Statistical Analysis of Nonlinear Processes Based on Penalty Factor

Yingwei Zhang, Chuanfang Zhang, and Wei Zhang

State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Liaoning 100819, China

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

Received 25 July 2014; Accepted 17 August 2014; Published 30 September 2014

Academic Editor: Ligang Wu

Copyright © 2014 Yingwei Zhang 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 process monitoring approach is proposed for handling the nonlinear monitoring problem in the electrofused magnesia furnace (EFMF). Compared to conventional method, the contributions are as follows: (1) a new kernel principal component analysis is proposed based on loss function in the feature space; (2) the model of kernel principal component analysis based on forgetting factor is updated; (3) a new iterative kernel principal component analysis algorithm is proposed based on penalty factor.

1. Introduction

In consideration of ensuring the safety of the equipment and quality of product, the monitoring of the process performance has become an indispensable issue. In order to enforce the rationality and effectiveness of monitoring, in the last few decades, multivariate statistical process monitoring (MSPM) has been intensively researched. Particularly, principal component analysis (PCA) and partial least squares (PLS) which are widely applied in the industrial processes have been important approaches for monitoring of the process performance and some improved methods, such as kernel principal component analysis (KPCA) and kernel partial least squares (KPLS), have achieved great success in process monitoring and fault diagnosis [1–5].

As the scale of modern industrial processes is expanding and the complexity of process is increasing, how to ensure the safety of process operation and improve product quality are two issues need to be solved in industrial production enterprises [6, 7]. Process monitoring technology is an effective way to solve these two issues. Since the complexity and fluctuation of industrial processes, accurate process models are difficult to build and apply [8]. Therefore, application of traditional process monitoring methods based on qualitative or quantitative models is subject to certain limitation.

Because of the developments of intelligent instrumentations and computer technology in industrial process applications, a large number of high dimensional and strongly correlated process data is collected and stored [9–11]. It is difficult to remove redundancy and interference to extract useful information. It is an efficient monitoring technology to deal with the correlation of multivariate statistical process [12, 13].

In this paper, the following work focused on the process change caused by aging equipment, process drift, and sensor measurement errors in nonlinear industrial process [14–16].

In practical industrial process, outliers are contained in the collected data, while traditional kernel principal component analysis method is based on the assumption that there are no outliers in the sample data [17, 18]. Outliers still exist even after mapping them into feature space. Even if the sample data contains only a small amount of outliers, great negative effect will be applied on process model [19, 20]. Therefore, an advanced kernel principal component analysis method is proposed in this paper, which defines a loss function in feature space in the sense of minimum reconstruction error [21]. Then iteration with penalty will be carried to obtain the principal components, which can eliminate the adverse effects of outliers. Whenever a new sample is available, reconstruct it with the previous transfer matrix and
calculate the reconstruct error [22–25]. If the new sample is an outlier, then update model with the reconstructed sample, otherwise update model with the original sample. Simulation results show that the advanced KPCA method can reduce the impact of outliers and improve the accuracy of the process monitoring model as well [26–28].

The rest of this paper is organized as follows. Kernel principal component analysis based on loss function in the feature space is proposed in Section 2. The model updating of kernel principal component analysis based on forgetting factor is proposed in Section 3. Improved kernel principal component analysis algorithm based on penalty factor is proposed in Section 4. Fault monitoring method is proposed in Section 5. The experiment results are given to show the effectiveness of the proposed method in Section 6. Finally, conclusions are summarized in Section 7.

2. Kernel Principal Component Analysis
Based on Loss Function in the Feature Space

In practical industrial process, outliers are contained in the collected data, while traditional kernel principal component analysis method is based on the assumption that there are no outliers in the sample data [29–32]. The so-called outliers usually refer to the samples whose reconstruction error is much larger than average values and proportions are very small. Outliers still exist even after mapping them into feature space. Even if the sample data contains only a small amount of outliers, great negative effect will be applied on process model. Therefore, an improved kernel principal component analysis method is proposed in this paper, which defines a loss function in feature space in the sense of minimum reconstruction error.

