Online Contribution Rate Based Fault Diagnosis for Nonlinear Industrial Processes

PENG Kai-Xiang[1] ZHANG Kai[1] LI Gang[2]

Abstract Over past decades, kernel principal component analysis (KPCA) appeared quite popularly in data-driven process monitoring area. Enormous work has been done to show its simplicity, feasibility, and effectiveness. However, the introduction of kernel trick makes it impossible to directly employ traditional contribution plots for fault diagnosis. In this paper, on the basis of revisiting and analyzing the existing KPCA-relevant diagnosis approaches, a new contribution rate based method is proposed which can explain the faulty variables clearly. Furthermore, a scheme for online nonlinear diagnosis is established. In the end, a case study on continuous stirred tank reactor (CSTR) benchmark is applied to access the effectiveness of the new methodology, where the comparisons with the traditional linear method are involved as well.

Key words Kernel principal component analysis (KPCA), nonlinear, fault detection, contribution rate, fault diagnosis

Citation Peng Kai-Xiang, Zhang Kai, Li Gang. Online contribution rate based fault diagnosis for nonlinear industrial processes. Acta Automatica Sinica, 2014, **40**(3): 423−430

DOI 10.3724/SP.J.1004.2014.00423

Online monitoring and diagnosis of process operation performance and conditions are crucial for the safety and reliability of industrial processes. As a data-driven methodology in process monitoring area, multivariate statistical process monitoring (MSPM) is quite popular and is overwhelmingly being focused in monitoring complex industrial processes, such as chemical and microelectronics manufacturing plants. Within MSPM community, multivariate projection techniques such as principal component analysis (PCA) and partial least squares or projection to latent structures (PLS) mainly focus on establishing the data-driven models for process monitoring. Generally, the schemes based on PCA and PLS are realized via training a benchmark model according to a series of normal operating measurements, then testing the online measurement[1−6].

For a linear process, PCA or PLS based approaches can be summarized sequentially in three steps:

Step 1. Acquire two low-dimensional subspaces, which reflect the most significant variations and the residuals, respectively.

Step 2. Detect the possible variations with some statistical indices in different subspaces.

Step 3. Isolate the potential faulty variables with contribution plots, which measure the variables′ contributions to the detection index.

As for nonlinear processes, the above steps should be modified and updated correspondingly. In essence, PCA performs well for linear data, while kernel principal component analysis (KPCA), as an advanced version of PCA functions better for nonlinear cases[7−10]. Naturally, KPCA model maps the process variables into a high-dimensional space, where a linear PCA structure is constructed. KPCA facilitates this procedure by introducing the idea of kernel, however, it relatively increases the complexity of overall model through producing some new parameters. Moreover, these parameters often play crucial roles towards the monitoring results. Various researches in this domain appear frequently[10−14]. Furthermore, a sequence of variants for KPCA models are considered in MSPM field to cope with the dynamics, time-varying and non-Gaussianity etc.[15−17].

Once a fault is detected, it is necessary to identify the faulty variables. The most popular methods for fault diagnosis are contribution plots, since they do not require any known faulty information. However, Alcala and Qin pointed out that the traditional method is easily affected by the smearing of the fault, and they proposed a reconstruction based contribution (RBC) plots which eases the problem to a large extent[18]. Fault detection based on KPCA is fairly similar to PCA based methods, while the stage of diagnosis is quite different and difficult. Cho et al. employed the gradient of kernel function explored by Rakotomamonjy to derive an analytical solution of contribution for KPCA based diagnosis, whereas its mechanism is quite confusing and not straightforward to understand[7, 9]. Choi et al. proposed a RBC based methodology which is similar to the one introduced by Dunia et al.[10, 19]. Nevertheless, their approach just borrowed the idea of KPCA based de-noising[20], and had not built the relationship between the diagnosis and detection indices. Recently, Alcala and Qin prompted the idea of RBC to KPCA successfully[11]. However, in their deduction, the estimation for the fault is biased and unconvincing. In this paper, a new contribution rate plot is proposed towards KPCA based process monitoring, which is partially motivated by Cho et al.′s work. Different from their idea, the proposed method is more simple to understand and explicit. Meanwhile, a more practical approach based on contribution rate denoted as accumulative relative contribution rate (ARCR) is also given and demonstrated.

The remaining sections of this paper are organized as follows. Section 1 revisits the KPCA model and KPCA based monitoring policy. The contribution rate based diagnosis is introduced to identify the faulty variables for nonlinear

Manuscript received May 29, 2012; accepted August 23, 2013
Supported by National Natural Science Foundation of China (61074085), Beijing Natural Science Foundation (4122029, 4142035), and the Fundamental Research Funds for the Central Universities (FRF-SD-12-008B, FRF-AS-11-004B)
Recommended by Associate Editor ZHONG Mai-Ying
1. Key Laboratory for Advanced Control of Iron and Steel Process, School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China 2. Department of Automation, Tsinghua National Laboratory for Information Science and Technology, Tsinghua University, Beijing 100084, China

processes in Section 2. Section 3 gives a benchmark case study to illustrate the feasibility of the new method. At last, the conclusions are drawn in Section 4.

1 KPCA based nonlinear process monitoring

1.1 KPCA model for nonlinear process

Given a nonlinear process consisting of m variables with n samples which can be described as: $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n]^{\mathrm{T}} \in \mathbf{R}^{n\times m}$. Define ϕ as a nonlinear mapping function to project the original variables into a high-dimensional feature space F: $\boldsymbol{x} \in \mathbf{R}^m \Rightarrow \phi(\boldsymbol{x}) \in \mathbf{R}^M$, where they are existing linearly. After the nonlinear mapping, the original input matrix $\boldsymbol{X}$ is transformed to $\boldsymbol{\Phi} = [\phi(\boldsymbol{x}_1), \phi(\boldsymbol{x}_2), \cdots, \phi(\boldsymbol{x}_n)]^{\mathrm{T}} \in \mathbf{R}^{n\times M}$. Define $\boldsymbol{K} \in \mathbf{R}^{n\times n}$ as the kernel matrix to represent $\boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{T}}$, where $\boldsymbol{K}_{ij} = K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \langle \phi(\boldsymbol{x}_i), \phi(\boldsymbol{x}_j)\rangle, i, j = 1, 2, \cdots, n$. The introduction of the kernel trick simplifies the overall algorithm without using an explicit nonlinear mapping function[7]. In order to center the feature data to zero mean, the following preprocessings are executed[21−22] : $\boldsymbol{\Phi} = \boldsymbol{\Phi}_{\mathrm{raw}} - 1_n\bar{\boldsymbol{\Phi}}_{\mathrm{raw}}$ where $\boldsymbol{\Phi}_{\mathrm{raw}}$ is the directly mapped matrix, $\bar{\boldsymbol{\Phi}}_{\mathrm{raw}}$ denotes the mean of $\boldsymbol{\Phi}_{\mathrm{raw}}$, 1_n represents the n-dimension column vector whose elements are all ones. The centered $\boldsymbol{K}$ matrix can be derived as: $\boldsymbol{K} = (\boldsymbol{I}_n - (1/n)\, 1_n 1_n^{\mathrm{T}})\, \boldsymbol{K}_{\mathrm{raw}}\, (\boldsymbol{I}_n - (1/n)\, 1_n 1_n^{\mathrm{T}})$, where $\boldsymbol{K}_{\mathrm{raw}}$ represents direct kernel matrix.

