Neurocomputing 202 (2016) 27–35

Contents lists available at ScienceDirect

Neurocomputing

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

Differential feature based hierarchical PCA fault detection method for dynamic fault ☆

Funa Zhou [a,b], Ju H. Park [b,*], Yajuan Liu [b]

[a] School of Computer and Information Engineering, Henan University, Kaifeng 475004, PR China
[b] Department of Electrical Engineering, Yeungnam University, 280 Daehak-Ro, Kyongsan 38541, Republic of Korea

ARTICLE INFO

Article history:
Received 5 November 2015
Received in revised form
24 December 2015
Accepted 12 March 2016
Communicated by Yang Tang
Available online 7 April 2016

Keywords:
Dynamic fault
Hierarchical PCA
Differential feature
Zero cross point

ABSTRACT

By sensor accuracy degradation or unwanted alternating current signals, sensor fault with zero cross point (ZCP) may occur in real systems and conventional data-driven fault detection methods could be invalid. In this regard, this paper proposes a hierarchical principal component analysis (PCA) fault detection method based on the differential features of dynamic faults to detect the fault with ZCPs. The main contribution of this work are as follows: (1) A new differential based feature extraction method is first proposed to well character the dynamic trend of the observation; (2) then, a hierarchical detection criterion is proposed according to the detection ability of each round of PCA anomaly detection; (3) it is convenient to extend the proposed method to other statistical based fault detection techniques whose detection criteria are also a distance defined by fault amplitude.

1. Introduction

At present, large scale systems are more popular in industry field. Since all parts of the large scale system are closely related, failures in some components may cause breakdown of the system or even lead to major accident, accuracy fault diagnosis and control of these systems is a challenging problem [1–7]. In this regard, research on fault diagnosis has received wide attention from experts both in academic and application [1–6,8–14]. In general, existing fault diagnosis method can be categorized into 3 classes: physical model based method, knowledge based method, and data-driven based method. Since the precise physical model is unavailable and quantity of prior knowledge is difficult to process which would invalidate the model based fault diagnosis methods and the knowledge based fault diagnosis methods, the data-driven techniques are widely applied in industry for process monitoring and fault diagnosis [9–14].

The complete fault diagnosis process includes 4 stages: fault detection, fault classification, fault location, and fault tolerant control [9–11]. Principal component analysis (PCA), Kullback–Leibler divergence, fisher discriminant analysis (FDA), support vector machine (SVM), and artificial neural network (ANN) are the popular data-driven techniques for fault diagnosis [15–22].

Multivariate statistical analysis based methods are the most common used data-driven method for faut diagnosis. Among these methods, PCA is a representative one. PCA based fault diagnosis methods and their variations are efficient for fault detection [16–29]. As is known to all, PCA is an efficient cross correlation feature extraction tool. PCA decomposition can be used to model a multivariate statistical model of the system concerned. When online multivariate statistical model is established, the residual characterizing difference between the historical normal model and the online model has been used to set up a fault detection model in the works [17–29]. In the last decades, variations of PCA are developed to make the detection model applicable to different situations. Cumulative sum technique based PCA (CUSUMPCA) is more adequate for small fault detection [16]. Dynamic PCA (DPCA) and multiple PCA are proper for fault detection when the observation at different sample time is not independent identically distributed (i.i.d.) [18,24]. Multi-way PCA (MPCA) is more suitable for fault detection of batch process [28]. Multi-scale PCA (MSPCA) can well detect faults occurring on different resolution since it combines PCA with multi-scale resolution analysis [21]. If the observation variables are nonlinearly correlated, kernel PCA (KPCA) can be used to establish the statistical monitoring model [22]. When statistical distribution of the observation is not normal distribution, independent component analysis (ICA) is an alternative method to extract statistical feature

☆This research was supported in part by the Natural Science Fund of China (Grant nos. 61174112 and 61203094) and Technical Innovation Talents Scheme of Henan Province (Grant no. 2012HASTIT005).

* Corresponding author.
E-mail addresses: zhoufn2002@163.com (F. Zhou), jessie@ynu.ac.kr (J.H. Park).

http://dx.doi.org/10.1016/j.neucom.2016.03.007
0925-2312/

of the observation [19]. When quality variable and process variable are processed separately, partial least square (PLS) based statistical analysis can be used to detect fault affecting quality variable. Compared with PCA, PLS is more adequate for quality prediction in most cases [25,27,29]. But, all the PCA based fault detection methods and their variations mentioned above share the common fact that they use only the magnitude of the observation to compute the difference between the historical normal data and the online data. Thus, only static feature of the fault is considered during the designing of the fault detection principle,and the dynamic character of the fault is ignored.

Machine learning based methods are also used as a data driven technique for faut diagnosis. ANN and SVM based fault diagnosis methods can diagnose fault occurring in the system by establishing a mapping modeling between fault source and symptom [15,30,32–34]. Just like FDA based fault diagnosis method, ANN and SVM based fault diagnosis methods and their variations are also optimal in terms of fault classification [15,30–35]. But, these methods are also invalid when dynamic fault occurs in the system due to the same shortcoming as PCA based methods. In addition, large amount of fault data for training are required for FDA, ANN and SVM based fault classifying methods. In the case when a large number of fault data is unavailable, timely fault detection using PCA is more preferable since fault detection is the first stage of fault diagnosis.

It can be seen from above that PCA based fault detection is effective when there is no fault data available. However, it is valid only in the case when there is nonzero deviation between the normal observation and the online observation. If fault signal occurring in the system has zero cross points (ZCPs) and the 1-order differential or high-order differential is not zero, for example,sine fault at $k\pi$, traditional PCA based fault detection method results in high missing detection rate. Since precise online detection is critical to the subsequent fault tolerant control, it is necessary to find a way to eliminate or decrease the miss detection in the case when dynamic fault occurs in the system.

In real world, sensor accuracy degradation is inevitable, which results in faults with many ZCPs having increased variances. On the other hand, if sensors' power supplies or other components are affected by unwanted alternating electrical currents, or vibrating signals, sensor faults with ZCPs may occur in systems. Hence, the development of an efficient data-driven fault detection method is of much significance.

In order to resolve the limitation stated above, a differential feature based hierarchical PCA (DFHPCA) fault detection method is developed in this paper to decrease miss detection rate when dynamic fault occurs in the system. State of the art of this research is to design a detection criterion by combining the differential information characterizing the dynamic feature. By this means, fault with many ZCPs can be well detected which is significant to fault tolerant control following fault detection. Also, by substituting PCA into other variations of PCA or PLS, it is convenient to extend our method to other statistical based fault detection techniques whose detection criteria are also a distance defined by fault amplitude.

The remainder of this paper is organized as follows: Section 2 reviews the PCA based fault detection methods. In Section 3, the dynamic fault detection problem is formulated, and the PCA fault detection ability for dynamic fault is analyzed in detail. Then, a new DFHPCA fault detection method is originally developed in Section 4. In Section 5, numerical simulation and application to Tennessee Eastman (TE) process illustrate the efficiency of the algorithm proposed. Section 6 is the conclusion and future work of this paper.