2.1. Kernel Principal Component Analysis. In 1909, Mercer demonstrated the concept of positive definite kernel function and regeneration kernel Hilbert space in terms of the math and listed the necessary and sufficient condition of existence and determination of the positive definite kernel function, which is called “Mercer kernel permit conditions.” Kernel method not only has been widely used in differential geometry, differential equation, group theory, and many other mathematical disciplines and in signal processing, machine learning, the Gaussian process analysis, and many other applications but also had the very big breakthrough. Kernel methods of theoretical research and practical application attract more and more attention of scholars and experts. Scholkopf combined the kernel method with principal component analysis and formed the theory of kernel principal component analysis method.

Kernel principal component analysis as the extension of principal component analysis maintains the various mathematical and statistical properties of linear principal component analysis. KPCA uses the nonlinear mapping of data to a high dimensional eigenspace to achieve kernel matrix diagonalization and then carries on the principal component analysis. There is no need to calculate inner product of the sample data of nonlinear transformation and we can easily get the nonlinear principal component of the mapping data by the kernel function value between two data points.

2.2. Loss Function in the Feature Space. The sample number is \( N \). \( \mathbf{X} \in \mathbb{R}^{N \times M} \) is mapped into high dimensional space \( \mathbf{F}: \mathbf{X} \rightarrow \Phi(\mathbf{X}) \), where \( \Phi(\mathbf{X}) = [\Phi(\mathbf{x}_1), \Phi(\mathbf{x}_2), \ldots, \Phi(\mathbf{x}_N)] \). \( \mathbf{X} \) is supposed to have been centralized processing. \( W \) is the transformation matrix, where \( \|W\| = 1 \). \( \langle \mathbf{W}, \Phi(\mathbf{x}) \rangle \mathbf{W} = \mathbf{W}^T \Phi(\mathbf{x}) \) is reconstruction vector of \( \Phi(\mathbf{x}) \); then the reconstruction error of \( \Phi(\mathbf{x}) \) in the feature space is defined as follows:

\[
\|\mathbf{e}(\Phi(\mathbf{x}))\|^2 = \|\Phi(\mathbf{x}) - \mathbf{W}^T \Phi(\mathbf{x})\|^2 = \Phi(\mathbf{x}) \cdot \Phi(\mathbf{x}) - 2\mathbf{W}^T \Phi(\mathbf{x}) \Phi(\mathbf{x}) + (\mathbf{W}^T)^2 \Phi(\mathbf{x}) \Phi(\mathbf{x})
\]

(1)

In order to minimize the reconstruction error, here the loss function in feature space is defined as follows:

\[
J_1(W) = \sum_{i=1}^{N} \|\Phi(\mathbf{x}_i) - \mathbf{W}^T \Phi(\mathbf{x}_i)\|^2.
\]

(2)

2.3. Kernel Principal Component Analysis Based on Loss Function. Formula (2) is expanded as follows:

\[
J_1(W) = \sum_{i=1}^{N} (\mathbf{W}^T \Phi(\mathbf{x}_i))^2 \|\mathbf{W}\|^2 - 2 \sum_{i=1}^{N} (\mathbf{W}^T \Phi(\mathbf{x}_i))^2 \\
+ \sum_{i=1}^{N} \|\Phi(\mathbf{x}_i)\|^2
\]

(3)

\[
= - \sum_{i=1}^{N} (\mathbf{W}^T \Phi(\mathbf{x}_i))^2 + \sum_{i=1}^{N} \|\Phi(\mathbf{x}_i)\|^2,
\]

where \( \sum_{i=1}^{N} \|\Phi(\mathbf{x}_i)\|^2 \) is constant. Therefore \( J_1(W) \), the loss function in the feature space, is the minimum, when \( \mathbf{W}^T \Phi(\mathbf{x}) \) is the maximum. This is equivalent to solving the following optimization problem:

\[
\max_{\mathbf{W},\mathbf{z}} \frac{1}{2} \sum_{i=1}^{N} (\mathbf{z}_i)^2 - \frac{1}{2} \mathbf{W}^T \mathbf{W}
\]

\[
\text{s.t.} \quad \mathbf{z}_i = \mathbf{W}^T \Phi(\mathbf{x}_i).
\]

(4)

