IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON DATA-DRIVEN MONITORING, FAULT DIAGNOSIS AND CONTROL OF CYBER-PHYSICAL SYSTEMS

Received December 14, 2017, accepted January 20, 2018, date of publication January 29, 2018, date of current version March 13, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2798278

Temporal-Spatial Global Locality Projections for Multimode Process Monitoring

BING SONG AND HONGBO SHI

Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China

Corresponding author: Hongbo Shi (hbshi@ecust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61703161, Grant 61374140, and Grant 61673173, in part by the Fundamental Research Funds for the Central Universities under Grant 222201714031 and Grant 222201717006, and in part by the China Postdoctoral Science Foundation under Grant 2017M611472.

ABSTRACT Multimode is an important feature of modern processes, since various manufacturing strategies are needed to satisfy different demands of markets. Direct application of traditional multivariate statistical process monitoring methods cannot obtain satisfactory results, as the data set collected from multimode processes always follows multimodal distribution. To construct a single model which can monitor multimode processes directly, this paper proposes an original algorithm named temporal–spatial global locality projections. First, given that both temporal and spatial neighbors can express the similarity, the determination of the neighborhood is conducted in both the temporal and spatial scale. Second, an optimization objective function which preserves not only the local structure but also the global structure is defined. Third, the monitoring statistic is established via the local outlier factor. To certify the effectiveness, a numerical example, the multimode Tennessee Eastman process, and the CE117 process which is proposed by TecQuipment for process control are studied.

INDEX TERMS Multimode process monitoring, temporal and spatial scale, local and global structure, temporal-spatial global locality projections.

I. INTRODUCTION

In modern industrial processes, to ensure production safety and improve product quality, process monitoring plays an important role [1]–[4]. With the extensive application of distributed control system (DCS), a large amount of data can be easily collected and stored. Therefore, MSPM methods gain more and more attention [5]–[9]. Due to different production plans, the change of raw materials, the diversification of market demands and the change of external environments, modern processes often have more than one operating condition. The dataset collected from the multimode process always follows multimodal distribution. Meanwhile, traditional MSPM methods such as principal component analysis (PCA) and partial least squares (PLS) require that the dataset follows unimodal distribution [10], [11].

Up to now, many methods have been proposed to monitor multimode processes. The most primitive idea for monitoring multimode processes is to construct one model for every mode. This kind of method is called multiple models method. Except establishing the model for every mode, multiple models method requires two extra steps. Specifically, in the offline modeling phase, the multimode dataset needs to be divided into multiple datasets corresponding to different modes. Moreover, in the online monitoring phase, a rule should be designed to determine the final result. To label the transition states, a clustering algorithm on the basis of dynamic k-principal component analysis-independent component analysis is designed [12]. Song *et al.* [13] proposed multi-subspace PCA with LOF to monitor multimode processes, where a novel clustering algorithm was designed. Through making an adequate projection, the mode-common subspace and multiple mode-specific subspaces are constructed [14]. Zhu *et al.* [15] developed a novel framework to construct process pattern and monitor multimode process. For a dynamic system with multiple modes, the state-space representation with different model parameters was established to characterize each mode [16]. To monitor multimode batch processes online, a concurrent phase partition and between-mode statistical modeling strategy was developed [17]. Tan *et al.* [18] developed a mode identification method on the basis of the resemblance of data characteristic. Based on the linear subspace and the Bayesian inference,

a two-dimensional Bayesian method was proposed for multimode processes with nonlinear characteristic [19]. Guo *et al.* [20] proposed the local neighbor normalized matrix method which can capture the nonlinear relationship between modes and within modes. According to the value of squared prediction error, Zhao *et al.* [21] selected the right model for the current data in multimode processes. For results combination in multimode processes monitoring, a novel probabilistic scheme was developed by Ge and Song [22].

In addition, the other multimode processes monitoring method is to establish a single model. This kind of method is called single model method. Compared with multiple models method, single model method never needs to conduct mode division and final result determination. The key of using the single model to monitor multimode processes is how to make the single model include different modes information. Ge and Song [23] applied the just-in-time learning strategy to deal with the multiple modes problem. Using the mode unfolding scheme, a single PCA model was established for processes with multiple modes [24]. For nonlinear multimode processes with between-mode transition, the LL-SVDD-MRDA method was developed where a local model was generated via the lazy learning algorithm [25]. To deal with the robust multimode process modeling, a novel Bayesian robust mixture factor analyzer was proposed [26]. Given that the single mode dataset can be described by the Gaussian distribution or the hidden state, the Gaussian mixture model (GMM) and hidden Markov model (HMM) have been applied to characterize multimode processes [27]–[29]. On the basis of the Expectation Maximization algorithm, a mixture model was developed for handling the multimode problem [30].

Apart from traditional MSPM methods, manifold learning algorithms which do not require data following unimodal distribution have been applied in process monitoring successfully. The widely used four algorithms are neighborhood preserving projections (NPE), locally linear embedding (LLE), locality preserving projections (LPP), and laplacian eigenmaps (LE) [31], [32]. Meanwhile, there are two issues in these original manifold learning algorithms. On one hand, the selection of the neighborhood is always limited in the spatial scale, and the temporal scale is neglected. On the other hand, these algorithms only focus on the local structure, and the global structure is undermined. Luo [33] developed a new algorithm called global-local preserving projections (GLPP) which solves a dual-objective optimization. Ma *et al.* [34] proposed the local and nonlocal embedding (LNLE) through minimizing distances among neighbors and maximizing distances among nonlocal samples. Based on the constructed dual weight matrix and enhanced objective function, the enhanced neighborhood preserving embedding (ENPE) was presented [35]. An unsupervised method called nonlocal structure constrained neighborhood preserving embedding (NSC-NPE) was proposed through minimizing the local scatter and maximizing the nonlocal scatter [36]. Zhang *et al.* [37] put forward the global-local structure analysis (GLSA) by combining advantages of LPP and PCA. Yu [38] proposed the local and global PCA (LGPCA) algorithm preserving both the local and global information. By employing the neighborhood embedding in both the global and local graph, the multi-manifold projection (MMP) algorithm was developed [39]. To explore the dynamic information, Miao *et al.* [40] presented the time neighborhood preserving embedding (TNPE). To solve the suppression of useful information, the weighted neighborhood preserving embedding and support vector data description (WNPE-SVDD) was developed [41]. Song *et al.* [42] proposed the improved dynamic neighborhood preserving embedding (IDNPE) to monitor multimode processes. In contrast to traditional MSPM methods, the greatest advantage of these manifold learning algorithms is that there is no assumption on the data distribution. Therefore, direct application of these algorithms to monitor multimode processes is reasonable.

To monitor multimode processes via one single model, an original dimensionality reduction algorithm named temporal-spatial global locality projections (TSGLP) is proposed in this paper. Firstly, to mine information in both the temporal and spatial scale, the neighborhood is constituted through aggregating temporal and spatial neighbors. Secondly, in multimode processes, the within-mode relationship can be represented by the local structure. Correspondingly, the mode-to-mode relationship can be represented by the global structure. An optimization objective function preserving both the local and global structure is established. Thirdly, considering that the dataset may disobey some particular distributions, the monitoring statistic is constructed using LOF which has no requirements on the data distribution. Compared to establishing multiple models for multimode processes, the advantage of establishing the single model is that the relevance between different modes can be taken into consideration. Finally, the effectiveness of the TSGLP method is proved through comparing with other process monitoring methods under a numerical example, the multimode TE process which is widely used to test process monitoring performance, and the CE117 process which is a fully integrated and self-contained bench top process control apparatus containing valves, pumps, power supplies and ancillaries to allow flow, level, temperature and pressure control strategies to be investigated individually and in combinations.

