

文章编号:1004-9037(2013)06-0000-00

# 基于 KICA-KFDA 的集成故障识别算法

许 洁<sup>1</sup> 赵 瑾<sup>1</sup> 刘如成<sup>1</sup> 胡寿松<sup>2</sup>

(1. 南京师范大学电气与自动化工程学院, 南京, 210042; 2. 南京航空航天大学自动化学院, 南京, 210016)

**摘要:**针对复杂的化工过程,提高过程监控能力,提出基于核独立成分分析(Kernel independent component analysis, KICA)和核 Fisher 判别分析(Kernel fisher discriminant analysis, KFDA)的过程监测与故障识别方法。通过利用核独立成分分析建立正常工况模型,得到检测故障信息。在发生故障的情况下,利用 Fisher 判别分析方法在高维的特征空间的特点和优势,可求出满足最大分离程度的核 Fisher 判别向量和特征向量,根据当前故障的判别向量和历史故障数据集中所含故障的最优核 Fisher 判别向量的相似度进行故障识别。仿真结果验证了所提方法的有效性。

**关键词:**过程监控;故障识别;核独立成分分析;核 Fisher 判别分析

**中图分类号:**TP277 **文献标志码:**A

## Integrated Fault Identification Algorithm Based on KICA and KFDA

Xu Jie<sup>1</sup>, Zhao Jin<sup>1</sup>, Liu Rucheng<sup>1</sup>, Hu Shousong<sup>2</sup>

(1. School of Electrical and Automation Engineering, Nanjing Normal University, Nanjing, 210042, China;

2. College of Automation Engineering, Nanjing University of Aeronautics & Astronautics, Nanjing, 210016, China)

**Abstract:** To improve the statistical monitoring performance of complex chemical process, a new statistical process monitoring and fault identification method having the character of non-linear based on kernel independent component analysis (KICA) and kernel fisher discriminant analysis (KFDA) is proposed. KICA is used to establish the normal operating conditions and identify the fault. If a fault occurs, the nuclear fisher discriminant vector and feature vector of the process data are extracted from the Fisher subspace. Thus, the batch normal or not can be detected by comparing distance with the predefined threshold. Comparing the present discriminant vector and the optimal discriminant vector of fault in historical data set, the similar degree can be identified. According to the similar degree, the perform fault can be diagnosed. The results of simulating demonstrate that the proposed method can efficient in detecting and diagnosing the malfunctions, with more accurate result.

**Key words:** process monitoring; fault identification; kernel independent component analysis; kernel fisher discriminant analysis

## 引 言

自 20 世纪 90 年代以来,多元统计过程控制成功应用于工业过程系统<sup>[1-5]</sup>。然而传统的统计性能监控方法均作了如下的假设:监测对象的测量数据满足高斯变量中的独立同分布,而实际的工业过程

中很难满足过程数据服从正态的前提,而且在复杂的工业过程中变量之间经常表现出较强的非线性。

由此本文提出了基于 KICA 和 KFDA 的过程监控与故障诊断方法。该方法通过核独立成分分析建立正常工况模型,监控统计量  $U^2$ ,  $U_s^2$  和 SPE, 并采用核密度估计方法确定相应的控制置信限,检测故障信息。如果一旦发生故障,则利用 Fisher

判别分析方法,在高维的特征空间中求出满足最大分离程度的核 Fisher 判别向量和特征向量,然后参照当前故障的判别向量和由历史故障数据集中所含故障的所得到的最优核 Fisher 判别向量的相似度实施故障诊断。最后以田纳西-伊斯曼化工过程为例,对提出的监控诊断方法进行仿真研究,验证了该方法在非线性过程监控与故障诊断中的有效性。

## 1 基于 KICA 的故障检测策略

### 1.1 KICA 算法

考虑观测数据序列  $\{x_1, x_2, \dots, x_N\}$  ( $N$  为观测样本数目),通过非线性核函数  $\Phi: R^m \rightarrow F$  把观测数据映射到高维特征空间,则数据在特征空间的协方差矩阵为  $S^F = \frac{1}{N} \sum_{i=1}^N \Phi(x_i) \Phi(x_i)^T$ ,若  $Q = [\Phi(x_1), \dots, \Phi(x_N)]$ ,那么  $S^F$  可进一步表示为  $S^F = (1/N)QQ^T$ 。定义  $N \times N$  矩阵  $K$

$$K = K_{ij} = \langle \Phi(x_i) \cdot \Phi(x_j) \rangle = K(x_i, x_j) \quad (1)$$

