

控制与决策

Control and Decision

基于交叉熵的改进NPE间歇过程故障检测算法

赵小强, 张和慧

引用本文:

赵小强, 张和慧. 基于交叉熵的改进NPE间歇过程故障检测算法[J]. *控制与决策*, 2021, 36(2): 411–417.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.0725>

您可能感兴趣的其他文章

Articles you may be interested in

[求解约束优化问题的改进果蝇优化算法及其工程应用](#)

Improved fruit fly optimization algorithm for solving constrained optimization problems and engineering applications

控制与决策. 2021, 36(2): 314–324 <https://doi.org/10.13195/j.kzyjc.2019.0557>

[基于混合差分遗传算法的Bouc–Wen迟滞模型辨识策略](#)

Bouc–Wen hysteresis model identification strategy based on hybrid differential genetic algorithm

控制与决策. 2021, 36(2): 371–378 <https://doi.org/10.13195/j.kzyjc.2019.0663>

[基于知识粒度特征的多目标粗糙集属性约简算法](#)

Multi objective rough set attribute reduction algorithm based on characteristics of knowledge granularity

控制与决策. 2021, 36(1): 196–205 <https://doi.org/10.13195/j.kzyjc.2019.0490>

[事件触发机制下分布时滞网络化控制系统 \$H_\infty\$ 故障检测](#)

Event-triggered H_∞ fault detection for networked control systems with distributed delays

控制与决策. 2020, 35(12): 3059–3065 <https://doi.org/10.13195/j.kzyjc.2019.0456>

[基于不变网络模型和故障注入的分布式信息系统故障溯源方法](#)

Fault source location algorithm for distributed information system based on invariant network and fault injection

控制与决策. 2020, 35(11): 2723–2732 <https://doi.org/10.13195/j.kzyjc.2019.0214>

基于交叉熵的改进NPE间歇过程故障检测算法

赵小强^{1,2,3†}, 张和慧¹

1. 兰州理工大学 电气工程与信息工程学院, 兰州 730050;
2. 兰州理工大学 甘肃省工业过程先进控制重点实验室, 兰州 730050;
3. 兰州理工大学 国家级电气与控制工程实验教学中心, 兰州 730050)

摘要: 针对具有复杂动态特性的间歇过程进行故障检测, 邻域保持嵌入 (neighborhood preserving embedding, NPE) 算法在保持数据局部几何结构时因忽略全局信息而造成检测率较低的问题, 提出一种基于交叉熵 (cross entropy, CE) 的邻域保持嵌入 (CEGLNPE) 算法. 首先, 将交叉熵保持全局结构的思想引入 NPE 的局部结构保持中, 得到兼顾全局-局部的目标函数; 然后, 利用滑动窗进一步解决间歇过程动态性问题并建立全局-局部的故障检测模型, 可以同时兼顾全局和局部数据的结构保持, 以提高故障检测效率和精度; 最后, 通过监测 SPE、 T^2 统计量判断是否发生故障. 利用人工数据集 Swiss-Roll 和青霉素发酵仿真过程与 KPCA、NPE 算法进行对比验证, 所得结果表明了所提出算法在故障检测中的有效性.

关键词: 间歇过程; 故障检测; 交叉熵; 邻域保持嵌入; 全局-局部; 滑动窗

中图分类号: TP277

文献标志码: A

DOI: 10.13195/j.kzyjc.2019.0725

开放科学(资源服务)标识码(OSID):



引用格式: 赵小强, 张和慧. 基于交叉熵的改进NPE间歇过程故障检测算法[J]. 控制与决策, 2021, 36(2): 411-417.

Improved NPE batch process fault detection algorithm based on cross entropy

ZHAO Xiao-qiang^{1,2,3†}, ZHANG He-hui¹

1. College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China;
2. Key Laboratory of Gansu Advanced Control for Industrial Processes, Lanzhou University of Technology, Lanzhou 730050, China;
3. National Experimental Teaching Center of Electrical and Control Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: The neighborhood preserving embedding (NPE) algorithm can preserve the data local structure, but ignore the data global structure for fault detection of complex dynamical batch process, which causes low detection rate. An improved NPE algorithm based on cross entropy global-local neighborhood preserving embedding (CEGLNPE) is proposed. Firstly, the idea of the global structure of cross entropy is introduced into NPE, and the objective function with global-local structure is obtained. Then the sliding window is used to further solve the dynamic problem of batch process and a global-local fault detection model is established to improve fault detection efficiency and accuracy, which can preserve the structures of both global and local data simultaneously. Finally, SPE and T^2 statistics are used to determine whether faults occur. The KPCA, NPE and CEGLNPE algorithms are used in the artificial dataset Swiss-Roll and penicillin fermentation simulation process. The results show the effectiveness of the proposed algorithm in fault detection.