II. PRELIMINARIES

A. LOCALITY PRESERVING PROJECTIONS (LPP)

LPP is one of manifold learning algorithms which can reduce the dimensionality of the original dataset through finding the manifold structure [43]. On the basis of the local information, the dataset $\boldsymbol{X}^T = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\} \in \boldsymbol{R}^{m\times n}$ (n is the number of data, m is the dimensionality) is transformed to the low-dimensional manifold $\boldsymbol{Y}^T \in \boldsymbol{R}^{d\times n}(d < m)$ using LPP. The steps are shown as follows.

1) Establishing the adjacency graph: Search k neighbors of $\boldsymbol{x}_i$ to compose the neighborhood. For the original dataset $\boldsymbol{X}$, an edge is set from i to j if $\boldsymbol{x}_j$ belongs to the neighborhood of $\boldsymbol{x}_i$.

2) Calculating the weight in each edge: Let W_{ij} represents the weight value in the edge between i to j. According to the Gaussian kernel, W_{ij} can be computed in (1).

$$W_{ij} = \begin{cases} e^{-\frac{||\boldsymbol{x}_i - \boldsymbol{x}_j||^2}{\sigma}} & \text{if } \boldsymbol{x}_j \text{ is a neighbor of } \boldsymbol{x}_i \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

3) Eigenmaps: Obtain the projection matrix $\boldsymbol{A}$ which is constituted by eigenvectors of (2).

$$\boldsymbol{X}^T \boldsymbol{L} \boldsymbol{X} \boldsymbol{\alpha} = \lambda \boldsymbol{X}^T \boldsymbol{D} \boldsymbol{X} \boldsymbol{\alpha} \quad (2)$$

$$D_{ii} = \sum_j W_{ji}, \quad \boldsymbol{L} = \boldsymbol{D} - \boldsymbol{W} \quad (3)$$

III. TSGLP FOR MULTIMODE PROCESS MONITORING

A. TSGLP

The objective of the proposed TSGLP algorithm is to obtain the projection matrix which can project the original space to the low-dimensional feature space. In the process of feature extraction based on TSGLP, the information contained in both the temporal and spatial scale can be excavated and employed. Moreover, both the local and global structure of the original dataset can be preserved. In order to mine information in both the temporal and spatial scale, a union form of temporal and spatial neighbors is determined to construct the final neighborhood. Considering that some temporal neighbors may belong to spatial neighbors, the number of neighbors for each data may be different. For the purpose of preserving both the local and global structure, an optimization function is established on the basis of the Gaussian kernel distance and the proposed weighted mutual information (WMI). The Gaussian kernel distance is applied to descript the distance between the current data and its neighbors. Correspondingly, the WMI which can express the correlation of the entire multimode dataset is developed to represent the global structure.

For multimode processes, the within-mode data can be represented using neighbors from the aspect of the spatial. The mode-to-mode data can be represented using neighbors from the aspect of the temporal. Thus, to construct the monitoring model for multimode processes, mining both the temporal and spatial information is necessary. In the temporal scale, the dataset collected from multimode process includes the sequence of sampling. In the spatial scale, the dataset collected from multimode process contains the information of topology and distance.

From the aspect of the temporal scale, the closer the sampling, the higher the similarity. Thus, choose k_1 adjacent data before the current data and k_1 adjacent data after the current data to establish the temporal neighborhood N_t. In order to make the current data and its temporal neighborhood N_t belong to the same mode, the k-means clustering or the clustering algorithm developed by our previous work [13] is used. Suppose the original dataset $\boldsymbol{X} \in R^{n \times m}$ is clustered into C sub-datasets as

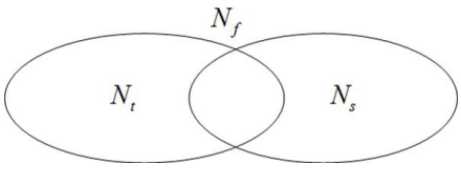

FIGURE 1. The final neighborhood in the TSGLP.

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}_1^T \\ \boldsymbol{x}_2^T \\ \vdots \\ \boldsymbol{x}_n^T \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \\ \vdots \\ \boldsymbol{X}_C \end{bmatrix} \in R^{n \times m} \quad (4)$$

$$\boldsymbol{X}_c \in R^{n_c \times m} \quad (c = 1, 2, \cdots, C) \quad (5)$$

Specifically, for a data $\boldsymbol{x}_i$ in $\boldsymbol{X}$, the temporal neighborhood N_t is $N_t(\boldsymbol{x}_i) = \{\boldsymbol{x}_{i-k_1}, \cdots, \boldsymbol{x}_{i-1}, \boldsymbol{x}_{i+1}, \cdots, \boldsymbol{x}_{i+k_1}\}$ if $\boldsymbol{x}_i$ does not belong to k_1 mode starting data and k_1 mode ending data. Otherwise, the temporal neighborhood N_t is composed by $2k_1$ nearest data from the view of sampling. Thus, the temporal neighborhood N_t is established based on those data which belong to the same mode with the current data. As a result, the number of data in N_t is $2k_1$.

Correspondingly, from the aspect of the spatial scale, the closer the distance, the higher the similarity. Therefore, select k_2 nearest data to construct the spatial neighborhood N_s.

Finally, the final neighborhood N_f in the TSGLP algorithm is determined as follows

$$N_f = N_t \cup N_s \quad (6)$$

Fig. 1 depicts the relationship between N_f and N_t, N_s. From this figure, we can know that the number of neighbors in N_f may be smaller than $2k_1 + k_2$.

Moreover, for multimode processes, different single modes are not independent of each other. Only considering the single mode may ignore the correlation between different modes and destroy the global structure of the dataset. To describe the multimode process more complete, in addition to extract the local structure information of every mode, the global structure information of the entire dataset should also be focused on.

To characterize the local structure, the Gaussian kernel is used to depict the distance between the data and its neighbors. In addition, the Gaussian kernel distance is maintained from the original space to the feature space. In other words, the neighbor relationship contained in the original dataset is preserved. The specific approach in TSGLP is to construct the weight matrix $\boldsymbol{W}$ as follows:

$$W_{ij} = \begin{cases} e^{-\frac{||\boldsymbol{x}_i - \boldsymbol{x}_j||^2}{\sigma}} & \text{if } \boldsymbol{x}_j \text{ is a neighbor of } \boldsymbol{x}_i \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

To describe the global structure, the WMI is proposed to represent the correlation of the entire dataset. Moreover,

the WMI relationship is hold from the original space to the feature space. In other words, the variable correlation contained in the original dataset is preserved. The details of MI are introduced as follows [44]:

$$I(\boldsymbol{x}^i, \boldsymbol{x}^j) = \int_{\boldsymbol{x}^i, \boldsymbol{x}^j} p(\boldsymbol{x}^i, \boldsymbol{x}^j) \log(\frac{p(\boldsymbol{x}^i, \boldsymbol{x}^j)}{p(\boldsymbol{x}^i)p(\boldsymbol{x}^j)}) d\boldsymbol{x}^i d\boldsymbol{x}^j \quad (8)$$

where $I(\boldsymbol{x}^i, \boldsymbol{x}^j)$ is the MI value between two variables $\boldsymbol{x}^i$ and $\boldsymbol{x}^j$, $p(\boldsymbol{x}^i, \boldsymbol{x}^j)$ is the joint probability density function, $p(\boldsymbol{x}^i)$ is the marginal probability density function. Then, the mutual information matrix of the dataset $\boldsymbol{X}_c$ $(c = 1, 2, \cdots, C)$ in the mode c is expressed as $MI(\boldsymbol{X}_c)$. It needs to be emphasized that the value of mutual information is large if the relationship between two different variables is close.