Then, KPCA attempts to find a unit vector w, which maximizes the following objective function:

$$\begin{aligned} &\max_w \left(\frac{1}{n-1}\right) w^{\mathrm{T}}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi} w \\ &\text{s.t.} \quad w^{\mathrm{T}} w = 1 \end{aligned} \tag{1}$$

Although $\boldsymbol{\Phi}$ is unavailable, it is able to acquire w via introducing a n-dimensional vector α. It should be a priori supposed that $w = \boldsymbol{\Phi}^{\mathrm{T}}\alpha$ holds. Thus, substituting $w = \boldsymbol{\Phi}^{\mathrm{T}}\alpha$ to (1):

$$\begin{aligned} &\max_\alpha \left(\frac{1}{n-1}\right) \alpha^{\mathrm{T}}\boldsymbol{K}^2\alpha \\ &\text{s.t.} \quad \alpha^{\mathrm{T}}\boldsymbol{K}\alpha = 1 \end{aligned} \tag{2}$$

Finally, the solution of α in above equation can be transformed into solving an eigen decomposition, which can be expressed as: $\lambda\alpha = (1/(n-1))\,\boldsymbol{K}\alpha$, where λ denotes the eigenvalue of $(1/(n-1))\,\boldsymbol{K}$ while α is corresponding eigenvector. Then, the ultimate α is normalized to $\|\alpha\|^2 = (1/(n-1))\,\lambda$, which guarantees $w^{\mathrm{T}}w = 1$ holds. Then the first score vector of $\boldsymbol{\Phi}$ is obtained by $\boldsymbol{t}_1 = \boldsymbol{K}\alpha_1$. Likewise, all stored scores are organized as $\boldsymbol{T} \in \mathbf{R}^{n\times A} = \boldsymbol{K}\boldsymbol{A}$, where $\boldsymbol{A} = [\alpha_1, \cdots, \alpha_A]$ contains the scaled eigenvectors corresponding to the A largest eigenvalues of $\boldsymbol{K}$, which can be described as $\lambda_1, \cdots, \lambda_A$. A represents the order of PCA, which is chosen according to [7−12]. Then, $\boldsymbol{W} = [w_1, \cdots, w_A] \in \mathbf{R}^{M\times A}$ is formed as the loading matrix regarding $\boldsymbol{\Phi}$, where $w_i = \boldsymbol{\Phi}^{\mathrm{T}}\alpha_i$.

After KPCA calculations, we can model $\boldsymbol{\Phi}$ as follows:

$$\boldsymbol{\Phi} = \hat{\boldsymbol{\Phi}} + \tilde{\boldsymbol{\Phi}} = \boldsymbol{T}\boldsymbol{W}^{\mathrm{T}} + \tilde{\boldsymbol{\Phi}} \tag{3}$$

where $\tilde{\boldsymbol{\Phi}}$ signifies the residual matrix that is used to account for the noise in $\boldsymbol{\Phi}$.

For a test data $\boldsymbol{x}_{\mathrm{new}} \in \mathbf{R}^m$, the directly mapped test feature vector is $\phi(\boldsymbol{x}_{\mathrm{new}})_{\mathrm{raw}}$, then the direct test inner dot vector is calculated as $(\boldsymbol{K}_{\mathrm{raw}}^{\mathrm{new}})_i = \phi(\boldsymbol{x}_i)_{\mathrm{raw}}\phi(\boldsymbol{x}_{\mathrm{new}})_{\mathrm{raw}} = K(\boldsymbol{x}_i, \boldsymbol{x}_{\mathrm{new}})$. Also the centered test data vector $\phi(\boldsymbol{x}_{\mathrm{new}})$ is mean-centered by $\phi(\boldsymbol{x}_{\mathrm{new}}) = \phi(\boldsymbol{x}_{\mathrm{new}})_{\mathrm{raw}} - \bar{\boldsymbol{\Phi}}_{\mathrm{raw}}^{\mathrm{T}}$. $\boldsymbol{K}_{\mathrm{new}}$ is also mean-centered by: $\boldsymbol{K}_{\mathrm{new}} = (\boldsymbol{I}_n - (1/n)\, 1_n 1^{\mathrm{T}}{}_n)\, (\boldsymbol{K}_{\mathrm{raw}}^{\mathrm{new}} - (1/n)\, \boldsymbol{K}_{\mathrm{raw}} 1_n)$. Ultimately, the new score of $\boldsymbol{x}_{\mathrm{new}}$ is calculated by

$$\boldsymbol{t}_{\mathrm{new}} = \boldsymbol{W}^{\mathrm{T}}\phi(\boldsymbol{x}_{\mathrm{new}}) = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}_{\mathrm{new}} \in \mathbf{R}^A \tag{4}$$

which plays a central role for process monitoring.

1.2 Fault detection based on KPCA

T^2 statistic and square prediction error (SPE) are often utilized for KPCA based detection. The residuals of $\phi(\boldsymbol{x}_{\mathrm{new}})$: $\tilde{\phi}(\boldsymbol{x}_{\mathrm{new}}) = \phi(\boldsymbol{x}_{\mathrm{new}}) - \boldsymbol{W}\boldsymbol{t}_{\mathrm{new}}$ cannot be calculated explicitly because of the unspecific map function $\phi(\cdot)$. After obtaining $\boldsymbol{t}_{\mathrm{new}}$ according to (4), two statistics T^2 and SPE can be calculated:

$$T^2 = \boldsymbol{t}_{\mathrm{new}}^{\mathrm{T}}\Lambda^{-1}\boldsymbol{t}_{\mathrm{new}} \tag{5}$$

$$SPE = \left\|\tilde{\phi}(\boldsymbol{x}_{\mathrm{new}})\right\|^2 \tag{6}$$

The popular approach for control limits of these two statistics are available in [1]. Although $\phi(\boldsymbol{x}_{\mathrm{new}})$ is unavailable, we calculate SPE by the kernel trick as follows,

$$\begin{aligned} SPE = \ & \left\|\tilde{\phi}(\boldsymbol{x}_{\mathrm{new}})\right\|^2 = \\ & \phi^{\mathrm{T}}(\boldsymbol{x}_{\mathrm{new}})\,\phi(\boldsymbol{x}_{\mathrm{new}}) - \\ & 2\boldsymbol{t}_{\mathrm{new}}^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}}\phi(\boldsymbol{x}_{\mathrm{new}}) + \boldsymbol{t}_{\mathrm{new}}^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{t}_{\mathrm{new}} = \\ & \phi^{\mathrm{T}}(\boldsymbol{x}_{\mathrm{new}})\,\phi(\boldsymbol{x}_{\mathrm{new}}) - \boldsymbol{t}_{\mathrm{new}}^{\mathrm{T}}\boldsymbol{t}_{\mathrm{new}} \end{aligned} \tag{7}$$

where $\phi^{\mathrm{T}}(\boldsymbol{x}_{\mathrm{new}})\,\phi(\boldsymbol{x}_{\mathrm{new}}) = 1 - (2/n)\sum_{i=1}^{n}\boldsymbol{K}_{\mathrm{raw}}^{\mathrm{new}}(i) + (1/n^2)\sum_{i=1}^{n}\sum_{j=1}^{n}\boldsymbol{K}_{\mathrm{raw}}(i,j)$.