2. Review of PCA based fault detection

Denote $Y_0 \in \mathcal{R}^{n\times m}$ as the historical normal observation data of a MIMO dynamic system operated in the steady state. m and n are the number of observation variable and samples, respectively. And the observation is sampled with sampling interval T_s. $Y_0(j,k) = y_{0j}(k)$ is the kth sample of the jth observation variable. $Y_0(k) = [y_{01}(k), y_{02}(k), \ldots, y_{0m}(k)]$ is the observation at sample time k.

PCA decomposition of Y_0 can be formulated by the following equation:

$$Y_0 = T_0 P^T + E_0 \tag{1}$$

where $P = [P_1, P_2, \ldots, P_v] \in \mathcal{R}^{m\times v}$ is the loading matrix constructed by the first v loading vectors P_i $(i = 1, 2, \ldots, m)$ which are the eigenvectors of $\Sigma = \frac{1}{n-1} Y_0^T Y_0$, and v is the number of key principal components. $T_0 \in \mathcal{R}^{n\times v}$ is the score matrix, and $E_0 = \sum_{j=v+1}^{m} T_{0j} P_j^T$ is the residual matrix.

For online fault detection, the statistics T^2 and square prediction error (SPE) are constructed in principal component space and residual space, respectively:

$$T^2 = YP\Lambda^{-1}P^TY^T, \tag{2}$$

$$SPE = \| E \|^2 = Y(I - PP^T)Y^T \tag{3}$$

where Λ is a diagonal matrix whose diagonal elements are the eigenvalue of Σ.

It is noted that T^2 and SPE are more sensitive to faults occurred in principal component space and residual space, respectively [3]. The control limit used for T^2 to fault detection, denoted as UCL, can be determined by

$$UCL = \frac{v(n^2-1)}{n(n-v)} F_\alpha(v, n-v) \tag{4}$$

where $F_\alpha(v, n-v)$ is the α quantile of $F-$distribution with parameter v and $n-v$, and α is the confidence limit.

The control limit used for SPE to fault detection, denoted as Q, can be determined by

$$Q = \theta_1 \left[\frac{h_0 C_\alpha \sqrt{2\theta_2}}{\theta_1} + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} + 1 \right]^{\frac{1}{h_0}} \tag{5}$$

where

$$\theta_j = \sum_{i=v+1}^{p} \lambda_i^j, \quad (j = 1, 2, 3), \tag{6}$$

$$h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}, \tag{7}$$

C_α is the α-quantile of normal distribution $N(0,1)$, and λ_i is the eigenvalue of Σ.

False detection rate (FDR) and miss detection rate (MDR) are commonly used as quantity indexes to assess the efficiency of different detection methods as defined below:

$$FDR = \frac{T_{fault}}{N_{normal}}, \tag{8}$$

$$MDR = \frac{T_{normal}}{N_{fault}}. \tag{9}$$

where T_{fault} is the number of sample points detected as fault, N_{normal} is the true number of normal sample points, T_{normal} is the number of sample points detected as normal, and N_{fault} is the true number of faulty sample points, respectively.

Here, it should be pointed out that conventional PCA based fault detection methods and their variations can only well detect faults with no ZCPs. When a dynamic fault with ZCP occurs in the system, these existing methods inevitably lead to high miss detection rate since they are invalid to detect those ZCPs of the dynamic fault since only amplitude of the fault used to design a fault detection criterion defined by a Mahalanobis distance.

3. Dynamic fault detection ability for PCA based method

Assuming that the dynamic fault is the solution of the following equation:

$$\dot{f}(t) = A_f f(t) + B_f u_f(t). \tag{10}$$

Without loss of generality, $u_f(t) = u_0 1(t)$ can be selected as a step input. The real part of the eigenvalue of A_f is denoted as λ. If $\lambda < 0$, the fault $f(t)$ is evolved as a constant fault $f(t) = f$ in the steady state, in which case the traditional PCA based fault detection methods can be used. If $\lambda = 0$, the fault has the form of $f(t) = \sin(\omega t)$, in which case $f(\frac{k\pi}{\omega}) = 0$ but the differential $\dot{f}(\frac{k\pi}{\omega}) \neq 0$, the traditional PCA based fault detection methods result in high miss detection rate. If $\lambda > 0$, the fault has the form of $f(t) = f + e^{\lambda t} \sin(\omega t)$, which is a divergent sine fault, and it cannot be detected by traditional PCA based fault detection methods either.

Fault similar to $f(t) = \sin(\omega t)$ is firstly considered in the following part of this paper. The continuous fault $f(t)$ is sampled by sampling interval T_s to get the sample $f(k)$. Its sampling interval T_s is the same as that of the observation.

In the case of $m = 1$, that is, single observation variable y is involved in the system, if a sine fault $f(k)$ occurs in the system, then the observation is of the form:

$$y(k) = y_0(k) + f(k) \tag{11}$$

where $y_0(k)$ and $y(k)$ are the normal observation and the online observation at time k, respectively. In this case, $3-\sigma$ criterion for univariate statistical fault detection principle can be used [3].

If $f(k) = 0$ as shown in point '**a**' in Fig. 1, then $y(k) = y_0(k)$ is localized in the $3-\sigma$ normal region. The detection result indicates that the system is normal at time k. In Fig. 1, the red and blue lines denote the fault signal $f(k) = \sin(10k)$, and $\dot{f}(k) = 10\cos(10k)$, respectively.

On the other hand, point '**c**' in Fig. 1 ($|\dot{f}(k)| \geqslant 0$), indicates that a fault has occurred in the system. So we are motivated to incorporate differential feature characterizing the dynamics of the fault into the designing of fault detection principle.

In the case $m > 1$, PCA can be used as a tool to extract cross correlation of the observation variables. Denoting the online observation involved fault $F(k)$ as below:

$$Y(k) = Y_0(k) + F(k) \tag{12}$$

where $Y(k)$ is the online observation at time k.

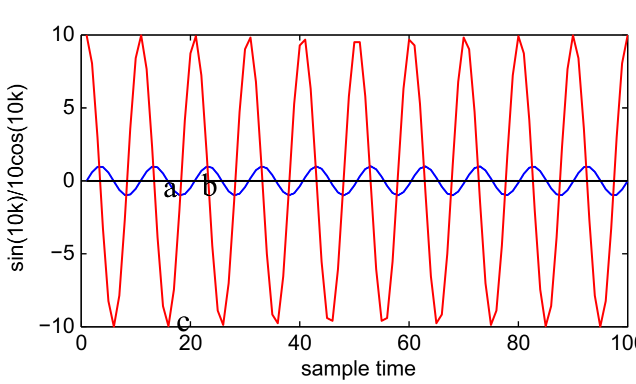

Fig. 1. Sine fault.

Once online observation $Y(k)$ at time k is obtained, $SPE(k)$ of PCA model can be used as an index to describe the deviation between $Y_0(k)$ and $Y(k)$ in the statistical sense. The online SPE value can be computed as

$$SPE(k) = \| E(k) \|^2 = [Y_0(k) + F(k)](I - PP^T)\left[Y_0(k) + F(k)\right]^T. \tag{13}$$

If $F(k) = 0$, then $Y(k) = Y_0(k)$, and the online $SPE(k)$ is equal to $SPE_0(k) = Y_0(k)(I - PP^T)Y_0^T(k)$, which is located in the acceptance region of the hypothesis test. This detection result indicates that the system is normal at time k which is not consistent with the true fact. Naturally, this leads to miss detection at time k.