By Lagrange multiplier method, it is obtained as follows:

\[
L(\mathbf{W}, \mathbf{z}, \alpha) = \frac{1}{2} \sum_{i=1}^{N} (\mathbf{z}_i)^2 - \frac{1}{2} \mathbf{W}^T \mathbf{W} - \sum_{i=1}^{N} \alpha_i (\mathbf{z}_i - \mathbf{W}^T \Phi(\mathbf{x}_i)).
\]

(5)
Then
\[
\frac{\partial L}{\partial W} = 0 \implies W = \sum_{i=1}^{N} \alpha_i \Phi(x_i)
\]
\[
\frac{\partial L}{\partial z_i} = 0 \implies \alpha_i = \gamma z_i
\]
\[
\frac{\partial L}{\partial \alpha_i} = 0 \implies ... reduce
\]
the influence on process model and even disappear, which
ensuresthatthemodelisadaptivetothetime-varyingsystem.

By eliminating \( W \) and \( z \), it is obtained as follows:
\[
1 - \frac{\alpha_i}{\gamma} - \sum_{i=1}^{N} \alpha_i \Phi(x_i)^T \Phi(x_i) = 0, \quad i = 1, 2, \ldots, N.
\]

It is defined as \( \lambda = 1/\gamma \) and then obtained as follows:
\[
K\alpha = \lambda \alpha,
\]
where \( K_{ij} = \langle \Phi(x_i), \Phi(x_j) \rangle, i, j = 1, 2, \ldots, N \). The result of \( \alpha \)
is the eigenvector of \( K \) in the above formula.

Because \( \alpha \) satisfies the normalization condition \( \lambda \alpha^k \alpha^k = 1 \), then it is obtained as follows:
\[
(W^k, W^k) = \sum_{i,j=1}^{N} \alpha_i^k \alpha_j^k \langle \Phi(x_i), \Phi(x_j) \rangle
\]
\[
= \sum_{i,j=1}^{N} \alpha_i^k \alpha_j^k K(x_i, x_j) = \lambda_k (\alpha^k, \alpha^k) = 1.
\]

In this way, the projection of \( \Phi(x) \) in the \( k \)th principal component \( W^k \) is as follows:
\[
t_k = \langle W^k, \Phi(x) \rangle = \sum_{i=1}^{N} \alpha_i^k \langle \Phi(x_i), \Phi(x) \rangle = \sum_{i=1}^{N} \alpha_i^k k(x_i, x).
\]

3. The Model Updating of Kernel
Principal Component Analysis
Based on Forgetting Factor

When the running state of the process changes in multivariate statistical process monitoring, regardless of the system changing slowly or fast, mean and covariance matrix of the model will change. Therefore, when the system changes, it needs to update mean and covariance of the sample data set.

Firstly, the method of updating PCA model based on forgetting factor is introduced. Then KPCA model updating method based on forgetting factor will be got by kernel method. When new samples are collected in the process, mean and covariance will change. These changes depend on the change degree of model structure, namely, the size of the forgetting factor. For time-varying Gaussian process, therefore, mean and covariance of forgetting factor in estimate time can be used. Its formula is as follows:
\[
m_t = \frac{\sum_{i=1}^{t} \alpha_{t-i} x_i}{\sum_{i=1}^{t} \alpha_{t-i}} = \frac{x_t + \alpha x_{t-1} + \cdots + \alpha^{t-1} x_1}{1 + \alpha + \cdots + \alpha^{t-1}},
\]
\[
S_t = \frac{\sum_{i=1}^{t} \beta_{t-i} (x_i - m_t) (x_i - m_t)^T}{\sum_{i=1}^{t} \beta_{t-i}},
\]
where \( \alpha, \beta \) are two forgetting factors, \( m_t \) and \( S_t \) are mean vector and covariance matrix of \( t \). Mean vector and covariance matrix of the sample data can be in a more convenient form, namely, the weighted sum of mean and covariance matrix in \( t \) moment and sample data in \( t \) moment. Its formula is as follows:
\[
m_t = 1 - \frac{\alpha}{1 - \alpha} x_t + \frac{1 - \alpha^{t-1}}{1 - \alpha} m_{t-1},
\]
\[
S_t = 1 - \frac{\beta}{1 - \beta} x_t x_t^T + \frac{1 - \beta^{t-1}}{1 - \beta} S_{t-1},
\]
where \( x_t = x_t - m_t \) is the new centralized sample of \( t \) moment. As \( t \) increases gradually, formula (12) can be further simplified as
\[
m_t = (1 - \alpha) x_t + \alpha m_{t-1},
\]
\[
S_t = (1 - \beta) x_t x_t^T + \beta S_{t-1}.
\]

