Research Article

Monitoring of Nonlinear Time-Delay Processes Based on Adaptive Method and Moving Window

Yunpeng Fan, Wei Zhang, and Yingwei Zhang

State Laboratory of Synthesis Automation of Process Industry, Northeastern University, Shenyang, Liaoning 110819, China

Correspondence should be addressed to Yingwei Zhang; zhangyingwei@mail.neu.edu.cn

Received 2 June 2014; Accepted 17 June 2014; Published 21 July 2014

Academic Editor: Ligang Wu

Copyright © 2014 Yunpeng Fan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A new adaptive kernel principal component analysis (KPCA) algorithm for monitoring nonlinear time-delay process is proposed. The main contribution of the proposed algorithm is to combine adaptive KPCA with moving window principal component analysis (MWPCA) algorithm, and exponentially weighted principal component analysis (EWPCA) algorithm respectively. The new algorithm prejudices the new available sample with MKPCA method to decide whether the model is updated. Then update the KPCA model using EWKPCA method. And also extend MPCA and EWPCA from linear data space to nonlinear data space effectively. Monitoring experiment is performed using the proposed algorithm. The simulation results show that the proposed method is effective.

1. Introduction

Fault detection and diagnosis are very important aspects in modern industrial processes because they concern the execution of planned operations and process productivity. In order to meet the need of production, data-based methods have been deeply developed, such as principal component analysis (PCA), partial least squares (PLS), and independent component analysis (ICA) [1–5].

PCA as a multivariable statistical method is widely used for damage detection and diagnosis of structures in industrial systems [6–8]. Because PCA is an orthogonal transformation of the coordinate system, it is a linear monitoring method. However, most industrial processes have strong nonlinearity characteristics [9, 10]. Document [11] shows that application of linear monitoring approaches to nonlinear processes may lead to unreliable process monitoring, because a linear method is inappropriate to extract the nonlinearities within the process variables. To solve the nonlinear problem, several nonlinear methods have been proposed in the past decades [12–17].

Kernel principal component analysis (KPCA), as a nonlinear version of PCA method which is developed by many researchers, is more proper in nonlinear process monitoring

[18–20]. KPCA can efficiently compute principal components (PCs) by nonlinear kernel functions in a high-dimensional feature spaces. The main advantages of KPCA are that it only solves an eigenvalue problem and does not involve nonlinear optimization [21]. However, a KPCA monitoring model requires kernel matrix, whose dimension is given by the number of reference samples. In addition, the KPCA with fixed model may lead to large error because of the gradual change of parameters with the operation of process [22–26]. Because the old samples are not representative of the current process status, the adaptation of KPCA model is necessary.

To date, MWPCA algorithm and EWPCA algorithm are two representative adaptive methods [27–31]. Performing a moving window approach, which was proposed by Hoegaerts et al., overcomes this problem and produces a constant scale of the kernel matrix and a fixed speed of adaptation [32, 33]. Choosing a proper weighting factor is an important issue, which determines the influence that the older data has on the model [31]. Paper [34] has proposed a moving window KPCA formulation, which has some advantages such as using the Gram matrix instead of the kernel matrix, incorporation of an adaptation method for the eigendecomposition of the Gram matrix.

In this paper, a new algorithm combining these two representative methods is proposed. The proposed algorithm mapped the sample set in feature space and prejudices the new available sample with MKPCA method to decide whether the model is updated. Then update the KPCA model using EWKPCA method. The proposed algorithm can reduce negative impact of outliers on model; the updating after prejudgment can reduce computational complexity and be more efficient. The remaining sections of this paper are organized as follows. The KPCA method based on loss function and the online monitoring strategy are introduced in Section 2. The iterative KPCA algorithm with penalty factor is presented in Section 3. The proposed adaptive KPCA method combining two representative methods is described in Section 4. The simulation results are presented in Section 5. Finally, the conclusion is given in Section 6.

2. Kernel Principal Component Analysis Based on Loss Function

2.1. Kernel Principal Component Analysis. KPCA is an extension of PCA, and it can be solved as an eigenvalue problem of its kernel matrix. KPCA algorithm makes use of the following idea.

