

DOI: 10.11949/j.issn.0438-1157.20181307

基于加权统计局部核主元分析的非线性化工过程微小故障诊断方法

邓佳伟¹, 邓晓刚¹, 曹玉苹¹, 张晓玲²(¹ 中国石油大学(华东) 信息与控制工程学院, 山东 青岛 266580; ² 中国石油大学胜利学院, 山东 东营 257061)

摘要: 传统统计局部核主元分析 (statistical local kernel principal component analysis, SLKPCA) 在构造改进残差时未考虑样本的差异性, 使得故障样本信息易于被其他样本所掩盖, 针对该问题, 提出一种基于加权统计局部核主元分析 (weighted statistical local kernel principal component analysis, WSLKPCA) 的非线性化工过程微小故障诊断方法。该方法首先利用KPCA获取过程的得分向量和特征值并构建初始残差。然后设计了一种基于测试样本与训练样本之间距离的加权策略构建加权改进残差, 对含有较强微小故障信息的样本赋予较大权值, 以增强故障样本的影响。最后, 采用基于测量变量与监控统计量之间的加权互信息构建贡献图以识别故障源变量。在连续搅拌反应釜和田纳西伊斯曼 (Tennessee Eastman, TE) 化工过程上的仿真结果表明, 所提方法具有良好的微小故障检测与识别性能。

关键词: 化工过程; 微小故障; 核主元分析; 统计局部方法; 故障诊断

中图分类号: TP 277

文献标志码: A

文章编号: 0438-1157 (2019) 07-2594-12

Incipient fault diagnosis method of nonlinear chemical process based on weighted statistical local KPCA

DENG Jiawei¹, DENG Xiaogang¹, CAO Yuping¹, ZHANG Xiaoling²¹Information and Control Engineering College, China University of Petroleum, Qingdao 266580, Shandong, China;²Shengli College, China University of Petroleum, Dongying 257061, Shandong, China)

Abstract: The traditional local kernel principal component analysis (SLKPCA) does not consider the difference of samples when constructing the improved residual, so that the fault sample information is easily covered by other samples. This paper proposes a new fault diagnosis method of nonlinear chemical process based on weighted statistical local kernel principal component analysis (WSLKPCA). Firstly, the score vectors and the eigenvalues are obtained using KPCA and the residual function is constructed. Then, a weighting strategy based on the distance between the test sample and the training sample is designed to construct the weighting improved residual, which assigns larger weights to samples with strong incipient fault information to enhance the impact of fault samples. Finally, the contribution graph is constructed based on the weighted mutual information between the measured variables and monitoring statistics to identify the fault source variables. Simulation results on continuous stirred tank reactor and TE process show that the proposed method can effectively detect incipient faults, and has better

收稿日期: 2018-11-12 修回日期: 2019-04-13

通信作者: 邓晓刚(1981—),男,博士,副教授, dengxiaogang@upc.edu.cn

第一作者: 邓佳伟(1994—),女,硕士研究生, 1440278056@qq.com

基金项目: 中央高校基本科研业务费专项资金(17CX02054); 国家自然科学基金项目(61403418, 21606256); 山东省自然科学基金项目(ZR2014FL016, ZR2016FQ21, ZR2016BQ14); 山东省重点研发计划项目(2018GGX101025); 山东省高等学校科技计划项目(J18KA359); 中国石油大学胜利学院科技计划项目(KY2017002); 浙江大学工业控制技术国家重点实验室开放课题(1900306)

引用本文: 邓佳伟, 邓晓刚, 曹玉苹, 张晓玲. 基于加权统计局部核主元分析的非线性化工过程微小故障诊断方法[J]. 化工学报, 2019, 70(7): 2594–2605

Citation: DENG Jiawei, DENG Xiaogang, CAO Yuping, ZHANG Xiaoling. Incipient fault diagnosis method of nonlinear chemical process based on weighted statistical local KPCA[J]. CIESC Journal, 2019, 70(7): 2594–2605

fault recognition performance.

Key words: chemical process; incipient fault; kernel principal component analysis; statistical local approach; fault diagnosis

引 言

由于现代工业过程日益复杂,人们对工业过程中的安全性和可靠性越来越重视。一旦复杂过程发生故障,将会造成无法估量的损失,因此如何运用故障诊断技术及时准确地检测并识别故障至关重要^[1-5]。随着存储技术的发展,大量生产过程数据被采集并记录。因此,基于数据驱动的故障诊断方法得到了广泛应用^[6-8]。典型的数据驱动故障诊断方法包括主元分析(principal component analysis, PCA)、偏最小二乘(partial least square, PLS)等,这些传统方法能够有效监控过程中的显著故障。然而对于幅值小、变化缓慢、故障特征不明显、容易被噪声或扰动所掩盖的微小故障监控性能较差。许多化工过程的故障(比如设备老化、元器件磨损、催化剂失活等)都存在着缓慢变化过程,其初期表现为微小故障。如果能在故障早期阶段及时进行检测并诊断,则有利于避免严重故障的出现,有助于保证产品质量^[9-11]。因此,研究微小故障诊断技术具有重要意义。

目前,许多专家针对微小故障诊断问题开展了一系列研究。Wold^[12]首次提出一种PCA和指数加权移动平均(exponentially weighted moving average, EWMA)相结合的微小故障检测方法。Harmouch等^[13]从概率密度函数的角度出发,将传统PCA与Kullback Leibler散度相结合,提高了微小故障检测的性能。进一步地,Chen等^[14]通过结合Kullback Leibler散度与PCA并定义对称评价函数,通过初始故障案例验证了方法的有效性。Ge等^[15]提出了一种基于故障特征提取和残差评估的两级微小故障检测方法,该方法将小波分析和最优奇偶空间分析相结合提取噪声中的早期故障特征,实现了对微小故障的检测。为了解决微小故障分离的问题,Ji等^[16]研究基于指数平滑重构的微小故障分离方法,通过实验验证了该方法比传统重构的方法具有更高的效率和准确性。尚骏等^[17]针对微小故障难以检测的问题,提出了一种基于变元统计分析的微小故障检测方法,通过提取变元并统计变元的高阶信息,可以改善故障的检测性能。为了挖掘高速列车

电气传动的精确信号特征,Chen等^[18]利用多个数据处理层设计了一种深层PCA方法。Zhao等^[19-20]针对微小故障检测问题也开展了深入研究,提出了一种两级故障检测方法和一种稀疏差异分析算法。上述研究均针对线性过程,未考虑实际工业过程的非线性特性。针对传统核最小二乘算法(kernel least square, KLS)难以有效监控非线性过程质量相关故障的问题,Zhao等^[21]研究基于加权双步特征提取的质量弱相关故障检测方法,通过对监控统计量进行加权以实现质量相关故障的检测。然而,该方法并未考虑到对微小故障的监控。针对非线性过程的微小故障检测问题,Ge等^[22]综合利用核主元分析(kernel principal component analysis, KPCA)和统计局部方法(statistical local approach, SLA),提出一种统计局部KPCA(SLKPCA)方法,具有较好的微小故障检测性能。

总体而言,现有方法在微小故障诊断问题中取得较好的效果,但是仍然存在一些问题。首先,大部分研究集中于线性方法,考虑过程非线性特性的方法较少,SLKPCA作为一种有效的非线性微小故障检测方法,值得深入研究;再者,目前的故障诊断方法在获取非线性成分时是以同等的地位对待每个样本,即每个样本具有相同的权值1,这可能导致含有故障信息样本的作用被其余样本所削弱,不利于微小故障检测性能的进一步提高;最后,对检测到微小故障后的故障变量识别问题研究较少。经典的故障变量识别方法是贡献图技术。现有的贡献图技术主要基于线性PCA^[23],但未有人利用统计局部KPCA方法设计贡献图进行故障变量识别,因此如何将贡献图技术与非线性统计局部KPCA监控模型相结合是一个值得深入研究的问题。