It can be seen from formula (14) that if only considering the sample covariance matrix, then it is obtained as
\[
D_t = (1 - \beta) x_t x_t^T + \beta S_{t-1},
\]
where \( D_t \) is a diagonal matrix, whose diagonal elements are the same as diagonal elements of \( S_t \). The correlation coefficient matrix can also be estimated as follows:
\[
R_t = \frac{\sum_{i=1}^{t} \beta_{t-i} D_i^{-1/2} x_i x_i^T D_i^{-1/2}}{\sum_{i=1}^{t} \beta_{t-i}}.
\]

As \( t \) increases gradually, formula (16) can be further simplified as
\[
R_t = (1 - \beta) D_t^{-1/2} x_t x_t^T D_t^{-1/2} + \beta R_{t-1}.
\]

As shown in formula (11), the updating of sample date set's mean and covariance matrix needs determining two weighting coefficients, which are called the forgetting factors. If both forgetting factors are 1, its mean vector and covariance matrix will have the highest similarity degree with mean vector and covariance matrix calculated by all the sample data. If forgetting factors less than 1, as the process running, the weight of adding on the old data will be smaller and smaller, until being eliminated automatically, without discarding the old data by human. Then old data will gradually reduce the influence on process model and even disappear, which ensures that the model is adaptive to the time-varying system.
When forgetting factors are closer to 1, there will be more number of sample data taking effect on the current process model.

So far, most of the model updating method is based on the constant value forgetting factor which is acquired by experience. However, the optimal value of forgetting factor depends on the degree of process change. Because process changes at different levels, the optimal value of forgetting factor will have a significant difference. When the process changes rapidly, the update rate of mode should be very large, namely, a few new data have the main influence on process model. When process changes slowly, update rate of the model should be smaller, namely, most of sample data has influence on process model. The basic process information will be in a long time to maintain its effectiveness. But the degree of process changes is changing with time in the actual industrial process, and then forgetting factor should be determined according to the actual situation of process changes. In order to deal with process which has nonconstant change degree, constant forgetting factor will not be used. Forgetting factor which is adjusted to different degree of process changes is being used. Here, Fortescue’s method is used to adjust the forgetting factor. The method adopts the process changes is being used. Here, Fortescue’s method is change degree, constant forgetting factor will not be used. The degree of process change will be in a long time to maintain its effectiveness. But influence on process model. The basic process information themodelshould bemore smallernamelymostofsampledatahas influence on process model. When process changes slowly, update rate of forgetting factor will have a significant difference. When the process depends on the degree of process change. Because process experience. However, the optimal value of forgetting factor which is acquired by method. Then the update KPCA method based on forgetting factor is obtained. Let kernel matrix be $\mathbf{K}_{t-1}$ at $t-1$ moment.

According to exponential weighting KPCA method, recursive update formula of kernel matrix at $t$ moment is as follows:

$$
\mathbf{K}_t = y_t \mathbf{K}_{t-1} + (1 - y_t) \mathbf{K}_{t-1}^{T} \mathbf{K}_t,
$$

(20)

where $y_t$ is the weighting factor; it can be calculated according to the formula as follows:

$$
y_t = \gamma_{max} - (\gamma_{max} - \gamma_{min}) \left[ 1 - \exp \left( -k \left( \| \Delta \mathbf{R} \| \right) \right) \right],
$$

(21)

where $\gamma_{min} = 0.1$, $\gamma_{max} = 0.99$, and $\| \Delta \mathbf{R} \|$ is the Euclidean norm of the difference between two consecutive correlation coefficient matrices. The sensitivity of model is controlled by $k$, whose default value is $k = 0.6931$.