Via a nonlinear mapping

$$\Phi: \mathbf{R}^N \longrightarrow F, \quad (1)$$

data set $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\} \in \mathbf{R}^N$ is mapped into a potentially much higher dimensional feature space F ; the mapped data set $\Phi(\mathbf{X}) = \{\Phi(\mathbf{x}_1), \Phi(\mathbf{x}_2), \dots, \Phi(\mathbf{x}_N)\}$ is obtained. $\Phi(\mathbf{X})$ is gotten by centering $\Phi(\mathbf{X})$ with the corresponding mean s.t. Consider

$$\frac{1}{N} \sum_{i=1}^N \Phi(\mathbf{x}_i) = 0. \quad (2)$$

A principal component $\mathbf{V}$ can be computed by solving the eigenvalue problem

$$\lambda \mathbf{V} = \frac{1}{N} \Phi(\mathbf{X}) \Phi(\mathbf{X})^T \mathbf{V} = \mathbf{C}^F \mathbf{V}, \quad (3)$$

where $\mathbf{C}^F = (1/N) \sum_{k=1}^N \Phi(\mathbf{x}_k) \Phi(\mathbf{x}_k)^T = (1/N) \Phi(\mathbf{X}) \Phi(\mathbf{X})^T$ is the sample covariance matrix of $\Phi(\mathbf{X})$ and $\lambda > 0$ is the eigenvalue corresponding to $\mathbf{V}$.

By defining the kernel matrix $\mathbf{K}$ of dimension $N \times N$,

$$[\mathbf{K}]_{ij} = \mathbf{K}_{ij} = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle = k(\mathbf{x}_i, \mathbf{x}_j), \quad (4)$$

the eigenvalue problem can be put in the form

$$N \lambda \boldsymbol{\alpha} = \mathbf{K} \boldsymbol{\alpha}. \quad (5)$$

Regarding the kernel function, radial basis function (RBF) is chosen in this paper as

$$k(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{c}\right), \quad (6)$$

where c is the width of the Gaussian kernel; it can be very small (<1) or quite large.

Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$ and $\alpha_1, \alpha_2, \dots, \alpha_N$ be the eigenvalues and the corresponding eigenvectors of the kernel matrix $\mathbf{K}$, respectively. As $\mathbf{V}^k = \sum_{i=1}^N \alpha_i^k \Phi(\mathbf{x}_i)$, the score of the k th eigenvector $\mathbf{V}^k$ can be calculated as follows:

$$\begin{aligned} \langle \mathbf{V}^k, \Phi(\mathbf{X}) \rangle &= \sum_{i=1}^N \alpha_i^k \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{X}) \rangle \\ &= \sum_{i=1}^N \alpha_i^k \cdot k(\mathbf{x}_i, \mathbf{X}). \end{aligned} \quad (7)$$

2.2. Loss Function in Feature Space. Reconstruct sample data set $\Phi(\mathbf{X})$ in feature space; the reconstructed data set is given by

$$\langle \mathbf{W}, \Phi(\mathbf{X}) \rangle \mathbf{W} = \mathbf{W} \mathbf{W}^T \Phi(\mathbf{X}), \quad (8)$$

where $\mathbf{W}$ is the transformation matrix, with $\|\mathbf{W}\| = 1$.

Between the original sample data set and the reconstructed data set exists the reconstructed error

$$\begin{aligned} \|\mathbf{e}(\Phi(\mathbf{x}_i))\|^2 &= \|\Phi(\mathbf{x}_i) - \mathbf{W} \mathbf{W}^T \Phi(\mathbf{x}_i)\|^2 \\ &= \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_i) - 2 \mathbf{W} \mathbf{W}^T \Phi(\mathbf{x}_i) \Phi(\mathbf{x}_i) \\ &\quad + (\mathbf{W} \mathbf{W}^T) (\mathbf{W} \mathbf{W}^T) \Phi(\mathbf{x}_i) \Phi(\mathbf{x}_i) \\ &= k(\mathbf{x}_i, \mathbf{x}_i) - 2 \mathbf{W} \mathbf{W}^T k(\mathbf{x}_i, \mathbf{x}_i) \\ &\quad + (\mathbf{W} \mathbf{W}^T)^2 k(\mathbf{x}_i, \mathbf{x}_i). \end{aligned} \quad (9)$$

Define the loss function in feature space as follows:

$$\begin{aligned} J_1(\mathbf{W}) &= \sum_{i=1}^N \|\Phi(\mathbf{x}_i) - \mathbf{W} \mathbf{W}^T \Phi(\mathbf{x}_i)\|^2 \\ &= \sum_{i=1}^N (\mathbf{W}^T \Phi(\mathbf{x}_i))^2 \|\mathbf{W}\|^2 - 2 \sum_{i=1}^N (\mathbf{W}^T \Phi(\mathbf{x}_i))^2 \\ &\quad + \sum_{i=1}^N (\Phi(\mathbf{x}_i))^2 \\ &= - \sum_{i=1}^N (\mathbf{W}^T \Phi(\mathbf{x}_i))^2 + \sum_{i=1}^N \|\Phi(\mathbf{x}_i)\|^2, \end{aligned} \quad (10)$$

where $\sum_{i=1}^N \|\Phi(\mathbf{x}_i)\|^2$ is a constant value. Thus, when the transformation matrix $\mathbf{W}$ maximizes $\mathbf{W}^T \Phi(\mathbf{x}_i)$, the loss function $J_1(\mathbf{W})$ reaches its minimum value. Similarly, the score of the k th PC $\mathbf{W}^k$ is calculated as follows:

$$\begin{aligned} \mathbf{t}_k &= \langle \mathbf{W}^k, \Phi(\mathbf{x}) \rangle \\ &= \sum_{i=1}^N \alpha_i^k \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{X}) \rangle = \sum_{i=1}^N \alpha_i^k k(\mathbf{x}_i, \mathbf{X}). \end{aligned} \quad (11)$$