Within the statistics, T^2 is applied for the systematic abnormalities, whereas SPE is for monitoring the process noise. According to the former study, two types of combined detection indices were proposed by Cho et al. and Alcala et al. as summarized in Table 1. By contrast, it can be concluded that, Alcala′s method balances T^2 and SPE into φ, while Cho incorporates T^2 and weighted SPE into φ. In addition, φ defined by Cho et al. can be treated as a measurement of the energy for a sample, which is quite simple to follow and apply[7]. Thus, Cho et al.′s index is an ideal choice in the subsequent parts of this paper. Meanwhile, the control limit for φ is obtained according to kernel density estimation (KDE) based method[10].

Table 1 Combined index based on KPCA model

Algorithm	φ
Cho et al. (2005)[7]	$\varphi = T^2 + \lambda_\perp^{-1} SPE$
Alcala and Qin (2010)[11]	$\varphi = \frac{T^2}{\delta^2} + \frac{SPE}{\tau^2}$

[1] $\lambda_\perp$ is a constant, representing the nonsignificant eigenvalues in KPCA algorithm[7].

[2] δ^2 and τ^2 are thresholds of T^2 and SPE, respectively[11].

With kernel idea, the whole procedure of KPCA model is considerably simplified. However, the kernel parameter has to be specified in the off-line training stage. Take Gaussian

kernel functions for example:

$$K(\boldsymbol{x}_i,\boldsymbol{x}_j)=\phi^{\mathrm{T}}(\boldsymbol{x}_i)_{\mathrm{raw}}\,\phi(\boldsymbol{x}_j)_{\mathrm{raw}}=\exp\left(-\frac{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|^2}{c}\right) \tag{8}$$

As $c>0$, if $c\mapsto 0$, it means $\phi(\boldsymbol{x}_i)\perp\phi(\boldsymbol{x}_j)$. Then, the order of PCA will be large, since all training data are treated without relations, and the false alarm rates will then be large, as many test samples which vary in tolerable scales may cause excessive alarms. If $c\mapsto\infty$, namely $\phi(\boldsymbol{x}_i)=\phi(\boldsymbol{x}_j)$ holds, then, the order A will be small respectively, since all training data are processed similarly, also the detection rates will be consequently small, because some faulty samples are wrongly seen as the normal data. To conclude, the choice of c is a trade-off between false alarm rate and fault detection rate. In addition, it also affects the complexity of KPCA model which is governed by order A[8−9, 13−14].

As for this paper, a cross-validation combined with bisection approach is used to search for c. First of all, we choose a tolerant false alarm rate zone Ω (5 %∼8 % in this paper) as a benchmark zone and $c_{\min}$ and $c_{\max}$ as the boundary values for c. Then bisection is implemented to search for the optimal c of which the false alarm rate drops into Ω[23].

2 Contribution rate based fault diagnosis

2.1 Problem formulation

After a fault is detected, the subsequent step of process monitoring is to diagnose the faulty variables[24]. For PCA model, traditional contribution plots based approaches function well. However, it cannot be utilized directly for nonlinear cases with an intangible $\phi(\boldsymbol{x})$. Intuitively, if we want to derive the contribution of each variable for KPCA based method, the kernel function should be concentrated.

Prior presenting contribution rate, the approach proposed by Cho et al. should be revisited at first. In their work, the Gussian kernel function is rewritten as:

$$K(\boldsymbol{x}_j,\boldsymbol{x}_k)=k(\boldsymbol{v}\cdot\boldsymbol{x}_j,\boldsymbol{v}\cdot\boldsymbol{x}_k)=\exp\left(-\frac{\|\boldsymbol{v}\cdot\boldsymbol{x}_j-\boldsymbol{v}\cdot\boldsymbol{x}_k\|^2}{c}\right) \tag{9}$$

where $\boldsymbol{v}=[v_1,v_2,\cdots,v_m]^{\mathrm{T}}$ represents a scaling factor and $v_i=1$ for $i=1,\cdots,m$. Then, the gradient of the new kernel function regarding the variable v_i defined by Rakotomamonjy was employed to calculate the contribution of the ith variable, which is presented as:

$$\frac{\partial K(\boldsymbol{x}_j,\boldsymbol{x}_k)}{\partial v_i}=-\frac{1}{c}(\boldsymbol{x}_{j,i}-\boldsymbol{x}_{k,i})^2K(\boldsymbol{x}_j,\boldsymbol{x}_k)\,|_{v_i=1} \tag{10}$$

Based on the preliminary knowledge, Cho et al. derived the contribution of each variable to the detection index as shown in [7]. The approach is applicable to the process monitoring using KPCA and has a precise theoretical analysis. However, its physical meaning is ambiguous which makes it difficult to understand.

2.2 The derivation of contribution rate

The contribution rate underlies an assumption that all variables are disturbed by the same scale. Namely, when the whole process variables are violated to the same extent, the variables which give larger influence to detection index are identified as the faulty variables. The violation is symbolized by $\boldsymbol{x}\odot\boldsymbol{v}$, where $\odot$ represents component-wise vector product, $\boldsymbol{v}$ is defined as the same as in (9), while it is treated as a variable vector instead of a constant vector with all unit elements. Furthermore, all elements of $\boldsymbol{v}$ are equally guarantee that $\boldsymbol{x}$ varies to the same extent. Then, for $\boldsymbol{x}_{\mathrm{new}}$, according to first-order Taylor series expansion, the following equation holds,

$$\varphi(\boldsymbol{x}_{\mathrm{new}}\odot\boldsymbol{v})\approx\varphi(\boldsymbol{x}_{\mathrm{new}})+\sum_{i=1}^{m}\frac{\partial\varphi(\boldsymbol{x}_{\mathrm{new}}\odot\boldsymbol{v})}{\partial v_i}\,|_{\boldsymbol{v}=\boldsymbol{1}_m}(v_i-1) \tag{11}$$

where $\boldsymbol{v}=\boldsymbol{1}_m$ means $v_i=1$, $i=1,\cdots,m$. The contribution rate of ith variable to φ is defined as:

$$C(\boldsymbol{x}_{\mathrm{new}},i)=\left|\frac{\partial\varphi(\boldsymbol{x}_{\mathrm{new}}\odot\boldsymbol{v})}{\partial v_i}\,|_{\boldsymbol{v}=\boldsymbol{1}_m}\right| \tag{12}$$

Conceptually, contribution rate defines the influence of the variable to φ. Given a measurement $\boldsymbol{x}$, Fig. 1 demonstrates the physical analysis for this definition, where there are two variables. The contribution rate of each variable to φ denotes the gradient of $\varphi(\boldsymbol{x}\odot\boldsymbol{v})$ with respect to its corresponding scale factor v.

