Fault diagnosis in chemical processes based on class-incremental FDA and PCA

Yang, Xiaohui; Rui, Songhong; Zhang, Xiaolong; Xu, Shaoping; Yang, Chunsheng; Liu, Peter X.

Publisher's version / Version de l'éditeur:
https://doi.org/10.1109/ACCESS.2019.2896056
IEEE Access, 7, pp. 18164-18171, 2019-01-31

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 14, 2019, accepted January 26, 2019, date of publication January 31, 2019, date of current version February 20, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2896056

Fault Diagnosis in Chemical Processes Based on Class-Incremental FDA and PCA

XIAOHUI YANG[1], SONGHONG RUI[1], XIAOLONG ZHANG[1], SHAOPING XU[1], CHUNSHENG YANG[1,2], AND PETER X. LIU [1,2,3], (Fellow, IEEE)

[1]College of Information Engineering, Nanchang University, Nanchang 330031, China
[2]National Research Council Canada, Ottawa, ON, Canada
[3]Department of Systems and Computer Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada

Corresponding author: Shaoping Xu (xushaoping@ncu.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 51765042, Grant 61463031, Grant 61662044, and Grant 61862044, and in part by the Educational Reform Project of the JiangXi Provincial Department of Education under Grant JXYJG-2017-02 and Grant JXJG-18-1-43.

ABSTRACT A class-incremental scheme of fisher discriminant analysis is proposed to improve the performance ofprocess fault diagnosis. Fisher discriminant analysis seeks directions which are efficient for discrimination and has excellent fault detection and diagnostic performance for the sample set with the tag. However, due to the property of the model, it has no detection and diagnostic capabilities for un-seen faults. In order to address this issue, the F direction, which is based on a partial F-values with the principle component analysis, is proposed in this paper. After a new fault being detected and added into the known fault collection, a class-incremental scheme is used to update the fisher discriminant analysis model to enhance the model's ability for continuous fault identification. The proposed approach is validated by the Tennessee Eastman process for the fault diagnosis. The results demonstrate that the proposed class-incremental fisher discriminant analysis method outperforms other conventional fisher discriminant analysis methods.

INDEX TERMS Fault diagnosis, class-incremental, fisher discriminant analysis, principle component analysis.

I. INTRODUCTION

With the development of computer-aided process control and instrumentation technology and the emergence of process systems, a large amount of data can be recorded. However, due to the complexity of the chemical industry production system, the collected process data often has high dimensional and complex characteristics, so there may be the problems called dimensional disasters. As one of multivariate projection techniques, principle component analysis (PCA) can extract the latent variables from the data of a highly correlated process and capture its main characteristics. PCA has had a widespread usage of monitoring continuous and batch processes in industry [1]–[8]. In addition, Dynamic Principal Component Analysis (DPCA) is used for monitoring of dynamic multivariate processes [7], and other data-driven methods such as discriminant partial least square (PLS) and independent component analysis (ICA) were also widely used for fault diagnosis of complex systems [4]. These methods play an important role in the variable extraction and fault diagnosis of complex systems. Two statistics, i.e. Hotelling's T^2 and the squared prediction error (SPE or Q), are regularly used to detect abnormal situations [5], [9].

Fisher discriminant analysis (FDA) is a classical technique for dimension reduction and classification [3], [10]. FDA seeks directions which are efficient for discrimination [2], [11], [12] and it has a better differentiation than PCA among fault data sets (i.e., fault diagnosis) [7] and between the normal data sets and the fault data sets (i.e., fault detections) [13]–[15]. FDA uses all of fault classification information, so in general it is more effective in separating the normal data sets from the fault ones [16]–[20], while FDA shows smaller error rate for detecting and diagnosing the known fault classes. However, FDA model may not be able to detect the faults that are not included in the training data set [21]–[25]. Furthermore, FDA for process monitoring requires different and available deterministic fault data sets, which is usually difficult in practice. On one hand, the fault data sets should be gathered as more as possible from

The associate editor coordinating the review of this manuscript and approving it for publication was Chuan Li.

historical data, plant tests or mechanistic knowledge, on the other hand, new fault data sets should be kept gathered in the future thereby enriching the known fault collection in FDA for fault diagnosis. Another challenge is how to identify a new fault from the known faults. The fault direction finds out the discriminant vector by maximizing separation between the normal data and fault data on the part of discriminant analysis. By adding a standardization named the standardized fault direction, the standardized discriminant vector makes all variables be compared and improves the explanation of the fault direction. Both the fault direction and the standardized fault direction can be used for identifying a new fault [17], [26]. However, it is difficult to correctly identify each fault class because of the fact that between-class variability is assessed relative to within-class variability, especially when a chemical process has a large scale and the process data masked with irrelevant information.

With orientation to process fault diagnosis, we propose a new scheme based on class-incremental FDA. For this scheme, the normal data set is only required to build the PCA model. The performance of the developed PCA model lies on its representative ability of the normal data sets and it can be used for detecting the unknown faults. In order to identify new faults from the known ones, the F direction, which is based on partial F-values with the cumulative percent variation (CPV), is introduced. While the conventional approach based on partial F-values works well in detecting faults in most cases [27]–[30], it still suffers from irrelevant variables and low computation efficiency [31]–[36]. In this paper, the CPV, based on the equivalent variation of each variable, is proposed to determine candidate variables. As a consequence, it not only eliminates redundant variables but also reduces the computational complexity, improving the performance of the fault detection. At the same time, once the F direction is obtained, the cosine values of the angles between the new fault samples and the fault classes are added into the library [18], [37]–[39].