### 4. Improved Kernel Principal Component Analysis Algorithm Based on Penalty Factor

KPCA algorithm based on eigenvalue decomposition is a batch-mode algorithm, which needs to know all the sample points before modeling. It is not suitable for online monitoring or samples increased gradually. And KPCA is often based on the assumption that sample is not contaminated by outliers. There are outliers in the actual samples, in this paper, an iterative kernel principal component analysis method based on penalty factor is presented to solve the problem of outliers in the sample.

#### 4.1. Iterative Kernel Principal Component Analysis Algorithm

For the loss function defined by formula (3), the stochastic gradient descent method is used to solve the optimization problem as follows:

$$
\frac{dJ_x(W)}{dW} = 2 \left( \mathbf{W} \mathbf{W}^T \mathbf{K} - \mathbf{K} \right) \mathbf{W}.
$$

(22)

Then iterative formula is as follows:

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

(23)

where $\mu_n$ is the iteration step length, $0 < \mu_n < 1$, and $\mathbf{W}_n$ is convergence to the first principal component.

Because the nonlinear principal components are orthogonal to each other, Schmidt orthogonal method is used to calculate the $j$th principal component $\mathbf{W}^j$:

$$
\mathbf{W}_{n+1}^j = \mathbf{W}_n^j + \mu_n \left( \mathbf{K}^j - \mathbf{W}_n^j \mathbf{W}_n^{Tj} \mathbf{K}^j \right) \mathbf{W}_n^j
$$

(24)

$$
\mathbf{K}^j = \mathbf{K} - \sum_{i=1}^{j-1} \mathbf{W}_n^i \mathbf{W}_n^{Tj}. \mathbf{K}.
$$

(25)

Steps of iterative KPCA algorithm can be summarized as follows.

1. Input: sample data set $\mathbf{X}$, maximum iterations $n_{max}$, the initial iteration step $n = 1$, and principal component $j = 1$. 
(2) Calculate the kernel matrix \( K \), where \( [K]_{ij} = K_{ij} = \langle \Phi(x_i), \Phi(x_j) \rangle = k(x_i, x_j) \). Then carry on the centralized processing \( \overline{K} = K - I_N K - K I_N + I_N K I_N \), where

\[
I_N = \frac{1}{N} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}
\]

(26)

(3) Calculate the \( j \)th principal component \( W_j \).

(4) \( |W_{n+1}^j - W_n^j| > \varepsilon \) and \( n < n_{\text{max}} n = n + 1 \) and return to step (3). \( |W_{n+1}^j - W_n^j| \leq \varepsilon \) output: \( W_{n+1}^j \).

(5) \( J_2(W) = \sum_{i=1}^N C_i \| \Phi(x_i) - WW^T \Phi(x_i) \|^2 + \eta(1-C_i) \geq \xi, j = j + 1 \). Return to step (3) and calculate next principal component.

(6) \( J_2(W) = \sum_{i=1}^N C_i \| \Phi(x_i) - WW^T \Phi(x_i) \|^2 + \eta(1-C_i) \leq \xi \). Terminate the iteration and output \( W \).

4.2. Iterative Kernel Principal Component Analysis Algorithm Based on Penalty Factor. Although there are little outliers in sample data in KPCA algorithm, it also has great influence on KPCA model. The calculated principal components are towards the direction of outliers in order to reduce the overall square errors in the process of calculating principal components. In order to reduce the influence of the outliers on KPCA model, the penalty factor is added to the square error formula in the feature space.

Penalty factor \( \eta(1-C_i) \) is added to formula (3):

\[
J_2(W) = \sum_{i=1}^N \left( C_i \left( \| \Phi(x_i) - WW^T \Phi(x_i) \|^2 + \eta(1-C_i) \right) \right), \tag{27}
\]

where \( \eta \) is the predefined threshold and \( \eta > 0 \). \( C_i \) is defined as follows:

\[
C_i = \begin{cases} 
1 & \| \Phi(x_i) - WW^T \Phi(x_i) \|^2 \leq \eta, \\
0 & \| \Phi(x_i) - WW^T \Phi(x_i) \|^2 > \eta.
\end{cases} \tag{28}
\]

By the above formula, after adding penalty factor, points which exceeded predefined threshold \( \eta \) are seen as outliers. After setting \( J_2(W) \) of outliers as \( \eta \), the influence on KPCA model is reduced. Noticing that \( C_i \) is discrete, in order to use the proposed iterative KPCA to calculate principal components, continuous Sigmoid function is adopted to approximate discrete variable \( C_i \).