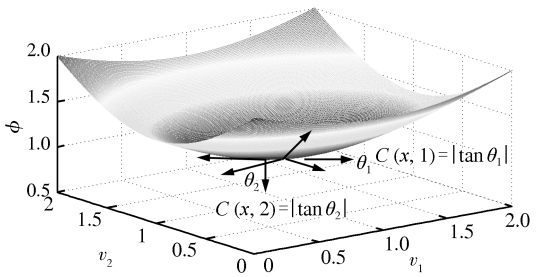


Fig. 1 Physical analysis of contribution rate

Equation (12) can be realized quantitatively in the following. First of all,

$$\varphi(\boldsymbol{x}_{\mathrm{new}}\odot\boldsymbol{v})=\boldsymbol{k}_{\mathrm{new}}^{\mathrm{T}}\boldsymbol{A}\Lambda^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{k}_{\mathrm{new}}+\lambda_{\perp}\left(\phi^{\mathrm{T}}(\boldsymbol{x}_{\mathrm{new}}\odot\boldsymbol{v})\,\phi(\boldsymbol{x}_{\mathrm{new}}\odot\boldsymbol{v})-\boldsymbol{k}_{\mathrm{new}}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{k}_{\mathrm{new}}\right) \tag{13}$$

where $\boldsymbol{k}_{\mathrm{new}}=\boldsymbol{\Phi}\phi(\boldsymbol{x}_{\mathrm{new}}\odot\boldsymbol{v})$. Then, (12) can be transformed to

$$C(\boldsymbol{x}_{\mathrm{new}},i)=\left|\mathrm{tr}\left(\frac{\partial(\boldsymbol{k}_{\mathrm{new}}\boldsymbol{k}_{\mathrm{new}}^{\mathrm{T}})}{\partial v_i}\,|_{\boldsymbol{v}=\boldsymbol{1}_m}\,\boldsymbol{A}\Lambda^{-1}\boldsymbol{A}^{\mathrm{T}}\right)+\lambda_{\perp}\left\{\frac{\partial\left(\phi^{\mathrm{T}}(\boldsymbol{x}_{\mathrm{new}}\odot\boldsymbol{v})\,\phi(\boldsymbol{x}_{\mathrm{new}}\odot\boldsymbol{v})\right)}{\partial v_i}-\mathrm{tr}\left(\frac{\partial(\boldsymbol{k}_{\mathrm{new}}\boldsymbol{k}_{\mathrm{new}}^{\mathrm{T}})}{\partial v_i}\,|_{\boldsymbol{v}=\boldsymbol{1}_m}\,\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\right)\right\}\right| \tag{14}$$

Subsequently, the pth row and qth column element of $\frac{\partial(\boldsymbol{k}_{\mathrm{new}}\boldsymbol{k}_{\mathrm{new}}{}^{\mathrm{T}})}{\partial v_i}\,|_{\boldsymbol{v}=\boldsymbol{1}_m}$ can be denoted as $\frac{\partial(\boldsymbol{k}_{\mathrm{new}}\boldsymbol{k}_{\mathrm{new}}^{\mathrm{T}})_{p,q}}{\partial v_i}|_{\boldsymbol{v}=\boldsymbol{1}_m}$.

To obtain the relevant partial gradient, it firstly requires to calculate $\frac{\partial K(\boldsymbol{x}_{\rm new}\odot\boldsymbol{v},\boldsymbol{x}_j)}{\partial v_i}|_{\boldsymbol{v}=\boldsymbol{1}_m}$:

$$\frac{\partial K\left(\boldsymbol{x}_{\rm new}\odot\boldsymbol{v},\boldsymbol{x}_j\right)}{\partial v_i}|_{\boldsymbol{v}=\boldsymbol{1}_m}= -\frac{2}{c}\boldsymbol{x}_{{\rm new},i}\left(\boldsymbol{x}_{{\rm new},i}-\boldsymbol{x}_{k,i}\right)K\left(\boldsymbol{x}_{\rm new},\boldsymbol{x}_k\right) \tag{15}$$

Remark 1. The fundamental difference between the present approach and Cho et al.′s work[7] could be found via comparing (15) with (10).

To the end, substitute (15) into $\frac{\partial\left(\boldsymbol{k}_{\rm new}\boldsymbol{k}_{\rm new}^{\rm T}\right)_{p,q}}{\partial v_i}|_{\boldsymbol{v}=\boldsymbol{1}_m}$ and $\frac{\partial\left(\phi^{\rm T}(\boldsymbol{x}_{\rm new}\odot\boldsymbol{v})\phi(\boldsymbol{x}_{\rm new}\odot\boldsymbol{v})\right)}{\partial v_i}$, the following equations hold.

$$\begin{aligned}\frac{\partial\left(\boldsymbol{k}_{\rm new}\boldsymbol{k}_{\rm new}^{\rm T}\right)_{p,q}}{\partial v_i}|_{\boldsymbol{v}=\boldsymbol{1}_m}= &-\frac{2}{c}(\boldsymbol{k}_{\rm new}\left(p\right)\boldsymbol{k}_{\rm new}^{\rm raw}\left(q\right)\boldsymbol{x}_{{\rm new},i}\left(\boldsymbol{x}_{q,i}-\boldsymbol{x}_{{\rm new},i}\right)+\\ &\boldsymbol{k}_{\rm new}\left(q\right)\boldsymbol{k}_{\rm new}^{\rm raw}\left(p\right)\boldsymbol{x}_{{\rm new},i}\left(\boldsymbol{x}_{p,i}-\boldsymbol{x}_{{\rm new},i}\right))+\frac{1}{nc}(\boldsymbol{k}_{\rm new}\left(p\right)+\\ &\boldsymbol{k}_{\rm new}\left(q\right))\boldsymbol{x}_{{\rm new},i}\sum_{k=1}^{n}\left(\boldsymbol{x}_{k,i}-\boldsymbol{x}_{{\rm new},i}\right)\boldsymbol{K}_{\rm new}^{\rm raw}\left(k\right)\end{aligned} \tag{16}$$

$$\frac{\partial\phi^{\rm T}\left(\boldsymbol{x}_{\rm new}\odot\boldsymbol{v}\right)\phi\left(\boldsymbol{x}_{\rm new}\odot\boldsymbol{v}\right)}{\partial v_i}|_{\boldsymbol{v}=\boldsymbol{1}_m}= -(2/n)\sum_{j=1}^{n}\boldsymbol{x}_{{\rm new},i}\left(\boldsymbol{x}_{{\rm new},i}-\boldsymbol{x}_{j,i}\right)K\left(\boldsymbol{x}_{\rm new},\boldsymbol{x}_j\right) \tag{17}$$

So far, (12) is calculable and can be applicable for fault diagnosis.