Keywords: batch process; fault detection; cross entropy; NPE; global-local; sliding window

0 引言

间歇过程是一种以生产更高质量、严格精细化产品为目标的应用十分广泛的生产方式^[1-2], 当间歇过程中的某个变量发生变化时, 就会影响整个工业生

产过程, 进而直接影响到产品的生产效率. 同时, 间歇过程具有强非线性、非高斯、多阶段和多模态共存等十分复杂的特性, 相比于连续过程发生的故障, 其故障随机性强、持续时间短以及重复发生频率高, 如

收稿日期: 2019-05-25; 修回日期: 2019-09-02.

基金项目: 国家自然科学基金项目(61763029, 61873116).

责任编委: 赵春晖.

†通讯作者. E-mail: xqzhao@lut.cn.

果没有高效地对故障做出回应,则将造成检测效率降低、速度较慢,对工业生产中的人员安全和财产起不到更好的保障作用,因此,对间歇过程进行故障检测是十分必要和迫切的^[3-5].

近年来,诸如主成分分析(principal component analysis, PCA)、独立主元分析(independent component analysis, ICA)、偏最小二乘(partial least squares, PLS)等传统算法在很多领域得到了应用^[6-8],其中在大型工业设备的故障检测等方面应用愈加广泛.这些算法在挖掘数据全局欧氏结构、解决数据非线性、考虑质量相关变量等方面有着不错的效果.然而,逐渐复杂化的工业大环境,使得上述算法无法对生产过程故障进行高效、快速、准确的检测,基于此,更多的流形算法被相继提出并被应用到各种工业生产中.邻域保持嵌入算法(neighborhood preserving embedding, NPE)是近年来提出的一种关注数据局部近邻信息的流形降维算法,其通过寻找样本的 k 个近邻保持邻域结构,对局部信息达到最大化利用,一定程度上提高了故障检测的精度.然而,由于NPE没有保持全局特性的能力,只对局部近邻加以利用,不能进一步提高故障检测的效率和精度,人们又提出了一些同时兼顾全局和局部的算法.Hui等^[9]提出了一种多向动态非线性全局邻域保持嵌入算法,有效兼顾了全局和局部的特征信息,在解决间歇过程故障检测问题上取得了较好的效果;马萍等^[10]基于 k 近邻提出了一种基于局部与全局结构保持的算法,用于滚动轴承故障诊断;Zhang等^[11]提出了考虑局部和全局的统计量模式分析(LGSPA)算法并有效用于故障检测;张晓涛等^[12]提出了一种结合PCA和LPP特点的半监督算法并用于故障辨识,该算法提高了辨识的精度与效果.这些算法虽然同时利用了全局和局部的信息,但都没有进行全局优化,仍间接损失了一些数据信息,同时,未考虑数据动态性特性对故障检测结果造成的不良影响.

针对上述问题,本文提出一种基于交叉熵的全局-局部邻域保持嵌入(CEGLNPE)算法,同时考虑全局和局部特性.利用交叉熵(cross entropy, CE)算法进行特征约简可以保持数据全局结构特征,而邻域保持嵌入(NPE)算法对数据进行低维映射时可以保证局部结构的不流失.为进一步提高故障检测效率,在线监测时利用滑动窗更新数据实现动态监测以解决数据动态性问题.所提出的CEGLNPE算法利用Swiss-Roll人工数据集对降维嵌入效果进行验证,并通过青霉素发酵仿真平台模拟间歇过程进行故障检测的进

一步验证,结果表明,CEGLNPE算法比KPCA和NPE算法效果都好,检测精度更高、速度更快.

1 基础算法

1.1 交叉熵算法(CE)

熵是一种在热力学中表征物质状态参量的概念,最早被克劳修斯(Clausius)于1854年提出.而最先被美国数学家香农(Shannon)提出的信息熵是一种相对抽象的概念,解决了对信息的量化度量问题.随后,科学家们开始对信息熵广泛研究,提出了解决属性约简问题并表达变量对条件期望的条件熵^[13]、用于衡量两个概率分布之间差异的相对熵^[14]、有效表达变量间相互关系的传递熵^[15]以及度量两个概率之间差异性的交叉熵.

交叉熵(CE)于1997年被Rubinstein^[16]提出用于稀疏事件的模拟,之后拓展到随机优化领域^[17].交叉熵代表两个概率密度函数的相似度,它在多目标全局优化上有着较高的效率,可以通过分析计算更新规则,使优化路径更加高效和快速,被广泛应用于电力系统故障诊断^[18]模型、图像匹配^[19]以及控制系统参数优化^[20]等多个领域.

交叉熵通过多次迭代进行全局优化,可以对三维数据矩阵全局寻优并特征约简.设有概率密度函数分别为 f 和 g ,则交叉熵可定义为

$$D(f(x)||g(x)) = \sum_{x \in X} f(x) \ln \frac{f(x)}{g(x)} = E_f(x) \ln \frac{f(x)}{g(x)}, \quad (1)$$

其中 $E(\cdot)$ 为期望.

由式(1)可知, $D(f(x)||g(x)) \geq 0$,因为交叉熵是一个下凸函数,所以当且仅当 $f(x) = g(x)$ 时, $D(f(x)||g(x)) = 0$.因此,交叉熵表征的是两个分布的相似程度,越相似,交叉熵越小^[21].