In this work, the global structure of mode c is represented by the mutual information matrix $MI(\boldsymbol{X}_c)$. For the entire multimode dataset, the global structure is represented based on every global structure of single mode. In addition, the scale of $MI(\boldsymbol{X}_c)$ $(c = 1, 2 \cdots, C)$ is different. In order to consider the scale of different single mode, the WMI of the original dataset $\boldsymbol{X}$ is defined as

$$\boldsymbol{WMI} = \sum_{i=1}^{C} \frac{\sum_{c=1}^{C} \rho\left[MI(\boldsymbol{X}_c)\right] - \rho\left[MI(\boldsymbol{X}_i)\right]}{\sum_{c=1}^{C} \rho\left[MI(\boldsymbol{X}_c)\right]} MI(\boldsymbol{X}_i) \quad (9)$$

where $\rho\left[MI(\boldsymbol{X}_c)\right]$ $(c = 1, 2, \cdots, C)$ is the spectral radius of $MI(\boldsymbol{X}_c)$. As shown in (9), the larger the scale of mode c, the smaller the weight coefficient.

The objective function of TSGLP is constructed as:

$$\begin{aligned} J_{local} &= \min \frac{1}{2} \sum_{ij} (\boldsymbol{y}_i - \boldsymbol{y}_j)^2 W_{ij} \\ &= \min \frac{1}{2} \sum_{ij} (\boldsymbol{A}^T \boldsymbol{x}_i - \boldsymbol{A}^T \boldsymbol{x}_j)^2 W_{ij} \\ &= \min \sum_{i} \boldsymbol{A}^T \boldsymbol{x}_i D_{ii} \boldsymbol{x}_i^T \boldsymbol{A} - \sum_{ij} \boldsymbol{A}^T \boldsymbol{x}_i W_{ij} \boldsymbol{x}_j^T \boldsymbol{A} \\ &= \min \boldsymbol{A}^T \boldsymbol{X}^T (\boldsymbol{D} - \boldsymbol{W}) \boldsymbol{X} \boldsymbol{A} \\ D_{ii} &= \sum_{j} W_{ji} \end{aligned} \quad (10)$$

$$J_{global} = \max \boldsymbol{A}^T (\boldsymbol{WMI}) \boldsymbol{A} \quad (11)$$

$$J = \frac{J_{glabal}}{J_{local}} \quad (12)$$

Then, the following constraint is imposed on the objective function:

$$\boldsymbol{A}^T \boldsymbol{X}^T (\boldsymbol{D} - \boldsymbol{W}) \boldsymbol{X} \boldsymbol{A} = \boldsymbol{I} \quad (13)$$

Based on the Lagrangian multiplier λ, the projection matrix $\boldsymbol{A}$ can be computed by solving (14).

$$(\boldsymbol{WMI}) \boldsymbol{A} = \lambda \boldsymbol{X}^T (\boldsymbol{D} - \boldsymbol{W}) \boldsymbol{X} \boldsymbol{A} \quad (14)$$

Then, the low-dimensional data $\boldsymbol{y}_i$ is determined as:

$$\boldsymbol{y}_i = \boldsymbol{A}^T \boldsymbol{x}_i \quad (15)$$

1) COMPARISONS

There are mainly two differences between the LPP and the proposed TSGLP. On one hand, in the process of constructing the neighborhood, LPP establishes the neighborhood based on the distance from the aspect of the spatial scale. Meanwhile, both the temporal and spatial neighborhood are built in the proposed TSGLP. Therefore, compared with LPP, the information mining in TSGLP is more adequate. On the other hand, in the process of constructing the objective function, LPP only focuses on the local structure of the original dataset. Meanwhile, not only the local structure but also the global structure is preserved in the proposed TSGLP.

Compared with GLSA [37] and LGPCA [38] which establish the global-local monitoring model, the proposed TSGLP method has three differences. Firstly, both GLSA and LGPCA methods are developed for monitoring the single mode process, the TSGLP method is proposed to monitor the multimode process. Secondly, the covariance matrix is used to represent the global structure in GLSA and LGPCA. The TSGLP method employs the developed WMI to represent the global structure. In contrast to the covariance matrix which only can express the linear correlation, the mutual information is not limited to linear relationship. Thirdly, only the spatial information is excavated in both GLSA and LGPCA. The TSGLP method applies temporal and spatial information. Moreover, the TNPE method employs the information in the scale of temporal. However, TNPE ignores the information in the scale of spatial and the global structure.

2) PARAMETERS

There are three important parameters $k1, k2, d$ in the proposed TSGLP. $2k_1$ is the number of neighbors in the temporal neighborhood. k_2 is the number of neighbors in the spatial neighborhood. If k_1 is large, the information contained in the temporal scale is emphasized. On the contrary, the information contained in the spatial scale is emphasized. d represents the dimension of the feature space, and it can be determined using the cross-validation method.

B. MONITORING STATISTIC AND ITS CONTROL LIMIT

Given that the training dataset $\boldsymbol{X}$ is collected from the multimode process, the obtained low-dimensional dataset $\boldsymbol{Y}$ always follows multimodal distribution. Traditional monitoring statistics such as T^2 and SPE cannot depict the data following multimodal distribution. In TSGLP, the LOF method is used to establish the monitoring statistic. The extent that a data is regarded as an outlier can be indicated via the LOF value [45]. The steps are listed as follows.

1) Constructing the neighborhood: For every data $\boldsymbol{y}_i$ in the dataset $\boldsymbol{Y} \in \boldsymbol{R}^{n \times d}$, find K nearest neighbors to constitute the neighborhood.
2) Defining the $K\_dis\tan ce$: Define the radius of the neighborhood as $K\_dis\tan ce\left(\boldsymbol{y}_i\right)$.

3) Determining the reachability distance: The reachability distance $reachd\left(\boldsymbol{y}_i, \boldsymbol{y}_i^f\right)$ is determined as

$$reachd\left(\boldsymbol{y}_i, \boldsymbol{y}_i^f\right) = \max\left\{K\_dis\tan ce(\boldsymbol{y}_i^f), d\left(\boldsymbol{y}_i, \boldsymbol{y}_i^f\right)\right\}, \quad f = 1, 2, \ldots, K \tag{16}$$

4) Computing the local reachability density: The local reachability density (LRD) of $\boldsymbol{y}_i$ is computed as

$$LRD\left(\boldsymbol{y}_i\right) = \frac{K}{\sum_{f=1}^{K} reach\_d\left(\boldsymbol{y}_i, \boldsymbol{y}_i^f\right)} \tag{17}$$

5) Calculating the local outlier factor: The LOF of $\boldsymbol{y}_i$ is given as

$$LOF\left(\boldsymbol{y}_i\right) = \frac{1}{K}\sum_{f=1}^{K}\frac{LRD\left(\boldsymbol{y}_i^f\right)}{LRD\left(\boldsymbol{y}_i\right)} \tag{18}$$

As shown above, the calculation of LOF is based on the density information. Therefore, the LOF statistic is not limited by the data distribution. For a testing data $\boldsymbol{x}_t$, the low-dimensional data is obtained according to $\boldsymbol{y}_t = \boldsymbol{A}^T\boldsymbol{x}_t$. Then, the neighborhood of $\boldsymbol{y}_t$ is constructed using the normal data in $\boldsymbol{Y}$. If the testing data is normal, the density of $\boldsymbol{y}_t$ is similar to that of data in $\boldsymbol{Y}$. Thus, the LOF value $LOF(\boldsymbol{y}_t)$ is close to 1. On the contrary, the density of $\boldsymbol{y}_t$ is different from that of data in $\boldsymbol{Y}$. Thus, the LOF value $LOF(\boldsymbol{y}_t)$ is larger than 1 to a large extent.