The paper is organized as follows. The concepts of PCA and FDA are reviewed in Section II. Section III presents the fundamental idea of the developed class-incremental FDA. And the class-incremental FDA for process fault diagnosis is illustrated in Section IV. Section V describes an application to Tennessee Eastman Process (TEP), and conclusion is given in Section VI.

II. REVIEW ON PCA AND FDA

Let $X_0 \in R^{n_0 \times m}$ (with mean of zero and variance of unit) be the normal data set with n_0 samples (rows) and m variables (columns). X_k $(k = 1, \cdots, c)$ is the k^{th} fault data set with n_k samples and m variables. c is the number of fault classes. All fault data are scaled with the mean and variance of the normal data set. Stack the normal data and all fault data sets into the matrix $X \in R^{n \times m}$. Thus, $X = \bigcup_{k=0}^{c} X_k$. The ith row of X represents the sample vector x_i.

A. PCA

PCA is an optimal dimensionality reduction technique in terms of capturing the variance of the data [6]. The matrix X_0 can be decomposed into a score matrix T and a loading matrix P.

$$X = TP^T + \tilde{X}_0 = TP^T + \tilde{T}\tilde{P}^T \tag{1}$$

where $\tilde{X}_0 = \tilde{T}\tilde{P}^T$ is the residual matrix and $P \in R^{m \times r}$ is the loading matrix. r is the number of the principal components sufficiently representing the variability of the data set X_0. The principal component subspace (PCS) is $S_p = span\{P\}$ and the residual subspace (RS) is $S_r = span\{\tilde{P}\}$. P can be calculated by the singular value decomposition (SVD) algorithm of the covariance matrix R_0, that is,

$$R_0 = \frac{1}{n_0 - 1} X_0^T X \approx P \Lambda P^T \tag{2}$$

where $\Lambda = diag\,(\lambda_1, \lambda_2, \cdots, \lambda_r)$, λ_i are the eigenvalues of the covariance matrix in descending order. Denote the PCA model for process monitoring as $\{b_0, n_0, P, \Lambda, r\}$ where b_0 is the mean vector of the normal data set.

B. FDA

FDA determines a set of Fisher optimal discriminant vectors by maximizing the scatter between the classes while minimizing the scatter within each class [13]. The transformation matrix consisting of these discriminant vectors is denoted by $W \in R^{m \times d}$ where $d \leq c-1$ is the dimension of the projection subspace. Then, a fault sample vector x_i is transformed to the projection subspace by

$$z_i = W^T x_i \tag{3}$$

The transformation matrix can be obtained by maximizing the Fisher criterion $J_F\,(W)$

$$J_F\,(W) = trace\{\left(W^T S_w W\right)^{-1}\left(W^T S_b W\right)\} \tag{4}$$

where

$$S_b = \sum_{k=0}^{c} n_k\,(b_k - b)\,(b_k - b)^T \tag{5}$$

And

$$S_w = \sum_{k=0}^{c} \sum_{x_i \in X_k} (x_i - b_k)\,(x_i - b_k)^T \tag{6}$$

S_b and S_w are the between-class scatter matrix and the within-class scatter matrix, respectively. Here, b is the total mean vector of all fault samples and b_k is the mean vector for the fault class k. Denote the FDA model for process monitoring as $\{b, n, S_w, S_b, W, d\}$.

III. CLASS-INCREMENTAL FDA

A. UPDATING S_W AND S_B

When the new fault class $c+1$ is detected, the fault data set $X_{c+1} \in R^{n_{c+1} \times m}$ is gathered and scaled with the normal mean and variance. Then, the mean of the $c+1^{th}$ fault data is given in the column vector

$$b_{c+1} = \frac{1}{n_{c+1}} X_{c+1}^T 1_{n_{c+1}} \tag{7}$$

where $1_{n_{c+1}} = [1, \cdots, 1]^T \in R^{n_{c+1}}$ The total mean vector of all data $X' = \begin{bmatrix} X^T & X_{c+1}^T \end{bmatrix}^T$ can be calculated by

$$b' = \frac{1}{n'} (X') 1_{n'} = \frac{n}{n'} b + \frac{n_{c+1}}{n'} b_{c+1} \tag{8}$$

where, $n' = n + n_{c+1}$ and $1_{n'} = [1, \cdots, 1]^T \in R^n$.

Then, the updated within-class matrix S'_w can be rewritten as

$$\begin{aligned} S'_w &= \sum_{k=0}^{c+1} \sum_{x \in X_k} (x - b_k)(x - b_k)^T \\ &= S_w + \sum_{x \in X_{c+1}} (x - b_{c+1})(x - b_{c+1})^T \end{aligned} \tag{9}$$

And the updated between-class matrix, S'_b has the following relationship:

$$\begin{aligned} S'_b &= \sum_{k=0}^{c+1} n_k (b_k - b')(b_k - b')^T \\ &= S_b + \frac{n * n_{c+1}}{n'} (b - b_{c+1})(b - b_{c+1})^T \end{aligned} \tag{10}$$

B. UPDATING FDA MODEL

Because the matrix $S_w^{-1} S_b$ does not need to be symmetric, the eigen system computation could be unstable. A stable batch eigen system computation method by diagonalizing two symmetric matrices [29], [31] is adopted. After S'_w and S'_b are calculated, let SVD decomposition of S_w is

$$S'_w = H'_w \Sigma'_w (H_{w'})^T \tag{11}$$

where H'_w is orthogonal and Σ'_w is diagonal.