2.3 Comparison and implementing remarks

Considering the two methods, the computational burden of present one seems less than Cho et al′s, as it merely calculates $\boldsymbol{x}_{\rm new}\odot\boldsymbol{v}$ instead of both $\boldsymbol{x}_{\rm new}\odot\boldsymbol{v}$ and $\boldsymbol{x}\odot\boldsymbol{v}$. Furthermore, the definition of the new contribution rate is clear physically. Also, the proposed technique provides a new diagnosis framework for other kernel-related models. The disadvantages lie in the aspect that (11) holds approximately, which entails that the new concept is given under a linear assumption. However, this approximation is reasonable and valid, since, to address the issue like nonlinearity, lots of similar treatments could be found[25−26].

Note that prior to construct KPCA model, the raw data are merely normalized to zero mean, and not scaled to unit variance. Thus, all variables′ contribution rates under normal condition may be different from each other, since the feature mapping may effect all variables unevenly. Therefore it is meaningful to normalize the raw contribution rates obtained by (14) to make sure that all variables contribute more or less the same to φ under normal operation. The ultimate contribution rate for both training and testing samples are expressed separately:

$$C\left(\boldsymbol{x},i\right)=\frac{C_{\rm raw}\left(\boldsymbol{x},i\right)}{m_c\left(i\right)} \tag{18}$$

$$C\left(\boldsymbol{x}_{\rm new},i\right)=\frac{C_{\rm raw}\left(\boldsymbol{x}_{\rm new},i\right)}{m_c\left(i\right)} \tag{19}$$

where $m_c\left(i\right)=(1/n)\sum_{j=1}^{n}C_{\rm raw}\left(\boldsymbol{x}_j,i\right)$.

2.4 Accumulative relative contribution rate

Up to now, the absolute contribution rates are available. However, directly implementing contribution rate on nonlinear fault diagnosis still needs to pay more attention on two problems: 1) When contribution rates are available for a faulty measurement, how could we identify the latent fault root-cause and 2) how can we eliminate the smearing caused by fault transmission effectively. Compared with the absolute contribution rate, the relative one is preferred to identify those variables whose absolute contribution rates are large. So, each contribution rate of a sample is firstly divided by the sum of all contribution rates, then the original relative contribution rate is obtained:

$$C_r\left(\boldsymbol{x}_{\rm new},i\right)=\frac{C\left(\boldsymbol{x}_{\rm new},i\right)}{\sum_{i=1}^{m}C\left(\boldsymbol{x}_{\rm new},i\right)} \tag{20}$$

where $C\left(\boldsymbol{x}_{\rm new},i\right)$ varies around $1/m$.

Finally, an accumulative relative contribution rate (ARCR) is established as follows:

A faulty sample $\boldsymbol{x}_{\rm new}$ is available,

1) Obtain the absolute contribution rate $C\left(\boldsymbol{x}_{\rm new},i\right)$, then calculate relative $C_r\left(\boldsymbol{x}_{\rm new},i\right)$, $\Sigma C_r(\boldsymbol{x}_{\rm new},i)=1$;

2) Rearrange $C_r\left(\boldsymbol{x}_{\rm new},i\right)$ in the descending sequence;

3) Orderly collect the corresponding variables whose accumulative relative contribution rates exceed θ.

ARCR attempts to isolate the variables according to their cumulative contribution rates. θ is a crucial parameter corresponding to the final control limit. An appropriate θ can deal with the smearing, while an extreme one may also result in wrong diagnosis. Generally, $\theta>1/m$, as, $C_r(\boldsymbol{x}_{\rm new})_{\max}>1/m$. However, a large control limit isolates more variables which may mislead the results. Practically, θ that satisfies $\theta<\|C_r\left(\boldsymbol{x}_{\rm new}\right)\|_2$ is a valid choice where $\|\cdot\|_2$ represents the operator of Euclid norm and $C_r(\boldsymbol{x}_{\rm new})_{\max}<\|C_r\left(\boldsymbol{x}_{\rm new}\right)\|_2<1$. Hence, the choice of θ is a trade-off between $1/m$ and $\|C_r\left(\boldsymbol{x}_{\rm new}\right)\|_2$. In this study, this problem can be reduced to choose $\theta=\frac{\|C_r(\boldsymbol{x}_{\rm new})\|_2+1/m}{2}$.

To sum up, the implementation of the proposed diagnosis scheme involves KPCA modeling, φ based detection and contribution rate based diagnosis. The whole procedure of online monitoring for nonlinear processes can be summarized step wise in Fig. 2.

3 Case study on CSTR benchmark

In this section the continuous stirred tank reactor (CSTR) with feedback control is applied to assess the performance of the new approach.

3.1 Description of CSTR

The CSTR process can be described by the following differential equations:

$$\begin{aligned}\frac{{\rm d}C_A}{{\rm d}t}&=\frac{q}{V}\left(C_{Af}-C_A\right)-k_0\exp\left(-\frac{E}{RT}\right)C_A+v_1\\ \frac{{\rm d}T}{{\rm d}t}&=\frac{q}{V}\left(T_f-T\right)+\frac{-H}{\rho C_p}k_0\exp\left(-\frac{E}{RT}\right)C_A+\\ &\quad\frac{UA}{V\rho C_p}\left(T_c-T\right)+v_2\end{aligned} \tag{21}$$

where C_A is the outlet concentration, T is the reaction temperature, T_c is the temperature of cooling water, q is the input fluent velocity of reactant, C_{Af} is the input reactant concentration, T_f is the input reactant temperature, and v_1, v_2 are process noise, where $v_i \sim \mathrm{N}\left(0, \sigma_v^2\right)$, $i = 1, 2$. In this case study, T_f, C_{Af}, T, q, T_c and C_A are referred as process variables $\boldsymbol{X}$. Measurement noises are also added to $\boldsymbol{X}$ in the formula of $\boldsymbol{x}(k) = \boldsymbol{x}_{\mathrm{raw}}(k) + \boldsymbol{e}(k)$, where $\boldsymbol{e}(k) \sim \mathrm{N}\left(0, \sigma_e^2\right)$, k is the observation number. The classic PID feedback control with the transfer function of $K\left(1 + T_d s + T_I\left(1/s\right)\right)\varepsilon$ is employed to fulfill the closed-loop control, where $\varepsilon = [C_A - C_A^*, T - T^*]^{\mathrm{T}}$, inputs are added to $[q, T_c]^{\mathrm{T}}$, C_A^* and T^* are the normal observations of the outlet concentration and reaction temperature. All relevant systematic parameters and operational conditions are available in [19]. To satisfy this application, a little changes are given as follows: $\sigma^2(T_f) = \sigma^2(C_{Af}) = 0.05$, $\sigma_v = 0.04$, $\sigma^2(C_A) = 1\mathrm{E} - 4$. Under normal circumstance, 500 samples are collected and recorded to form KPCA model, where $c_{\min} = 0.1$, $c_{\max} = 1\,000$, $c = 500$ is obtained according to the end part of Subsection 1.2, $\lambda_\perp = 0.02$, $A = 7$ is kept for this simulation. Absolute contribution rate plot of a normal sample is presented in Fig. 3 which is consistent with the conclusion in Subsection 2.3. Three types of fault scenarios are formulated with 1 000 samples:

1) Fault I: a step bias in $\boldsymbol{x}_2$ at 501st sample: $\begin{cases} \boldsymbol{x}_2(k) = 1 & k < 500 \\ \boldsymbol{x}_2(k) = 1.2 & k \geq 501 \end{cases}$.