Because the LOF value of normal data is close to 1, the control limit can be determined as $1\pm\varepsilon$ where ε is approximately equal to 0. In order to make the control limit more accurate, the kernel density estimation (KDE) method is applied to estimate the control limit. A univariate kernel estimator is shown as:

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

where $\hat{f}_h(\boldsymbol{y})$ is the estimated probability density, n is the number of data and h is the smoothing parameter. K is the kernel function, where Gaussian kernel is used in this work.

C. THE PROCEDURE OF TSGLP

Offline modeling phase:

1) Acquire the normal training dataset $\boldsymbol{X}$ which is collected from the multimode process.
2) Preprocess $\boldsymbol{X}$ to zero mean and unit variance via the z-score method.
3) Construct the temporal neighborhood N_t for every data in $\boldsymbol{X}$.
4) Establish the spatial neighborhood N_s for every data in $\boldsymbol{X}$.
5) Determine the final neighborhood N_f for every data in $\boldsymbol{X}$ according to (6).
6) Compute the weight matrix $\boldsymbol{W}$ based on (7) to describe the local structure.
7) Calculate the WMI to depict the global structure.
8) Build the objective function according to (13).
9) Obtain the projection matrix $\boldsymbol{A}$ via (14).
10) Construct the LOF statistic and estimate the control limit.

Online monitoring phase:

1) Acquire the testing data $\boldsymbol{x}_t$.
2) Preprocess $\boldsymbol{x}_t$ using the mean and the variance of the normal training dataset.
3) Project $\boldsymbol{x}_t$ onto the feature space applying $\boldsymbol{y}_t = \boldsymbol{A}^T\boldsymbol{x}_t$.
4) Calculate the monitoring statistic $LOF(\boldsymbol{y}_t)$.
5) Judge $LOF(\boldsymbol{y}_t)$ exceeds the control limit or not.

IV. EXAMPLES AND APPLICATIONS

To show the advantage of the proposed TSGLP, a numerical example, the multimode TE process, and the multimode CE117 process are applied to test the monitoring performance. The TSGLP is compared with the LPP and GLSA [37] methods. The monitoring results of all three methods are calculated based on the 99% control limit.

A. NUMERICAL EXAMPLE

A numerical example is designed in this work as follows:

$$\begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix} = \begin{bmatrix} 0.811 & -0.226 \\ 0.477 & 0.415 \end{bmatrix} \cdot \begin{bmatrix} u_1(k-1) \\ u_2(k-1) \end{bmatrix} + \begin{bmatrix} 0.193 & 0.689 \\ -0.320 & -0.749 \end{bmatrix} \cdot \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix}; \tag{20}$$

$$\begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} = \begin{bmatrix} 0.118 & -0.191 & 0.287 \\ 0.847 & 0.264 & 0.943 \\ -0.333 & 0.514 & -0.217 \end{bmatrix} \cdot \begin{bmatrix} x_1(k-1) \\ x_2(k-1) \\ x_3(k-1) \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & -4 \\ -2 & 1 \end{bmatrix} \cdot \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix}; \tag{21}$$

$$\begin{bmatrix} y_1(k) \\ y_2(k) \\ y_3(k) \end{bmatrix} = \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} v_1(k) \\ v_2(k) \\ v_3(k) \end{bmatrix}; \tag{22}$$

It includes eight variables $u_1, u_2, x_1, x_2, x_3, y_1, y_2, y_3$ which are produced via five factors v_1, v_2, v_3, w_1, w_2. In this work, two different operating modes are set as follows.

$$\text{Mode 1}\ \begin{matrix} v_1\ N(-0.1, 0.1); \\ v_2\ N(-0.1, 0.1); \\ v_3\ N(-0.1, 0.1); \\ w_1\ N(1, 2); \\ w_2\ N(1, 2); \end{matrix} \tag{23}$$

$$\text{Mode 2}\ \begin{matrix} v_1\ N(0.1, 0.1); \\ v_2\ N(0.1, 0.1); \\ v_3\ N(0.1, 0.1); \\ w_1\ N(-1, 2); \\ w_2\ N(-1, 2); \end{matrix} \tag{24}$$

The normal training dataset is simulated according to (20)–(24). It contains 400 mode 1 data and 400 mode 2 data. In the proposed TSGLP, the number of neighbors in the

TABLE 1. False detection rates (%) of three methods in the testing dataset.

Testing dataset	LPP (T^2)	LPP (SPE)	GLSA (T^2)	GLSA (SPE)	TSGLP (LOF)
1	0.5	1	1.5	0.5	0
2	0	0	0	0	1

TABLE 2. Missed detections rates (%) of three methods in the testing dataset.

Testing dataset	LPP (T^2)	LPP (SPE)	GLSA (T^2)	GLSA (SPE)	TSGLP (LOF)
1	90.5	99.75	14.25	99.25	**4**
2	79.75	97.5	21.5	98.25	**16.5**

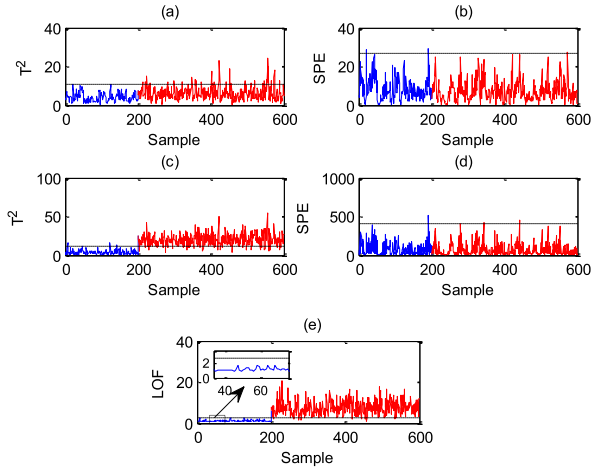

FIGURE 2. The monitoring results of three methods for the first testing dataset. (a) LPP. (b) LPP. (c) GLSA. (d) GLSA. (e) TSGLP.

temporal neighborhood $2k_1$ is 10 and that in the spatial neighborhood k_2 is 10. The dimension of the feature space d is selected as 4. For fair comparisons, the number of neighbors in the neighborhood k is 20 and the dimension of the feature space d is selected as 4 in the LPP and GLSA methods.

For the purpose of testing the monitoring performance of TSGLP, two testing datasets are simulated in the following way where first 200 data are normal and the remaining 400 data are abnormal. In the first testing dataset, the process runs under mode 2, a step of 0.5 in v_3 occurs from data 200 to the end. In the second testing dataset, the process runs under mode 2, a drift of 0.005 in v_3 occurs from data 200 to the end.

Table 1 and Table 2 list false detection rates and missed detection rates for LPP, GLSA, and TSGLP, respectively. In Table 2, the smallest missed detection rates are marked with bold. To show the monitoring performance visually, the monitoring results of LPP, GLSA and TSGLP for two testing datasets are plotted in Fig. 2 and Fig. 3.