Similarly,

$$\left(H'_w (\Sigma'_w)^{-\frac{1}{2}} \right)^T S'_b H'_w (\Sigma'_w)^{-\frac{1}{2}} = H'_b \Sigma'_b (H'_b)^T \tag{12}$$

where H'_b is orthogonal and Σ'_b is diagonal. Let $H' = H'_w (\Sigma'_w)^{-\frac{1}{2}} H'_b$.

Then,

$$(S'_w)^{-1} S'_b = H' \Sigma'_b (H')^{-1} \tag{13}$$

In the above equation, H' consists of the eigenvectors of $S_w^{-1} S_b$ and Σ'_b contains the eigenvalues of $S_w^{-1} S_b$. Next, we denote the updated FDA model as $\{b', n', S'_w, S'_b, W', d'\}$

IV. PROCESS MONITORING USING CLASS-INCREMENTAL FDA

A. FAULT DETECTION USING CLASS-INCREMENTAL FDA

As mentioned above, the merit of the FDA model takes full advantages of fault classification information, so the performance of fault detection is improved. FDA shows small error rate for detecting the known fault classes in comparison with PCA.

When $c \neq 0$, there are fault data sets to construct FDA model, the monitoring statistic T_{FDA}^2 can be defined as

$$T_{FDA}^2 = x_{new}^T W (W R_0 W)^{-1} W x_{new} \tag{14}$$

where x_{new} is a new sample vector. The process is considered as "in control" if $T_{FDA}^2 \leq T_{FDA,\alpha}^2$. $T_{FDA,\alpha}^2$ denotes the upper control limit with a given significance level α [1].

The monitoring performance depends on the similarity between the normal data and the fault data in the training data set. As a result, T_{FDA}^2 may not correctly detect the new faults. Unlike the FDA model, the PCA model represents the normal data sufficiently and the normal data set is only required to build the PCA model. The PCA model can be used for detecting and identifying those unknown faults. In other words, the FDA model and PCA model are complementary to each other in terms of process monitoring, so it will be better if we could combine them and take advantage of both method to improve the monitoring performance.

For process monitoring, typically, Hotelling T_{PCA}^2 and Q_{PCA} statistics are used to represent the variability in the Principal Component Subspace (PCS) and (Residual Subspace)RS, respectively [28].

$$T_{PCA}^T = x_{new}^T P \Lambda^{-1} P^T x_{new} \tag{15}$$

$$Q_{PCA} = \| \left(I - PP^T \right) x_{new} \|^2 \tag{16}$$

where x_{new} is a new sample vector, I is a unit matrix. The process is considered normal if $T_{PCA}^2 \leq T_{PCA,\alpha}^2$ and $Q_{PCA} \leq Q_{PCA,\alpha}$, where $T_{PCA,\alpha}^2$ and $Q_{PCA,\alpha}$ denotes the upper control limits for T_{PCA}^2 and Q_{PCA} respectively [17], [32].

When $c = 0$, there are no fault data sets and the statistic T_{FDA}^2 loses its meaning. In fact, process fault diagnosis based on class-incremental FDA turns to be based on the conventional PCA.

B. FAULT DIAGNOSIS USING THE DIRECTION

After a fault being detected, fault diagnosis can aims at determining the root cause of the fault [35]. As far as discriminant analysis is concerned, the F direction based the partial F-values with CPV finds out the discriminant vector by maximizing separation between the normal data sets and fault ones. With continuous value, each element of the discriminant vector shows the corresponding variable's responsibility for the fault.

The partial F-values assess each variable's contribution to Hotelling's T^2 for class separation. In the case of two classes:

X_0 and X_k, the partial F value of the variable i is given by [37]

$$F(i) = (v - m + 1)\frac{T_m^2 - T_{m-1,i}^2}{v + T_{m-1,i}^2} \quad i = 1, \cdots, m \tag{17}$$

where $v = n_0 + n_k - 2$, T_m^2 is the two-class Hotelling's T^2 with all m variables. It is defined as:

$$T_m^2 = \frac{n_0 n_k}{n_0 + n_k}(b_0 - b_k)^T S_k^{-1}(b_0 - b_k) \tag{18}$$

where, S_k is the covariance matrix of the stacked matrix $X_{0+k} = \begin{bmatrix} X_0^T & X_k^T \end{bmatrix}^T$. $T_{m-1,i}^2$ is Hotelling T^2 with all variables except for the variable i .The F statistics is distributed as $F_{l,v-m}$. If $F(i) \geq F_{\delta,l,v-m}$, the variable offers useful information for classification. Here, $F_{\delta,l,v-m}$ is the critical value with a significant level δ .

A chemical process always has characteristics of large scale and process data masked with irrelevant information, which leads to the difficulty of variable weighting [23], [35]. Although the partial F-values show better interpretation of the single discriminant vector than the fault direction and the standardized fault direction, they still suffer from the irrelevant variables and lower computation efficiency.