2.3. Online Monitoring Strategy of KPCA. A measure of the variation within the KPCA model is given by Hotelling's T^2 statistic. The measure of goodness of fit of a sample to the model is the squared prediction error (SPE), also known as the Q statistic. T^2 is given by

$$T^2 = [\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_A] \Delta^{-1} [\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_A]^T, \quad (12)$$

where A is the PC number, N is the number of samples, $\mathbf{t}_A$ is the A th nonlinear principal component which can be calculated with (11), Δ^{-1} is the inverse matrix of the diagonal matrix composed of eigenvalues, and $\Delta = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_A\} = (1/(N-1))\mathbf{T}^T\mathbf{T} = (1/(N-1))\mathbf{W}^T\mathbf{K}^T\mathbf{K}\mathbf{W}$.

Control limit of Hotelling's T^2 is as follows:

$$T_{A,N,\alpha}^2 \sim \frac{A(N-1)}{N-A} F_{N,N-A,\alpha}. \quad (13)$$

Once a new sample $\Phi(\mathbf{x}_{\text{new}})$ is available, T^2 new can be calculated with the following equation:

$$T_{\text{new}}^2 = \mathbf{t}_{\text{new}} \Delta^{-1} \mathbf{t}_{\text{new}}^T. \quad (14)$$

The SPE and the corresponding control limit are obtained using the following equations:

$$\text{SPE} = \|\overline{\Phi}(\mathbf{x}) - \widehat{\Phi}(\mathbf{x})\|^2 = \sum_{i=1}^N \mathbf{t}_i^2 - \sum_{i=1}^A \mathbf{t}_i^2, \quad (15)$$

$$\text{SPE}_\alpha \sim g\chi_h^2,$$

where g and h are two relevant parameters. For new sample $\Phi(\mathbf{x}_{\text{new}})$, SPE is given by

$$\begin{aligned} \text{SPE} &= \|\Phi(\mathbf{x}_{\text{new}}) - \widehat{\Phi}(\mathbf{x}_{\text{new}})\|^2 \\ &= \Phi(\mathbf{x}_{\text{new}}) \Phi(\mathbf{x}_{\text{new}})^T - 2\Phi(\mathbf{x}_{\text{new}}) \widehat{\Phi}(\mathbf{x}_{\text{new}})^T \\ &\quad + \widehat{\Phi}(\mathbf{x}_{\text{new}}) \widehat{\Phi}(\mathbf{x}_{\text{new}})^T \\ &= k(\mathbf{x}_{\text{new}}, \mathbf{x}_{\text{new}}) - 2\Phi(\mathbf{x}_{\text{new}}) \mathbf{W} \mathbf{t}_{\text{new}}^T + \mathbf{t}_{\text{new}}^T \mathbf{W} \mathbf{W}^T \mathbf{t}_{\text{new}} \\ &= k(\mathbf{x}_{\text{new}}, \mathbf{x}_{\text{new}}) - 2\mathbf{k}_{\text{new}} \mathbf{T} \mathbf{t}_{\text{new}}^T + \mathbf{t}_{\text{new}}^T \mathbf{T}^T \mathbf{K} \mathbf{T} \mathbf{t}_{\text{new}}^T, \end{aligned} \quad (16)$$

where $k(\mathbf{x}_{\text{new}}, \mathbf{x}_{\text{new}}) = 1 - (2/N) \sum_{i=1}^N k_{\text{new},i} + (1/N) \sum_{i=1}^N \sum_{j=1}^N k_{ij}$ and $k_{\text{new},i}$ is the i th element of $\mathbf{k}_{\text{new}}$.

3. Iterative Kernel Principal Component Analysis

KPCA algorithm based on eigenvalue decomposition is a batch learning method. This algorithm is not available for online monitoring as all sample data has to be known before modeling. In addition, no outlier is allowed in modeling process; this requires all sample data to be normal. But the requirement is hard to reach in actual production process. Outliers always exist in sample set even after mapping

in feature space. In the sense of least squared error, an iterative KPCA algorithm is proposed in this paper. Furthermore, penalty factor is added to solve the outlier problem.