2) Fault II: a ramp disturbance in $\boldsymbol{x}_1$ at 501st sample: $\begin{cases} \boldsymbol{x}_1(k) = 400 & k < 500 \\ \boldsymbol{x}_1(k) = 400 + (k - 400)/20 & k \geq 501 \end{cases}$.

3) Fault III: a random disturbance in $\boldsymbol{x}_1$ at 501st sample: $\begin{cases} \boldsymbol{x}_1(k) = 400 & k < 500 \\ \boldsymbol{x}_1(k) = 400 + \xi & k \geq 501 \end{cases}$, where $\xi \sim \mathrm{N}\left(0, 25^2\right)$.

There are also artificial noises that should be added to the samples under the fault scenarios. Meanwhile for this simulation, we also give the following remark.

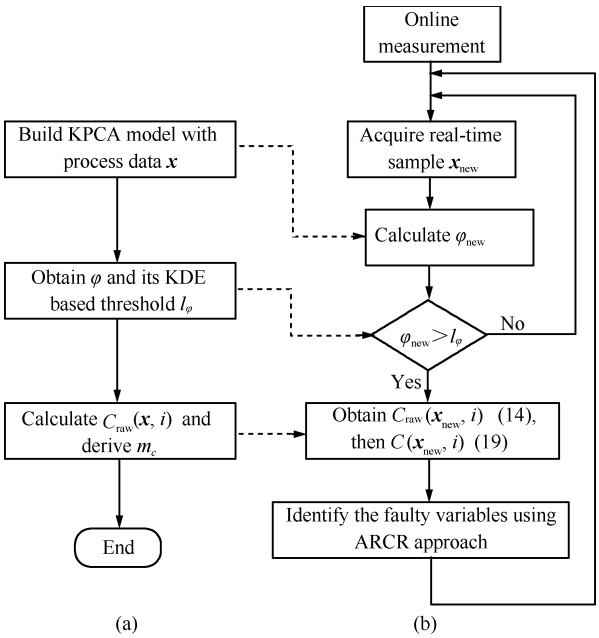


Fig. 2 Implementation of contribution rate ((a) Off-line training model; (b) Online testing model)

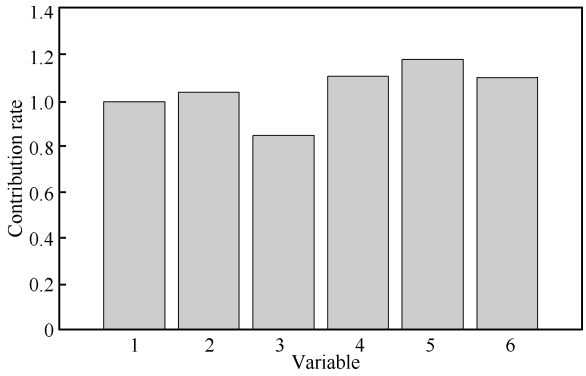


Fig. 3 Diagnosis result under normal situation

Remark 2. For comparison with the linear method, a PCA based diagnosis approach is also employed, where the detection index and contribution plot are constructed by referring refs. [19] and [27], respectively.

3.2 Results and discussion

Fault I involves a step change in C_{Af} ($\boldsymbol{x}_2$), which leads to abnormalities in other variables due to the close-loop feedback control. The detection result with φ is shown in Fig. 4, where the fault is detected from 501st time instant. PCA based diagnosis for this fault is shown in Fig. 5. To give a valid result compared with the present method, Fig. 5 utilizes the average contributions of all variables regarding a time scope from the detecting time to the end of the simulation, which is likewise in the following experiments. From the result, it can be concluded that the contributions of $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_5$ and $\boldsymbol{x}_6$ all exceed the relative threshold 1, so it is hard to identify the ultimate faulty source. The real time diagnosis utilizing ARCR for this fault is given in Fig. 6, where "o" is marked as an effective diagnosis, and different values of θ are included in the simulation. $\theta = 0.4$ in the first subfigure is chosen empirically, whereas $\theta = 0.26$ is given according to Subsection 2.4. From the two subfigures, it is obvious that the new real-time diagnosis scheme can successfully locate the faulty origination, namely $\boldsymbol{x}_2$. Meanwhile, it can be found that ARCR with an appropriate θ is able to deal with the smearing effect caused by the transmission of the fault.

Fault II, an incipient fault, is triggered by a slow ramp drift T_f ($\boldsymbol{x}_1$). As can be seen from Fig. 7, there is a time-delay for the detection, since it is successfully detected from the 589th sample. This phenomenon is caused by the dynamics of CSTR process and the nature of incipient fault. The online diagnosis starts from the 589th sample that is presented in Fig. 9, where the new scheme identifies $\boldsymbol{x}_1$ as the origin of this fault straightforwardly. By contrast, PCA based contribution plot is sketched in Fig. 8. With the confusing result, it is hard to determine which component is truly responsible for this fault. Indeed, the performances could be explained according to the nature of both methods. When a change happens in a nonlinear system like CSTR, it will smear to other components from the occurrence point. The traditional contribution plot measures each variable′s contribution to the detection index at a time point, thus, with the smearing effect, other variables will be wrongly labeled as the fault source. Instead, contribution rate accounts for the fault origin by analyzing each variable′s sensitivity to the the detection index, which could

successfully locate the faulty component.