On the other hand, if 1-order differential of the fault signal $S_F(k) \neq 0$, we can implement PCA fault detection to the online 1-order differential $S(k) = S_0(k) + S_F(k)$, where $S_0(k)$ is the 1-order differential of historical normal observation. The detection result may indicate that the system is abnormal at time k.

If $S_F(k) = 0$ as shown in point '**b**' in Fig. 1, implement PCA fault detection to the 2-order differential $C(k) = C_0(k) + C_F(k)$. The detection result may indicates that the system is abnormal at time k, where $C(k)$, $C_0(k)$ and $C_F(k)$ are the 2-order differential of the online observation, the historical normal observation and the fault signal, respectively.

Remark 1. If and only if 1-order differential and high-order differential as well as the fault signal are all equal to 0, the system can be really a normal system.

4. Differential feature based hierarchical PCA fault detection method

Motivated by the analysis in Section 3, the DFHPCA fault detection method is presented in this section. The complete fault detection algorithm is as follows:

4.1. Offline modeling

Step 1: Compute the 1-order differential $S_0(k)$ and the 2-order differential $C_0(k)$ of the historical normal observation at time k by the following equations:

$$S_0(k) = [s_{01}(k)\ s_{02}(k)\ \cdots s_{0m}(k)], \tag{14}$$

$$C_0(k) = [c_{01}(k)\ c_{02}(k)\ \cdots c_{0m}(k)], \tag{15}$$

where

$$s_{0j}(k) = \frac{y_{0j}(k+1) - y_{0j}(k)}{T_s}, \tag{16}$$

$$c_{0j}(k) = \frac{s_{0j}(k+1) - s_{0j}(k)}{T_s}. \tag{17}$$

Step 2: As given in Eqs. (16)–(18), establish 3 PCA fault detection models to the normal historical Y_0, S_0 and C_0, respectively:

$$Y_0 = T_0 P^T + E_0, \tag{18}$$

$$S_0 = T_0^S P^{S^T} + E_0^S, \tag{19}$$

$$C_0 = T_0^C P^{C^T} + E_0^C, \tag{20}$$

where the superscript S and C represent the score and loading matrices corresponding to the 1-order differential and the 2-order differential, respectively. Then, the corresponding control limit of SPE statistics Q, Q^S and Q^C can also be computed via Eq. (5).

4.2. Online detection

For online detection, the DFHPCA based detection method is used to detect fault occurs at time k.

Step 3: Implement the first round of PCA fault detection to $Y(k)$.

Once online observation $Y(k)=[y_1(k), y_2(k), \ldots, y_m(k)]$ at time k is obtained, project $Y(k)$ to the PCA model established in Eq. (18), and compute the online SPE value by

$$SPE(k) = \|E(k)\|^2 = Y(k)(I-PP^T)Y^T(k). \tag{21}$$

Then, uniformization of the online SPE value by the control limit Q using Eq. (22) can result in the fact that $\tilde{SPE}(k)$ can be used to define a detection ability index:

$$\tilde{SPE}(k) = \frac{SPE(k)}{Q}, \tag{22}$$

$$\tilde{Q} = 1, \tag{23}$$

where $\tilde{Q} = \frac{Q}{Q}$ is the normalized control limit. The detection ability index can be defined as below:

$$\delta(k) = \left|\tilde{SPE}(k) - 1\right|. \tag{24}$$

Remark 2. As will be given in Section 5, we can roughly determine the beginning time of fault occurring at time k by the following modified fault detection criterion:

(1) $\tilde{SPE}(k) > \tilde{Q}$,
(2) the subsequent several $\tilde{SPE}(k) > \tilde{Q}$.

In general, after the first round of PCA to $Y(k)$, we can determine the beginning time of fault occurring at time k. But, miss detection rate may be very high since it cannot detect the fault whose amplitude is equal to zero. On the other hand, 1-order differential information is not adequately used

Step 4: For those miss detected $Y(k)$ or those $Y(k)$ with small detection ability index, we can wait for the 1-order differential information involved in $Y(k+1)$.

Once $Y(k+1)$ is obtained, compute the 1-order differential $S(k)=[s_1(k), s_2(k), \ldots, s_m(k)]$ by

$$s_j(k) = \frac{y_j(k+1) - y_j(k)}{T_s}. \tag{25}$$

Step 5: Implement the 2nd round of PCA fault detection to $S(k)$, whose detection ability is weak in Step 3.

Project $S(k)$ to the PCA model given in (17), compute the online SPE value by

$$SPE^S(k) = \|E^S(k)\|^2 = S(k)(I - P^S P^{S^T})S^T(k). \tag{26}$$

Then, uniformization of the online SPE value by the control limit Q^S can be computed as

$$\tilde{SPE}^S(k) = \frac{SPE^S(k)}{Q^S}. \tag{27}$$

After the second round of PCA to $S(k)$, miss detection rate may be still a little high since it cannot well detect the fault whose detection ability index is small in Step 5.

Step 6: Wait to compute the 2-order differential $C(k)=[c_1(k), c_2(k), \ldots, c_m(k)]$ via Eq. (25) until $Y(k+2)$ is obtained:

$$c_j(k) = \frac{s_j(k+1) - s_j(k)}{T_s}. \tag{28}$$

Step 7: Implement the 3rd round of PCA fault detection to $C(k)$, whose detection ability is weak in Step 5.

Project $C(k)$ to the PCA model given in (18), compute the online SPE value by

$$SPE^C(k) = \|E^C(k)\|^2 = C(k)(I - P^C P^{C^T})C^T(k). \tag{29}$$

Then, uniformization of the online SPE value by the control limit Q^C can be computed as

$$\tilde{SPE}^C(k) = \frac{SPE^C(k)}{Q^C}. \tag{30}$$

After the third round of PCA to $C(k)$, sine fault occurring at time k can be well detected. If fault signal is a complex function, higher-order differential can be used in a similar hierarchical way.

Flow chart of DFHPCA fault detection algorithm is illustrated in Fig. 2.

Remark 3. During the hierarchical detection process, whether the detection result should be updated in a new round of differential based PCA fault detection is decided by whether the detection ability index in a new round is much larger or not. The detection ability is defined as the absolute deviation of the normalized detection statistics value and the control limit as in Eq. (24).

5. Simulation and case study

In this section, we first test the DFHPCA fault detection algorithm with MATLAB numerical simulation data. Then, case study of TE process is also illustrated. It should be noted that sensor fault rather than system fault is considered in our research.

Remark 4. With the assumption that each sample of an observation variable in the steady state of a dynamic system can be i.i.d., random function "*randn*" in MATLAB is used to generate the observation. The number '−2' and '1' in Eqs. (31) and (32) mean that the steady value of y_1 and y_2 are -2 and 1, respectively.