For Hotelling's T^2, generally, the mean matrix and the covariance matrix change two essential factors representing the process change from the normal situation to some fault situation [27]. There are two portions of the result in each variable's contribution to the process change. One is the mean or variance change of the variable itself and the other is the relation with other variables responsible for the process change. Therefore, we define the equivalent variation of the i^{th} variable for the k^{th} fault as:

$$\Delta V_k(i) = \sum_{i=1}^{m} |R_k(j,i)\Delta b_k(i)| + \|\Delta S_k(:,i)\|_1 \tag{19}$$

Δb_k and ΔS_k are the mean vectors and the covariance matrix and they change between the normal data set X_0 and the k^{th} fault data set X_k, respectively. Moreover, $\Delta b_k = b_k - b_0$ and $\Delta S_k = S_k - S_0$. $\Delta(i)$ is the i^{th} element of the vector Δb_k, and $\Delta S_k(:,i)$ is the i^{th} column of the matrix ΔS_k. $\|\cdot\|_1$ denotes the 1-norm of a vector. $R_k(j,i)$ is the j^{th} row and the i^{th} column element of the correlation matrix of X_k. And $\sum_{j=1}^{m} |R_k(j,i)\Delta b_k(i)^2|$ indicates the equivalent mean change of the i^{th} variable, which includes its own mean change and the mean change induced by other variables. Likewise, $\|\Delta S_k(:,i)\|_1$ represents the equivalent variance change of the i^{th} variable. Each variable's equivalent variation $\Delta V_k(i)$ comes from the mean and variance changes of either the variable itself or the relative other variables. When these variables neither experience their mean or variance change nor have the relation with other contribution variables, they have no contribution to the two-class separation and their partial F values are directly set to be zero. For the partial F-values, finding out these irrelevant variables and then eliminating the redundant variables not only increase the computation efficiency but also improve the variable weighting performance.

After obtaining all variable's equivalent variations, we can rank them. Then, CPV can be calculated by [22]

$$CPV(l_k) = \frac{\sum_{i=1}^{l_k} \Delta V_k(i)}{\sum_{i=1}^{m} \Delta V_k(i)} \times 100\% \tag{20}$$

CPV is a measure of the percent variation captured by the first l_k candidate variables. The number of these candidate variables can be defined when CPV reaches a predetermined limit, such as 85%

The F direction is based on the partial F-values with CPV. For a fault class, it can be obtained from the normal data and a class of fault data. Let F_k be the F direction for the fault k, $F_k = [F_k(1), F_k(2), \cdots, F_k(i), \cdots, F_k(m)]^T$. The ith element $F_k(i)$ is the F value of the variable and represents the corresponding variable contribution to Hotelling's T^2 statistic. This process is not repeated until the analysis of all faults.

After they are normalized as $\overline{F}_k = \frac{F_k}{norm(F_k)}$, the fault signature library consists of all known fault F direction as follows:

$$\overline{F} = \{\overline{F}_1 \quad \overline{F}_2 \quad \cdots \quad \overline{F}_c\} \tag{21}$$

When the i th new fault data set X_{new} is available, the F direction F_{new} can be obtained in the same way. After F_{new} normalized into $\overline{F}_{new}$, the Euclidean Distance D between the new fault signature vector and the corresponding fault signature vector in the known fault signature library is used for fault diagnosis. So,

$$\begin{aligned} D_k &= \|\overline{F}_k - \overline{F}_{new}\|^2 = \sum_{i=1}^{m}\left(\overline{F}_k(i) - \overline{F}_{new}(i)\right)^2 \\ k &= 1, \cdots, c\frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \end{aligned} \tag{22}$$

where $D_k \in [0, 1]$. The closer the distance values are to zero, the more nearly collinear the new samples are to the known fault. The new fault samples should be classified to the one whose distance value is closest to zero. Then, the distance value is used for process diagnosis.

C. A COMPLETE MONITORING PROCEDURE USING CLASS-INCREMENTAL FDA

By off-line learning, construct PCA model $\{b_o, n_0, P, \Lambda, r\}$ for the normal data set X_0, FDA model $\{b, n, S_w, S_b, W, d\}$ for the data sets X, and calculate the three monitoring limits: $T_{PCA,\alpha}^2$, $Q_{PCA,\alpha}$ and $T_{FDA,\alpha}^2$. Then, the monitoring procedure using class-incremental FDA with PCA is expressed as follows:

Calculating T_{PCA}^2, Q_{PCA} and $T_{FDA,\alpha}^2$ for the new sample normalized x_{new}.

1. Check the condition: If $T_{FDA}^2 > T_{FDA,\alpha}^2$, $T_{PCA}^2 > T_{PCA,\alpha}^2$ or $Q_{PCA} > Q_{PCA,\alpha}$, a fault is detected.

2. Obtain the F direction for the new fault by calculating CPV and partial F-value. Then, the Euclidean Distance D is used for fault diagnosis. If the fault is new, jump to Step 4. Otherwise, jump to Step 1.

3. Update S'_w and S'_b according to Eqs. (7) and (8) respectively after the new fault data set X_{c+l} is gathered.
4. Obtain the updated FDA model $\{b', n', S'_w, S'_b, W', d'\}$.
5. Calculate $T^2_{FDA,\alpha}$.
6. Set $\{b, n, S_w, S_b, W, d\} = \{b', n', S'_w, S'_b, W', d'\}$.
7. Go to Step 1.

V. APPLICATION IN THE TE PROCESS

In order to provide a realistic testbed for control and condition monitoring tasks in a chemical engineering context, a software simulator of a production plant was proposed [23]. A reactor, condenser, stripper, compressor and separator constitute the main components of the system where two liquid products and a liquid byproduct in two parallel reactions are obtained. The original code is written in Fortran, a Matlab/Simulink adaptation is also available [40]. Recent publications from the field of fault diagnose corroborate the efficiency of the TE simulator [18], [34]–[38]. In these references, the data produced by the simulator in [4] and [13] are used as input to the diagnosis system. Fig. 1 outlines the schema of the Tennessee Eastman simulator.