Use stochastic gradient descent algorithm to solve the problem proposed in (10):

$$\frac{dJ_1(\mathbf{W})}{d\mathbf{W}} = 2(\mathbf{W}\mathbf{W}^T\mathbf{K} - \mathbf{K})\mathbf{W}. \quad (17)$$

Iterative formula is given by

$$\mathbf{W}_{n+1} = \mathbf{W}_n + \mu_n (\mathbf{K} - \mathbf{W}\mathbf{W}^T\mathbf{K}) \mathbf{W}_n, \quad (18)$$

where μ_n is the iteration step, s.t. $0 < \mu_n < 1$, and $\mathbf{W}_n$ converges to the first nonlinear principal component.

Because nonlinear principal components are mutually orthogonal, the Schmidt orthogonal method is used to calculate the other nonlinear principal components:

$$\begin{aligned} \mathbf{W}_{n+1}^j &= \mathbf{W}_n^j + \mu_n (\mathbf{K}^j - \mathbf{W}_n^j \mathbf{W}_n^{jT} \mathbf{K}^j) \mathbf{W}_n^j, \\ \mathbf{K}^j &= \mathbf{K} - \sum_{i=1}^{j-1} \mathbf{W}^i \mathbf{W}^{iT} \mathbf{K}, \end{aligned} \quad (19)$$

where $\mathbf{W}^j$ is the j th nonlinear principal component ($j = 2, 3, \dots, A$) and A is the PC number.

To solve the outlier problem, penalty factor is added in (10):

$$J_2(\mathbf{W}) = \sum_{i=1}^N \left(C_i \|\overline{\Phi}(\mathbf{x}_i) - \mathbf{W}\mathbf{W}^T\overline{\Phi}(\mathbf{x}_i)\|^2 + \eta(1 - C_i) \right), \quad (20)$$

where η is the predefined threshold, $\eta > 0$, and C_i satisfies the following formula:

$$C_i = \begin{cases} 1 & \|\overline{\Phi}(\mathbf{x}_i) - \mathbf{W}\mathbf{W}^T\overline{\Phi}(\mathbf{x}_i)\|^2 \leq \eta, \\ 0 & \|\overline{\Phi}(\mathbf{x}_i) - \mathbf{W}\mathbf{W}^T\overline{\Phi}(\mathbf{x}_i)\|^2 > \eta. \end{cases} \quad (21)$$

As is shown above, when $\|\overline{\Phi}(\mathbf{x}_i) - \mathbf{W}\mathbf{W}^T\overline{\Phi}(\mathbf{x}_i)\|^2 > \eta$, $\overline{\Phi}(\mathbf{x}_i)$ is regarded as outlier. The value of $J_2(\mathbf{W})$ is set to η to reduce impact on KPCA model. Note that C_i in the penalty factor is discrete; thus the continuous Sigmoid function is used to approach C_i . Iteration formula with penalty factor is shown as follows:

$$\mathbf{W}_{n+1} = \mathbf{W}_n + \mu_n \frac{1}{1 + e^{\beta(\|\mathbf{e}_n(\overline{\Phi}(\mathbf{X}))\|^2 - \eta)}} (\mathbf{K} - \mathbf{W}\mathbf{W}^T\mathbf{K}) \mathbf{W}_n, \quad (22)$$

where μ_n is the iteration step, $0 < \mu_n < 1$, and $1/(1 + e^{\beta(\|\mathbf{e}_n(\overline{\Phi}(\mathbf{X}))\|^2 - \eta)})$ is the continuous Sigmoid function. As (16) shows, the number of outliers is determined by the value of η in the penalty factor. The smaller the threshold value is, the more the sample points which will be treated as outlier are. Therefore, the value of η can be determined by the proportion of outliers. Sort the sample reconstruction error in descending order, and set the percentage of outliers to r .

Select the largest r percent reconstruction error data to be outliers. Thus, the value of η can be set to the smallest reconstruction error value in all the outliers.

Similarly, the other nonlinear principal components are given by

$$\begin{aligned} \mathbf{W}_{n+1}^j &= \mathbf{W}_n^j + \mu_n \frac{1}{1 + e^{\beta(\|\mathbf{e}_n(\bar{\Phi}(\mathbf{x}))\|^2 - \eta)}} (\mathbf{K}^j - \mathbf{W}\mathbf{W}^T\mathbf{K}^j) \mathbf{W}_n^j, \\ \mathbf{K}^j &= \mathbf{K}_0 - \sum_{i=1}^{j-1} \mathbf{W}^i \mathbf{W}^{iT} \mathbf{K}_0. \end{aligned} \quad (23)$$

The iterative kernel principal component analysis algorithm with penalty factor is summarized as follows.