5.1. Numerical simulation study

Considering the case that $m_i=4$ and $m_o=1$ where m_i is the number of process or manipulated variable and m_o is the number of quality variable, $m=m_i+m_o$ is the observation variables involved in a system, and 200 samples of each variable ($n=200$) are collected at a sample interval $h=0.0125$ h, then generate the normal observation by the random function" *randn*" as below:

$$y_{01} = randn(n,1) - 2 \tag{31}$$

$$y_{02} = 1.5\ randn(n,1) + 1 \tag{32}$$

$$y_{03} = 0.2y_{01} + 0.8y_{02} + 0.5\ randn(n,1) \tag{33}$$

$$y_{04} = (\sqrt{2}/2)y_{01} - (\sqrt{2}/2)y_{02} + 0.5\ randn(n,1) \tag{34}$$

$$z_0 = 0.1y_{01} + 0.4y_{02} + 0.2y_{03} + 0.3y_{04} \tag{35}$$

where y_{01}, y_{02}, y_{03} and y_{04} are the process variables, and z_0 is the corresponding quality variable. The mean of the 1st observation variable is -2 and the noise variance of the 1st variable is about 1. The mean of the 2nd observation variable is 1 and the noise variance of the 2nd variable is about 1.5. Three kinds of fault observation are generated for comparison purpose to show the detection efficiency of our algorithm for detecting faults with different amplitude and number of ZCPs. Online observation y_1, y_2, y_3 and y_4 are first generated by Eqs. (31)–(34). For fault Case 1, a sine fault $f_1(k) = 6\sin(100k)$ is added in y_2 from sample time $k=51$ to $k=200$. For fault Case 2, a sine fault $f_2(k) = 4\sin(50k)$ is added in y_2 from sample time $k=51$ to $k=200$. For fault Case 3, a variance increasing fault corresponding to the 2nd sensor's measurable

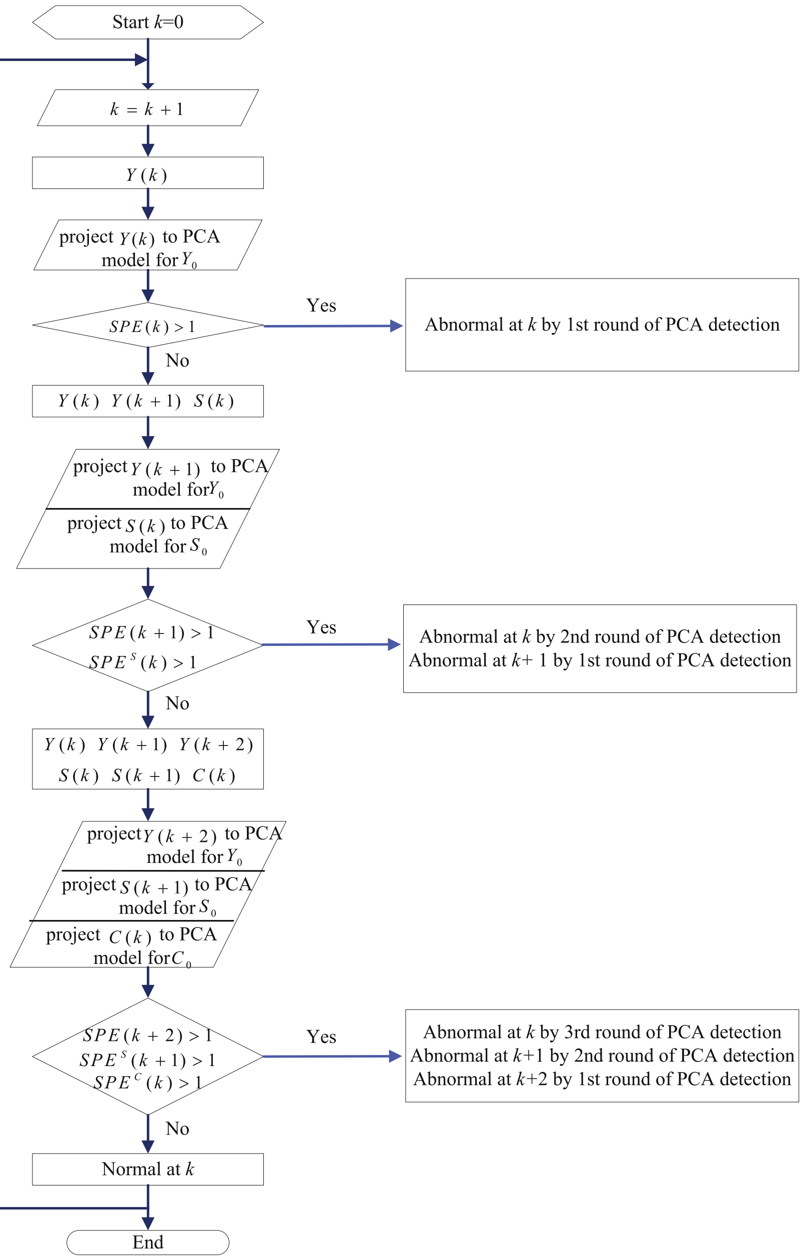

Fig. 2. Flowchart of DFHPCA fault detection algorithm.

precision degradation has occurred from sample time k=51 to k=200. Then, z_0 is generated by Eq. (35). Fault Case 1—fault Case 3 correspond that the 2nd sensor's power supply is affected by an unwanted alternating current power supply with different frequencies and amplitudes.

SPE value of traditional PCA based fault detection for fault case 1 with $f = 6\sin(100\ k)$ is displayed in Fig. 3. The solid line denotes the online SPE value, and the dotted line denotes the control limit. It can be seen from Fig. 3 that the system is abnormal from sample time 51 by using the modified fault detection criterion mentioned in Section 4. On the other hand, it can be obviously seen that missing detection rate is very high since traditional PCA method is invalid for detecting faults having more ZCPs.

The SPE value of DFHPCA fault detection method is displayed in Fig. 4. The blue solid line denotes online SPE value and the red dotted line denotes the control limit. Compared with Fig. 3, it can be seen from Fig. 4 that in the case when a fault is really detectable in the sense that the maximum amplitude of the fault 6 is larger than 4.5 (3 times of the variance of observation noise), miss detection rate can be significantly reduced since dynamic character of the observation is extracted to detect all possible dynamic fault occurring in the system. It can also be concluded that the detection ability is increased due to the fact that the amplitude of fault's differential is larger than the original fault signal which can be seen in Fig. 1.

Remark 5. The miss detection rate is affected by the frequency and duration time of the sine fault occurring in the system. And the simulation experiment is operated in the case that signal-to-noise rate SNR of fault signal is large enough to secure that the fault can be detected. In the case that SNR of fault signal is small, or “small” sine fault occurs in the system, necessary denoising preprocessing is required to increase the SNR value.