Minimum of error function in formula (25) is calculated, and iterative formula is obtained as follows:

\[
W_{n+1} = W_n + \frac{1}{1 + e^{(1+\eta)|x_i \cdot (\Phi(x_i))|^2}} (K - WW^T K) W_n, \tag{29}
\]

where \( 0 < \mu_n < 1 \) is iteration step length, and \( 1/(1 + e^{(1+\eta)|x_i \cdot (\Phi(x_i))|^2}) \) is continuous Sigmoid function, which can adjust parameters according to the current input values and eliminate the influence of the outliers on KPCA model. So the smaller the threshold value \( \eta \) is, the more the sample points will be treated as outliers.

Because the nonlinear principal components are orthogonal to each other, Schmidt orthogonal method is used to calculate the \( j \)th principal component \( W_j \):

\[
W_{n+1}^j = W_n + \frac{1}{1 + e^{(1+\eta)|x_i \cdot (\Phi(x_i))|^2}} (K - WW^T K) W_n, \tag{30}
\]

\[
K^j = K_0 - \sum_{i=1}^{j-1} W_i^j W_i^T K_0.
\]

Steps of iterative KPCA algorithm based on penalty factor can be summarized as follows.

(1) Input: sample data set \( X \), maximum iterations \( n_{\text{max}} \), the initial iteration step \( n = 1 \), and principal component \( j = 1 \).

(2) Calculate the kernel matrix \( K \), where \( [K]_{ij} = K_{ij} = \langle \Phi(x_i), \Phi(x_j) \rangle = k(x_i, x_j) \). Then carry on the centralized processing \( \overline{K} = K - I_N K - K I_N + I_N K I_N \), where

\[
I_N = \frac{1}{N} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}
\]

(26)

(3) Calculate the \( j \)th principal component \( W_j \). If \( j = 1 \), formula (29) is used to calculate the first principal component \( W^1 \). From the second iteration, sample reconstruction error is calculated in each time. Take points with ratio of \( r \) as outliers, determine the threshold \( \eta \), and calculate penalty factor for iteration. Then the first principal component is obtained.

(4) If \( j = 2, 3, \ldots, l \), formula (30) is used to calculate the rest of principal components \( W^2, W^3, \ldots, W^l \).

(5) \( |W_{n+1}^j - W_n^j| > \varepsilon \) and \( n < n_{\text{max}} n = n + 1 \) and return to step (3). \( |W_{n+1}^j - W_n^j| \leq \varepsilon \) output: \( W_{n+1}^j \).

(6) \( J_2(W) = \sum_{i=1}^N C_i \| \Phi(x_i) - WW^T \Phi(x_i) \|^2 + \eta(1-C_i) \geq \xi, j = j + 1 \). Return to step (3) and calculate next principal component.

(7) \( J_2(W) = \sum_{i=1}^N C_i \| \Phi(x_i) - WW^T \Phi(x_i) \|^2 + \eta(1-C_i) \leq \xi \). Terminate the iteration and output \( W \).

5. Fault Monitoring Method

This section provides fault monitoring method using the proposed iterative kernel principal component analysis algorithm based on penalty factor. It can be broadly divided into offline modeling phase and online monitoring phase.
Establish KPCA model based on historical data, standardize the kernel matrix, and calculate the statistic control limits.

Calculate the mean, covariance matrix, and statistic control limits at $t - 1$ moment.

Collect new sample data and calculate forgetting factor and mean and covariance matrix at $t$ moment.

Calculate the kernel vector of new sample and standardize it.

Use iterative KPCA method based on penalty factor to update model and calculate the statistic control limits.

Use forgetting factor to update the kernel matrix

$$K_t = \gamma_t K_{t-1} + (1 - \gamma_t) \bar{K}_t \bar{K}_t^T$$

$$\gamma_t = \gamma_{max} - (\gamma_{max} - \gamma_{min}) \left[ 1 - \exp \left( -k \left( \| \Delta R \| \right) \right) \right]$$

Figure 1: Flow chart of improved kernel principal component analysis method.

