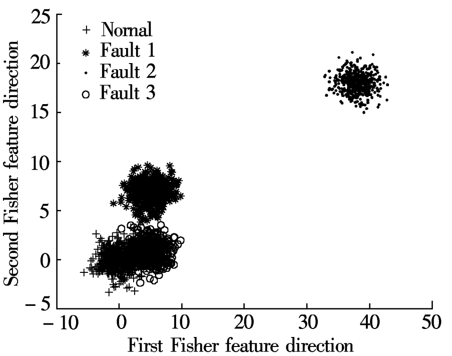

Fig. 1 Scatter plot in original feature space

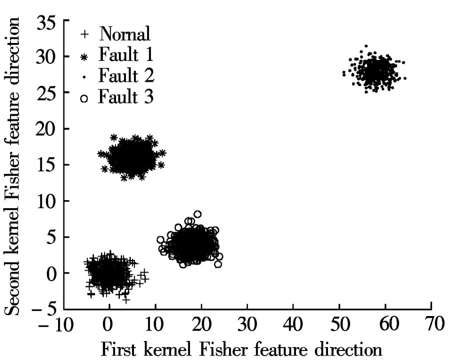

Fig. 2 Scatter plot in high-dimensional feature space

From the above simulation results we know that the kernel PCA plus FDA method (we also can call it the kernel FDA method) can effectively extract the feature vector of nonlinear data and differentiate data belonging to different classes. So it is possible for us to perform process monitoring using Euclidean distance between the optimal kernel Fisher feature vectors of different data sets as a statistic.

In order to demonstrate the predominance of the KPCA plus FDA monitoring scheme, three monitoring approaches were investigated using the recorded data set. PCA and KPCA were applied first, followed by a comparison to the KPCA plus FDA monitoring method.

When we do online monitoring using KPCA plus FDA. We first calculate a 99% confidence limit D^* via kernel density estimation. The PCA and KPCA monitoring charts are shown in Fig. 3 and Fig. 4, respectively. From Fig. 3, we know that the T^2 and Q statistic of PCA can detect the occurring fault, but they fall below the 99% control limit after about sample 390 in the T^2 chart and sample 410 in the Q chart. However, the monitoring results of applying KPCA to the same process data are given in Fig. 4, which shows relatively fast and clear fault detection results in comparison to PCA, but it is still not satisfactory.

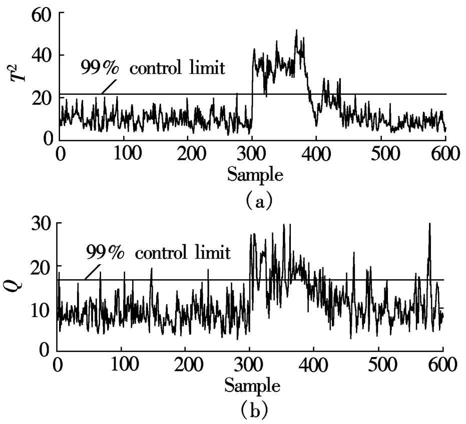

Fig. 3 PCA monitoring result of case 2

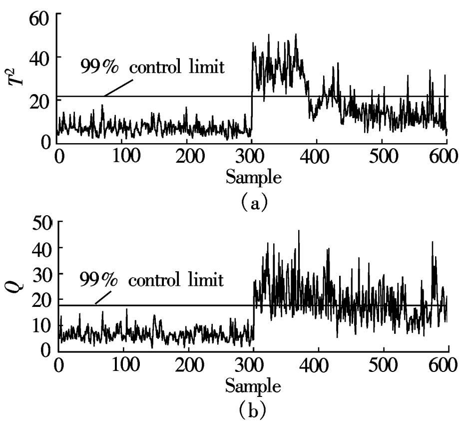

Fig. 4 KPCA monitoring result of case 2

The monitoring result of fault 2 using kernel PCA plus FDA is shown in Fig. 5. From the figure, we know that the statistical distance increases sharply when the fault occurs at sample 300 and exceeds the 99% control limit. Then it calculates the similar degree between the present optimal discriminant vector and the optimal vector of fault in the historical dataset to perform fault diagnosis. As shown in Fig. 6, the present optimal kernel Fisher discriminant vector is more similar to fault 2 than others and the similar value is 0. 95. So, we can determine that the fault is from water filling in the condenser (fault 2).

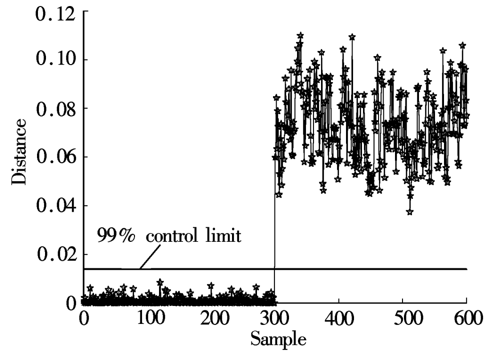

Fig. 5 Monitoring result of case 2

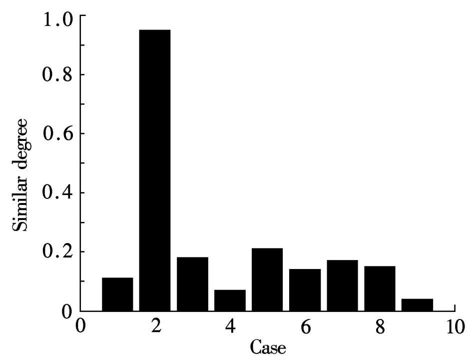

Fig. 6 Similar degree of case 2 in historical database

4 Conclusion

A novel process monitoring and fault diagnosis method of a turbine generator based on kernel PCA and FDA was proposed. It first mapped the data from the original space into high-dimensional KPCA-transform space via nonlinear kernel function and then extracted the optimal kernel Fisher feature vector and discriminant vector to perform process monitoring and fault diagnosis. The proposed method can effectively capture the nonlinear relationship in process variables. It is evaluated by the application to the turbine generator historical fault data set and its effectiveness is demonstrated.