Fault Case 2 means that the maximum amplitude of a sine fault is 4 which is smaller than 4.5 (3 times of the noise variance) and the frequency of the sine fault is 50 Hz, which implies that there

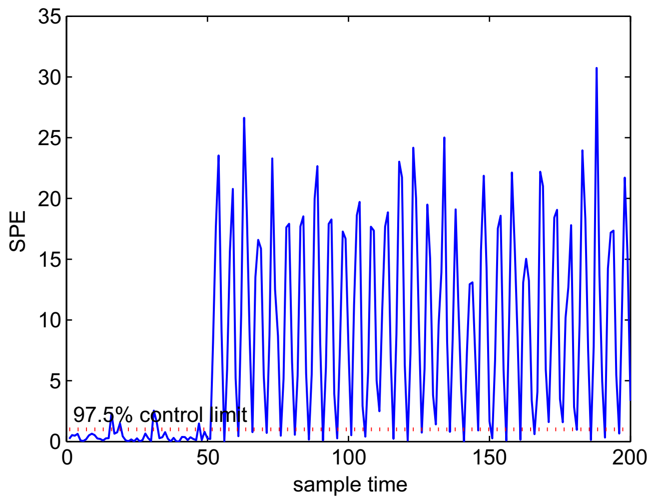

Fig. 3. Fault Case 1 SPE for traditional PCA fault detection.

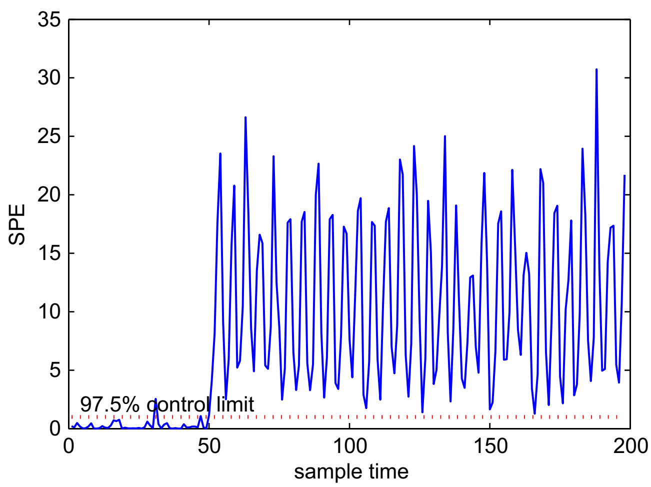

Fig. 4. Fault Case 1 SPE for DFHPCA fault detection.

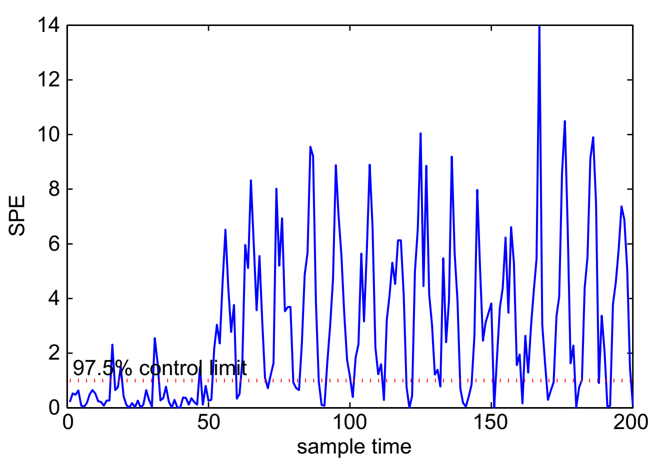

Fig. 5. Fault Case 2 SPE for traditional PCA fault detection.

are less ZCPs than fault Case 1. The SPE value of PCA and DFHPCA is displayed Figs. 5 and 6, respectively.

Comparing Fig. 6 with Fig. 5, it can be concluded that, even in the case when the fault size is small, incorporating differential feature into the design of a fault detection criterion characterized by a Mahalanobis distance defined in Eq. (4) can reduce the missing detection rate as well as the false detection rate.

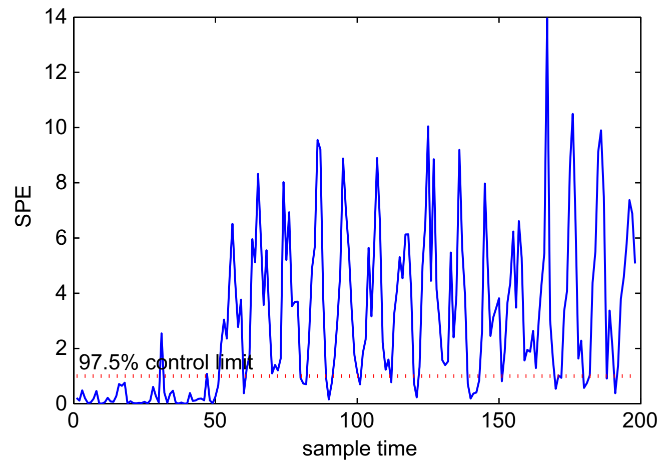

Fig. 6. Fault Case 2 SPE for DFHPCA fault detection.

Table 1
PCA/DFHPCA False and miss detection rate.

Fault type	Fault size	Method used	FDR (%)	MDR (%)
Fault Case 1	Fault amplitude $>3\sigma_n$	PCA	10	17.33
		DFHPCA	6	0
Fault Case 2	Fault amplitude $<3\sigma_n$	PCA	10	36
		DFHPCA	6	10.81
Fault Case 3	Variance increasing to $5\sigma_n$	PCA	10	21.33
		DFHPCA	6	16.22

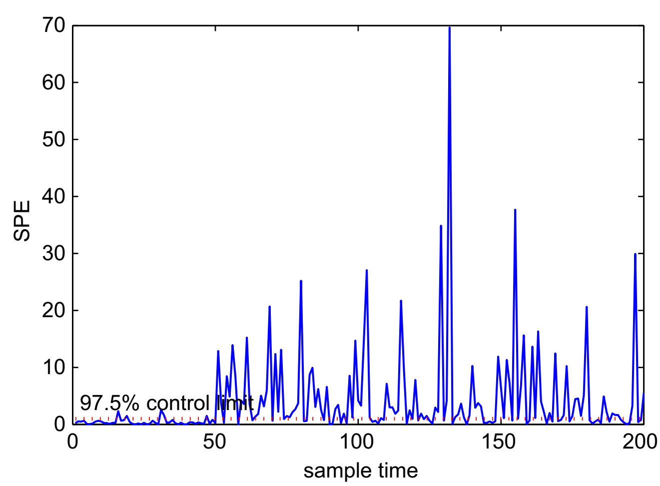

Fig. 7. Fault Case 3 SPE for traditional PCA fault detection.

In addition, comparison result of Fig. 3,4,5 and 6 tells us that when small fault occurs in the system, the detection rate of DFHPCA is also affected by the number of ZCPs which can be more clearly seen in Table 1.

With the lapse of time in real systems, the measurable precision of sensors may be degraded which can be characterized by a larger noise variance of the observation data. Fault Case 3 is used in our research to characterize a sensor fault of measurable precision dropping.

It can be seen from Figs. 7 and 8 that when the measurable precision of the sensor is degraded, which means that the variance increasing fault also has many ZCPs, and dynamic feature of the observation data is useful in well detection of this kind of the variance increasing fault.