式中:  $K_{ij}$  为满足 Mercer 条件的核函数,有  $K = QQ^T$ 。可由下式得到核矩阵的中心化

$$\tilde{K} = K - E_N K - K E_N + E_N K E_N \quad (2)$$

式中:  $E_N = \frac{1}{N} \begin{bmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{bmatrix} \in R^{N \times N}$ 。归一化处理

$\tilde{K}$ ,可得到  $\hat{K} = \frac{\tilde{K}}{\text{trace}(\tilde{K})/N}$ ,通过求解如下特征方

程:  $\lambda \alpha = \hat{K} \alpha$ ,可以得到前  $d$  个最大特征值  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$  所对应的特征向量为  $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_d$ ;进一步可得到  $S^F$  的特征值为  $\lambda_1/N, \lambda_2/N, \dots, \lambda_d/N$ ,与之相对应的正交特征向量  $v_1, v_2, \dots, v_d$  可表示为

$$v_i = \frac{1}{\sqrt{\lambda_i}} Q \alpha_i \quad (i=1, 2, \dots, d) \quad (3)$$

其矩阵形式为  $V = QHA^{-1/2}$ ,其中  $H = [\alpha_1, \alpha_2 \dots \alpha_d]$ ,  $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_d)$ 。

为使协方差矩阵  $S^F$  变为对角矩阵

$$S^F = V \text{diag}\left(\frac{\lambda_1}{N}, \frac{\lambda_2}{N}, \dots, \frac{\lambda_d}{N}\right) V^T = \frac{1}{N} V \Lambda V^T \quad (4)$$

令  $G = V\left(\frac{1}{N}\Lambda\right)^{-1/2} = \sqrt{N}QHA^{-1}$ ,则  $G^T S^F G = I$ 。

于是可以得到转换矩阵  $G$ ,采用  $y = G^T \Phi(x)$  进行白化处理,可以进一步整理为

$$y = G^T \Phi(x) = \sqrt{N} \Lambda^{-1} H^T Q^T \Phi(x) =$$

$$\sqrt{N} \Lambda^{-1} H^T [\Phi(x_1), \dots, \Phi(x_N)]^T \Phi(x) =$$

$$\sqrt{N} \Lambda^{-1} H^T [\tilde{K}(x_1, x), \dots, \tilde{K}(x_N, x)]^T = \sqrt{N} \Lambda^{-1} H^T \tilde{K} \quad (5)$$

其中  $H = [\alpha_1, \alpha_2, \dots, \alpha_d]$ ,  $\Lambda = [\lambda_1, \lambda_2, \dots, \lambda_d]$ ,  $z$  满足  $E\{yy^T\} = I^{[6]}$ 。

寻找分离矩阵  $W$  是独立成分分析的主要任务,这样得到的变换后的信号是源信号的估计。本文利用由 Hyvarinen 提出的快速 ICA 算法,在计算过程中为避免因随机选择初值而导致的计算结果不收敛的问题,增加初值的选择次数,另外,按照通过计量所得到的独立成分的负熵值,进行其按非高斯分布的原则进行大小的排列,从中选取独立成分分量的个数。其算法见下面的步骤过程:

(1)对测量数据进行基于标准化和白化的处理;

(2)选取独立成分的个数  $n=m$ ,令  $i=1$ ;

(3)对  $w(i)$  (分离矩阵的组成向量)选择单位标准化的初始随机向量值;

(4)迭代更新  $w(i)$ :

$$w(i) = E\{zg(w(i)^T z)\} - E\{g'(w(i)^T z)\} w(i)$$

式中:  $E$  表示数学期望,可由样本值进行估计获得,  $z$  是均值为 0,方差为 1 的高斯变量,  $g$  为函数  $G$  的一阶导数,  $g'$  为函数  $G$  的一阶导数;

(5)正交化

$$w(i) = w(i) - \sum_{j=1}^{i-1} (w(i)^T w(j)) w(j)$$

(6)标准化  $w(i) = \frac{w(i)}{\|w(i)\|}$ ;

(7)如果  $w(i)$  不收敛,返回第(4)步;