- (1) Given an initial standardized block of sample data set $\bar{\Phi}(\mathbf{X})$ and maximum iterations N and ϵ , set initial iteration steps $n = 1$ and principal component number $j = 1$.
- (2) Construct the kernel matrix and scale it to obtain $\mathbf{K}_0$.
- (3) Calculate the j th nonlinear principal component ($j = 1, 2, 3, \dots, A$). If $j = 1$, calculate the first PC $\mathbf{W}^1$ with (22); if $j = 2, 3, \dots, A$, calculate the j th PC $\mathbf{W}^j$ with (23).
- (4) If $|\mathbf{W}_{n+1}^j - \mathbf{W}_n^j| > \epsilon$ as well as $n < N$, make $n = n + 1$ and go back to Step (3) to do the next iteration; otherwise, $\mathbf{W}^j$ is obtained.
- (5) Make $j = j + 1$; go back to Step (3) to calculate the next PC.

4. Adaptive Kernel Principal Component Analysis

In practice, the size of sample data set is gradually increased during dynamic process. Using the obtained KPCA model for online monitoring may cause large error. Therefore, in order to improve the ability to adapt new samples, dynamic performance is added in KPCA algorithm.

In this paper, an adaptive kernel PCA algorithm combining two representative methods is proposed. Kernel matrix and corresponding control limits of two statistics can be updated in real time.

In MKPCA, once a new sample data is available, a data window of fixed length is moving in real time to update the KPCA model.

In EWKPCA algorithm proposed by Li et al., the updating of the covariance matrix is as follows:

$$(\mathbf{X}^T \mathbf{X})_t = \lambda_t (\mathbf{X}^T \mathbf{X})_{t-1} + (1 - \lambda_t) (\mathbf{x}_t^T \mathbf{x}_t), \quad (24)$$

where $\mathbf{x}_t$ is a new sample and λ_t is the corresponding weighting factor, $0 < \lambda_t < 1$. As is shown above, the weight on the latest observation is increasing with the decrease of λ_t . So the latest observation contributes most to the model. Introduce the idea to KPCA algorithm. Once a new sample is available, a new kernel vector $\mathbf{k}_t$ can be calculated. In this paper,

$\mathbf{k}_t$ is used to update the kernel matrix with the following formula:

$$\mathbf{K}_{\text{rec}}^t = \gamma_t \mathbf{K}_{\text{rec}}^{t-1} + (1 - \gamma_t) \mathbf{k}_t^T \mathbf{k}_t, \quad (25)$$

where $\mathbf{K}_{\text{rec}}^{t-1}$ is the kernel matrix calculated at time $t - 1$, γ_t is the weighting factor, $0 < \gamma_t < 1$, and $\mathbf{k}_t$ is the standardized kernel vector at time t . As t increases, old samples affect the model less and less until the affection can almost be negligible. Therefore, old samples can be abandoned automatically instead of manually.

One important step of the EWKPCA algorithm is determination of the weighting factor γ_t . Fixed forgetting factor is not applicable because industrial process does not change in a constant rate. In the algorithm proposed by Choi et al., two forgetting factors α and β are used to update the sample mean and covariance matrix, respectively. The forgetting factor used to update the covariance matrix is given by

$$\beta_t = \beta_{\max} - (\beta_{\max} - \beta_{\min}) \left[1 - \exp \left(-k \left(\frac{\|\Delta \mathbf{R}_{t-1}\|}{\|\Delta \mathbf{R}_{\text{nor}}\|} \right)^n \right) \right], \quad (26)$$

where $\beta_{\max}$ and $\beta_{\min}$ are the maximum and minimum values of forgetting factor, respectively, $\|\Delta \mathbf{R}\|$ is the Euclidean matrix norm of the difference between two consecutive correlation matrices, and $\|\Delta \mathbf{R}_{\text{nor}}\|$ is the average $\|\Delta \mathbf{R}\|$ obtained using historical data. k and n are two parameters which control the sensitiveness needed to be determined.

Let $\mathbf{X}_{t-1} = \{\mathbf{x}_t, \mathbf{x}_{t+1}, \dots, \mathbf{x}_{N+t-1}\} \in \mathbf{R}^N$ be the sample data set, let $\mathbf{K}_{\text{rec}}^{t-1}$ be the kernel matrix and $\mathbf{K}_{\text{rec}}^0 = \mathbf{K}_0$, let $\mathbf{W}_{t-1}$ be the transformation matrix, let Δ_{t-1} be the principal eigenvalue diagonal matrix, and let $T_{\text{lim},t-1}^2$ and $\text{SPE}_{\text{lim},t-1}$ be the control limits at time $t - 1$, respectively.

The adaptive kernel principal component analysis algorithm can be concluded as follows.