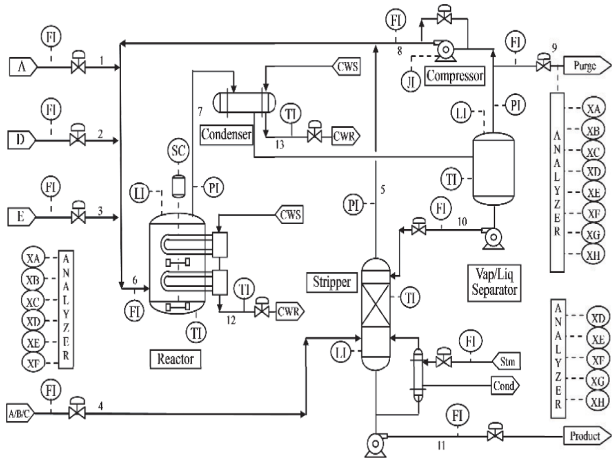

FIGURE 1. A flow sheet of the Tennessee Eastman process.

Tennessee Eastman process (TEP) is used to evaluate the new process monitoring approach with class-incremental FDA. TEP is based on an industrial process where the components, kinetics, and operating conditions are modified for proprietary reasons [13], [30]. It has been a well-known benchmark process for evaluating process monitoring methods [4].

Faults 2, 5, 8 and 12 are considered. Each of them runs for 45 hours. There are no faulty conditions at the beginning of the simulation. Fault 2 is introduced at $t = 35$ hours (Observation 701). Fault 5 firstly occurs at $t = 15$ hours (Observation 301) and lasts in 15 hours. After the fault is removed and the process goes back to the normal condition, Fault 8 and 12 are introduced at $t = 15$ hours (Observation 301). Besides, the sampling interval is set to be 3 minutes. The observations from 1 to 200 are normal data and used to build PCA monitoring model. Assure that the fault data sets

TABLE 1. Selected faults.

ID	Fault description	Type
Fault 2	B composition,A/C ratio constant	Step change
Fault 5	Condenser cooling water inlet temperature	Step change
Fault 8	A,B,C feed composition	Random variation
Fault 12	Condenser cooling water inlet temperature	Random variation

for Faults 2, 8 and 12 have been gathered. So the initial FDA model can be built on the normal and these fault data sets. Table 1 lists the selected fault descriptions.

Fault 2 is a step change of gaseous inert B in Stream 4. Because of the close loop reactor, the inert B composition in Stream 6, 9 and 11 returns to the normal situation after experiencing a transient increase. And the byproduct F composition decreases in those streams. Total 4 variables show abnormal behaviors. In Figure 2, (a) to (c) shows the monitoring results using the initial PCA-enabled FDA model. Although all threes statistics T^2_{FDA}, T^2_{PCA} and Q_{PCA} show good performance, T^2_{FDA} detects the fault at Observation 704 and is 30 minutes and 15 minutes earlier than T^2_{PCA} and Q_{PCA}, respectively.

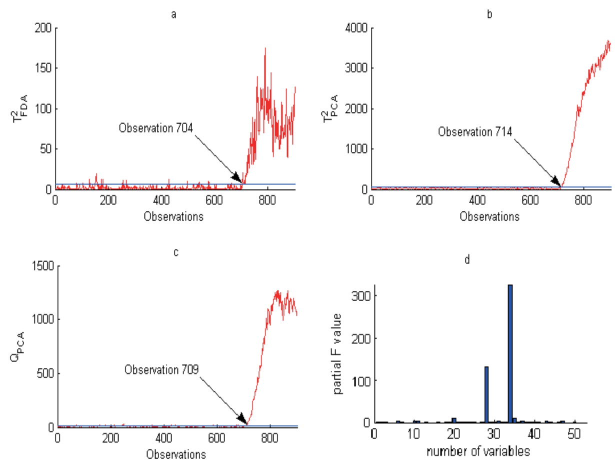

FIGURE 2. Monitoring results for Fault 2: a) T^2_{FDA}, b) T^2_{PCA}, c) Q_{PCA} and d) the F fault direction for Fault 2.

After the fault being detected, fault diagnosis should be performed according to the monitoring procedure discussed. Figure 2 (d) is the F fault direction for the fault. When the predetermine limit is set to be 85%, CPV selects 37 candidate variables from total 52 measured variables. The F fault direction puts in the paper weights these 16 fault variables exactly, especially variables 28 and 34. The Euclidean Distance D between these new faults and the corresponding faults vectors is 0.03, close to zero. It indicates that the fault has a same root cause as Fault 2. The fault is not a new one, FDA model does not need to be updated in this case.

Fault 5 is a step change in the unmeasured condenser cooling water inlet temperature. When Fault 5 occurs at $t = 15$ hours (Observation 301), a step change in the reactor cooling water flow (variable 52) is directly induced. The variable 17 lies in a same control loop with the variable 52.

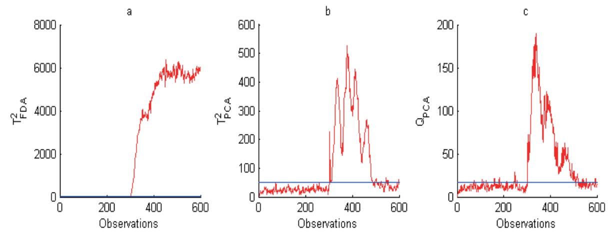

FIGURE 3. Monitoring results for Fault 5: a) T^2_{FDA}, b) T^2_{PCA}, c) Q_{PCA}.