(8)如果  $w(i)$  收敛,则输出向量  $w(i)$ 。如果  $i \leq m$ ,令  $i=i+1$  返回第(3)步。

关于 FastICA 算法的详细描述可参见文献 [7]。

### 1.2 监控统计量和控制限

给定观测数据为  $x_1, x_2, \dots, x_N \in R^M$ ,在用核独立成分分析方法进行运算的过程中,先将数据信息映射到高维空间进行标准化的处理后再进行运算及监控。设  $X_{\text{normal}}$  为在 KPCA 空间中经过标准化处理后的数据,那么可由 FastICA 算法得到转换矩阵  $W$  和特征空间中的独立主元  $S_{\text{normal}}$ 。将  $W$  矩阵按一定的范数即 Euclidean 范数 ( $L_2$ ) 进行大小的排列,可用下式表示  $W$  矩阵每一行  $w_i$  的  $L_2$  范数:  $\arg, \text{Max} \|w_i\|$  [8]。独立成分的排序便完成了,接下来就是确定独立主元个数,根据负熵的最小化原则选择独立成分维数  $d$ ,选择  $W$  前  $d$  列构

成矩阵  $\mathbf{W}_d$  ( $\mathbf{W}$  的主要部分),  $\mathbf{W}$  的剩余部分构成  $\mathbf{W}_e$ 。新采样时刻的数据为  $\mathbf{X}_{\text{new}}(k)$ , 其在主元和残差空间中的对应独立主元分量为<sup>[8]</sup>  $\hat{\mathbf{S}}_{\text{newd}}(k) = \mathbf{W}_d \mathbf{X}_{\text{new}}(k)$  和  $\hat{\mathbf{S}}_{\text{newe}}(k) = \mathbf{W}_e \mathbf{X}_{\text{new}}(k)$ , 与 ICA 和 PCA 中类似, 在核主元特征空间中, 定义采样时刻  $k$  的  $U^2$  统计量<sup>[9]</sup>

$$\mathbf{U}^2(k) = \hat{\mathbf{S}}_{\text{newd}}(k)^T \hat{\mathbf{S}}_{\text{newd}}(k) \quad (6)$$

通过计算独立主元向量的平方求得  $U^2$  统计量, 从主元模型内部的波动反映了过程的变化情况。SPE 统计量表示采样点与模型空间的距离, 它反映了测量值对模型的偏离程度, 即

$$\text{SPE}(k) = \mathbf{e}(k)^T \mathbf{e}(k) = (\mathbf{x}(k) - \hat{\mathbf{x}}(k))^T (\mathbf{x}(k) - \hat{\mathbf{x}}(k)) \quad (7)$$

其中  $\hat{\mathbf{x}}(k) = \mathbf{Q}^{-1}(\mathbf{W}_d \mathbf{Q}^{-1})^T \mathbf{W}_d \mathbf{x}(k)$ 。

基于残差空间的  $T^2$  统计量来监控系统的状态由 Simoglou 等<sup>[10]</sup> 提出。本文采用了类似于 Simoglou 的方法, 定义基于残差空间核独立主元 ( $\hat{\mathbf{S}}_{\text{newe}}$ ) 的统计量  $U_e^2$

$$\mathbf{U}_e^2(k) = \hat{\mathbf{S}}_{\text{newe}}(k)^T \hat{\mathbf{S}}_{\text{newe}}(k) \quad (8)$$

从上式可以看出,  $U_e^2$  可以对由于独立成分选择恰当对  $U^2$  统计量的影响起补偿作用, 两者相互补充, 从而可以更好地反映系统的运行状态。

在采用 ICA 方法实施过程监控时, 因为变量分布并不满足多变量高斯分布特征, 则不能采用 PCA 中采用的置信限确定方法, 此时采用即核密度估计方法<sup>[11]</sup> 确定统计量的控制限, 一个单变量统计量的核密度估计可表示为<sup>[9]</sup>

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K \left\{ \frac{x - x_i}{h} \right\} \quad (9)$$

式中,  $n$  为样本数,  $h$  为平滑参数,  $K$  为核函数,  $x$  是待求的控制界限值,  $x_i$  是由样本数据集中所得到的观测值。核函数的形式有多种, 但事实上核密度估计的精确度与核函数的类型关系不是很大, 反而与平滑参数  $h$  有很大的关系<sup>[11]</sup>, 本文采用高斯核函数。

## 2 基于 KFDA 的故障诊断策略

### 2.1 KFDA 算法

假设, 数据矩阵为  $\mathbf{X}$ , 它是由数量为  $N$  的观测变量  $x \in \mathbf{R}^n$  组成,  $M$  代表观察变量中不同的模式类,  $N_i (i=1, 2, \dots, M)$  代表模式类中的分量。  $\Phi(x) \in \mathbf{H}$  为模式向量, 它由数据非线性映射到  $\Phi$  后产生, 可计算出样本数据在高维特征空间  $H$  中的类内离散度矩阵、类间离散度矩阵和总体离散度矩