- (1) At time t , a new sample is available. $\Phi(\mathbf{x}_{N+t})$ is obtained after scaling it with the mean and covariance gotten at time $t - 1$.
- (2) Calculate kernel vector $\mathbf{k}_t = \Phi(\mathbf{x}_{N+t})^T \Phi(t - 1)$, and scale it, that is, $\bar{\mathbf{k}}_t$.
- (3) Hotelling's T^2 and SPE statistics of $\Phi(\mathbf{x}_{N+t})$ are obtained with (14) and (16). Compare them with the control limits $T_{\text{lim},t-1}^2$ and $\text{SPE}_{\text{lim},t-1}$ at time $t - 1$. $\Phi(\mathbf{x}_{N+t})$ is accepted if control limits are not exceeded; go to Step (4) for model updating. Otherwise, go to Step (6) for further judgment.
- (4) Kernel matrix $\mathbf{K}_{\text{rec}}^t$ is calculated with (25); the forgetting factor is

$$\gamma_t = \gamma_{\max} - (\gamma_{\max} - \gamma_{\min}) \left[1 - \exp(-k \|\Delta \mathbf{R}_{t-1}\|^n) \right], \quad (27)$$

where $\gamma_{\max} = 0.99$, $\gamma_{\min} = 0.80$, and k and n are adjusted to control the model sensitiveness.

- (5) Update KPCA model, and transformation matrix $\mathbf{W}_t$ is gotten. Control limits of Hotelling's T^2 and SPE statistics are updated at the same time. Then go back to Step (1) for the next adaptation.

FIGURE 1: Sample points for training the proposed algorithm.

FIGURE 2: Sample points for testing the proposed algorithm.

- (6) Model updating will not conduct if the control limits are exceeded. According to the number of continuous transfinite samples, they are judged to be outlier or fault. Let i be the number of continuous transfinite samples, initially $i = 0$. The value of i keeps adding by 1 until the next normal sample is obtained. Transfinite samples are judged to be outlier if $i < 3$; otherwise, they are judged to be fault. After the judgment, go back to Step (1) for the next sample.

FIGURE 3: T^2 statistics of testing data.

FIGURE 4: SPE statistics of testing data.

5. Experimental Results

5.1. Electrofused Magnesium Furnace Process. The purpose of the section is to test the performance of the proposed algorithm. We choose the electrofused magnesium furnace (EFMF) process to be monitoring object, because the working conditions of the EFMF are complex and changed frequently such as strong nonlinearity [35]. The shell of EFMF is round and slightly tapered, which can facilitate melting process. There are rings on the furnace wall and trolley under the furnace. When melting process has completed, move the trolley to cool. The control object is to ensure the temperature of EFMF can meet the set value. The average time of the whole EFMF multimode processes is 10 h. The current value and voltage value of three phases and the temperature of furnace all can be online measured, which provide abundant process information. The “healthy” process data is used for modeling. The “faulty” process data is used for monitoring.

In this experiment, the data set for training has 600×6 sample points and the data set for testing has 1000×6 sample points, which contains the current value and voltage value, respectively. The two data sets are printed in Figures 1 and 2. The size of the moving window is 50; the length of every step is 1. According to the steps which are given in Section 4, Hotelling's T^2 and SPE statistics are shown in Figures 3 and 4. The results indicate that the algorithm which this paper proposed can follow the variation of the object well.

5.2. Continuous Annealing Process. Continuous annealing process is a highly efficient heat treatment process after cold

FIGURE 5: Schematic diagram of continuous annealing process.

FIGURE 6: Data sample points for monitoring.

FIGURE 7: Data sample for training the algorithm.

rolling in steel works. As there are furnace zones, stable operation is necessary to product quality and continuous processing of the upstream and downstream. Process monitoring and fault diagnosis have always been the primary concern [36].

The strip in the continuous annealing line is heated in order to arrange the internal crystal of the strip. The material for annealing is a cold-rolled strip coil, which is put on a payoff reel on the entry side of the line. The head end of

the coil is then pulled out and welded with the tail end of the preceding coil. Then the strip runs through the process with a certain line speed. On the delivery side, the strip is cut into a product length by a shear machine and is coiled again by a tension reel. The schematic diagram of continuous annealing process is shown in Figure 5, in which some abbreviations are listed in Table 1.

The strip-break is a frequent fault in continuous annealing process [37]. So a strip-break fault is considered here. Four

FIGURE 8: Hotelling's T^2 statistics for the algorithm.

FIGURE 9: SPE statistics for the algorithm.

TABLE 1: Meanings of some abbreviations in Figure 5.

Abbreviation	Meaning
POR	Payoff reel
BR	Bridle roll
ELP	Entrance loop
DCR	Dancer roll
HF	Heating furnace
SF	Soaking furnace
SCF	Slow cooling furnace
C	Cooling furnace
RF	Reheating furnace
DLP	Delivery loop
TPM	Temper rolling machine
TR	Roll type reel
OA	Overageing furnace
TM	Tension model

current data sets with faults are measured from SF3R-SF6R; every data set contains 800 sample points, which is shown in Figure 6. And we choose 300 normal data (Figure 7) to train the algorithm. For the sake of easiness to calculate, normalization has been performed before. The size of the moving window is 60; the length of every step is 1. Hotelling's T^2 and SPE statistics are shown in Figures 8 and 9.