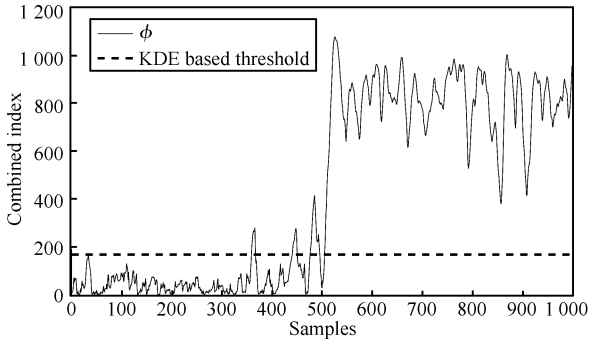


Fig. 4 Detection result of Fault I with index φ

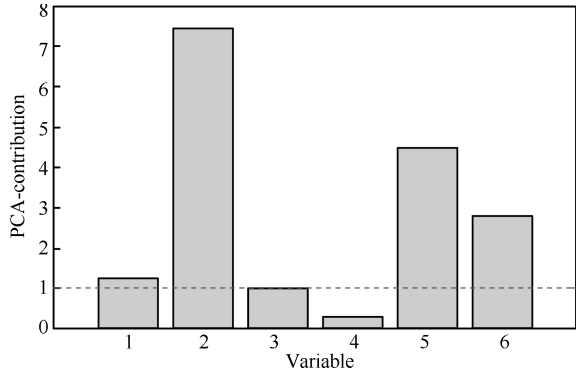


Fig. 5 PCA based diagnosis for Fault I

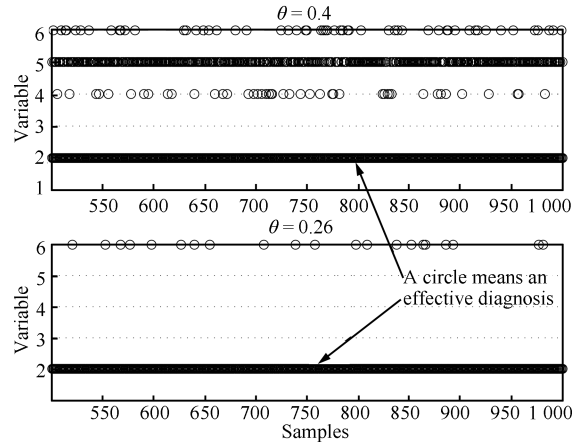


Fig. 6 Diagnosis result of Fault I with ARCR

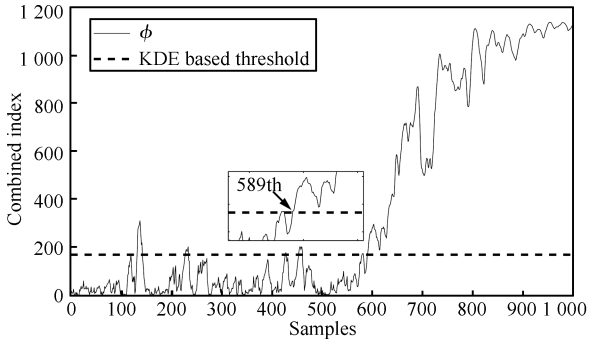


Fig. 7 Detection result of Fault II with index φ

Fault III is random disturbance that occurs in T_f. In general, the mathematical description of this fault resembles a declination in precision of sensor 1 physically, and it may result in severe disturbances in other process variables, of which T_c, namely $\boldsymbol{x}_5$ is intuitively included due to the function of the closed-loop control. φ detects this fault immediately in Fig. 10. PCA based approach picks out $\boldsymbol{x}_1$, while $\boldsymbol{x}_5$ and $\boldsymbol{x}_6$ contribute a lot as well. As for ARCR based diagnosis in Fig. 12, it isolates $\boldsymbol{x}_1$ directly. Different from Fault II, $\boldsymbol{x}_4$ is also isolated subsequently in this fault, that is because the abnormal variations occurring in $\boldsymbol{x}_1$ cause significant threat to $\boldsymbol{x}_4$ that cannot be resisted by ARCR. However, the result from another side demonstrates that ARCR can not only find the faulty source, but also point out the fault sequence as well.

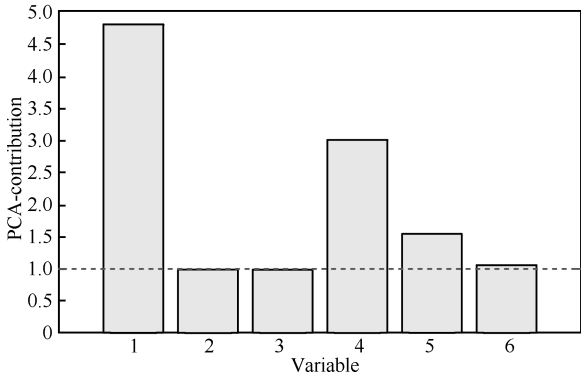


Fig. 8 PCA based diagnosis for Fault II

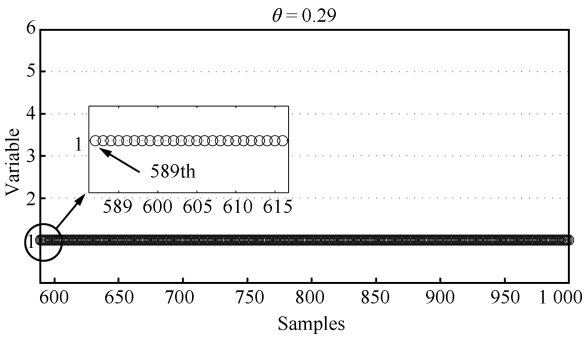


Fig. 9 Diagnosis result of Fault II with ARCR

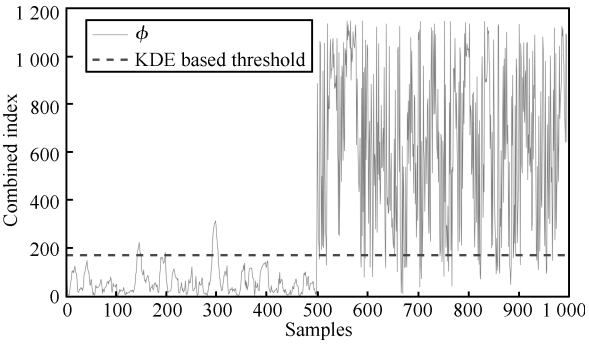


Fig. 10 Detection result of Fault III with index φ

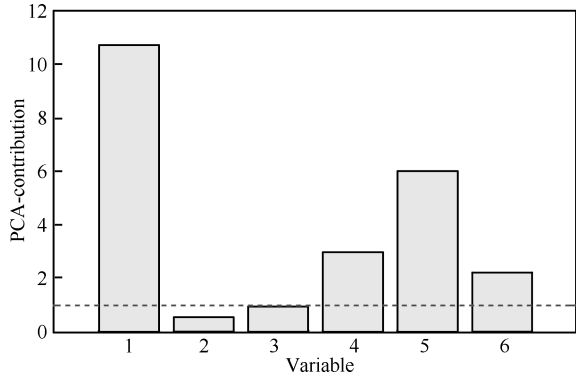


Fig. 11 PCA based diagnosis for Fault III

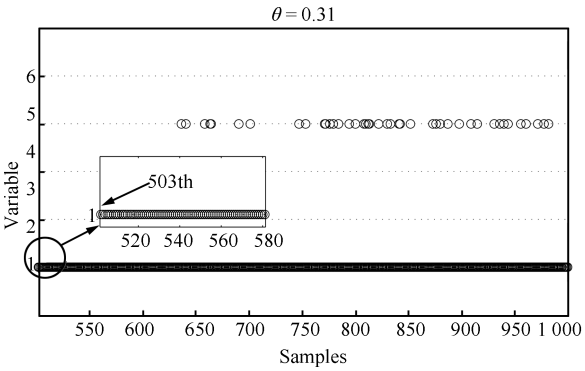


Fig. 12 Diagnosis result of Fault III with ARCR

4 Conclusion

In this paper, KPCA model and KPCA based process monitoring are revisited at first. Then a contribution rate based diagnosis scheme is established to address the difficulties existing in KPCA based diagnosis area which successfully overcomes the demerits introduced by kernel function. Instead of focusing on the contribution to the detection index as shown in linear models, contribution rate primarily concerns with how to measure the variable′s influence to the increment of the fault detection. From this viewpoint, the present approach is derived physically and theoretically. Meanwhile, a more effective method based on contribution rate called accumulative relative contribution rate is also developed from the practical motivation. In the simulation, a case study on CSTR demonstrates the effectiveness of the contribution rate based strategy. To conclude, the present scheme is powerful enough to address the online nonlinear fault diagnosis issue. It also provides a universal paradigm to handle other kernel-relevant data-driven fault diagnosis problem.