阵分别为

$$\mathbf{S}_w^\Phi = \frac{1}{N} \sum_{i=1}^M \sum_{j=1}^{N_i} (\Phi(x_j^i) - \mathbf{m}_i^\Phi) (\Phi(x_j^i) - \mathbf{m}_i^\Phi)^T \quad (10)$$

$$\mathbf{S}_b^\Phi = \sum_{i=1}^M \frac{N_i}{N} (\mathbf{m}_i^\Phi - \mathbf{m}_o^\Phi) (\mathbf{m}_i^\Phi - \mathbf{m}_o^\Phi)^T \quad (11)$$

$$\mathbf{S}_t^\Phi = \frac{1}{N} \sum_{j=1}^N (\Phi(x_j) - \mathbf{m}_o^\Phi) (\Phi(x_j) - \mathbf{m}_o^\Phi)^T = \mathbf{S}_w^\Phi + \mathbf{S}_b^\Phi \quad (12)$$

式中:  $\Phi(x_j^i) (i=1, 2, \dots, M; j=1, 2, \dots, N_i)$  为采样向量, 即特征空间  $H$  中第  $i$  类中的第  $j$  个向量。  $\mathbf{m}_i^\Phi$  表示其中第  $i$  类采样值的均值,  $\mathbf{m}_i^\Phi = \frac{1}{N_i} \sum_{j=1}^{N_i} \Phi(x_j^i)$ 。  $\mathbf{m}_o^\Phi$  表示特征空间  $H$  中全体采样值的均值,  $\mathbf{m}_o^\Phi = \sum_{i=1}^M p(M_i) \mathbf{m}_i^\Phi$ 。式中,  $p(M_j)$  是第  $i$  类采样数据的先验概率。由式(10~12)可以看出,  $\mathbf{S}_w^\Phi$ ,  $\mathbf{S}_b^\Phi$  和  $\mathbf{S}_t^\Phi$  都是非负定对称矩阵。

在特征空间  $H$  中, 其 Fisher 判别准则可以定义为

$$\mathbf{J}(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_b^\Phi \mathbf{w}}{\mathbf{w}^T \mathbf{S}_w^\Phi \mathbf{w}} \quad (13)$$

式中:  $\mathbf{w}$  代表任一非零的列向量, 根据该式可以求出最优判别向量  $\mathbf{w}$ , 将数据映射到  $\mathbf{w}$  上的投影则是其 Fisher 特征向量。

对于高维的特征空间, 其最优判别量的计算不可能直接完成。所以通过下面的公式进行变换, 将其最后简化到内积运算的模式, 原始空间定义的核函数为  $\mathbf{K}(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$ , 再生核理论指出, 高维空间中的所有数据  $\Phi(x_1), \Phi(x_2), \dots, \Phi(x_N)$  所张成的空间一定包含了最优化准则函数的解向量即  $\mathbf{w} = \sum_{i=1}^N \alpha_i \Phi(x_i) = \Phi \alpha$ 。

通过在特征空间  $H$  进行采样值  $\Phi(x_i)$  在最优判别向量  $\mathbf{w}$  上的投影可以得到以下计算式

$$\mathbf{w}^T \Phi(x_i) = \alpha^T \Phi^T \Phi(x_i) = \alpha^T (\Phi(x_1), \Phi(x_2), \dots, \Phi(x_N))^T \Phi(x_i) = \alpha^T \xi_{x_i} \quad (14)$$

对原始空间的采样数据  $x \in \mathbf{R}^n$ , 核采样向量  $\xi_x = (K(x_1, x), \dots, K(x_N, x))^T$ , 同样也可以将  $\mathbf{m}_i^\Phi$  和  $\mathbf{m}_o^\Phi$  在特征空间  $H$  的最优判别向量  $\mathbf{w}$  上投影为  $\mathbf{w}^T \mathbf{m}_i^\Phi = \alpha^T \Phi^T \frac{1}{N_i} \sum_{j=1}^{N_i} \Phi(x_j^i)$  和  $\mathbf{w}^T \mathbf{m}_o^\Phi = \alpha^T \Phi^T \times \frac{1}{N} \sum_{j=1}^N \Phi(x_j)$ 。

类内的核采样均值  $\mu_i$  ( $i=1, 2, \dots, M$ ) 和总体的核采样均值  $\mu_0$  的计算式如下

$$\begin{aligned}\mu_i &= \left( \frac{1}{N_i} \sum_{j=1}^{N_i} K(x_1, x_j^i), \dots, \left( \frac{1}{N_i} \sum_{j=1}^{N_i} K(x_N, x_j^i) \right)^T \right. \\ \mu_0 &= \left( \frac{1}{N} \sum_{j=1}^N K(x_1, x_j), \dots, \left( \frac{1}{N} \sum_{j=1}^N K(x_N, x_j) \right)^T \right.\end{aligned}\quad (15)$$