针对非线性化工过程微小故障检测及识别问题,本文提出一种基于加权统计局部核主元分析的微小故障诊断方法。该方法立足于SLKPCA这种非线性故障监控模型,一方面通过设计一种加权策略,对故障样本施加较大权值以突出故障样本的作用,另一方面基于变量与监控统计量之间的加权互信息构建贡献图以识别故障变量。最后,本文通过连续搅拌反应釜仿真系统和TE过程进行方法的验

证,以分析方法的有效性。

1 基本方法

1.1 核主元分析

核主元分析(KPCA)是一种典型的非线性 PCA 方法。给定 n 个样本 m 个变量的训练集 $\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_n]^T$, 其中 $\mathbf{x}_i \in \mathbf{R}^m$, ($i = 1, 2, \cdots, n$), n 为样本个数, m 为测量变量数, KPCA 通过一个非线性映射 $\mathbf{R}^m \xrightarrow{\Phi(\cdot)} \mathbf{F}^h$ 构建特征空间, $\Phi(\cdot)$ 为非线性映射函数, h 为特征空间的维数。映射数据 $\Phi(\mathbf{x}_i)$ 的协方差矩阵为

$$\mathbf{C} = \frac{1}{n} \sum_{i=1}^n \Phi(\mathbf{x}_i) \Phi^T(\mathbf{x}_i) \quad (1)$$

对上述协方差阵进行特征值分解^[24-25]

$$\lambda_i \mathbf{v}_i = \mathbf{C} \mathbf{v}_i \quad (2)$$

其中, λ_i 和 $\mathbf{v}_i$ 分别为协方差矩阵的第 i 个特征值和特征向量。 $\mathbf{v}_i$ 可用特征空间的样本线性表示

$$\mathbf{v}_i = \sum_{j=1}^n \alpha_{ij} \Phi(\mathbf{x}_j) \quad (3)$$

其中, α_{ij} 为特征空间展开系数。

将式(1)、式(3)带入式(2),可得

$$n \lambda_i \boldsymbol{\alpha}_i = \mathbf{K} \boldsymbol{\alpha}_i \quad (4)$$

其中, $\boldsymbol{\alpha}_i = [\alpha_{i1} \ \alpha_{i2} \ \cdots \ \alpha_{in}]^T$, $\mathbf{K}$ 表示核矩阵, 其第 (i, j) 个元素为 $k_{ij} = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle$ 。

对于第 h 个采样时刻的测试样本 $\mathbf{x}_h$, 其在第 k 个特征向量的投影为

$$t_{h,k} \leq v_k, \Phi(\mathbf{x}_h) \geq \sum_{j=1}^n \alpha_j^k < \Phi(\mathbf{x}_j), \Phi(\mathbf{x}_h) > \quad (5)$$

通过式(5), 构造监控统计量 T^2 和 SPE 来监控过程变化如下^[26]

$$T_h^2 = [t_{h,1}, t_{h,2}, \cdots, t_{h,a_1}] \Lambda^{-1} [t_{h,1}, t_{h,2}, \cdots, t_{h,a_1}]^T \quad (6)$$

$$\text{SPE}_h = \sum_{j=1}^{a_1} t_{h,j}^2 - \sum_{j=1}^a t_{h,j}^2 \quad (7)$$

其中, a_1 为通过 99% 的累计贡献率所确定的特征空间的维数。

1.2 统计局部核主元分析

Kruger^[27]在 2007 年将该方法用于线性 PCA 并对早期故障进行检测。随后又有研究人员将其与非线性 KPCA 相结合^[22], 以改进多元统计监控方法的监控效果。统计局部 KPCA (SLKPCA) 方法通过核变换构建基于得分向量和特征值的残差函数, 以增强 KPCA 的微小故障检测能力。对原始训练集矩阵 $\mathbf{X}$, 应用核主元分析后获得核成分矩阵 $\mathbf{T}$, 基于核成

分 $t_{h,k}$ 和特征值 λ_k , SLKPCA 构建相应的初始残差^[22]

$$r_{h,k} = 2t_{h,k}^2 - 2\lambda_k \quad (8)$$

在式(8)基础上可利用滑动窗技术构建如下改进残差^[22]

$$\bar{r}_{h,k} = \frac{1}{\sqrt{q}} \sum_{l=h-q+1}^h r_{l,k} \quad (9)$$

其中, q 为窗宽, l 为当前观测样本数, h 为采样时刻。对第 h 个采样时刻的样本 $\mathbf{x}_h$, 其改进残差向量被定义为 $\bar{\mathbf{r}}_h = [\bar{r}_{h,1}, \bar{r}_{h,2}, \cdots, \bar{r}_{h,a_1}]$ 。改进残差 $\bar{\mathbf{r}}_h$ 可以被分为主系统变化 $\bar{\mathbf{r}}_h^{(s)}$ 和噪声 $\bar{\mathbf{r}}_h^{(n)}$ 两部分

$$\bar{\mathbf{r}}_h^{(s)} = [\bar{r}_{h,1}, \bar{r}_{h,2}, \cdots, \bar{r}_{h,a_1}] \quad (10)$$

$$\bar{\mathbf{r}}_h^{(n)} = [\bar{r}_{h,a_1+1}, \bar{r}_{h,a_1+2}, \cdots, \bar{r}_{h,a_1}] \quad (11)$$

针对式(9)中的改进残差, 分别建立 T^2 和 SPE 统计量监控过程变化, 可以有效监控非线性过程的微小故障, 统计量的计算公式如下

$$T_h^2 = (\bar{\mathbf{r}}_h^{(s)})^T \boldsymbol{\Sigma}_s^{-1} (\bar{\mathbf{r}}_h^{(s)}) \quad (12)$$

$$\text{SPE}_h = (\bar{\mathbf{r}}_h^{(n)})^T \boldsymbol{\Sigma}_n^{-1} (\bar{\mathbf{r}}_h^{(n)}) \quad (13)$$

其中, $\boldsymbol{\Sigma}_s, \boldsymbol{\Sigma}_n$ 为训练数据初始残差的协方差。

2 加权统计局部核主元分析

为了更好地检测微小故障, 本文考虑残差样本的不同, 采用加权策略来突出故障样本的影响, 提出一种加权 SLKPCA (WSLKPCA) 方法。该方法在 SLKPCA 方法的基础上, 利用加权策略重新构建改进残差, 并建立贡献图用于故障变量的识别。

2.1 加权策略

传统 SLKPCA 方法在构建改进残差时每个样本的地位是等同的, 即 1.2 节中的式(9)滑动窗内每个样本具有相同的权值 1。事实上, 当故障发生时, 所采集的故障数据相对于正常工况下采集的数据距离较远, 应当对该类数据给予较大的权值, 以更好地突出样本维度上的故障信息。为了评估测试样本所携带的故障信息, 本文考虑测试样本相对于正常样本之间的距离, 将原始特征空间的数据映射到样本距离空间, 从而计算当前样本与所有正常样本的平均间距, 记为 d_h , 其表达式为

$$d_h = \frac{1}{n} \sum_{i=1}^n \text{dist}(\mathbf{r}_h, \mathbf{r}_i^*) \quad (14)$$

其中, $\mathbf{r}_i^*$ 为第 i 个训练样本 $\mathbf{x}_i$ 相对应的初始残差行向量, $\text{dist}(\mathbf{r}_h, \mathbf{r}_i^*)$ 代表向量 $\mathbf{r}_h$ 和向量 $\mathbf{r}_i^*$ 的欧式距离。

考虑到所采集数据的随机性, 构建一种修正的滑动窗距离如下

$$\bar{d}_h = \frac{1}{q} \sum_{j=h-q+1}^h d_j \quad (15)$$

根据正常数据的滑动窗距离 $\bar{d}_h$ 设定一阈值 $d_{\lim}$, 当测试数据的 $\bar{d}_h$ 超过该阈值, 意味着检测到了故障信息。 $\bar{d}_h$ 值越大, 故障信息越强, 相应的样本应赋予较大权值。当测试样本的滑动窗距离 $\bar{d}_h$ 偏离正常数据越明显, 则认为该测试样本所携带较强的故障信息, 此时应对该样本赋予较大的权值增强其影响力, 这样将更加有利于对微小故障的检测。基于上述分析, 构建一种加权初始残差

$$\mathbf{w}r_h = r_h w_h \quad (16)$$

其中, w_h 为基于欧式距离的权值因子, 其表达式为

$$w_h = \begin{cases} \beta \left(\frac{\bar{d}_h - d_{\lim}}{d_{\lim}} \right), & \bar{d}_h \geq d_{\lim} \\ 1, & \text{otherwise} \end{cases} \quad (17)$$