6. Conclusion

This paper has studied an algorithm for monitoring nonlinear time-delay process. In the paper, we propose an adaptive KPCA algorithm, which is combined with MWPCA and EWPCA. The proposed algorithm mapped the sample set in feature space and prejudices the new available sample with MKPCA method to decide whether the model is updated. Then update the KPCA model using EWKPCA method. This algorithm can reduce negative impact of outliers on model; the updating after prejudgment can reduce computational complexity and be more efficient. The experimental results show that the proposed algorithm is effective.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] S. Li and J. Wen, "A model-based fault detection and diagnostic methodology based on PCA method and wavelet transform," *Energy and Buildings*, vol. 68, part A, pp. 63–71, 2014.
- [2] B. S. Chen, H. N. Wu, and S. F. Y. Li, "Development of variable pathlength UV-vis spectroscopy combined with partial-least-squares regression for wastewater chemical oxygen demand (COD) monitoring," *Talanta*, vol. 120, pp. 325–330, 2014.
- [3] S. Stubbs, J. Zhang, and J. Morris, "Multiway interval partial least squares for batch process performance monitoring," *Industrial & Engineering Chemistry Research*, vol. 52, no. 35, pp. 12451–12462, 2013.
- [4] A. Ajami and M. Daneshvar, "Data driven approach for fault detection and diagnosis of turbine in thermal power plant using Independent Component Analysis (ICA)," *International Journal of Electrical Power and Energy Systems*, vol. 43, no. 1, pp. 728–735, 2012.
- [5] 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 and Intelligent Laboratory Systems*, vol. 115, pp. 44–58, 2012.
- [6] R. Dunia, S. J. Qin, T. F. Edgar, and T. J. McAvoy, "Identification of faulty sensors using principal component analysis," *AIChE Journal*, vol. 42, no. 10, pp. 2797–2811, 1996.
- [7] V. H. Nguyen and J. C. Golinval, "Fault detection based on Kernel Principal Component Analysis," *Engineering Structures*, vol. 32, no. 11, pp. 3683–3691, 2010.
- [8] L. Zhang, W. Dong, D. Zhang, and G. Shi, "Two-stage image denoising by principal component analysis with local pixel grouping," *Pattern Recognition*, vol. 43, no. 4, pp. 1531–1549, 2010.
- [9] Y. Zhang, L. Zhang, and R. Lu, "Fault identification of nonlinear processes," *Industrial and Engineering Chemistry Research*, vol. 52, no. 34, pp. 12072–12081, 2013.