From Table 1, we can know that false detection rates of LPP, GLSA and TSGLP for two testing datasets are smaller than 2%, and can be accepted. From Table 2, the proposed TSGLP method can obtain the best missed detection rates among three methods. Since the information in the temporal

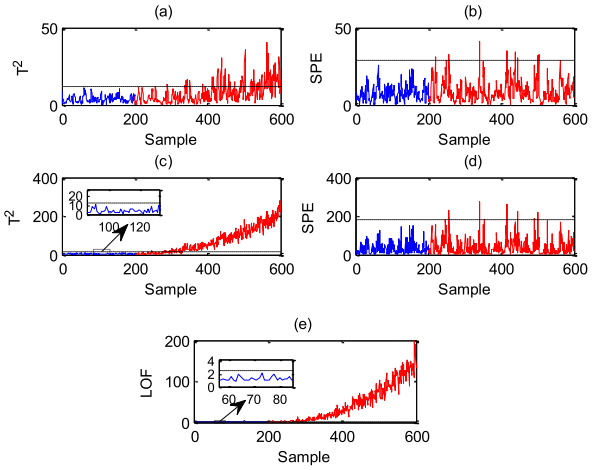

FIGURE 3. The monitoring results of three methods for the second testing dataset. (a) LPP. (b) LPP. (c) GLSA. (d) GLSA. (e) TSGLP.

scale is not mined and the global structure of the original dataset is ignored, the LPP method hardly detects the fault. Compared with LPP, the monitoring results of GLSA have been improved because both the local and global structure are preserved. Considering that the information in the temporal and spatial scale is excavated and both the local and global structure are preserved, the proposed TSGLP can acquire the best monitoring result for two testing datasets.

For the first testing dataset, almost all fault data are wrongly classified as normal ones where the missed detection rate is 90.5% and 99.75% in Fig. 2(a) and Fig. 2(b). In other words, the monitoring results of LPP (T^2) and LPP (SPE) cannot satisfy the detection requirement. From Fig. 2(c), the T^2 value of more than 10% fault data are similar to those of normal data. The missed detection rate of GLSA (T^2) is 14.25% since both the local and global structure are preserved. Compared with LPP and GLSA, the LOF statistic of TSGLP can detect this fault timely and effectively in Fig. 2(e) with the missed detection rate 4%. For the second testing dataset, LPP (SPE) cannot detect the occurrence of the fault where the missed detection rate is 97.5%. There is some progress in LPP (T^2) with the missed detection rate 79.75%. In Fig. 3(c), the T^2 of GLSA can detect this fault successfully and the missed detection rate is 21.5% because both the local and global structure is considered. In contrast to LPP and GLSA, the proposed TSGLP method can obtain the lowest missed detection rate for the second testing dataset in Fig. 3(e).

B. TE

The TE process is the simulation of an industrial chemical process [46]–[48]. Fig. 4 is the schematic graph of it. From this figure, we can know that it contains five units: a reactor, a condenser, a vapor separator, a recycle compressor, and a product stripper [49], [50]. According to the mass ratio of two desired products G/H, the TE process has six modes. In this article, 9 manipulated variables and 22 continuous process variables are chosen as monitoring variables.

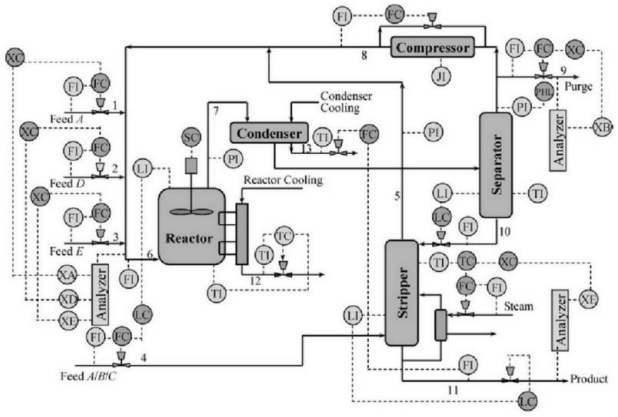

FIGURE 4. The schematic graph of TE.

TABLE 3. Missed detection rates (%) of 20 faults in TE Mode 1.

Fault	LPP (T^2)	LPP (SPE)	GLSA (T^2)	GLSA (SPE)	TSGLP (LOF)
1	0.375	0.625	0.125	1.5	**0**
2	0.875	0.875	**0.5**	1.125	**0.5**
3	99.625	99.625	91.375	100	**88.75**
4	**0**	**0**	**0**	**0**	**0**
5	99.875	99.625	97.5	99.875	**96.875**
6	**0**	**0**	**0**	77.622	**0**
7	**0**	97.125	**0**	0.625	**0**
8	3.125	9.75	**2.5**	58.125	**2.5**
9	99.5	99.5	94.25	99.875	**93.375**
10	34.5	97.625	6.75	95.125	**5.5**
11	5.875	12.25	**5.375**	51.375	7.125
12	82.625	89	68.75	98.375	**54.625**
13	**3.125**	8.125	3.625	45.25	3.375
14	9.5	32.375	1.25	56.875	**1**
15	99.75	99.625	96.5	100	**94.875**
16	99.5	99.375	96.25	100	**94.5**
17	15	18.625	**0.75**	99.625	0.875
18	41.875	96.625	22.375	99.625	**21.125**
19	19.25	96.75	1.375	91.25	**1.125**
20	13.625	30.875	**8.5**	100	8.625

In this section, the TE process under mode 1 and mode 3 are simulated. A total of 1000 data including 500 normal data of mode 1 and 500 normal data of mode 3 are collected as the training dataset. In addition, there are 20 faults in the TE process. Then, 20 datasets produced under mode 1 and 20 datasets produced under mode 3 are collected as the testing datasets. In each testing dataset, data 1 to data 200 are normal and data 201 to data 1000 are abnormal.

In the proposed TSGLP method, the number of neighbors in the temporal neighborhood $2k_1$ is 10 and that in the spatial neighborhood k_2 is 10. The dimension of the feature space d is selected as 9. For fair comparisons, the number of neighbors in the neighborhood k is 20 and the dimension of the feature space d is selected as 9 in the LPP and GLSA methods.

Table 3 and Table 4 list missed detection rates of 20 faults in mode 1 and those in mode 3, respectively. In these tables, the smallest missed detection rates are marked with bold. Moreover, for computing false detection rates of normal process, the normal dataset collected under mode 1 and mode 3 is

TABLE 4. Missed detection rates (%) of 20 faults in TE Mode 3.

Fault	LPP (T^2)	LPP (SPE)	GLSA (T^2)	GLSA (SPE)	TSGLP (LOF)
1	0.75	1.875	0.125	100	**0**
2	7.375	99.125	2.125	100	**2**
3	98.25	98.5	88.5	99.5	**86.5**
4	**0**	**0**	**0**	100	**0**
5	1.625	95.625	**0**	100	**0**
6	**0**	**0**	**0**	97.58	**0**
7	**0**	98.125	**0**	100	**0**
8	3.75	8.625	**2.125**	72.875	**2.125**
9	88.625	85.875	79.75	95.875	**78.875**
10	31.875	93.125	7.875	93.125	**5.375**
11	**7.125**	13.5	9	58.5	8.375
12	1.875	5.25	0.875	70.375	**0.375**
13	18.25	29	10.875	82.5	**8.5**
14	10.75	37.5	1.625	57.875	**1.125**
15	98.625	98.375	97.25	98.875	**96.5**
16	98.125	97.625	96	99.125	**92.75**
17	22.125	25.375	**0.875**	23.5	**0.875**
18	11.75	21.375	11.25	98.25	**10.75**
19	14.875	77.125	1.875	99	**1.625**
20	28.375	54	10.625	93.625	**10.125**