这里 β 为权值调节因子, $\text{sat}()$ 为饱和函数, 由式 (18) 来表示。

$$\text{sat}(x) = \begin{cases} x, & \text{if } 0 < x < x_{\lim} \\ x_{\lim}, & \text{if } x \geq x_{\lim} \\ 0, & \text{if } x \leq 0 \end{cases} \quad (18)$$

其中, $x_{\lim}$ 为饱和函数的上限, 在本文中设置为 10。

基于上述分析, 加权改进残差表示如下

$$\overline{\mathbf{w}r}_h = \frac{1}{\sqrt{q}} \sum_{j=h-q+1}^h \mathbf{w}r_j \quad (19)$$

通过引入距离权值因子, $\bar{d}_h$ 值大的样本权值给予大于 1 的权值, 用于增强早期故障样本的影响。但是, 如果仅考虑当前样本的滑动窗距离 $\bar{d}_h$ 是否大于 $d_{\lim}$, 将会由于噪声的存在可能导致较大的误报。故本文定义当 $\bar{d}_j (h-c+1 \leq j \leq h)$ 连续 c 个样本超过 $d_{\lim}$ 时, 才会赋予当前样本大于 1 的权值。 $d_{\lim}$ 是通过核密度估计正常样本的滑动窗距离所得到的阈值。

构建 T^2 和 SPE 统计量监控过程中的变化, 其计算公式如下

$$T_h^2 = (\overline{\mathbf{w}r}_h^{(s)})^T \Sigma_s^{-1} (\overline{\mathbf{w}r}_h^{(s)}) \quad (20)$$

$$\text{SPE}_h = (\overline{\mathbf{w}r}_h^{(n)})^T \Sigma_n^{-1} (\overline{\mathbf{w}r}_h^{(n)}) \quad (21)$$

这里 $\overline{\mathbf{w}r}_h^{(s)} = [\overline{wr}_{h,1}, \overline{wr}_{h,2}, \dots, \overline{wr}_{h,a}]$, $\overline{\mathbf{w}r}_h^{(n)} = [\overline{wr}_{h,a+1}, \overline{wr}_{h,a+2}, \dots, \overline{wr}_{h,a_1}]$

2.2 基于加权互信息的故障贡献图

当检测到故障发生时, 便需要维护人员快速地对分析产生故障的原因, 即找到引起故障的根本性变量, 这便是故障变量识别的任务^[28]。通过故障变量

识别, 找到可能引起故障发生的故障变量源, 有利于维护人员将更多的精力和时间放在局部子系统上, 从而可以避免维护人员盲目地去找故障变量, 进而可以缩短故障诊断的时间。因此, 检测到故障后对故障变量进行识别是非常有必要的。为此, 本文通过测量变量与监控统计量之间的加权互信息构建贡献图对故障变量进行识别。

对加权统计局部核主元分析 (WSLKPCA) 方法的过程分析可以发现, 要想识别故障变量, 首先需要检测到该故障的发生, 而检测故障的指标为统计量。统计量作为一个标量, 是多个成分统计量的加和, 原始测量变量和监控统计量存在一定关系, 即故障变量和监控统计量之间的相关性较大, 而非故障变量与其相关性较小。互信息作为一种典型的非线性相关性分析工具, 既可以描述变量间的线性相关关系, 也可以描述非线性相关关系^[29-30], 故本文采用基于变量与每个非线性成分监控统计量之间的加权互信息提出一种新的故障变量识别方法。首先介绍 T^2 的贡献图构建方法如下。

对于测试样本 $\mathbf{x}_h$, WSLKPCA 模型获得测试数据 T^2 统计量, 表达如下

$$T_h^2 = (\overline{\mathbf{w}r}_h^{(s)})^T \Sigma_s^{-1} (\overline{\mathbf{w}r}_h^{(s)}) = \frac{\overline{wr}_{h,1}^2}{\Sigma_s(1,1)} + \frac{\overline{wr}_{h,2}^2}{\Sigma_s(2,2)} + \dots + \frac{\overline{wr}_{h,a}^2}{\Sigma_s(a,a)} \quad (22)$$

定义 $T_{hj} = \frac{\overline{wr}_{hj}}{\Sigma_s(j,j)}$ 为测试样本 $\mathbf{x}_h$ 的第 j 个 T^2 成分。

对于测试数据集 $\mathbf{X}$, 其对应的 $\mathbf{T}$ 矩阵表示如下

$$\mathbf{T} = \begin{bmatrix} T_{1,1} & \dots & T_{1,a} \\ \vdots & T_{h,j} & \vdots \\ T_{n,1} & \dots & T_{n,a} \end{bmatrix} \quad (h = 1, 2, \dots, n; j = 1, 2, \dots, a) \quad (23)$$

测试数据第 k 列变量 $\mathbf{X}(:,k)$ 与第 j 个加权改进残差的 T^2 统计量所组成的列向量 $\mathbf{T}_j$ 之间的互信息被定义为

$$MI(\mathbf{X}(:,k), \mathbf{T}_j) = \sum \sum p(\mathbf{X}(:,k), \mathbf{T}_j) \lg \left(\frac{p(\mathbf{X}(:,k), \mathbf{T}_j)}{p(\mathbf{X}(:,k)) p(\mathbf{T}_j)} \right) \quad (24)$$

这里, $p(\mathbf{X}(:,k), \mathbf{T}_j)$ 为联合概率密度函数, $p(\mathbf{X}(:,k))$ 和 $p(\mathbf{T}_j)$ 为边缘概率密度函数, 互信息可通过直方图或核密度估计来获得^[31-32]。通过计算变量与非线性主成分的 T^2 统计量之间的互信息可得到互信息矩阵

$$\mathbf{MI} = \begin{bmatrix} mi_{1,1} & \dots & mi_{1,a} \\ \vdots & mi_{k,j} & \vdots \\ mi_{m,1} & \dots & mi_{m,a} \end{bmatrix} \quad (k = 1, 2, \dots, m) \quad (25)$$

在互信息矩阵中,每行代表同一变量与每个非线性主成分的 T^2 统计量之间的互信息,互信息越大,说明变量与该统计量的关系密切。则第 k 列原始测量变量 $X(:,k)$ 对 T^2 统计量的贡献量被定义为式(25)中每行互信息的加权和

$$CT_k = \sum_{j=1}^a s_j mi_{kj}, \quad k = 1, 2, \dots, m \quad (26)$$

其中, s_j 为测试样本 x_h 的第 j 个 T^2 成分的方差占所有 T^2 成分方差的比重,其表达式为

$$s_j = \frac{\frac{1}{n} \sum_{h=1}^n (T_{hj} - \bar{T}_j)^2}{\sum_{j=1}^a \frac{1}{n} \sum_{h=1}^n (T_{hj} - \bar{T}_j)^2} \quad (27)$$

这里 $\bar{T}_j$ 为式(23)中每列的均值。当发生故障或 T^2 统计量超限时,可以认为对统计量影响较大的变量为故障变量,即具有较大贡献的变量可能是导致故障的变量。

与 T^2 的贡献图公式相似,可以进一步构建SPE统计量的贡献图 $CSPE_k$ 。

3 基于WSLKPCA的非线性化工过程微小故障诊断应用实例

本节详细描述了本文所提方法WSLKPCA的故障诊断流程以及流程图,并以典型的非线性化工过程控制系统连续搅拌反应釜^[33-34](continuous stirred tank reactor, CSTR)和田纳西伊斯曼(Tennessee Eastman, TE)过程作为仿真对象,验证所提算法的有效性。