则在高维特征空间  $H$  中式(20)Fisher 判别准则函数与下式等价

$$J(w) = \frac{w^T S_b^\Phi w}{w^T S_w^\Phi w} = \frac{\alpha^T K_b \alpha}{\alpha^T K_w \alpha} = J(\alpha) \quad (16)$$

式中:  $K_b = \sum_{i=1}^M \frac{N_i}{N} (\mu_i - \mu_0)(\mu_i - \mu_0)^T$  为核类间离

散度矩阵,  $K_w = \frac{1}{N} \sum_{i=1}^M \sum_{j=1}^{N_i} (\xi_{x_j^i} - \mu_i)(\xi_{x_j^i} - \mu_i)^T$  表示核类内离散度矩阵。这种最优的核 Fisher 判别向量是与判别准则函数最大相对应的判别向量。它可以通过求解下列广义特征方程:  $K_b \alpha = \lambda K_w \alpha$  得到, 式中  $\alpha$  是任意的非零向量 ( $N$  维)。其中, 选用适当的核函数将数据映射到高维特征空间后, 核 Fisher 判别方法可以很容易区分不同类型的数据。仿真时采用高斯核函数可以使得数据类有较大的分离。

## 2.2 基于 KFDA 的故障诊断

在 KFDA 的监控中, 监控统计量选择为核 Fisher 特征向量的欧式距离。将参考数据的核 Fisher 特征向量设为  $T_r$  ( $t_{r,1}, t_{r,2}, \dots, t_{r,n}$ ), 采样数据的核 Fisher 特征向量设为  $T_c$  ( $t_{c,1}, t_{c,2}, \dots, t_{c,n}$ ), 则可以用下式表示欧氏距离的统计量  $D$

$$D = \|T_r - T_c\| = \left[ \sum_{i=1}^n (t_{r,i} - t_{c,i})^2 \right]^{1/2} \quad (17)$$

采用核密度估计方法来计算统计量  $D$  的控制限  $D^*$ 。在 KFDA 对不同故障进行分析的基础上, 可以建立一个最优核 Fisher 判别向量库  $F$ , 将所有的故障集进行归纳总结, 其中一共包含  $k$  类故障。

$$F = \{\alpha_{\text{opt},1}, \alpha_{\text{opt},2}, \dots, \alpha_{\text{opt},k}\} \quad (18)$$

式中:  $\alpha_{\text{opt},i}$  表示对应于的故障  $i$  类的最优核 Fisher 判别向量。在进行故障诊断的过程中, 计算当前的  $\alpha_{\text{cur}}$ , 将其与之前的  $\alpha_{\text{opt},i}$  进行比较, 利用两者的相似度, 可以对发生故障的类型进行判别。

$$S_i = \frac{(\alpha_{\text{opt},i})(\alpha_{\text{cur}})^T}{\alpha_{\text{opt},i} \cdot \alpha_{\text{cur}}} \quad (i=1, 2, \dots, k) \quad (19)$$

由上式可知, 也就是说相似程度是通过两个核 Fisher 判别向量交角的余弦值来表示的, 如果相似程度与 1 比较接近, 则说明这两个向量相近, 由此便

可以得出发生故障的类型即为  $i$  类。在实际应用中, 也可以设置一个诊断阈值, 当相似度大于等于该阈值时, 可以认定第  $i$  类故障发生。如果两者的相似度小于设定的阈值, 那么就不能认定所发生的故障为第  $i$  类故障, 而是不认识的一类故障。这时, 需要将这个新的故障添加到故障数据集的最优的核 Fisher 判别向量库中, 并根据实际工艺确定故障属于哪一类, 然后就相当于对新的故障进行了一定程度的了解和学习了。

## 3 仿真研究

将上述方法应用到 Tennessee Eastman 化工仿真模型, 检验方法的有效性。TE 过程的原型是 Eastman 化学公司的一个实际的工艺流程。TE 过程包括 20 种预先设定好的故障模式, 分别代表阶跃、随机变化、慢漂移和粘滞等故障类型, 其工艺流程图和详细的过程见文献[12]。文中采用文献[13]控制方案, 进行闭环控制, 得到仿真数据, 该数据为一个  $960 \times 52$  的矩阵, 52 代表过程变量数, 960 代表每个变量的采样点数, 仿真运行 48 h, 故障是在第 8 小时引入的。