Due to the scope of this paper, some other aspects such as dynamics, non-Gaussianity and time-varying characteristic etc. are not considered for industrial processes. These fields will be considered in the future work.

References

1 Qin S J. Statistical process monitoring: basics and beyond. *Journal of Chemometrics*, 2003, **17**(8−9): 480−512

2 Yin S, Ding S X, Abandan Sari H A, Hao H Y, Zhang P Y. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. *Journal of Process Control*, 2012, **22**(9): 1567−1581

3 Yin S, Ding S X, Haghani A, Hao H Y. Data-driven monitoring for stochastic systems and its application on batch process. *International Journal of Systems Science*, 2013, **44**(7): 1366−1376

4 Liu J L. Fault diagnosis using contribution plots without smearing effect on non-faulty variables. *Journal of Process Control*, 2012, **22**(9): 1609−1623

5 Zhou Dong-Hua, Wei Mu-Heng, Si Xiao-Sheng. A survey on anomaly detection, life prediction and maintenance decision for industrial processes. *Acta Automatica Sinica*, 2013, **39**(6): 711−722 (in Chinese)

6 Qin S J, Zheng Y Y. Quality-relevant and process-relevant fault monitoring with concurrent projection to latent structures. *AIChE Journal*, 2013, **59**(2): 496−504

7 Cho J H, Lee J M, Choi S W, Lee D K, Lee I B. Fault identification for process monitoring using kernel principal component analysis. *Chemical Engineering Science*, 2005, **60**(1): 279−288

8 Choi S W, Morris J, Lee I B. Nonlinear multiscale modelling for fault detection and identification. *Chemical Engineering Science*, 2008, **63**(8): 2252−2266

9 Rakotomamonjy A. Variable selection using SVM-based criteria. *Journal of Machine Learning Research*, 2003, **3**: 1357−1370

10 Choi S W, Lee C Y, Lee J M, Park J H, Lee I B. Fault detection and identification of nonlinear processes based on kernel PCA. *Chemometrics and Intelligent Laboratory Systems*, 2005, **75**(1): 55−67

11 Alcala C F, Qin S J. Reconstruction-based contribution for process monitoring with kernel principal component analysis. *Industrial & Engineering Chemistry Research*, 2010, **49**(17): 7849−7857

12 Cremers D, Kohlberger T, Schnörr C. Shape statistics in kernel space for variational image segmentation. *Pattern Recognition*, 2003, **36**(9): 1929−1943

13 Ding M T, Tian Z, Xu H X. Adaptive kernel principal component analysis. *Signal Processing*, 2010, **90**(5): 1542−1553

14 Nguyen V H, Golinval J C. Fault detection based on kernel principal component analysis. *Engineering Structures*, 2010, **32**(11): 3683−3691

15 Choi S W, Lee I B. Nonlinear dynamic process monitoring based on dynamic kernel PCA. *Chemical Engineering Science*, 2004, **59**(24): 5897−5908

16 Liu X Q, Kruger U, Littler T, Xie L, Wang S Q. Moving window kernel PCA for adaptive monitoring of nonlinear processes. *Chemometrics and Intelligent Laboratory Systems*, 2009, **96**(2): 132−143

17 Zhang Y W, Li S, Hu Z Y, Song C H. Dynamical process monitoring using dynamical hierarchical kernel partial least squares. *Chemometrics and Intelligent Laboratory Systems*, 2012, **118**: 150−158

18 Alcala C F, Qin S J. Reconstruction-based contribution for process monitoring. *Automatica*, 2009, **45**(7): 1593−1600

19 Li G, Qin S J, Ji Y D, Zhou D H. Reconstruction based fault prognosis for continuous processes. *Control Engineering Practice*, 2010, **18**(10): 1211−1219

20 Mika S, Schölkopf B, Smola A, Müller K R, Scholz M, Rätsch G. Kernel PCA and de-noising in feature spaces. In: Proceedings of the 1998 Conference on Advances in Neural Information Processing Systems II. Cambridge, MA, USA: MIT Press, 1999. 536−542

21 Schölkopf B, Mika S, Burges C J C, Knirsch P, Muller K, Ratsch G, Smola A J. Input space versus feature space in kernel-based methods. *IEEE Transactions on Neural Networks*, 1999, **10**(5): 1000−1017

22 Müler K R, Mika S, Räsch G, Tsuda K, Schökopf B. An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks*, 2001, **12**(2): 181−201

23 Peng K X, Zhang K, Li G, Zhou D H. Contribution rate plot for nonlinear quality-related fault diagnosis with application to the hot strip mill process. *Control Engineering Practice*, 2013, **21**(4): 360−369

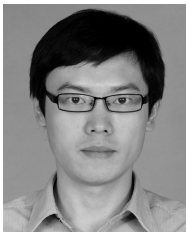
24 Zhang Y W, Zhou H, Qin S J. Decentralized fault diagnosis of large-scale processes using multiblock kernel principal component analysis. *Acta Automatica Sinica*, 2010, **36**(4): 593−597

25 Baffi G, Martin E B, Morris A J. Non-linear projection to latent structures revisited: the quadratic PLS algorithm. *Computers and Chemical Engineering*, 1999, **23**(3): 395−411

26 Kruger U, Wang X, Chen Q, Qin S J. An alternative PLS algorithm for the monitoring of industrial process. In: Proceedings of the 2001 American Control Conference. Arlington, VA, USA: IEEE, 2001. 4455−4459

27 Li G, Qin S Z, Ji Y D, Zhou D H. Total PLS based contribution plots for fault diagnosis. *Acta Automatica Sinica*, 2009, **35**(6): 759−765

PENG Kai-Xiang Professor at the School of Automation and Electrical Engineering, University of Science and Technology Beijing. He received his bachelor, master, and Ph. D. degrees from University of Science and Technology Beijing, in 1995, 2002, and 2007, respectively. From January to July in 2012, he visited University of Duisburg-Essen as a visiting scholar at the Institute of Automatic Control and Complex Systems. His research interest covers statistical process monitoring and fault diagnosis, process modeling and control for the iron and steel industry. E-mail: kaixiang@ustb.edu.cn



ZHANG Kai Ph. D. candidate at the Institute of Automatic Control and Complex Systems, University of Duisburg-Essen. He received his master degree from University of Science and Technology Beijing in 2012. His research interest covers process monitoring, fault diagnosis and prediction for industrial processes. Corresponding author of this paper.
E-mail: kai.zhang@uni-due.de

LI Gang Provisional research assistant in the Department of Automation, Tsinghua University. He received his bachelor degree from the Department of Precision Instruments and Mechanology, Tsinghua University in 2004, master and Ph. D. degrees from the Department of Automation, Tsinghua University in 2008 and 2011, respectively. His research interest covers dynamic and nonlinear modeling with multivariate statistical techniques, soft sensing and quality prediction with partial least squares, and other issues in data-driven modeling, monitoring and diagnosis for industrial processes.
E-mail: dlg00@mails.tsinghua.edu.cn