3.1 非线性化工过程微小故障诊断流程

WSLKPCA方法的监控过程可分为离线建模和在线监控两部分。图1为WSLKPCA的故障诊断流程图,详细步骤描述如下。

离线建模阶段

(1)收集正常数据 $X \in R^{n \times m}$ 并标准化;

(2)计算核矩阵并用式 $K' = K - I_n K - K I_n +$

$I_n K I_n$ 中心化处理,这里 $I_n = \frac{1}{n} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix} \in R^{n \times n}$;

(3)用式(2)求取协方差阵 C 的特征值 λ 和特征向量 v ;

(4)根据累计贡献率来求取核主元个数 a ;

(5)用式(5)、式(7)和式(8)求取历史数据的核成分、初始残差和改进残差;

(6)用式(11)和式(12)求正常工况下的两个统计量并用核密度估计法求取控制限。

在线监控阶段

(1)获取新观测数据 x_h ,用训练数据的均值和方差进行标准化处理;

(2)将新数据 x_h 按照建模的特征向量进行投影得到新数据的得分向量;

(3)计算新数据 x_h 的初始残差并执行样本加权策略;

(4)计算 T^2 和SPE统计量并判断统计量是否超限,若没有超限,则没有发生故障,继续步骤(1),若超限,则出现故障,继续步骤(5);

(5)采用2.2节中的故障变量识别方法计算每个原始测量变量 $X(:,k)$ 对 T^2 统计量的贡献,并根据贡献量的大小确定故障变量。

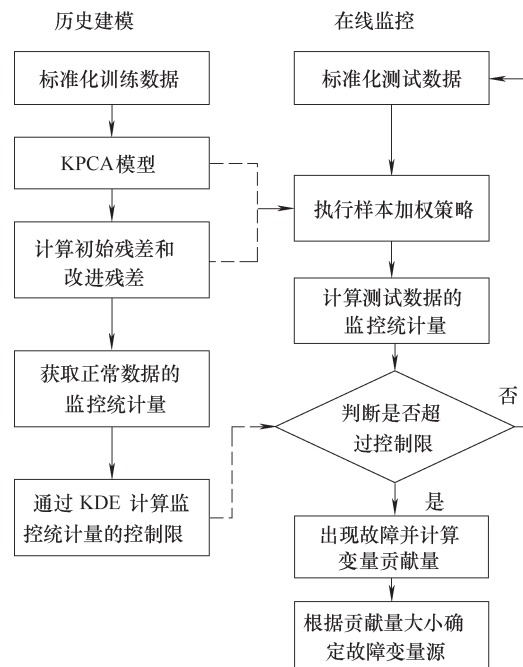


图1 WSLKPCA的故障诊断流程图

Fig.1 Flow chat of fault diagnosis based WSLKPCA

3.2 基于CSTR控制系统的方法仿真验证

为了评估本文所研究方法的有效性,以CSTR控制系统为仿真对象,分别采用KPCA、SLKPCA和WSLKPCA三种不同的方法进行仿真验证。

CSTR控制系统在反应过程中,反应物A在反应器内发生一级不可逆的放热反应并生成物质B,其控制系统图如图2所示。该系统在仿真过程中测量10个变量,分别为4个状态变量 $[C_a, T, T_c, h]$ 和6个输入变量 $[Q, Q_f, C_{af}, T_f, Q_c, T_{cf}]$,具体描述见表1。采集

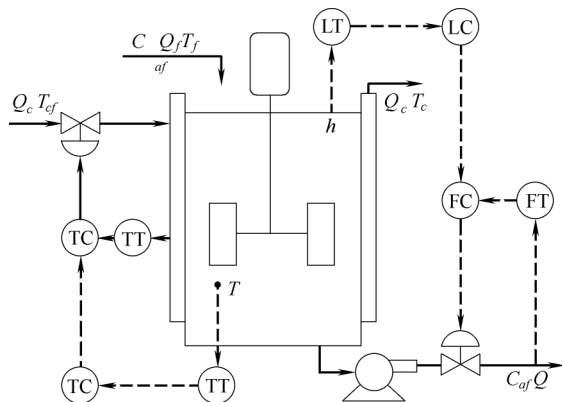


图2 CSTR控制系统

Fig.2 CSTR control system

1000个正常样本作为训练样本。另生成表2中的五种故障数据测试,其中每种故障各有1000个样本,均从第301个采样时刻加入故障。

表1 CSTR控制系统变量列表

Table 1 Variable table for CSTR control system

变量	说明
C_a	反应物A从反应釜流出时的浓度
T	反应釜的温度
T_c	夹套出口冷却剂的温度
h	反应釜液位高度
Q	反应釜流出物料的浓度
Q_f	进料A的流量
Q_{af}	反应釜进料A的浓度
T_f	进料A的温度
Q_c	夹套内冷却剂的流量
T_{cf}	夹套入口冷却剂温度

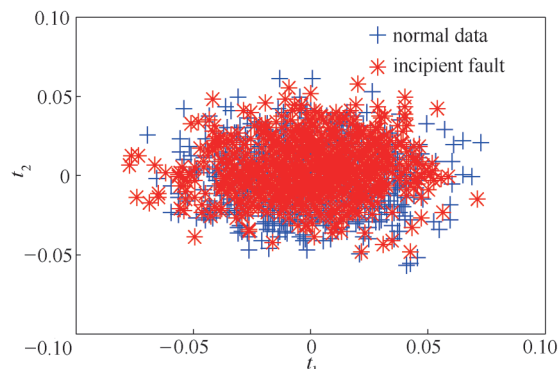
表2 CSTR控制系统故障列表

Table 2 Fault patterns for CSTR control system

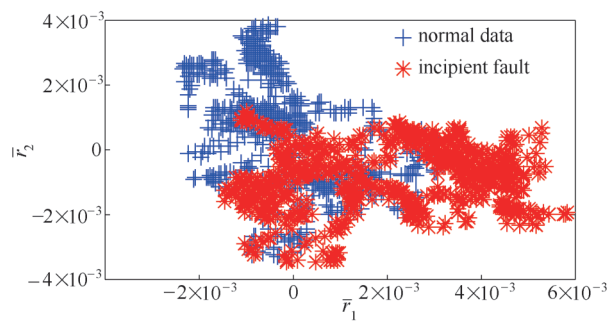
故障	描述
1	反应釜进料A的浓度 C_{af} 发生斜坡变化,幅值为 10^{-5}
2	催化剂的活性以斜坡变化失活,幅值为0.7
3	传热量 U_{ac} 发生斜坡变化,幅值为-10
4	反应器内温度传感器出现偏差,幅值为0.06
5	夹套温度传感器出现偏差,幅值为0.8

分别采用KPCA、SLKPCA和本文所提出的WSLKPCA对CSTR反应过程进行仿真实验。其中,核主元个数根据95%的贡献率选择,控制限采用99%的置信限;核函数选择高斯核函数 $K(\mathbf{x}, \mathbf{y}) = \exp\left(\frac{-\|\mathbf{x} - \mathbf{y}\|^2}{b}\right)$,通过交叉验证法选择核参数 $b=7000$ ^[5,22];构建改进残差时,选择的窗宽 $q=200$ 。

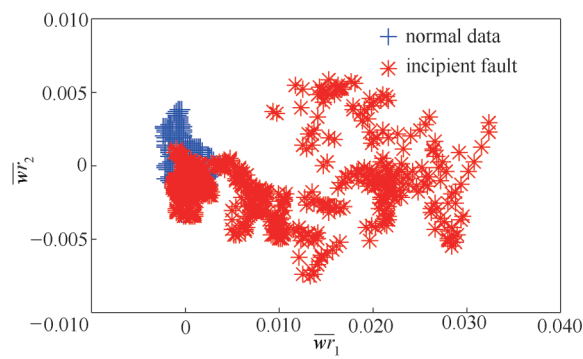
为体现本文方法故障检测和故障变量识别的有效性,KPCA、SLKPCA和WSLKPCA对故障1和故障4的详细仿真对比如下。图3为3种方法对故障1中正常数据和微小故障数据前两个非线性主成分的散点图,3种方法的故障监控图如图4~图6所示。图中虚线代表控制限,实线表示统计量。本文定义连续5个样本的统计量超出控制限,认为第1个超出的时刻为故障检出时刻。