### 过程监控与故障诊断的步骤

(1) 对数据采用小波包变换消噪, 采用 PCA 方法进行白化处理;

(2) 在核独立主元分析的基础上建立正常的工况模型, 进行统计量  $U^2$ 、 $U_c^2$  和 SPE 的计算;

(3) 采用核密度估计方法确定相应的控制置信限, 监控统计量  $U^2$ 、 $U_c^2$  和 SPE 是否超限, 检测故障信息;

(4) 计算正常的工况数据集的核 Fisher 特征向量间欧氏距离, 然后用核密度估计方法计算其控制限  $D^*$ ;

(5) 为故障数据集建立最优核 Fisher 判别向量库  $F$ , 公式如(18)所示;

(6) 在选择适当的核函数的前提下, 将原始空间中的参考数据  $X_r$  和新采样数据  $X_c$  采用非线性核函数映射到高维特征空间来得到核参考数据  $\xi_{X_r}$  和新的核采样数据  $\xi_{X_c}$ ;

(7) 把  $\xi_{X_r}$  和  $\xi_{X_c}$  设定为不同的类, 然后进行 Fisher 判别分析, 计算核类间离散度矩阵  $K_b$  和类内离散度矩阵  $K_w$ ;

(8) 计算出最优的核判别向量  $\alpha_{\text{opt}}$ ;

(9) 进行特征矢量的计算, 即分别将参考数据  $\xi_{X_r}$  和采样数据  $\xi_{X_c}$  投射到最优判别向量  $\alpha_{\text{opt}}$  的方向, 从而得到二者的特征矢量  $T_r$  和  $T_c$ ;

(10)进行两类数据核 Fisher 特征向量间的欧式距离  $D$  的计算,公式参照式(17);

(11)当统计量  $D > \text{控制限 } D^*$ ,说明有故障产生,同时可以通过式(19)进行故障类型的判别。

首先对数据采用小波包变换降噪,采用 PCA 方法进行白化处理,再利用核独立主元分析,建立正常工况模型。确定监控统计量  $U^2$ 、 $U_e^2$  和 SPE,并采用核密度估计方法确定相应的控制置信限,检测故障信息。以故障 11 为例,验证该方法的有效性。其中,核函数采用高斯核函数,参数在仿真中根据经验选取  $\sigma = 1.6$ ,数据类之间有较程度的分离。故障 11 和故障 4 类似,它引起的故障是使反应器的冷却水入口的温度波动。这种故障导致反应器中的冷却水的流速波动明显,这样导致反应器的温度波动。

MPCA 方法和 KICA 方法的监控结果如图 1 和 2 所示,其中,MPCA 方法中的 SPE 统计量监控图可以检测到故障,但监控效果不理想,漏报现象很明显,而  $T^2$  监控图没有检测到故障的发生。KICA 方法从图中可以看出,故障发生后, $U^2$ 、 $U_e^2$  和 SPE 统计量均很快超过了置信限,监控效果明显好于 MPCA 方法。在实际的生产中,操作人员要及时地发现故障的存在,尽早采取措施,将损害降低到最小,可见故障检测的重要性。