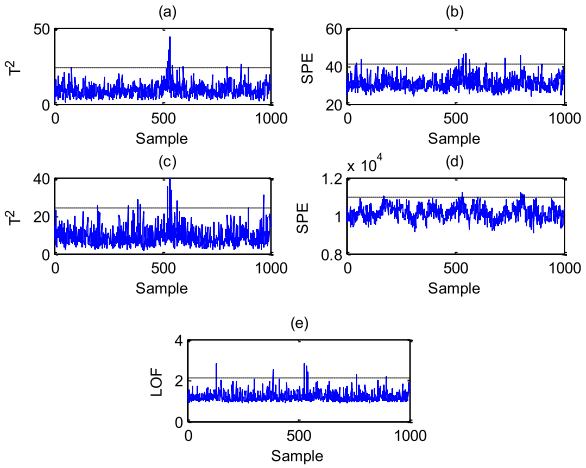

FIGURE 5. The monitoring results of three methods for normal process. (a) LPP. (b) LPP. (c) GLSA. (d) GLSA. (e) TSGLP.

tested, which is shown in Fig. 5. In addition, to show the monitoring performance visually, the monitoring results of LPP, GLSA and TSGLP for Fault 10 in mode 3 are plotted in Fig. 6. In contrast to LPP and GLSA, the proposed TSGLP method can obtain the best monitoring results for 16 faults in mode 1. For 20 faults in mode 3, as shown in Table 3, the proposed TSGLP method can obtain the smallest missed detection rates for 19 faults compared with LPP and GLSA. Fig. 5 shows the monitoring performance of normal process where first 500 data are collected from mode 1 and the remaining data are collected from mode 3. In this figure, no fault happens and the false detection rate is acceptable.

Fault 10 is a random change occurring in the temperature of C feed. The monitoring results of this fault under mode 3 are presented in Fig. 6. As shown in Fig. 6(a), although the LPP(T^2) method can detect this fault, more than 30% fault data are judged as normal data with the missed detection

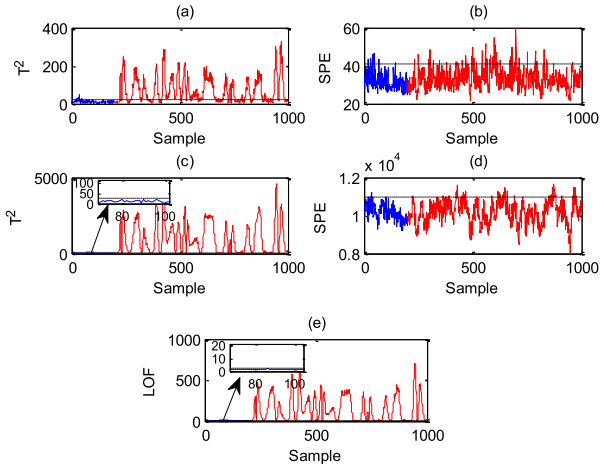

FIGURE 6. The monitoring results of three methods for Fault 10 in mode 3. (a) LPP. (b) LPP. (c) GLSA. (d) GLSA. (e) TSGLP.

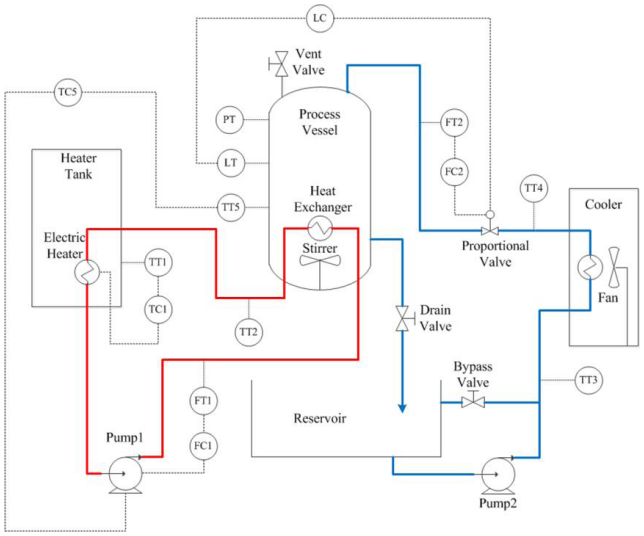

FIGURE 7. The schematic graph of CE117.

rate 31.875%. Compared with LPP which only considers the local structure, the GLSA method preserves both the local and global structure. The monitoring result of GLSA(T^2) has been progressed to a large extent where the missed detection rate is 7.875%. In addition, in contrast to GLSA, the TSGLP method can obtain better monitoring result since not only the spatial information but also the temporal information is applied.

C. CE117

The CE117 process trainer is proposed by TecQuipment for chemical engineering control. Fig. 7 is the schematic graph of it. As shown in this figure, the meanings of symbols connected by the dashed arrows are listed in Table 5, where the variables with bold are selected as the monitoring variables. The CE117 process trainer can be controlled by means of the built-in computer interface and CE2000 software or any other suitable analogue or digital controller that may be available. In this study, the CE117 process trainer is linked to computer by the interface provided by the control modules integrated on a control panel.

TABLE 5. The meanings of symbols attached to the module.

Symbol	Meaning	Location
PT	Pressure Transmitter	Process Vessel
LT	**Level Transmitter**	**Process Vessel**
FT1	Flow Transmitter	Heater Loop
FT2	**Flow Transmitter**	**Process Loop**
TT1	**Temperature Transmitter**	**Heater Tank**
TT2	Temperature Transmitter	Heater Exchanger Output
TT3	**Temperature Transmitter**	**Process Loop Cooler Inlet**
TT5	**Temperature Transmitter**	**Process Vessel**
TT4	**Temperature Transmitter**	**Process Loop Cooler Outlet**
TC1	Temperature Controller	Heater Tank
TC5	Temperature Controller	Process Vessel
FC1	Flow Controller	Heater Loop
FC2	Flow Controller	Process Loop
LC	Level Controller	Process Vessel

In this study, the CE117 runs under two modes. In the first mode, LT is set as 8cm and TT5 is set as 37 ^oC. In the second mode, LT is set as 9cm and TT5 is set as 35.2 ^oC. A total of 600 data including 300 normal data of mode 1 and 300 normal data of mode 2 are collected as the training dataset. Two testing datasets are simulated as follows:

In the first testing dataset, the CE117 process runs under mode 1 firstly, then, it is switched to mode 2. After 600 normal data are collected, the fan can not work properly.

In the second testing dataset, the CE117 process runs under mode 1 firstly, then, it is switched to mode 2. After 600 normal data are collected, the flow transmitter of FT2 occurs a hard-over fault [51], where the value of FT2 and the actual value have an error of 0.3.

In the proposed TSGLP method, the number of neighbors in the temporal neighborhood $2k_1$ is 10 and that in the spatial neighborhood k_2 is 10. The dimension of the feature space d is selected as 2. For fair comparisons, the number of neighbors in the neighborhood k is 20 and the dimension of the feature space d is selected as 2 in the LPP and GLSA methods.

In the first testing dataset, the fan can not work after the 600th data. Firstly, TT4 rises suddenly. Compensated by the closed-loop control, this fault influences the hot water system. Then, TT2 and FT1 change. Fig. 8 presents monitoring results for the first testing dataset. As shown in this figure, although LPP (T^2) and GLSA (T^2) can detect this fault, the monitoring results are unsatisfactory. Moreover, the monitoring result of LPP (SPE) has been improved to a large extent, where the missed detection rate is 5%. In contrast to LPP and GLSA, TSGLP can obtain the best monitoring result. Once the fault occurs, the statistic of TSGLP exceeds the control limit and can maintain all the time.