(a) KPCA



(b) SLKPCA



(c) WSLKPCA

图3 正常和微小故障数据前两个非线性主成分散点图

Fig.3 Two nonlinear principal scatter points for normal and incipient fault

故障1是反应釜进料浓度 C_{af} 发生斜坡变化,且变化缓慢,需经较长时间的累积才对过程产生明显的影响,属于非线性缓变微小故障,故对其检测较

为困难。从图3(a)KPCA方法中前两个非线性主成分的散点图可以看出,正常数据和微小故障数据混合在一起,传统KPCA方法无法提取微小故障特征。图3(b)为通过SLKPCA方法提取的非线性特征散点图,从该图可以看出有部分微小故障特征偏离正常数据,与KPCA方法相比,具有较好的区分微小故障和正常数据特征的能力。通过图3(c)可以清楚地看出,更多的数据偏离正常数据范围,具有比前两种方法更好的区分能力,这也验证了本文所提方法WSLKPCA通过采用样本加权策略,更容易提取微小故障特征,有利于提高对该类故障的检测性能。

通过图4(a)可以看出,传统KPCA对该故障的

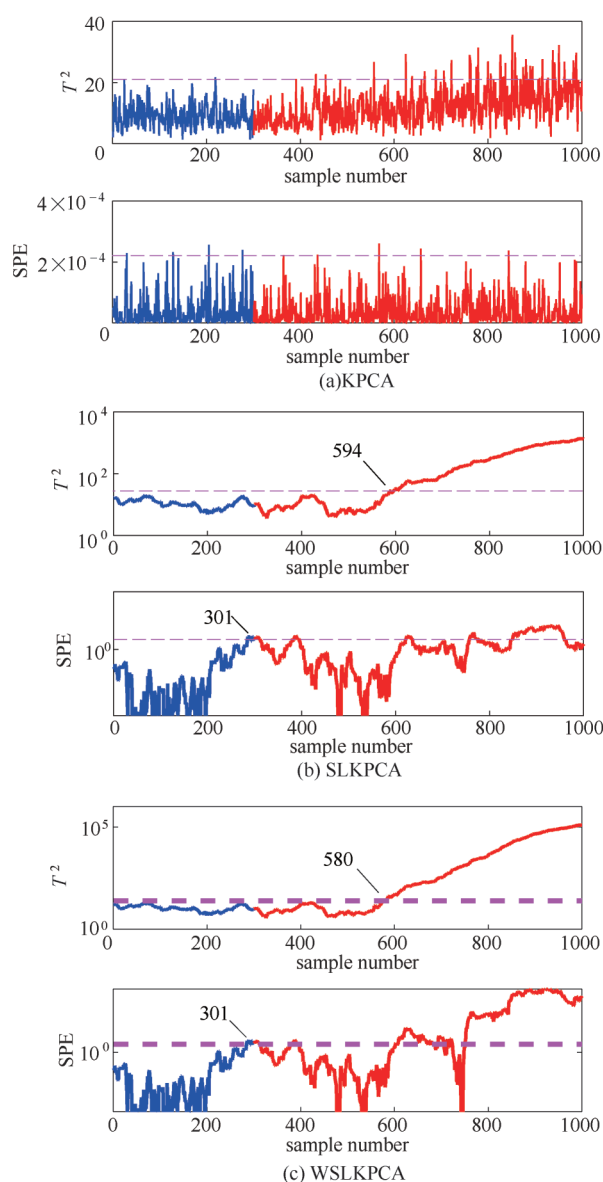


图 4 故障 1 的监控结果

Fig.4 Monitoring results of fault 1

检测结果并不理想, T^2 和SPE统计量大多在控制线以下,其检出率分别为7.14%和0.71%,误报率分别为0.33%和1.33%;从图4(b)可以看出,SLKPCA与KPCA相比,对故障1的检测性能得到了提高,具有较高的故障检出率。其中 T^2 统计量检测故障的时间为第594个采样点,故障检出率为58.71%,误报率为0;而SPE统计量从第301个采样时刻便给出了报警信号,检测更及时。从图4(c)可以看出本文所提方法的两个统计量的检测时间分别为第580和第301个采样点, T^2 统计量相较于SLKPCA方法提前了14个样本检测到该故障,故障检出率为60.29%。SPE统计量同SLKPCA方法一样也能够及时给出报警信号,其故障检出率由25.71%提高到54%。总的来说,在CSTR控制系统的故障1验证了本文方法相较于KPCA和SLKPCA方法对故障检测的优越性。

图5给出了故障1的样本权值,从图5可以清楚地看出,样本权值从第547个采样时刻开始便得到了加强,这显然是因为样本从第547个样本开始偏离了正常数据,但SLKPCA却不能准确地检测出这种变化。通过样本加权,SPE统计量在后期得到了加强,故障检测性能得到了提高。

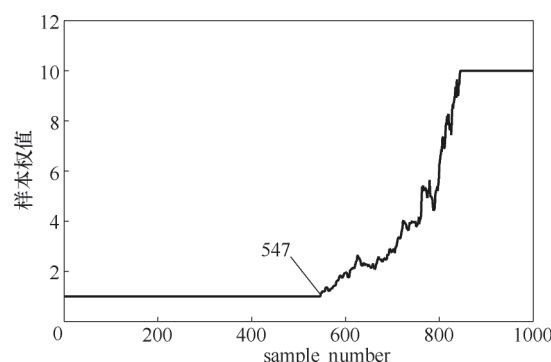


图 5 故障 1 的样本权值

Fig.5 Sample weights of fault 1

对故障1的变量识别结果如图6所示,从该图可以看出, T^2 贡献图和SPE贡献图均表明变量 C_{af} 对统计量的贡献量最大,对应于反应釜进料A的浓度。事实上,故障1是进料浓度 C_{af} 发生斜坡变化,因此该故障的发生受到影响最大的便是变量 C_{af} 。故该方法对故障1的故障变量识别是合理的。与前两种方法相比,本文方法可以很好地捕捉到过程中的微弱变化,并能够准确识别故障变量,对故障的监控性能得到了进一步提高。从故障1的仿真结果可知,WSLKPCA方法具有更好的微小故障监控效果。

故障4反应器内温度传感器发生阶跃变化,幅

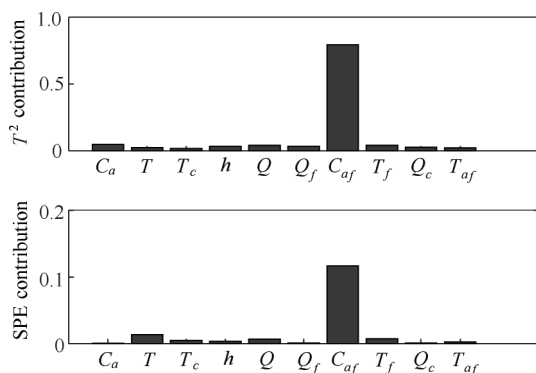


图6 故障1发生时 WSLKPCA 故障变量识别结果

Fig.6 Variable identification result of fault 1 based WSLKPCA

值大小为0.06,属于幅值小的突变微小故障,故对其检测同样较为困难。KPCA、SLKPCA 和 WSLKPCA 三种方法对故障4的监控结果如图7~图9所示。从图7(a)可以看出,KPCA的两个监控统计量大多都在控制线以下,基本检测不到该故障,检出率只有1%和0.71%;而SLKPCA方法通过采用统计局部分析,可以从图7(b)很清楚地看出 T^2 和SPE统计量均可检测到该故障,两个统计量检测到异常点的时间为第712和第301个采样时刻,故障检出率提高到25.29%和22.57%,与KPCA方法相比较,检测效果得到了提高;从图7(c)可以看出,在采取加权策略后,本文所提方法的 T^2 统计量检测该故障的时间提前到第625个时刻,且会持续报警到仿真结束,具有较高的故障检出率,SPE统计量的监控性能也得到了改善,两个统计量的检出率达到53.71%和37.29%。与KPCA和SLKPCA相比,本文所提方法对微小故障的变化更为敏感,检测效果更好。