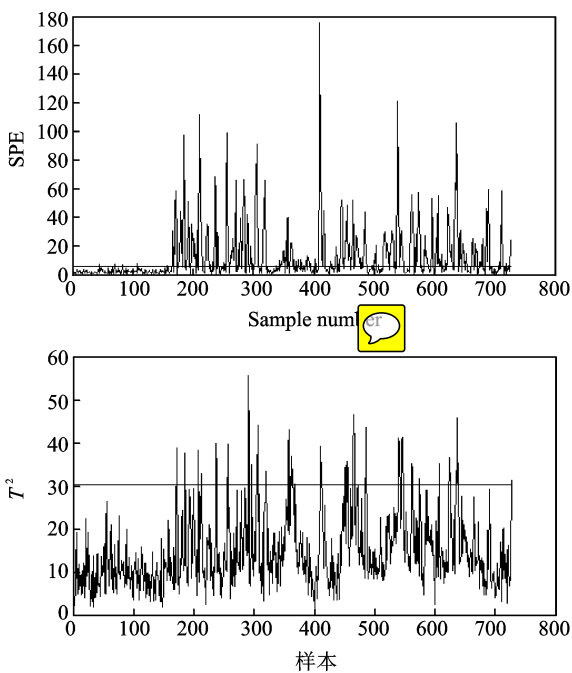


图 1 故障 11 的 MPCA 性能监控图

故障诊断的方法就是将当前的数据和历史故障数据的最优核 Fisher 判别向量进行比较,二者

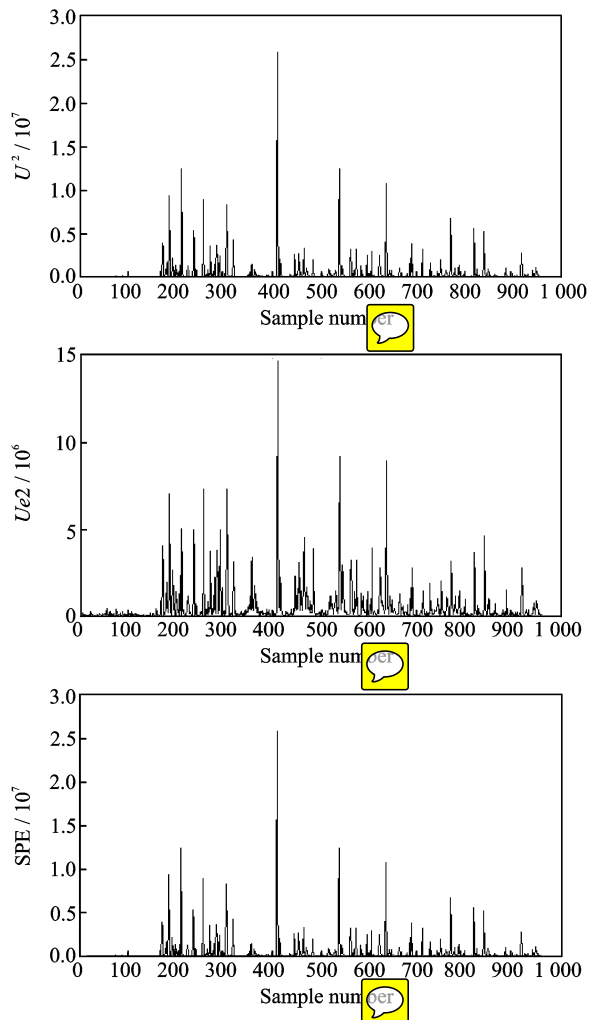


图 2 故障 11 的 KICA 性能监控图

的相似度越接近于 1,则说明当前数据越接近历史故障数据,就可以认定属于该类故障,比如图 3 所示,当前数据与第 11 类故障最优核 Fisher 判别向量的相似度为 0.92,接近于 1,因此可以判断故障来源于反应器冷却水的入口温度发生变化(故障 11)。

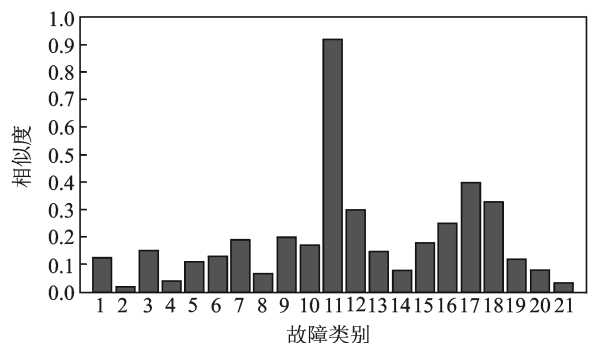


图 3 故障 11 判别向量的相似度分析

为进一步验证该方法的有效性,针对 TE 化工仿真模型的多种不同故障,分别采用 KPCA、KI-

CA、KICA-KFDA方法进行比较研究,比较结果如表 1 所示。从表中可以看出,该方法可以有效地辨

识故障,对于故障 3、7、9、15 等难以诊断的故障,效果也有改善。

表 1 故障识别率对比分析表

典型故障	1	2	3	4	6	7	8	9	11	12	13	14	15
KPCA	45.5	91.3	22.8	12.8	97.9	55.1	66.9	34.9	96.2	88.3	78.6	97.3	52.8
KICA	47.8	91.7	33.5	45.6	98.1	62.5	77.5	38.9	95.2	89.2	85	92.5	45.9
KICA-KFDA	88.9	94.8	42.6	97.7	99.5	68.8	89.2	55.8	96.8	99.2	93.5	98.5	77.8