They are correlated. Because of control loops, 34 variables in the rest of monitoring variables experience transient behaviors and take about 8 hours to reach their steady states. Because FDA model takes fault information into account, T^2_{FDA} shows not only earlier fault detection and smaller error rate but also better detection duration than T^2_{PCA} and Q_{PCA} (see Figure 3 a-c)

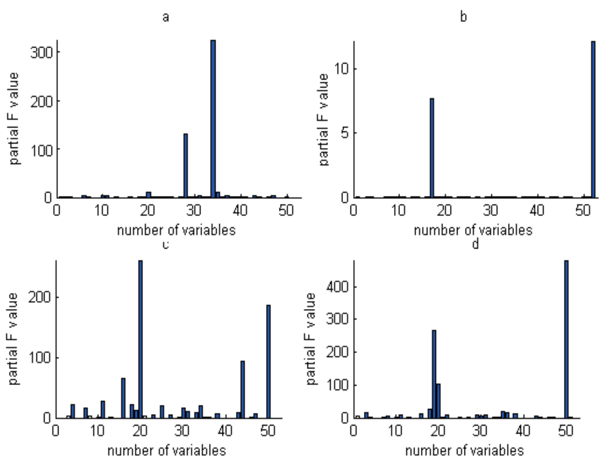

FIGURE 4. The F fault direction based on the partial F-values with CPV for: (a) Fault 5, (b) Fault 2, (c) Fault 8, and (d) Fault 12.

After the fault is detected, the fault direction should be performed according to the monitoring procedure discussed in. Figure 4(a) indicates the F fault direction using partial F-values with CPV for Fault 5. Fault 8 and Fault 12 have a same fault type, namely random variation, but they are associated with different fault variables. For Fault 8, there is a random change of A, B and C feed composition at $t = 30h$, which causes 36 variables to experience obvious changes. For the weighting data set, the F fault direction based on the partial F-values with CPV correctly interprets the fault and identifies all fault variables (see Figure 4(c)). For fault 12, total of 28 variables show the different behavior from the normal data. Its fault direction can be seen in Figure 4.(d). Then, the Euclidean Distances D between the new fault and the corresponding fault directions are larger than 0.5, so the fault can't match any fault direction in the history library, we can conclude that the fault is a new one. So, the fault data are gathered to update FDA monitoring model. At the same time, the fault direction is added into the library.

VI. CONCLUSION

This paper proposes a novel approach to the fault diagnosis of industrial processes based on class-incremental fisher discriminant analysis. The method includes the class-incremental based on the partial F-values with cumulative percent variation into fisher discriminant analysis. The developed class-incremental fisher discriminant analysis extracts discriminative features more effectively than the conventional fisher discriminant analysis from overlapping fault data. This proposed approach is applied to the Tennessee Eastman process. The results obtained for the Tennessee Eastman process demonstrated that the proposed method outperforms other conventional fisher discriminant analysis.

The fault detection experiment of the model is currently only carried out in the Tennessee Eastman process, and it is difficult to evaluate the fault diagnosis reliability of the method in other aspects. The future work is further optimizing the model and try to apply the model to the system with more complex processes and more diverse fault types to improve the practical possibilities of the model.

REFERENCES

[1] V. Venkatasubramanian, R. Rengaswamy, S. N. Kavuri, and K. Yin, "A review of process fault detection and diagnosis: Part III: Process history based methods," *Comput. Chem. Eng.*, vol. 27, no. 3, pp. 327–346, 2003.

[2] T. Kourti, "Application of latent variable methods to process control and multivariate statistical process control in industry," *Int. J. Adapt. Control Signal Process.*, vol. 19, no. 4, pp. 213–246, 2005.

[3] M. A. A. Rad and M. J. Yazdanpanah, "Designing supervised local neural network classifiers based on EM clustering for fault diagnosis of Tennessee Eastman process," *Chemometrics Intell. Lab. Syst.*, vol. 146, pp. 149–157, Aug. 2015.

[4] H. Zhao, J. Liu, W. Dong, X. Sun, and Y. Ji, "An improved case-based reasoning method and its application on fault diagnosis of Tennessee Eastman process," *Neurocomputing*, vol. 249, pp. 266–276, Aug. 2017.

[5] A. Bathelt, N. L. Ricker, and M. Jelali, "Revision of the Tennessee Eastman process model," *IFAC-PapersOnLine*, vol. 48, no. 8, pp. 309–314, 2015.

[6] S. J. Qin, "Statistical process monitoring: Basics and beyond," *J. Chemometrics*, vol. 17, nos. 8–9, pp. 480–502, 2003.

[7] G. Li, S. J. Qin, and D. Zhou, "A new method of dynamic latent-variable modeling for process monitoring," *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6438–6445, Nov. 2014.

[8] B. Chebel-Morello, S. Malinowski, and H. Senoussi, "Feature selection for fault detection systems: Application to the Tennessee Eastman process," *Appl. Intell.*, vol. 44, no. 1, pp. 111–122, 2016.

[9] S. Cho and J. Jiang, "Optimal fault classification using Fisher discriminant analysis in the parity space for applications to NPPs," *IEEE Trans. Nucl. Sci.*, vol. 65, no. 3, pp. 856–865, Mar. 2018.

[10] Q. Jiang and B. Huang, "Distributed monitoring for large-scale processes based on multivariate statistical analysis and Bayesian method," *J. Process Control*, vol. 46, pp. 75–83, Oct. 2016.