故障4的样本权值如图8所示,从图8可以看出,该故障从第560个采样时刻的样本距离开始偏离正常数据,通过采用本文的样本加权策略,强化偏离正常数据的样本,从而使得故障样本的作用更加明显的体现出以达到更快更准确地监控微小故障的目的。

每个变量对 T^2 统计量的贡献图如图9所示,从 T^2 贡献图可以看出,变量对统计量拥有相似的贡献量,基本无法识别故障变量,但SPE贡献图表明,变量 T 反应釜的温度对统计量的贡献最大,同时变量 Q (反应釜流出物料的浓度)和变量 T_c (夹套出口冷却剂的温度)对统计量的贡献也较大,根据贡献量的大小,可以判断该故障的发生和这3个变量有关。事实上,该化学反应为一级不可逆的放热反应,该故障是反应器内温度传感器出现偏差,受到最直接

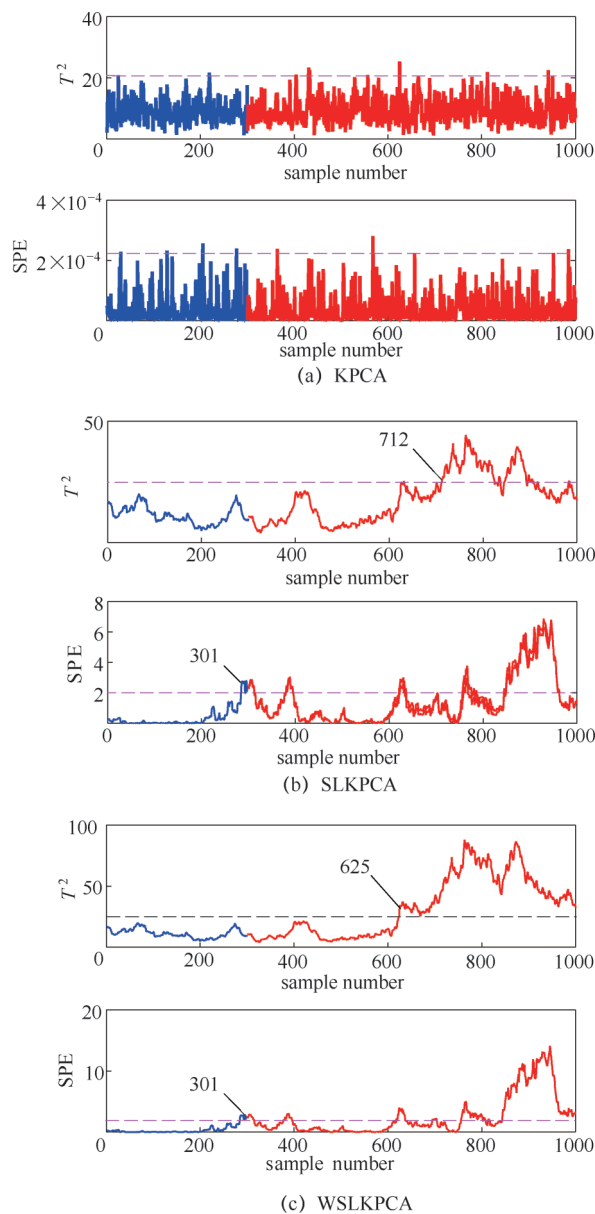


图7 故障4的监控结果

Fig.7 Monitoring results of fault 4

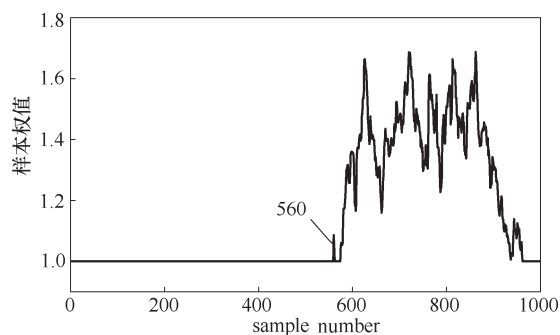


图8 故障4的样本权值

Fig.8 Sample weights of fault 4

影响的便是变量 T ,该故障的发生会导致实际的温度要比测量温度低,从而反应釜内的反应速率会降

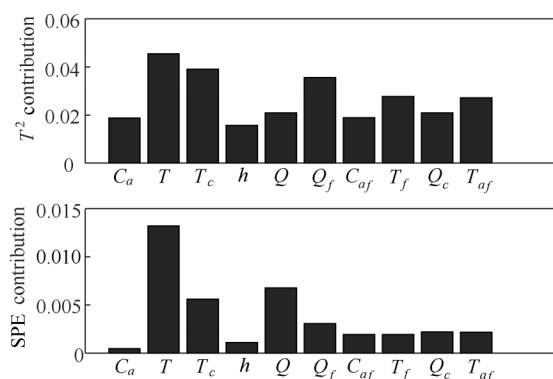


图9 故障发生时WSLKPCA故障变量识别结果

Fig.9 Variable identification result of fault 4 based WSLKPCA

低,进而引起了变量 T 和 T_c 的变化,故本文方法对该故障的变量识别是合理的。

CSTR控制系统的5种故障的检出率对比见表3。从表3可以看出,WSLKPCA方法具有最高的平均故障检出率,与KPCA和SLKPCA相比,对微小故障的检测效果更好。因此,WSLKPCA方法所设计的加权策略对微小故障的变化更为敏感,更加有助于改善故障的监控性能。

表3 五种故障的检出率

Table 3 Fault detection rates of 5 faults/%

故障	KPCA		SLKPCA		WSLKPCA	
	T^2	SPE	T^2	SPE	T^2	SPE
1	7.14	0.71	58.71	25.71	60.29	54
2	2.71	1.71	51.57	61.29	56	62.57
3	1.29	3.57	7.14	60.86	15.71	61
4	1	0.71	25.29	22.57	53.71	37.29
5	1.43	3.71	5.86	90.86	16.14	90.86
平均	2.71	2.00	29.71	52.26	40.37	61.14

虽然本文所提出的“基于加权统计局部核主元分析的非线性化工过程微小故障诊断方法”是针对微小故障而设计的,但在发生显著故障时仍具有良好的故障检测能力。为了验证这一点,另外采集了CSTR控制系统的两组显著故障数据集故障6和故障7,两种故障的描述同故障4和故障5,但幅值大小分别为25和30。

表4和表5分别给出了三种方法对CSTR控制系统显著故障的检出率和检测时间。通过表4和表5监控结果的对比可以看出,传统KPCA方法能够及时检测到显著故障,平均故障检出率达到99.86%。当采用统计局部方法后,SLKPCA方法的 T^2 统计量在第303个采样时刻检测到故障7。与传统KPCA方法相比,延迟了2个采样时刻。这是因为SLKPCA

方法涉及到了基于滑动窗技术的改进残差的构造,在一个数据窗内监控每一个样本,因此产生故障检测延迟是合理的。WSLKPCA方法对这两种故障的检测结果与SLKPCA相同,即平均故障检出率同样达到较好的效果,分别为99.79%和99.93%。这是因为样本加权策略在该显著阶跃故障并未起到效果。但是对于微小故障的监控依然能够得到较好的效果。通过比较这三种方法可以看出,本文方法能够在提高对微小故障监控性能的同时,也适用于显著故障的检测。

表4 显著故障的故障检出率

Table 4 Fault detection rates of significant faults/%

故障	KPCA		SLKPCA		WSLKPCA	
	T^2	SPE	T^2	SPE	T^2	SPE
6	99.86	99.86	99.86	99.86	99.86	99.86
7	99.86	99.86	99.71	100	99.71	100
平均	99.86	99.86	99.79	99.93	99.79	99.93