When fault signal with ZCPs is involved in the system, false detection rate and miss detection rate of traditional PCA based

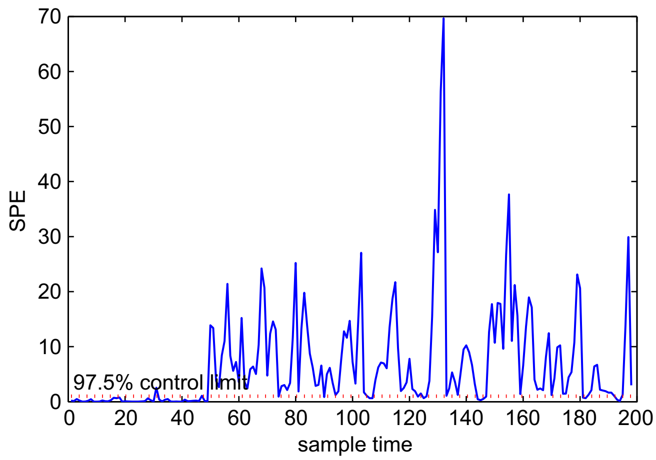

Fig. 8. Fault Case 3 SPE for DFHPCA fault detection.

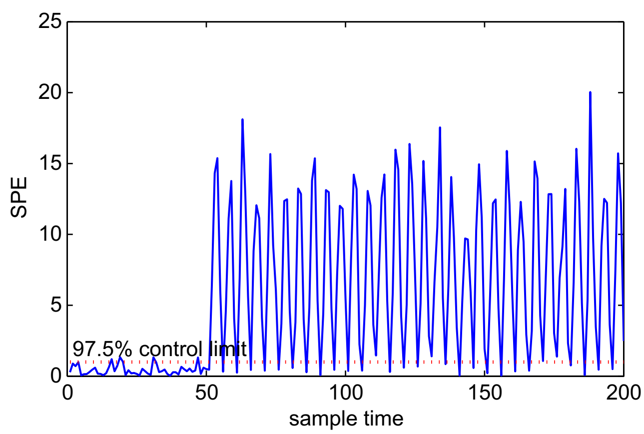

Fig. 9. Fault Case 1 SPE for traditional PLS fault detection.

fault detection method and DFHPCA fault detection method is listed in Table 1, where σ_n is the noise variance of the corresponding observation.

Since different faults are added into the same set of normal test observation, it can be seen from column 3 that the detection result before fault occurring is identical for three kind of fault cases. Comparing row 2 with row 3, row 4 with row 5, and row 6 with row 7, it can be concluded that both false detection rate and miss detection rate can be largely reduced by using DFHPCA fault detection method since the final detection result of multi-hierarchical detection is always the detection of the special hierarchical with largest detection ability. Comparing with row 2–5, it can be seen that, if the fault is small, then the miss detection rate is relatively high. On the other hand, comparing row 4 with row 5, it is obvious that, for a certain small fault size, the miss detection of our DFHPCA method is significantly smaller than that of traditional PCA method.

The main innovation of this research is to present a detection criterion by using of the differential information characterizing the dynamic feature. By this means, fault signals with many ZCPs can also be well detected which is significant to the fault tolerant control following fault detection.

The following simulation result of PLS detection for fault Case 1 is implemented by simply substituting PCA into PLS to get the DFHPLS method.

From Figs. 9 and 10, it is clear that the DFHPLS method is also suitable for PLS and other variations of PCA/PLS.

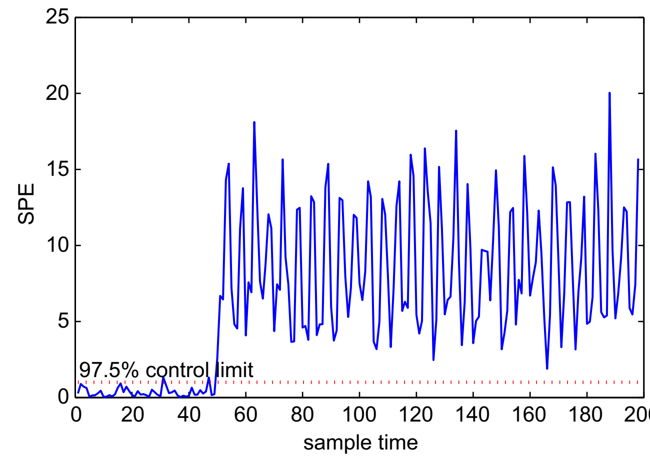

Fig. 10. Fault Case 1 SPE for DFHPLS fault detection.

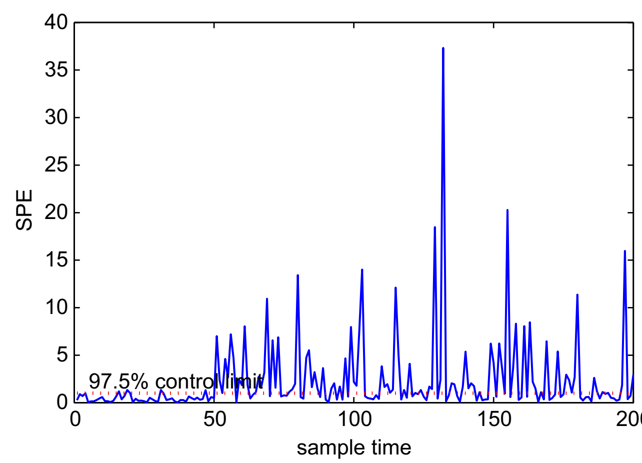

Fig. 11. Fault Case 3 SPE for traditional PLS fault detection.

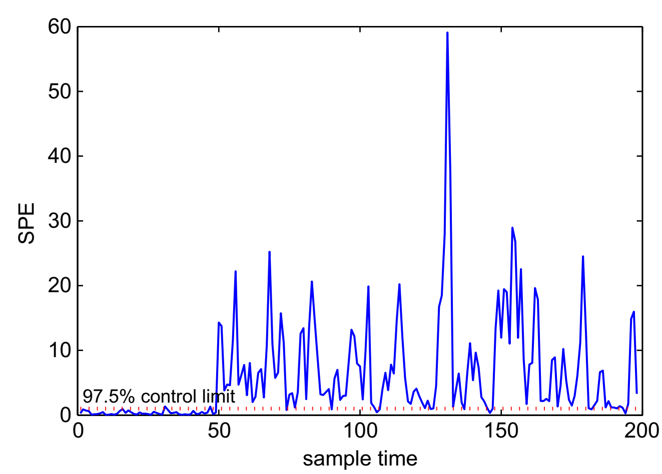

Fig. 12. Fault Case 3 SPE for DFHPLS fault detection.

Table 2
PLS/ DFHPLS False and miss detection rate.