In the second testing dataset, a transmitter fault occurs. Once this fault happens, apart from the value of FT2, the other parts of the process are operated under normal condition. As shown in Fig. 9(a) and Fig. 9(c), the LPP(T^2) and GLSA(T^2) methods cannot detect this fault, where all fault data are regarded as normal data. From Fig. 9(b) and Fig. 9(d),

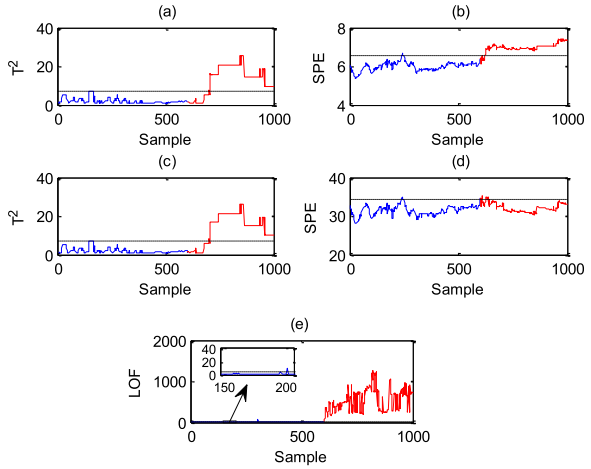

FIGURE 8. The monitoring results of three methods for the first testing dataset. (a) LPP. (b) LPP. (c) GLSA. (d) GLSA. (e) TSGLP.

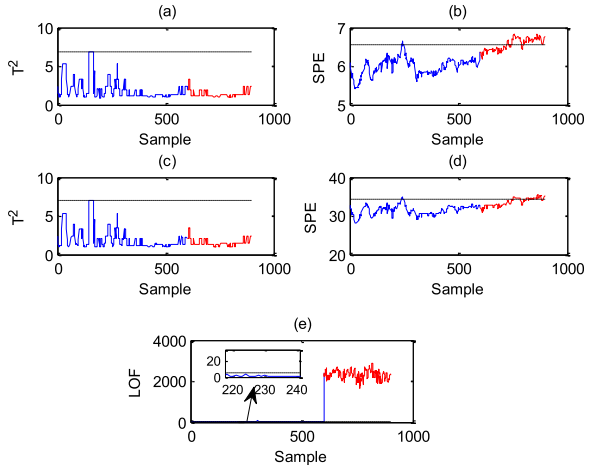

FIGURE 9. The monitoring results of three methods for the second testing dataset. (a) LPP. (b) LPP. (c) GLSA. (d) GLSA. (e) TSGLP.

the monitoring results of LPP (SPE) and GLSA (SPE) have been improved with the missed detection rate 48.3% and 68%, respectively. Compared with LPP and GLSA, the proposed TSGLP method can obtain the best monitoring result, where the missed detection rate is 0.

V. CONCLUSION

In this article, a novel algorithm TSGLP is proposed to monitor multimode processes. For mining both the temporal and spatial information, the neighborhood construction in TSGLP is implemented using neighbors in both temporal and spatial scale. In addition, not only the local structure but also the global structure is preserved in TSGLP. In contrast to LPP and GLSA, the TSGLP method can acquire the best monitoring performance in three case studies. Considering that one batch in the batch process always contains multiple phases, it would be worth some research efforts to extend the proposed method to batch processes.

REFERENCES

[1] S.-F. Yang and S.-H. Wu, "A double sampling scheme for process mean monitoring," *IEEE Access*, vol. 5, pp. 6668–6677, May 2017.
[2] S. Yin, S. X. Ding, X. Xie, and H. Luo, "A review on basic data-driven approaches for industrial process monitoring," *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6418–6428, Nov. 2014.
[3] Y. Zhang, Y. Fan, and W. Du, "Nonlinear process monitoring using regression and reconstruction method," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 3, pp. 1343–1354, Jul. 2016.
[4] Z. Ge and Z. Song, "Bagging support vector data description model for batch process monitoring," *J. Process Control*, vol. 23, no. 8, pp. 1090–1096, Sep. 2013.
[5] C. Sankavaram, A. Kodali, K. R. Pattipati, and S. Singh, "Incremental classifiers for data-driven fault diagnosis applied to automotive systems," *IEEE Access*, vol. 3, pp. 407–419, 2015.
[6] Y. Qin, C. Zhao, and F. Gao, "An iterative two-step sequential Phase partition (ITSPP) method for batch process modeling and online monitoring," *AIChE J.*, vol. 62, no. 7, pp. 2358–2373, Jul. 2016.
[7] H. Zhang, X. Tian, and X. Deng, "Batch process monitoring based on multiway global preserving kernel slow feature analysis," *IEEE Access*, vol. 5, pp. 2696–2710, Mar. 2017.
[8] L. M. Elshenawy, S. Yin, A. S. Naik, and S. X. Ding, "Efficient recursive principal component analysis algorithms for process monitoring," *Ind. Eng. Chem. Res.*, vol. 49, no. 1, pp. 252–259, 2010.
[9] S. Yin, X. Li, H. Gao, and O. Kaynak, "Data-based techniques focused on modern industry: An overview," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 657–667, Jan. 2015.
[10] X. Deng, X. Tian, S. Chen, and C. J. Harris, "Fault discriminant enhanced kernel principal component analysis incorporating prior fault information for monitoring nonlinear processes," *Chemometrics Intell. Lab. Syst.*, vol. 162, pp. 21–34, Mar. 2017.
[11] J. Wang, B. Zhong, and J. L. Zhou, "Quality-relevant fault monitoring based on locality-preserving partial least-squares statistical models," *Ind. Eng. Chem. Res.*, vol. 56, no. 24, pp. 7009–7020, 2017.
[12] Z. Zhu, Z. Song, and A. Palazoglu, "Transition process modeling and monitoring based on dynamic ensemble clustering and multiclass support vector data description," *Ind. Eng. Chem. Res.*, vol. 50, no. 24, pp. 13969–13983, 2011.
[13] B. Song, H. Shi, Y. Ma, and J. Wang, "Multisubspace principal component analysis with local outlier factor for multimode process monitoring," *Ind. Eng. Chem. Res.*, vol. 53, no. 42, pp. 16453–16464, Sep. 2014.
[14] C. Zhao, Y. Yao, F. Gao, and F. Wang, "Statistical analysis and online monitoring for multimode processes with between-mode transitions," *Chem. Eng. Sci.*, vol. 65, no. 22, pp. 5961–5975, Nov. 2011.
[15] Z. Zhu, Z. Song, and A. Palazoglu, "Process pattern construction and multi-mode monitoring," *J. Process Control*, vol. 22, no. 1, pp. 247–262, 2013.
[16] A. Haghani, T. Jeinsch, and S. X. Ding, "Quality-related fault detection in industrial multimode dynamic processes," *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6446–6453, Nov. 2014.
[17] C. H. Zhao, "Concurrent phase partition and between-mode statistical analysis for multimode and multiphase batch process monitoring," *AIChE J.*, vol. 60, no. 2, pp. 559–573, 2014.
[18] S. Tan, F. Wang, J. Peng, Y. Chang, and S. Wang, "Multimode process monitoring based on mode identification," *Ind. Eng. Chem. Res.*, vol. 51, no. 1, pp. 374–388, 2012.
[19] Z. Ge, F. Gao, and Z. Song, "Two-dimensional Bayesian monitoring method for nonlinear multimode processes," *Chem. Eng. Sci.*, vol. 66, no. 21, pp. 5173–5183, 2011.
[20] J. Guo, T. Yuan, and Y. Li, "Fault detection of multimode process based on local neighbor normalized matrix," *Chemometrics Intell. Lab. Syst.*, vol. 154, pp. 162–175, May 2016.
[21] S. J. Zhao, J. Zhang, and Y. M. Xu, "Monitoring of processes with multiple operating modes through multiple principle component analysis models," *Ind. Eng. Chem. Res.*, vol. 43, no. 22, pp. 7025–7035, 2004.
[22] Z. Ge and Z. Song, "Mixture Bayesian regularization method of PPCA for multimode process monitoring," *AIChE J.*, vol. 56, no. 11, pp. 2838–2849, 2010.
[23] Z. Ge and Z. Song, "Online monitoring of nonlinear multiple mode processes based on adaptive local model approach," *Control Eng. Pract.*, vol. 16, no. 12, pp. 1427–1437, Jun. 2008.
[24] C. Tong, A. Palazoglu, and X. Yan, "An adaptive multimode process monitoring strategy based on mode clustering and mode unfolding," *J. Process Control*, vol. 23, no. 10, pp. 1497–1507, 2013.