表5 显著故障的故障检测时间

Table 5 Fault detection times of significant faults/%

故障	KPCA		SLKPCA		WSLKPCA	
	T^2	SPE	T^2	SPE	T^2	SPE
6	302	302	302	302	302	302
7	301	301	303	301	303	301
平均	301.5	301.5	302.5	301.5	302.5	301.5

3.3 基于TE过程的方法有效性验证

为了进一步验证本文方法对微小故障检测的有效性,下面以田纳西伊斯曼(Tennessee Eastman, TE)过程为仿真对象进行仿真验证。TE过程是一种著名的化工过程,它被广泛地应用于评价各种故障诊断方法。图10为该过程的结构流程图,对TE过程的具体描述参考文献[35-37]。本文中TE过程参与故障诊断的变量为52个,其中不包括搅拌速率。TE过程中共有21个故障,其中故障3、9和15的变化对过程的影响微乎其微,对其检测也是相当困难^[37],对三种故障的描述见表6。在本文所用到的TE过程仿真中,采集了1460个正常数据作为训练集,其余21种故障均包含960个样本作为测试集,且从第161个采样点开始为故障数据。所用到的数据可以从 <http://web.mit.edu/braatzgroup/links.html> 获取。

表7给出了KPCA^[22]、SLKPCA^[22]和WSLKPCA对TE过程难以检测的故障(故障3、故障9和故障15)的仿真结果。其中KPCA和SLKPCA的故障检测结

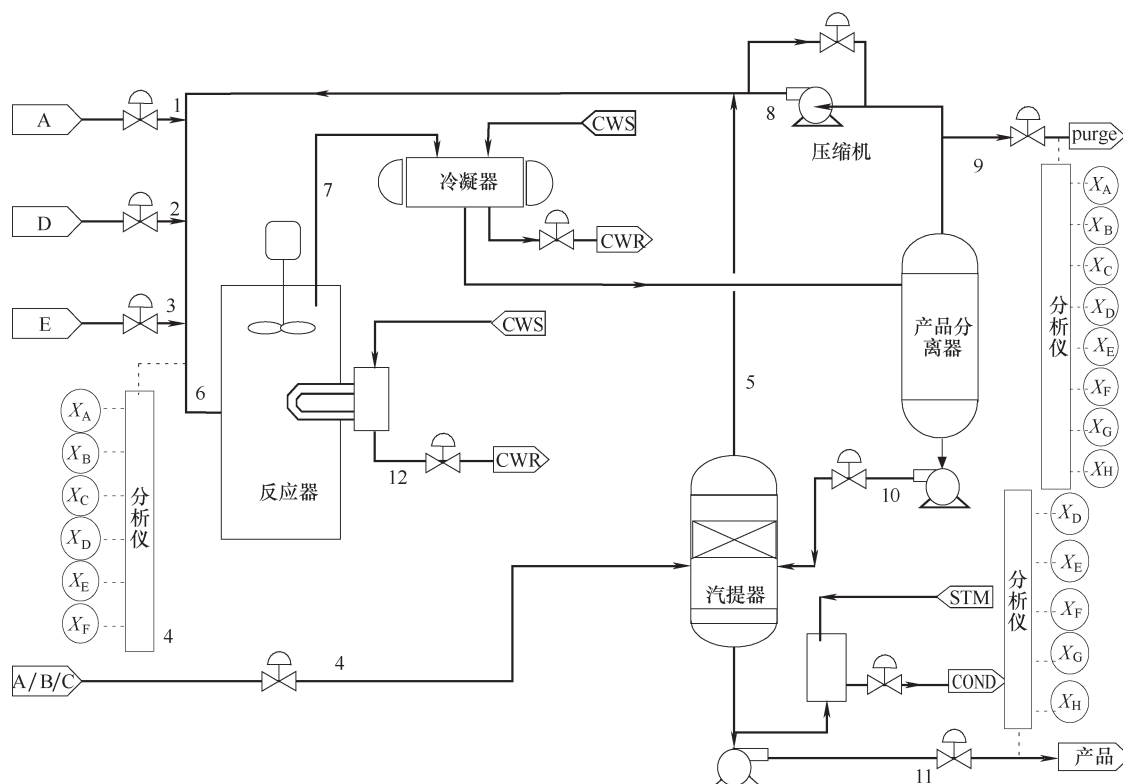


图10 TE过程结构流程图

Fig.10 TE process structure flow chart

表6 TE过程故障列表

Table 6 Fault patterns for TE process

故障名称	描述	类型
3	D进料温度	阶跃
9	D进料温度	随机
15	冷凝器冷却水阀门	粘滞

表7 TE过程的故障检出率

Table 7 Fault detection rates of TE process/%

故障	KPCA ^[22]		SLKPCA ^[22]		WSLKPCA	
	T^2	SPE	T^2	SPE	T^2	SPE
3	3.2	0.4	25.5	0.6	32	2.13
9	3.7	0.5	16.9	0	26.5	23.63
15	8.9	1.1	25.6	14	20.75	20.13
平均	5.27	0.67	22.67	4.87	26.42	15.30

果来自于文献[22]。文献[22]是统计局部核主元分析方法研究的原始文献,也是本文所研究的基础,与其进行结果比较能够充分验证本文方法的有效性。从表7可以看出,传统KPCA^[22]方法难以有效检测该类型的故障,具有最低的平均故障检出率,分别为5.27%和0.67%。而SLKPCA^[22]通过采用统计局部方法,提高了传统KPCA^[22]方法检测微小故障的能力,尤其是 T^2 统计量的故障检出率得到较大提

升。本文所提方法WSLKPCA对微小故障的检测性能进一步得到改善,两个统计量的平均故障检出率分别提高了4%和10%左右,这也验证了本文基于样本距离加权策略的有效性。总的来说,SLKPCA^[22]的微小故障监控性能要超过KPCA^[22],WSLKPCA方法要优于前两种方法。

4 结 论

针对传统KPCA方法难以有效检测化工过程中的微小故障和准确识别故障变量的问题,本文提出了一种基于样本加权统计局部核主元分析(WSLKPCA)的化工过程微小故障诊断方法。该方法通过采用KPCA方法获取数据的核成分,并采用统计局部方法初始残差,通过分析故障数据与正常数据之间的距离差异设计样本加权策略提高对微小故障的敏感性。最后构建基于样本加权改进残差的贡献图用于识别过程中的故障变量。在典型的非线性化工过程CSTR控制系统和TE过程的仿真结果表明,本文所提方法具有良好的微小故障检测性能并能够准确识别故障变量,为化工过程中的微小故障诊断提供了新的思路。