References

[1] Hou Guolian, Sun Xiaogang, Zhang Jianhua, et al. Research on fault diagnosis of condenser via nonlinear principal component analysis and probabilistic neural networks [J]. *Proceedings of the CSEE*, 2005, **25**(18): 104 − 108.

(in Chinese)
[2] Wang Peihong, Zhu Yuna, Jia Junying, et al. Application of fuzzy pattern recognition to fault diagnosis of condensers[J]. *Proceedings of the CSEE*, 1999, **19**(10): 46-49. (in Chinese)
[3] Ma Liangyu, Wang Bingshu, Tong Zhensheng, et al. Fuzzy pattern recognition and artificial neural networks used for fault diagnosis of the double channel condenser[J]. *Proceedings of the CSEE*, 2001, **21**(8): 68-73. (in Chinese)
[4] Zhang Jianhua, Hou Guolian, Zhang Wei, et al. A research on fuzzy rules and genetic algorithm based method for a condenser fault diagnosis[J]. *Proceedings of the CSEE*, 2004, **24**(4): 205-209. (in Chinese)
[5] Yu Zhiwei, Su Baoku, Zeng Ming. Application of wavelet packet in fault diagnosis system of large scale DC motor rotor[J]. *Proceedings of the CSEE*, 2005, **25**(22): 158-162. (in Chinese)
[6] Zhang Guoyun, Zhang Jing. Fuzzy SVM-based multilevel binary tree classifier for fault diagnosis of hydro turbine speed regulating system[J]. *Proceedings of the CSEE*, 2005, **25**(8): 100-104. (in Chinese)
[7] Jiang Dongxiang, Diao Jinhui, Zhao Gang, et al. Study on methods of vibration fault diagnosis based on time-frequency contour map for turbine generator unit[J]. *Proceedings of the CSEE*, 2005, **25**(6): 146-151. (in Chinese)
[8] Bakshi B R. Multiscale PCA with application to multivariate statistical process monitoring[J]. *AIChE Journal*, 1998, **44**(7): 1596-1610.
[9] Simoglou A, Martin E B, Morris A J. Statistical performance monitoring of dynamic multivariate processes using state space modeling[J]. *Computers and Chemical Engineering*, 2002, **26**(6): 909-920.
[10] Negiz A, Cinar A. Statistical monitoring of multivariable dynamic processes with state-space models[J]. *AIChE Journal*, 1997, **43**(8): 2002-2020.
[11] Yang Jian, Jin Zhong, Yang Jingyu, et al. Essence of kernel Fisher discriminant: KPCA plus LDA[J]. *Pattern Recognition*, 2004, **37**(10): 2097-2100.
[12] Yang Jian, Frangi Alejandro F, Yang Jinyu, et al. KPCA plus LDA: A complete kernel Fisher discriminant framework for feature extraction and recognition[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2005, **27**(2): 230-244.
[13] Yang Jian, Gao Xiumei, Zhang David, et al. Kernel ICA: An alternative formulation and its application to face recognition[J]. *Pattern Recognition*, 2005, **38**(10): 1784-1787.
[14] Lee Jong-Min, Yoo Chang-Kyoo, Choi Sang-Wook, et al. Nonlinear process monitoring using kernel principal component analysis[J]. *Chemical Engineering Science*, 2004, **59**(1): 223-234.
[15] Baudat G, Anouar F. Generalized discriminant analysis using a kernel approach[J]. *Neural Computation*, 2000, **12**(10): 2385-2404.
[16] Schölkopf B, Smola A, Müller K. Nonlinear component analysis as a kernel eigenvalue problem[J]. *Neural Computation*, 1998, **10**(5): 1299-1319.
[17] Johnson R A, Wichern D W. *Applied multivariate statistical analysis*[M]. 3rd ed. Englewood Cliffs, New Jersey: Prentice-Hall, 1992.
[18] Martin E B, Morris A J. Non-parametric confidence bounds for process performance monitoring charts[J]. *Journal of Process Control*, 1996, **6**(6): 349-358.
[19] Chen Q, Wynne R J, Goulding P, et al. The application of principal component analysis and kernel density estimation to enhance process monitoring[J]. *Control Engineer Practice*, 2000, **8**(5): 531-543.

基于核 PCA 和 FDA 的凝汽器非线型过程监控和故障诊断

张　曦　阎威武　赵　旭　邵惠鹤

(上海交通大学自动化系,上海 200240)

摘要:提出了基于核主元分析和 Fisher 判别分析相结合的非线性统计过程监控和故障诊断新方法. 该方法首先利用非线性核函数将数据从原始空间映射到高维空间,然后在高维空间中利用线性 Fisher 判别分析法提取数据最优的核 Fisher 判别矢量和特征矢量,通过计算特征矢量之间的欧式距离来实现过程监控. 若系统发生故障,则根据当前故障的判别矢量和历史故障数据集中所含故障的最优核 Fisher 判别矢量的相似度进行故障诊断. 所提出的方法能有效的捕获过程变量之间的非线性关系,汽轮机特征故障数据集仿真试验验证了该方法的有效性.

关键词:非线性;核 PCA;FDA;过程监控;故障诊断;凝汽器

中图分类号:TK267;TP277