## 4 结束语

本文针对工业过程提出了一种基于 KICA 与 KFDD 的集成监控方法。为克服传统过程监控方法需要假设过程特征信号服从多元正态分布的缺陷,采用核独立成分分析建立正常工况模型,监测故障信息。由于通过计算每个监控变量对  $U^2$ 、SPE 和  $U^2$  的贡献度,可以更好地监控系统的运行情况,大大缩小了故障分类的范围;若发生故障,在高维特征空间中再运用 KFDD 方法,计算满足最大分离程度的核 Fisher 特征向量和判别向量,最后根据当前故障的判别向量和历史故障数据集 中所含故障的最优核 Fisher 判别向量的相似度进行故障诊断。采用 KFDD 方法对故障进行诊断,准确性有所提高,而且具有对新故障的学习能力。若当前数据与故障数据集中的任一判别向量的相似度小于诊断阈值时,可以结合工艺只是确定新的故障类型,将其添加至故障数据集的最优核 Fisher 判别向量库中。通过对 TE 过程的仿真研究,验证了所提方法的有效性和合理性。

### 参考文献:

- [1] 尹雪岩,刘飞. 基于 DIFA 的动态非高斯过程监控方法及其应用[J]. 化工学报,2011,62(5):1345-1351.  
Yin Xuayan, Liu Fei. Dynamic non-Gaussian process monitoring method based on DIFA and its application [J]. CIESC Journal, 2011,62(5):1345-1351.
- [2] 肖应旺. 动态隐变量法及其在动态过程监控中的应用[J]. 仪器仪表学报,2012,33(1):113-119.  
Xiao Yingwang. Dynamic latent variable method and its application in dynamic process monitoring [J]. Chinese Journal of Scientific Instrument, 2012, 33 (1):113-119.
- [3] 王姝,冯淑敏,常玉清,等. 改进的基于数据重构的 KPCA 故障识别方法[J]. 东北大学学报(自然科学版),2012,33(4):500-503.

- Wang Shu, Feng Shumin, Chang Yuqing, et al. Improved KPCA fault identification method based on data reconstruction[J]. Journal of Northeastern University (Natural Science), 2012,33(4):500-503.
- [4] 薄翠梅,柏杨进,杨海荣,等. 多切面分类改进独立成份与支持向量机集成故障诊断方法[J]. 控制理论与应用,2012,29(2):229-234.  
Bo Cuimei, Bai Yangjin, Yang Hairong, et al. Multi-section classification improving integrated fault diagnosis method based on independent component analysis and support-vector-machines [J]. Control Theory & Applications, 2012,29(2):229-234.
- [5] 彭红星,陈祥光,徐巍. PCA 特征抽取与 SVM 多类分类在传感器故障诊断中的应用[J]. 数据采集与处理, 2010,25(1):111-116.  
Peng Hongxing, Chen Xiangguang, Xu Wei. Application of PCA feature extraction and SVM multi-classification on sensor fault diagnosis[J]. Journal of Data Acquisition and Processing, 2010,25(1):111-116.
- [6] Yang J, Gao X, Zhang D, et al. Kernel ICA: An alternative formulation and its application to face recognition[J]. Pattern Recognition, 2005, 38 (10): 1784-1787.
- [7] Hyvärinen A, Karhunen J, Oja E. Independent component analysis [M]. New York, USA: John Wiley & Sons, Inc, 2001.
- [8] Cardoso J F, Soulomica A. Blind beamforming for non-Gaussian signals[J]. IEEE Proc F, 1993, 140 (6):362-370.
- [9] Lee J M, Yoo C, Lee I B. Statistical process monitoring with independent component analysis[J]. Journal of Process Control, 2004,14(5):467-485.
- [10] Simoglou A, Martin E B, Morris A J. Statistical performance monitoring of dynamic multivariate processes using state space modeling[J]. Computers and Chemical Engineering, 2002, 26(6):909-920.
- [11] Chen Q, Wynne R J, Goulding P, et al. The application of principal component analysis and kernel density estimation to enhance process monitoring [J].

Control Engineering Practice, 2000,8:531-543.

- [12] Chiang L H, Russell E L, Braatz R D. Fault detection and diagnosis in industrial systems [M]. London: Springer-Verlag London Limited, 2001.
- [13] Chen J H, Liao C M. Dynamic process fault monitoring based on neural network and PCA[J]. Journal of Process Control, 2002,12(2):277-289.

**作者简介:**许洁(1976-),女,博士,讲师,研究方向:统计过程控制与故障诊断,E-mail: 63038@njnu.edu.cn;赵瑾(1961-),女,博士,副教授,研究方向:统计过程控制与故障诊断;刘如成(1960-),男,高级实验师,研究方向:自动检测与自动控制;胡寿松(1937-),男,教授,研究方向:故障诊断、自修复控制和复杂系统控制。