参考文献

- [1] Zhang S, Zhao C. Stationarity test and Bayesian monitoring strategy for fault detection in nonlinear multimode processes[J]. *Chemometrics and Intelligent Laboratory Systems*, 2017, **168**: 45–61.
- [2] 韩敏, 张占奎. 基于改进核主成分分析的故障检测与诊断方法[J]. *化工学报*, 2015, **66**(6): 2139–2149.
Han M, Zhang Z K. Fault detection and diagnosis method based on modified kernel principal component analysis[J]. *Journal of Chemical Industry and Engineering*, 2015, **66**(6): 2139–2149.
- [3] Ge Z, Song Z, Ding S X, et al. Data mining and analytics in the process industry: the role of machine learning[J]. *IEEE Access*, 2017; **5**: 20590–20616.
- [4] 王磊, 邓晓刚, 徐莹, 等. 基于变量子域 PCA 的故障检测方法[J]. *化工学报*, 2016, **67**(10): 4300–4308.
Wang L, Deng X G, Xu Y, et al. Fault detection method based on variable sub-region PCA[J]. *CIESC Journal*, 2016, **67**(10): 4300–4308.
- [5] Deng X, Tian X, Chen S, et al. Fault discriminant enhanced kernel principal component analysis incorporating prior fault information for monitoring nonlinear processes[J]. *Chemom. Intell. Lab. Syst.*, 2017, **162**: 21–34.
- [6] 文成林, 吕菲亚, 包哲静, 等. 基于数据驱动的微小故障诊断方法综述[J]. *自动化学报*, 2016, **42**(9): 1285–1299.
Wen C L, Lyu F Y, Bao Z J, et al. A review of data driven-based incipient fault diagnosis[J]. *Acta Automatica Sinica*, 2016, **42**(9): 1285–1299.
- [7] 李娟, 周东华, 司小胜, 等. 微小故障诊断方法综述[J]. *控制理论与应用*, 2012, **29**(12): 1517–1529.
Li J, Zhou D H, Si X S, et al. Review of incipient fault diagnosis methods[J]. *Control Theory and Applications*, 2012, **29**(12): 1517–1529.
- [8] Shen Y, Steven X D, Adel H, et al. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process[J]. *Journal of Process Control*, 2012, **22**(9): 1567–1581.
- [9] 李元, 宋玉成. MWMPCA 在过程故障检测与诊断中的应用[J]. *沈阳化工大学学报*, 2010, **24**(2): 170–174.
Li Y, Song Y C. Application of MWMPCA in process fault detection and diagnosis[J]. *Journal of Shenyang University of Chemical Technology* 2010, **24**(2): 170–174.
- [10] 王政, 孙锦程, 王迎春, 等. 基于复杂网络理论的符号有向图(SDG)化工故障诊断[J]. *化工进展*, 2016, **35**(5): 1344–1352.
Wang Z, Sun J C, Wang Y C, et al. Research on chemical process signed directed graph(SDG) fault diagnosis based on complex network[J]. *Chemical Industry and Engineering Progress*, 2016, **35**(5): 1344–1352.
- [11] 郑鑫, 田学民, 张汉元. 基于动态稀疏保局投影的故障检测方法[J]. *化工学报*, 2016, **67**(3): 833–838.
Zheng X, Tian X M, Zhang H Y. Fault detection method based on dynamic sparse locality preserving projections[J]. *CIESC Journal*, 2016, **67**(3): 833–838.
- [12] Wold S. Exponentially weighted moving principal components analysis and projections to latent structures[J]. *Chemometrics & Intelligent Laboratory Systems*, 1994, **23**(1): 149–161.
- [13] Harmouche J, Delpha C, Diallo D. Incipient fault detection and diagnosis based on Kullback–Leibler divergence using principal component analysis: Part I[J]. *Signal Processing*, 2014, **94**(1): 278–287.
- [14] Chen H, Jiang B, Lu N. An improved incipient fault detection method based on Kullback–Leibler divergence[J]. *ISA Transaction*, 2018, **79**: 127–136.
- [15] Ge W, Wang J, Zhou J, et al. Incipient fault detection based on fault extraction and residual evaluation[J]. *Industrial & Engineering Chemistry Research*, 2015, **54**(14): 3664–3677.
- [16] Ji H, He X, Shang J, et al. Exponential smoothing reconstruction approach for incipient fault isolation[J]. *Ind. Eng. Chem. Res.*, 2018, **57**(18), 6353–6363.
- [17] 尚骏, 陈茂银, 周东华. 基于变元统计分析的微小故障检测[J]. *上海交通大学学报*, 2015, **49**(6): 799–805.
Shang J, Chen M Y, Zhou D H. Incipient fault detection using transformed component statistical analysis[J]. *Journal of Shanghai Jiaotong University*, 2015, **49**(6): 799–805.
- [18] Chen H, Jiang B, Lu N, et al. Deep PCA based real-time incipient fault detection and diagnosis methodology for electrical drive in high-speed trains[J]. *IEEE Trans. Veh. Technol.*, 2018, **67**(6): 4819–4830.
- [19] Zhang S, Zhao C, Gao F. Incipient fault detection for multiphase batch processes with limited batches[J]. *IEEE Transactions on Control Systems Technology*, 2019, **27**(1): 1–15.
- [20] Zhao C, Gao F. A sparse dissimilarity analysis algorithm for incipient fault isolation with no priori fault information[J]. *Control Engineering Practice*, 2017, **65**: 70–82.
- [21] Zhao B, Shi H, Song B, et al. Quality weakly related fault detection based on weighted Dual-Step feature extraction[J]. *IEEE Access*, 2019, PP(99): 1–1.
- [22] Ge Z, Yang C, Song Z. Improved kernel PCA-based monitoring approach for nonlinear processes[J]. *Chemical Engineering Science*, 2009, **64**(9): 2245–2255.
- [23] 吴希军. 基于主元分析方法的空调系统传感器故障诊断研究[D]. 秦皇岛: 燕山大学, 2005.
Wu X J. Study of the sensor fault diagnosis in air conditioning system based on PCA[D]. Qinhuangdao: Yanshan University, 2005.
- [24] 张成, 李元. 基于统计模量分析间歇过程故障检测方法研究[J]. *仪器仪表学报*, 2013, **34**(9): 2103–2110.
Zhang C, Li Y. Study on the fault detection method in batch process based on statistical pattern analysis[J]. *Chinese Journal of Scientific Instrument*, 2013, **34**(9): 2103–2110.
- [25] Lee J M, Yoo C K, Lee I B. Fault detection of batch processes using multiway kernel principal component analysis[J]. *Computers & Chemical Engineering*, 2004, **28**(9): 1837–1847.
- [26] 赵忠盖, 刘飞. 基于稀疏核主元分析的在线非线性过程监控[J]. *化工学报*, 2008, **59**(7): 1773–1777.
Zhao Z G, Liu F. On-line nonlinear process monitoring based on sparse kernel principal component analysis[J]. *Journal of Chemical Industry and Engineering(China)*, 2008, **59**(7): 1773–1777.
- [27] Krugera U. Improved principal component monitoring using the local approach[J]. *Automatica*, 2007, **43**(9): 1532–1542.
- [28] 胡金海, 谢寿生, 骆广琦, 等. 一种基于贡献率图的 KPCA 故障识别方法[J]. *系统工程与电子技术*, 2008, **30**(3): 572–576.

- Hu J H, Xie S S, Luo G Q, et al. Fault identification method of kernel principal component analysis based on contribution plots and its application[J]. Systems Engineering and Electronics, 2008, **30**(3): 572–576.
- [29] Chen H M, Varshney P K, Arora M K. Performance of mutual information similarity measure for registration of multitemporal remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2003, **41**(11): 2445–2454.
- [30] Keys D, Kholikov S, Pevtsov A A. Application of mutual information methods in time distance helioseismology[J]. Solar Physics, 2015, **290**(3): 659–671.
- [31] Qian W, Shu W. Mutual information criterion for feature selection from incomplete data[J]. Neurocomputing, 2015, **168**: 210–220.
- [32] Haeri M A, Ebadzadeh M M. Estimation of mutual information by the fuzzy histogram[J]. Fuzzy Optimization and Decision Making, 2014, **13**(3): 287–318.
- [33] 周东华, 李钢, 李元. 数据驱动的工业过程故障诊断技术: 基于主元分析与偏最小二乘的方法[M]. 北京: 科学出版社, 2011.
- Zhou D H, Li G, Li Y. Data Driven Industrial Process Fault Diagnosis Technology: Method Based on Principal Component Analysis and Partial Least Squares[M]. Beijing: Science Press, 2011.
- [34] Deng X, Tian X. Multimode process fault detection using local neighborhood similarity analysis[J]. Chinese Journal of Chemical Engineering, 2014, **22**(11/12): 1260–1267.
- [35] Deng X, Wang L. Modified kernel principal component analysis using double-weighted local outlier factor and its application to nonlinear process monitoring[J]. ISA Transactions, 2018, **72**: 218–228.
- [36] Li N, Yang Y. Ensemble kernel principal component analysis for improved nonlinear process monitoring[J]. Industrial & Engineering Chemistry Research, 2015, **54**(1): 318–329.
- [37] Tong C, Song Y, Yan X. Distributed statistical process monitoring based on four-subspace construction and Bayesian inference[J]. Industrial & Engineering Chemistry Research, 2013, **52**(29): 9897–9907.