- [10] L. Wu, X. Su, and P. Shi, "Sliding mode control with bounded $\mathcal{L}_2$ gain performance of Markovian jump singular time-delay systems," *Automatica*, vol. 48, no. 8, pp. 1929–1933, 2012.
- [11] Y. W. Zhang and C. Ma, "Fault diagnosis of nonlinear processes using multiscale KPCA and multiscale KPLS," *Chemical Engineering Science*, vol. 66, no. 1, pp. 64–72, 2011.
- [12] A. Maulud, D. Wang, and J. A. Romagnoli, "A multi-scale orthogonal nonlinear strategy for multi-variate statistical process monitoring," *Journal of Process Control*, vol. 16, no. 7, pp. 671–683, 2006.
- [13] Z. Q. Ge and Z. H. Song, "Online monitoring of nonlinear multiple mode processes based on adaptive local model approach," *Control Engineering Practice*, vol. 16, no. 12, pp. 1427–1437, 2008.
- [14] Y. Zhang, "Enhanced statistical analysis of nonlinear processes using KPCA, KICA and SVM," *Chemical Engineering Science*, vol. 64, no. 5, pp. 801–811, 2009.
- [15] P. P. Odiwei and Y. Cao, "Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations," *IEEE Transactions on Industrial Informatics*, vol. 6, no. 1, pp. 36–45, 2010.
- [16] X. Liu, K. Li, M. McAfee, and G. W. Irwin, "Improved nonlinear PCA for process monitoring using support vector data description," *Journal of Process Control*, vol. 21, no. 9, pp. 1306–1317, 2011.
- [17] J. Yu, "A nonlinear kernel Gaussian mixture model based inferential monitoring approach for fault detection and diagnosis of chemical processes," *Chemical Engineering Science*, vol. 68, no. 1, pp. 506–519, 2012.
- [18] S. W. Choi, C. Lee, J. Lee, J. H. Park, and I. Lee, "Fault detection and identification of nonlinear processes based on kernel PCA," *Chemometrics and Intelligent Laboratory Systems*, vol. 75, no. 1, pp. 55–67, 2005.
- [19] M. Žvokelj, S. Zupan, and I. Prebil, "Non-linear multivariate and multiscale monitoring and signal denoising strategy using Kernel Principal Component Analysis combined with Ensemble Empirical Mode Decomposition method," *Mechanical Systems and Signal Processing*, vol. 25, no. 7, pp. 2631–2653, 2011.
- [20] L. G. Wu and W. X. Zheng, "Weighted H_∞ model reduction for linear switched systems with time-varying delay," *Automatica*, vol. 45, no. 1, pp. 186–193, 2009.
- [21] J. M. Lee, C. K. Yoo, S. W. Choi, P. A. Vanrolleghem, and I. B. Lee, "Nonlinear process monitoring using kernel principal component analysis," *Chemical Engineering Science*, vol. 59, no. 1, pp. 223–234, 2004.
- [22] Y. Zhang, S. Li, and Y. Teng, "Dynamic processes monitoring using recursive kernel principal component analysis," *Chemical Engineering Science*, vol. 72, no. 1, pp. 78–86, 2012.
- [23] T. Ogawa and M. Haseyama, "Adaptive example-based super-resolution using kernel PCA with a novel classification approach," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, article 138, 2011.
- [24] D. S. Lee, M. W. Lee, S. H. Woo, Y. J. Kim, and J. M. Park, "Multivariate online monitoring of a full-scale biological anaerobic filter process using kernel-based algorithms," *Industrial and Engineering Chemistry Research*, vol. 45, no. 12, pp. 4335–4344, 2006.
- [25] L. G. Wu, X. Z. Yang, and H.-K. Lam, "Dissipativity analysis and synthesis for discrete-time T-S fuzzy stochastic systems with time-varying delay," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 2, pp. 380–394, 2014.
- [26] L. G. Wu, X. J. Su, P. Shi, and J. B. Qiu, "A new approach to stability analysis and stabilization of discrete-time T-S fuzzy time-varying delay systems," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 41, no. 1, pp. 273–286, 2011.
- [27] J. Jeng, "Adaptive process monitoring using efficient recursive PCA and moving window PCA algorithms," *Journal of the Taiwan Institute of Chemical Engineers*, vol. 41, no. 4, pp. 475–481, 2010.
- [28] X. B. He and Y. P. Yang, "Variable MWPCA for adaptive process monitoring," *Industrial and Engineering Chemistry Research*, vol. 47, no. 2, pp. 419–427, 2008.
- [29] S. R. Ryu, I. Noda, and Y. M. Jung, "Moving window principal component analysis for detecting positional fluctuation of spectral changes," *Bulletin of the Korean Chemical Society*, vol. 32, no. 7, pp. 2332–2338, 2011.
- [30] Q. Jiang and X. Yan, "Weighted kernel principal component analysis based on probability density estimation and moving window and its application in nonlinear chemical process monitoring," *Chemometrics and Intelligent Laboratory Systems*, vol. 127, pp. 121–131, 2013.
- [31] S. Lane, E. B. Martin, A. J. Morris, and P. Gower, "Application of exponentially weighted principal component analysis for the monitoring of a polymer film manufacturing process," *Transactions of the Institute of Measurement and Control*, vol. 25, no. 1, pp. 17–35, 2003.
- [32] L. Hoegaerts, L. de Lathauwer, I. Goethals, J. A. K. Suykens, J. Vandewalle, and B. de Moor, "Efficiently updating and tracking the dominant kernel principal components," *Neural Networks*, vol. 20, no. 2, pp. 220–229, 2007.
- [33] X. Wang, U. Kruger, and G. W. Irwin, "Process monitoring approach using fast moving window PCA," *Industrial & Engineering Chemistry Research*, vol. 44, no. 15, pp. 5691–5702, 2005.
- [34] X. Liu, U. Kruger, T. Littler, L. Xie, and S. Wang, "Moving window kernel PCA for adaptive monitoring of nonlinear processes," *Chemometrics and Intelligent Laboratory Systems*, vol. 96, no. 2, pp. 132–143, 2009.
- [35] Y. Zhang, T. Chai, Z. Li, and C. Yang, "Modeling and monitoring of dynamic processes," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 2, pp. 277–284, 2012.
- [36] Q. Liu, T. Chai, and S. J. Qin, "Fault diagnosis of continuous annealing processes using a reconstruction-based method," *Control Engineering Practice*, vol. 20, no. 5, pp. 511–518, 2012.
- [37] R. Bai, Z. Zhang, and T. Chai, "Modeling and simulation for strip tension in continuous annealing process," *Journal of System Simulation*, vol. 12, pp. 5477–5481, 2007.