Figure 2: Statistics of $T^2$ and SPE process monitoring using improved KPCA in normal working condition.
5.1. Offline Modeling Phase

(1) KPCA model is established based on historical data, and the initial standardization of kernel matrix is obtained.

(2) Set the ratio of outliers $r$, determine the threshold $\eta$, and calculate penalty factor and principal components $W$.

(3) Calculate $T^2$ and SPE statistics and the corresponding control limits.

5.2. Online Monitoring Phase

(1) Calculate the mean $m_{t-1}$, covariance matrix $S_{t-1}$, $T^2_{\text{lim},t-1}$, and SPE$_{\text{lim},t-1}$ at $t-1$ moment.

(2) Collect new samples data of $t$ moment; calculate forgetting factors $\alpha_t$, $\beta_t$, the mean $m_t$, and covariance matrix $S_t$.

(3) Obtain kernel matrix $K_t$ of new samples and standardize it.

(4) Use the method of iteration kernel principal component to update KPCA model and calculate $T^2$ and SPE statistics and the corresponding control limits $T^2_{\text{lim},t}$, SPE$_{\text{lim},t}$.

(5) Collect new sample data and return to step (3).

The flow chart of improved KPCA algorithm is shown in Figure 1.

6. Experiment and Discussion

With the development of technology of melting, electrofused magnesia furnace has already gotten extensive application in the industry. Electrofused magnesia furnace refining technology can enhance the quality and increase the production variety. The working conditions of the electrofused magnesia furnace are changed frequently and have complex characteristics such as strong nonlinearity and multiple modes. Electrofused magnesia furnace production process is used for fault diagnosis to verify the effectiveness of the proposed statistical analysis of nonlinear processes based on penalty factor. The improved KPCA method is used to monitor the normal and failure condition, respectively. Process fault is introduced from the 700th sample. It is caused by abnormal electrode actuators. Current of electrofused magnesia furnace plunges sharply. Temperatures become abnormal.

800 pieces of sample data in normal working condition is used to test the improved KPCA process monitoring method proposed in this paper. Then $T^2$ and SPE statistics of improved KPCA method are obtained in normal working condition, as shown in Figure 2. It can be seen from Figure 2 that the changes of $T^2$ and SPE statistics in improved KPCA method are reduced. This is because penalty factors are used to punish deviation larger samples in the iterative calculation process of principal components. The distance between sample points and original points of principal components is reduced. Therefore, $T^2$ and SPE statistic fluctuation decrease. But only iterative KPCA is used to model without updating the control limits; fault alarms still exist in the process. If $T^2$ and SPE control limits are updated at the same time, the statistics will not overrun the limits obviously. Compared with traditional methods, the proposed method has better accuracy and lower fault alarm rate. Simulation results verify the feasibility of this method to eliminate outliers.

In order to monitor the process of fault condition, faults are added to sample data in the normal working condition. Process faults are introduced from the 700th sample. The parameters start to drift and change faster at this time.
Improved KPCA and conventional KPCA are used to monitor the process of fault condition. And process monitoring charts are shown in Figure 3. Figure 3 is statistics of $T^2$ and SPE process monitoring using improved KPCA in fault condition. From Figure 3 you can see under the condition of process parameter drift, when faults do not occur in the process, the improved KPCA method can eliminate the influence of outliers and better describe the process of change. When faults occur in the process, the improved KPCA method can accurately and timely find them. Compared with the traditional method, the improved KPCA method has better accuracy and lower fault alarm rate.

7. Conclusion

In order to solve the problem of outliers in the sample data, an improved KPCA method is proposed in this paper. The method is based on loss function in feature space. And forgetting factor is introduced into recursive update of kernel matrix. Then penalty factor is added to calculate the process monitoring model. Compared with conventional KPCA method, improved KPCA method proposed in this paper does further research on eliminating outliers. Iterative KPCA method is more suitable for online monitoring of the process. Adding the penalty factor has good effect in eliminating outliers. In this paper, MATLAB software is used to do simulation experiments, and the simulation results verify the feasibility of the method. The improvement of KPCA method is more useful in the process of monitoring contained outliers.

Conflict of Interests

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

References