[11] N. M. Nor, M. A. Hussain, and C. R. C. Hassan, "Fault diagnosis and classification framework using multi-scale classification based on kernel Fisher discriminant analysis for chemical process system," *Appl. Soft Comput.*, vol. 61, pp. 959–972, Dec. 2017.

[12] Q. Jiang, B. Huang, and X. Yan, "GMM and optimal principal components-based Bayesian method for multimode fault diagnosis," *Comput. Chem. Eng.*, vol. 84, pp. 338–349, Jan. 2016.

[13] L. H. Chiang, E. L. Russell, and R. D. Braatz, "Fault diagnosis in chemical processes using Fisher discriminant analysis, discriminant partial least squares, and principal component analysis," *Chemometrics Intell. Lab. Syst.*, vol. 50, no. 2, pp. 243–252, Mar. 2000.

[14] L. Xie, X. Lin, and J. Zeng, "Shrinking principal component analysis for enhanced process monitoring and fault isolation," *Ind. Eng. Chem. Res.*, vol. 52, no. 49, pp. 17475–17486, 2013.

[15] Y. Zhou, J. Hahn, and M. S. Mannan, "Process monitoring based on classification tree and discriminant analysis," *Rel. Eng. Syst. Saf.*, vol. 91, no. 5, pp. 546–555, 2006.
[16] T. Rato, M. Reis, E. Schmitt, M. Hubert, and B. De Ketelaere, "A systematic comparison of PCA-based statistical process monitoring methods for high-dimensional, time-dependent processes," *AIChE J.*, vol. 62, no. 5, pp. 1478–1493, 2016.
[17] Y. Wang and C. Zhao, "Probabilistic fault diagnosis method based on the combination of nest-loop fisher discriminant analysis and analysis of relative changes," *Control Eng. Pract.*, vol. 68, pp. 32–45, Nov. 2017.
[18] F. de Assis Boldt, T. W. Rauber, and F. M. Varejão, "Cascade feature selection and ELM for automatic fault diagnosis of the Tennessee Eastman process," *Neurocomputing*, vol. 239, pp. 238–248, May 2017.
[19] D. L. Swets and J. Weng, "Using discriminant eigenfeatures for image retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 8, pp. 831–836, Aug. 1996.
[20] J. E. Jackson and G. S. Mudholkar, "Control procedures for residuals associated with principal component analysis," *Technometrics*, vol. 21, no. 3, pp. 341–349, Aug. 1979.
[21] Z. Zhang and J. Zhao, "A deep belief network based fault diagnosis model for complex chemical processes," *Comput. Chem. Eng.*, vol. 107, pp. 395–407, Dec. 2017.
[22] X. B. He, W. Wang, Y. P. Yang, and Y. H. Yang, "Variable-weighted Fisher discriminant analysis for process fault diagnosis," *J. Process Control*, vol. 19, no. 6, pp. 923–931, 2009.
[23] J. J. Downs and E. F. Vogel, "A plant-wide industrial process control problem," *Comput. Chem. Eng.*, vol. 17, no. 3, pp. 245–255, Mar. 1993.
[24] G. Li, K. Peng, T. Yuan, and M. Zhong, "Kernel dynamic latent variable model for process monitoring with application to hot strip mill process," *Chemometrics Intell. Lab. Syst.*, vol. 171, pp. 218–225, Dec. 2017.
[25] F. Akbaryan and P. R. Bishnoi, "Fault diagnosis of multivariate systems using pattern recognition and multisensor data analysis technique," *Comput. Chem. Eng.*, vol. 25, nos. 9–10, pp. 1313–1339, 2001.
[26] C. K. Lau, K. Ghosh, M. A. Hussain, and C. R. C. Hassan, "Fault diagnosis of Tennessee Eastman process with multi-scale PCA and ANFIS," *Chemometrics Intell. Lab. Syst.*, vol. 120, pp. 1–14, Jan. 2013.
[27] G. S. Chadha and A. Schwung, "Comparison of deep neural network architectures for fault detection in Tennessee Eastman process," in *Proc. 22nd IEEE Int. Conf. Emerg. Technol. Factory Autom. (ETFA)*, Sep. 2017, pp. 1–8.
[28] M. F. S. V. D'Angelo, R. M. Palhares, M. C. O. C. Filho, R. D. Maia, J. B. Mendes, and P. Y. Ekel, "A new fault classification approach applied to Tennessee Eastman benchmark process," *Appl. Soft Comput.*, vol. 49, pp. 676–686, Dec. 2016.
[29] H. Sun, S. Zhang, C. Zhao, and F. Gao, "A sparse reconstruction strategy for online fault diagnosis in nonstationary processes with no *a priori* fault information," *Ind. Eng. Chem. Res.*, vol. 56, no. 24, pp. 6993–7008, 2017.
[30] L. H. Chiang and R. J. Pell, "Genetic algorithms combined with discriminant analysis for key variable identification," *J. Process Control*, vol. 14, no. 2, pp. 143–155, 2004.
[31] C. Bouveyron and C. Brunet, "Probabilistic fisher discriminant analysis: A robust and flexible alternative to Fisher discriminant analysis," *Neurocomputing*, vol. 90, pp. 12–22, Aug. 2012.
[32] X. Gao and J. Hou, "An improved SVM integrated GS-PCA fault diagnosis approach of Tennessee Eastman process," *Neurocomputing*, vol. 174, pp. 906–911, Jan. 2016.
[33] W.-S. Chen, P. C. Yuen, J. Huang, and D.-Q. Dai, "Kernel machine-based one-parameter regularized fisher discriminant method for face recognition," *IEEE Trans. Syst., Man, Cybern. B. Cybern.*, vol. 35, no. 4, pp. 659–669, Aug. 2005.
[34] Y. Xu and X. Deng, "Fault detection of multimode non-Gaussian dynamic process using dynamic Bayesian independent component analysis," *Neurocomputing*, vol. 200, pp. 70–79, Aug. 2016.
[35] C. Xinyi and Y. Xuefeng, "Fault diagnosis in chemical process based on self-organizing map integrated with Fisher discriminant analysis," *Chin. J. Chem. Eng.*, vol. 21, no. 4, pp. 382–387, 2013.
[36] C. Bouveyron and C. Brunet, "Simultaneous model-based clustering and visualization in the Fisher discriminative subspace," *Statist. Comput.*, vol. 22, no. 1, pp. 301–324, 2012.
[37] Z. Zhang, G. Dai, and M. I. Jordan, "A flexible and efficient algorithm for regularized fisher discriminant analysis," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*. Springer, 2009, pp. 632–647.
[38] J. Wang, Q. Li, J. You, and Q. Zhao, "Fast kernel Fisher discriminant analysis via approximating the kernel principal component analysis," *Neurocomputing*, vol. 74, no. 17, pp. 3313–3322, 2011.
[39] K. Ghosh, M. Ramteke, and R. Srinivasan, "Optimal variable selection for effective statistical process monitoring," *Comput. Chem. Eng.*, vol. 60, pp. 260–276, Jan. 2014.
[40] N. L. Ricker, "Tennessee eastman challenge archive," Tech. Rep., 1999.