Fault type	Fault size	Method used	FDR (%)	MDR(%)
Fault Case 1	Fault amplitude $> 3\sigma_n$	PLS	10	18
		DFHPLS	6	0
Fault Case 3	Variance increasing to $5\sigma_n$	PLS	10	30.67
		DFHPLS	6	4.73

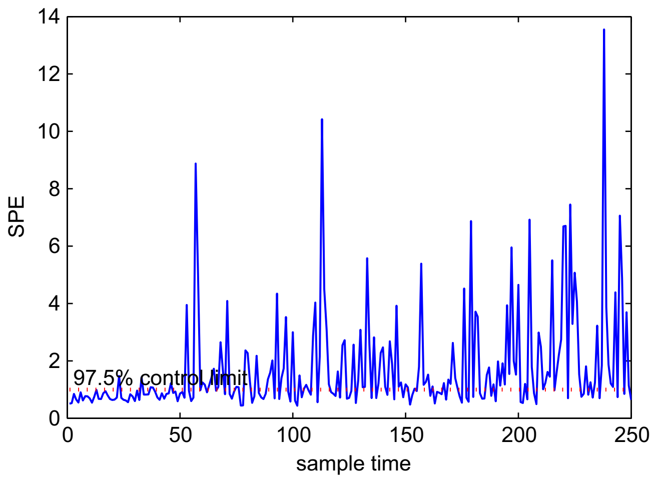

Fig. 13. SPE for traditional PCA to TEP.

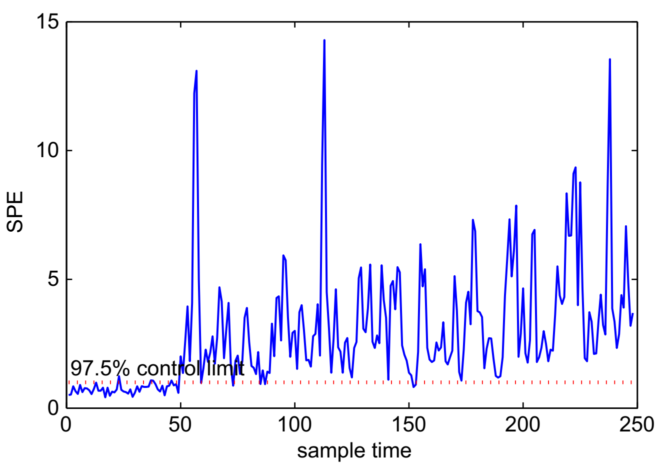

Fig. 14. SPE for DFHPCA to TEP.

When sensor fault relative to sensor accuracy degrading occurs, DFHPLS is still an efficient fault detection method, which is verified from Figs. 11 and 12.

Table 2 lists the false detection rate and miss detection rate of PLS and DFHPLS fault detection method. It is implemented by simply substituting PCA into PLS. Therefore, it is clear that our proposed method, DFHPLS, can easily generalize any statistical based fault detection methods.

5.2. Representative results of TE benchmark process study

In this subsection, we apply the DFHPCA fault detection to the representative Tennessee Eastman (TE) process. TE is a benchmark process to test the efficiency of anomaly monitoring algorithm [15,20,36]. There are 41 measured variables and 12 manipulated variables involved in the TE process. The flowchart and more detail information of TE process can be seen in the literatures [15,20,36]. The simulink code used to generate normal observation is available at the website,http://depts.washington.edu/control/LARRY/TE/download.html As illustrated in Section 5.1, DFHPCA algorithm can be also simulated by simply substituting PCA into other variation forms of PCA/PLS. Hence, the TE case study result is just illustrated for DFHPCA.

All 960 samples of trained normal data is used to establish the PCA/PLS and the DFHPCA fault detection model. The tested normal data comprise 250 samples, and the variance increasing fault is

Table 3
PCA/ DFHPCA False and miss detection rate to TEP.

Fault type	Fault size	Method used	FDR (%)	MDR (%)
Fault Case 3	Variance increasing to $5\sigma_n$	PCA	16	33
		DFHPCA	10	2

occurred. The variance of the 5th variable is increased to 1.5 from sample time $k=51$ to $k=250$ for fault Case 4.

Fig. 13 displays the SPE value of traditional PCA based fault detection. From the figure, it is easy to see that the fault with CZPs caused by sensor accuracy degradation can be detected from sample time 51 to 250, but there is a large amount of miss detection.

Comparing with Fig. 13, it can be seen from Fig. 14 that DFHPCA fault detection method can well detect the variance increasing fault occurring in TE process.

When sensor measurable accuracy degrading fault occurs in the process, it can be clearly seen from Table 3 that miss detection rate of traditional PCA is obviously larger than that of the DFHPCA method.

6. Conclusion and future work

In this paper, we have analyzed the detection ability of conventional PCA based fault detection methods, and have arrived at the conclusion that traditional method does not take dynamic character of fault signal into account during the designing of fault detection principle. Motivated by our intention to well detect dynamic fault with many ZCPs, a DFHPCA fault detection method is developed in detail. Differential feature is used in a hierarchical way to detect those miss detected points by simply implementing traditional PCA based detection. Using the DFHPCA method, miss detection rate can be significantly decreased.

For the case that dynamic fault is a divergent signal, such as divergent sine fault, future works will be conducted to avoid disastrous fault because divergent faults make the system completely fail in a short time.

References

[1] R.K. Mehra, J. Reschon, An innovation approach to fault detection and diagnosis in dynamics, Automatica 7 (5) (1971) 637–640.
[2] P.M. Frank, Fault diagnosis in dynamics systems using analytical and knowledge-based redundancy: a survey and some new results, Automatica 26 (3) (1990) 459–474.
[3] J.V. Kresta, J.F. MacGregor, T.E. Marlin, Multivariate statistical monitoring of process performance, Can. J. Chem. Eng. 69 (1) (1991) 35–47.
[4] Xu Xiaobin, Feng Haishan, Wang Zhi, Wen Chenglin, An information fusion method of fault diagnosis based on interval basic probability assignment, Chinese J. Electron. 20 (2) (2011) 255–260.
[5] Wen ChengLin, Qiu AiBing, Jiang Bin, An output delay approach to fault estimation for sampled-data systems, Sci. China Inf. Sci. (Science in China Series F) 55 (9) (2012) 2128–2138.
[6] H. Shen, X. Song, Z. Wang, Robust fault tolerant control of uncertain fractional-order systems against actuator faults, IET Control Theory Appl. 7 (9) (2013) 1233–1241.
[7] S. Yin, P. Shi, and H.Y. Yang. Adaptive fuzzy control of strict-feedback nonlinear time-delay systems with unmodeled dynamics. IEEE Trans. Cybern. 2015, available online, http://dx.doi.org/10.1109/TCYB.2015.2457894.
[8] D. Zhang, Q.G. Wang, L. Yu, H.Y. Song, Fuzzy-model-based fault detection for a class of nonlinear systems with networked measurements, IEEE Trans. Instrum. Meas. 62 (12) (2013) 3418–3159.
[9] S. Yin, S.X. Ding, X. Xie, H. Luo, A review on basic data-driven approaches for industrial process monitoring, IEEE Trans. Ind. Electron. 61 (11) (2014) 6418–6428.
[10] S. Yin, X.W. Li, H.J. Gao, O. Kaynak, Data-based techniques focused on modern industry: an overview, IEEE Trans. Ind. Electron. 62 (1) (2015) 657–667.