[25] W. Du, Y. Tian, and F. Qian, "Monitoring for nonlinear multiple modes process based on LL-SVDD-MRDA," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 4, pp. 1133–1148, Oct. 2014.
[26] J. Zhu, Z. Ge, and Z. Song, "Multimode process data modeling: A Dirichlet process mixture model based Bayesian robust factor analyzer approach," *Chemometrics Intell. Lab. Syst.*, vol. 142, pp. 231–244, Mar. 2015.
[27] J. Yu, "A nonlinear kernel Gaussian mixture model based inferential monitoring approach for fault detection and diagnosis of chemical processes," *Chem. Eng. Sci.*, vol. 68, no. 1, pp. 506–519, Oct. 2011.
[28] M. M. Rashid and J. Yu, "Hidden Markov model based adaptive independent component analysis approach for complex chemical process monitoring and fault detection," *Ind. Eng. Chem. Res.*, vol. 51, no. 15, pp. 5506–5514, Mar. 2012.
[29] G. Wang and M. Liu, "Fault detection for discrete-time systems with fault signal happening randomly: The Markov approach," *IEEE Access*, vol. 5, pp. 14680–14689, Aug. 2015.
[30] J. Zhu, Z. Ge, and Z. Song, "Robust supervised probabilistic principal component analysis model for soft sensing of key process variables," *Chem. Eng. Sci.*, vol. 122, pp. 573–584, Jan. 2015.
[31] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Comput.*, vol. 15, no. 6, pp. 1373–1396, 2003.
[32] L. K. Saul and S. T. Roweis, "Think globally, fit locally: Unsupervised learning of low dimensional manifolds," *J. Mach. Learn. Res.*, vol. 4, pp. 119–155, Jun. 2003.
[33] L. J. Luo, "Process monitoring with global–local preserving projections," *Ind. Eng. Chem. Res.*, vol. 53, no. 18, pp. 7696–7705, 2014.
[34] Y. Ma, B. Song, H. Shi, and Y. Yang, "Fault detection via local and nonlocal embedding," *Chem. Eng. Res. Des.*, vol. 94, pp. 538–548, Feb. 2015.
[35] B. Song, S. Tan, and H. Shi, "Process monitoring via enhanced neighborhood preserving embedding," *Control Eng. Pract.*, vol. 50, pp. 48–56, May 2016.
[36] A. Miao, Z. Ge, Z. Song, and F. Shen, "Nonlocal structure constrained neighborhood preserving embedding model and its application for fault detection," *Chemometrics Intell. Lab. Syst.*, vol. 142, pp. 184–196, Mar. 2015.
[37] M. Zhang, Z. Ge, Z. Song, and R. Fu, "Global–local structure analysis model and its application for fault detection and identification," *Ind. Eng. Chem. Res.*, vol. 50, no. 11, pp. 6837–6848, 2011.
[38] J. Yu, "Local and global principal component analysis for process monitoring," *J. Process Control*, vol. 22, no. 7, pp. 1358–1373, Jul. 2012.
[39] C. Tong and X. Yan, "Statistical process monitoring based on a multimanifold projection algorithm," *Chemometrics Intell. Lab. Syst.*, vol. 130, pp. 20–28, Jan. 2014.
[40] A. Miao, Z. Ge, Z. Song, and L. Zhou, "Time neighborhood preserving embedding model and its application for fault detection," *Ind. Eng. Chem. Res.*, vol. 52, no. 38, pp. 13717–13729, 2013.
[41] Q. Jiang and X. Yan, "Probabilistic weighted NPE-SVDD for chemical process monitoring," *Control Eng. Pract.*, vol. 28, pp. 74–89, Jul. 2014.
[42] B. Song, Y. Ma, and H. Shi, "Multimode process monitoring using improved dynamic neighborhood preserving embedding," *Chemometrics Intell. Lab. Syst.*, vol. 135, pp. 17–30, Jul. 2014.
[43] X. He and P. Niyogi, "Locality preserving projection," in *Proc. Adv. Neural Inf. Process Syst.*, vol. 16. 2003, pp. 186–197.
[44] M. M. Rashid and J. Yu, "A new dissimilarity method integrating multidimensional mutual information and independent component analysis for non-Gaussian dynamic process monitoring," *Chemometrics Intell. Lab. Syst.*, vol. 115, pp. 44–58, Jun. 2012.
[45] M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "LOF: Identifying density-based local outliers," in *Proc. ACM SIGMOD Int. Conf. Mange. Data*, 2000, pp. 93–104.
[46] 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.
[47] T. Rato, M. Reis, E, Schmitt, M. Hubert, and B. D. 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, May 2016.
[48] Y. Ma and H. Shi, "Multimode process monitoring based on aligned mixture factor analysis," *Ind. Eng. Chem. Res.*, vol. 53, no. 2, pp. 786–799, 2014.
[49] Q. Jia and Y. Zhang, "Quality-related fault detection approach based on dynamic kernel partial least squares," *Chem. Eng. Res. Des.*, vol. 106, pp. 242–252, Feb. 2016.
[50] S. Yin, S. X. Ding, A. Haghani, H. Hao, and P. Zhang, "A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process," *J. Process Control*, vol. 22, no. 9, pp. 1567–1581, 2012.
[51] S. U. Jan, Y.-D. Lee, J. Shin, and I. Koo, "Sensor fault classification based on support vector machine and statistical time-domain features," *IEEE Access*, vol. 5, pp. 8682–8690, Jun. 2017.

BING SONG received the B.E. degree in automation and the Ph.D. degree in control theory and control engineering from the East China University of Science and Technology, Shanghai, China, in 2012 and 2017, respectively.

He is currently a Post-Doctoral Researcher with the Department of Automation, East China University of Science and Technology. His research interests include feature extraction, fault detection, fault diagnosis, and multimode process monitoring.

HONGBO SHI received the B.E. degree in chemical automation and the Ph.D. degree in control theory and control engineering from the East China University of Science and Technology, China, in 1986 and 2000, respectively.

He was the 2003 Shu Guang Scholar of Shanghai. He is currently a Professor with the East China University of Science and Technology. His research interest covers modeling of industrial process and advanced control technology, theory and methods of integrated automation systems, condition monitoring, and fault diagnosis of industrial process.