XIAOHUI YANG received the B.Sc. M.Sc., and Ph.D. degrees from Nanchang University, Nanchang, China, in 2003, 2006, and 2015, respectively, where he has been with the Department of Electronic Information Engineering, School of Information Engineering, since 2006, and is also an Associate Professor. He has published over 30 research articles. His current interests include intelligent control, process control, fault diagnosis, and stochastic nonlinear systems.

SONGHONG RUI received the B.E. degree in electrical engineering and its automation from Xinyu University, China, in 2018. He is currently pursuing the M.S. degree in electrical engineering with Nanchang University. His research interest includes motor control strategy.

XIAOLONG ZHANG received the B.E. degree in electrical engineering and its automation from Hefei Normal University, Hefei, China, in 2017. He is currently pursuing the M.E. degree in electrical engineering with Nanchang University. His research interests include dynamics and simulation of mechanical systems, robotics, and constrained systems.

SHAOPING XU received the M.S. degree in computer application from the China University of Geosciences, Wuhan, China, in 2004, and the Ph.D. degree in mechatronics engineering from Nanchang University, Nanchang, China, in 2010, where he is currently a Professor with the Department of Computer Science, School of Information Engineering. He has published more than 30 research articles and serves as a Reviewer for several journals including the IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT. His current research interests include digital image processing and analysis, computer graphics, virtual reality, and surgery simulation.

CHUNSHENG YANG received the B.Sc. degree (Hons.) in electronic engineering from Harbin Engineering University, China, the M.Sc. degree in computer science from Shanghai Jiao Tong University, China, and the Ph.D. degree from National Hiroshima University, Japan. He was with Fujitsu Inc., Japan, as a Senior Engineer, and engaged on the development of ATM Network Management Systems. He was an Assistant Professor with Shanghai Jiao Tong University, from 1986 to 1990, working on hypercube distributed computer systems. He is interested in data mining, machine learning, prognostic health management, reasoning technologies such as case-based reasoning, rule-based reasoning and hybrid reasoning, multi-agent systems, and distributed computing. He was a Program Co-Chair for the 17th International Conference on Industry and Engineering Applications of Artificial Intelligence and Expert Systems. He has served as a Program Committee for many conferences and institutions, and has been a Reviewer for many conferences, journals, and organizations, including *Applied Artificial Intelligence*, NSERC, IEEE TRANSACTIONS, ACM KDD, PAKDD, and AAMAS. He is a Guest Editor for the *International Journal of Applied Intelligence*.

PETER X. LIU received the B.Sc. and M.Sc. degrees from Northern Jiaotong University, China, in 1992 and 1995, respectively, and the Ph.D. degree from the University of Alberta, Canada, in 2002. He has been with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada, since 2002, where he is currently a Professor and the Canada Research Chair with the Department of Systems and Computer Engineering. He is also with the School of Information Engineering, Nanchang University, as an Adjunct Professor. He has published more than 280 research articles. His research interests include interactive networked systems and teleoperation, haptics, micromanipulation, robotics, intelligent systems, context-aware intelligent networks, and their applications to biomedical engineering. He is a Licensed Member of the Professional Engineers of Ontario, and a Fellow of IEEE and the Engineering Institute of Canada. He received the 2003 Province of Ontario Distinguished Researcher Award, the 2006 Province of Ontario Early Researcher Award, the 2006 Carty Research Fellowship, the Best Conference Paper Award of the 2006 IEEE International Conference on Mechatronics and Automation, and the 2007 Carleton Research Achievement Award. He serves as an Associate Editor for several journals including the IEEE TRANSACTIONS ON CYBERNETICS, the IEEE/ASME TRANSACTIONS ON MECHATRONICS, the IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING, and the IEEE ACCESS.