[11] V. Venkat, R. Raghunathan, Y. Kewn, N.K. Surya, A review of process fault detection and diagnosis Part I: quantitative model-based methods, Comput. Chem. Eng. 27 (3) (2003) 293–311.
[12] D. Zhang, W.A. Zhang, L. Yu, Q.G. Wang, Distributed fault detection for a class of large-scale systems with multiple incomplete measurements, J. Frankl. Inst. 352 (9) (2015) 3730–3749.
[13] J.S. Zeng, U. Kruger, J. Geluk, X. Wang, L. Xie, Detecting abnormal situations using the Kullback–Leibler divergence, Automatica 50 (11) (2014) 2777–2786.
[14] M. Adil, M. Abid, A.Q. Khan, G. Mustafa, N. Ahmed, Exponential discriminant analysis for fault diagnosis, Neurocomputing 171 (2016) 1344–1353.
[15] C. Jing, J. Hou, SVM and PCA based fault classification approaches for complicated industrial process, Neurocomputing 167 (1) (2015) 636–642.
[16] M.A. Bin Shams, H.M. Budman, T.A. Duever, Fault detection, identification and diagnosis using CUSUM based PCA, Chem. Eng. Sci. 66 (2011) 4488–4498.
[17] Z.D. Zhang, B.C. Peng, L.B. Xie, L. Peng, Process monitoring based on recursive probabilistic PCA for Multi-mode Process, IFAC-Papers OnLine 48 (8) (2015) 1294–1299.
[18] K. Pollanen, A. Hakkinen, S. Reinikainen, J. Rantanen, P. Minkkinen, Dynamic PCA-based MSPC charts for nucleation prediction in batch cooling crystallization processes, Chemometrics Intell. Lab. Syst. 84 (1) (2006) 126–133.
[19] L.F. Cai, X.M. Tian, S. Chen, A process monitoring method based on noisy independent component analysis, Neurocomputing 127 (15) (2014) 231–246.
[20] Z.Q. Ge, F.R. Gao, Z.H. Song, Mixture probabilistic PCR model for soft sensing of multimode processes, Chemometrics Intell. Lab. Syst. 105 (2011) 91–105.
[21] C.L. Wen, F.N. Zhou, An extended multi-scale principal component analysis and application in fault detection, Chinese J. Electron. 21 (3) (2012) 471–479.
[22] M. Navi, M.R. Davoodi, N. Meskin, Sensor fault detection and isolation of an industrial gas turbine using partial kernel PCA, IFAC-Papers OnLine 48 (21) (2015) 1389–1396.
[23] L.F. Cai, X.M. Tian, S. Chen, A process monitoring method based on noisy independent component analysis, Neurocomputing 127 (15) (2014) 231–246.
[24] D.Z. Peng, Y. Zhang, Dynamics of generalized PCA and MCA learning algorithms, IEEE Trans. Netw. 18 (6) (2007) 1777–1784.
[25] S. Yin, G. Wang, H.J. Gao. Data-driven process monitoring based on modified orthogonal projections to latent structures, IEEE Trans. Control Syst. Technol. available online, 2015, http://dx.doi.org/10.1109/TCST.2015.2481318.
[26] F. Harrou, F. Kadri, S. Chaabane, C. Tahon, Y. Sun, Improved principal component analysis for fault detection: application to an emergency department, Comput. Ind. Eng. 88 (2015) 63–77.
[27] J. Hu, C.L. Wen, P. Li, T.Q. Yuan, Direct projection to latent variable space for fault detection, J. Frankl. Inst. 351 (3) (2014) 1226–1250.
[28] Y. Gao, X. Wang, Z.L. Wang, Fault detection for a class of industrial processes based on recursive multiple models, Neurocomputing 169 (2) (2015) 430–438.
[29] S. Yin, Z.H. Huang, Performance monitoring for vehicle suspension system via fuzzy positivistic C-Means clustering based on accelerometer measurements, IEEE/ASME Trans. Mechatronics 20 (5) (2015) 2613–2620.
[30] L. Liu, Z.S. Wang, H.G. Adaptive NN fault-tolerant control for discrete-time systems in triangular forms with actuator fault, Neurocomputing 152 (25) (2015) 209–221.
[31] M.H. Gharavian, F. Almas Ganj, A.R. Ohadi, H. Heidari Bafroui, Comparison of FDA-based and PCA-based features in fault diagnosis of automobile gearboxes, Neurocomputing 121 (9) (2013) 150–159.
[32] A. Asuhaimi, M. Zin, M. Saini, M.W. Mustafa, A.R. Sultan, New algorithm for detection and fault classification on parallel transmission line using DWT and BPNN based on Clarkes transformation, Neurocomputing 168 (30) (2015) 983–993.
[33] Y. Tian, M.Y. Fu, F. Wu, Steel plates fault diagnosis on the basis of support vector machines, Neurocomputing 151 (3) (2015) 296–303.
[34] Y.Q. Xiao, Y.G. He, A novel approach for analog fault diagnosis based on neural networks and improved kernel PCA, Neurocomputing 74 (7) (2011) 1102–1115.
[35] F. Li, J.X. Wang, B.Q. Tang, Weak fault diagnosis of rotating machinery based on feature reduction with supervised orthogonal local Fisher discriminant analysis, Neurocomputing 168 (30) (2015) 505–519.
[36] J.J. Downs, E.F. Vogel, A plant-wide industrial process control problem, Comput. Chem. Eng. 17 (1993) 245–255.

F. Zhou received the Ph.D degree at Shanghai Maritime University in 2009. She is now an associate professor and a supervisor of master students in School of Computer and Information Engineering at Henan University. Her research interest covers the areas of fault diagnosis and information fusion.

Ju H. Park received the Ph.D. degree in Electronics and Electrical Engineering from POSTECH, Pohang, Republic of Korea, in 1997. From May 1997 to February 2000, he was a Research Associate in ERC-ARC, POSTECH. In March 2000, he joined Yeungnam University, Kyongsan, Republic of Korea, where he is currently the Chuma Chair Professor. From December 2006 to December 2007, he was a Visiting Professor in the Department of Mechanical Engineering, Georgia Institute of Technology, USA. His research interests include robust control and filtering, neural networks, complex networks, multi-agent systems, and chaotic systems. He has published a number of papers in these areas. He serves as an Editor of International Journal of Control, Automation and Systems. He is also a subject Editor/Associate Editor/Editorial Board Member for several international journals, including IET Control Theory and Applications, Nonlinear Dynamics, Cogent Engineering, Applied Mathematics and Computation, Journal of The Franklin Institute, and Journal of Applied Mathematics and Computing.

Y. Liu received the B.S. degree in Mathematics and Applied Mathematics from Shanxi Normal University, Linfen, China, in 2010, M.S. degree in Applied Mathematics, University of Science and Technology Beijing, Beijing, China, in 2012, and Ph.D. degree at the Division of Electronic Engineering in Daegu University, Deagu, Republic of Korea, in 2015. From September, 2015, she has been doing postdoctoral research at the Department of Electrical Engineering in Yeungnam University, Kyongsan, Republic of Korea. Her research focus is control of dynamic systems including neural networks and complex systems.