根据二维图像熵的概念^[19],将其引申到二维矩阵中,对于矩阵 $A(M \times N)$, $f(i, j) (\geq 0)$,矩阵二维熵可定义为

$$H = - \sum_{i=1}^M \sum_{j=1}^N p_{ij} \log p_{ij}, \quad (2)$$

$$p_{ij} = \frac{f(i, j)}{\sum_{i=1}^M \sum_{j=1}^N f(i, j)}. \quad (3)$$

因此,矩阵的二维熵可以表达矩阵间的信息量,并根据设定的概率密度函数不断更新样本,移动所要提取特征的样本点的中心来提高所需样本的生成概率.通过交叉熵迭代寻优可以大大提高矩阵降维后

的全局结构的保持性,并以此提高故障的检测率。

1.2 邻域保持嵌入算法(NPE)

相比于传统的PCA和ICA算法,邻域保持嵌入算法的特点是“以局部观整体”,它更多地关注数据的局部结构,通过一个样本来联系其他与该点欧氏距离相近的近邻点,并在之间构造重构权向量以保持数据的局部特性,得到局部结构分布,最终获取到数据整体的流形特征。He等^[22]认为,NPE构造的权重向量可以掌握那些在平移、旋转、缩放中保持不变的性质。近年来,有大量学者开始关注并将其应用到故障监测领域^[23-25]。

首先,对于给定的训练数据 $X(x_1, x_2, \dots, x_n) \in R^{m \times n}$ 中利用 k 近邻法 (k -nearest neighbors, k NN) 为每个样本点确定与其欧氏距离最近的 k 个点,将这些点作为近邻点,利用找到的近邻点构造重构系数 W , W 可通过最小化下式得到:

$$\phi(W) = \sum_i \left\| x_i - \sum_{j=1}^n w_{ij} x_j \right\|^2. \quad (4)$$

其中: $\sum_{j=1}^n w_{ij} = 1, j = 1, 2, \dots, m$. 若 i 到 j 有一条路径可以通过,则权值为 w_{ij} ; 若没有,则权值为0。

然后,根据式(4)求得的权重系数矩阵,将高维空间 R^m 中的 x_i 对应地通过投影矩阵 A 重构到低维空间 y_i . 投影矩阵 $A(a_1, a_2, \dots, a_d)$ 的求解为最小化下式的代价函数:

$$\phi(A) = \min \left\| \sum_{i=1}^n A^T x_i - A^T \sum_{j=1}^n w_{ij} x_j \right\|^2. \quad (5)$$

最后,利用拉格朗日乘子法,将有约束问题转化为无约束问题,得到下式所示的广义特征值问题,进而通过求解得到投影矩阵 A :

$$XMX^T a = \lambda XX^T a, \quad (6)$$

其中 $M = (I - W)^T(I - W)$ 是一个对称半正定矩阵。

由于 XMX^T 和 XX^T 为半正定矩阵,则由最小的 d 个特征值 ($\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_d$) 所对应的特征向量即可组成投影矩阵 A , 通过 $Y = A^T X$ 得到 X 在低维空间 R^d 的低维表示, Y 为降维后的数据矩阵。

2 基于交叉熵的全局-局部邻域保持嵌入(CEGLNPE)算法

2.1 全局结构保持

为了有效保持全局结构,可以通过寻找方差最大的方向,对于样本 $X = [x_1, x_2, \dots, x_n]^T$, 找到投影矩阵 A 并将其映射至低维空间. 全局保持的目标函数^[26]为

$$\begin{aligned} J &= \max \sum_{j=1}^n \|y_j - \bar{y}\|^2 = \\ &= \max \sum_{i=1}^n A^T(x_i - \bar{x})(x_i - \bar{x})^T A = \\ &= \max A^T C A. \end{aligned} \quad (7)$$

$$\text{其中: } C = \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})^T, \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

结合交叉熵算法,在产生带参数且服从某个概率分布函数的训练集里,计算各个样本所对应的目标函数值,从小到大进行排序,并以最大的函数值样本的特征更新概率分布函数中的参数,由此在下一次迭代中产生更“优质”的样本. 通过这种方式,使得样本质量在全局性上不断提高,最后向着最优值靠拢,直到达到最优值。

由式(2)、(3)和(7)可得到在方差最大的方向寻求全局最优的交叉熵目标函数,即

$$\begin{aligned} J_g(A) &= \max \sum_{j=1}^N \|y_j - \bar{y}\|^2 S_{ij}^p = \\ &= \frac{1}{2} \sum_{j=1}^N (A^T(x_i - \bar{x})(x_i - \bar{x})^T A) S_{ij}^p = \\ &= A^T X D^p X^T A - A^T X S^p X^T A = \\ &= A^T S A. \end{aligned} \quad (8)$$

其中: S_{ij}^p 为所对应的交叉熵概率矩阵, D^p 、 S^p 为所对应的系数矩阵。

2.2 局部结构保持

局部数据结构的保持通过寻求 k 近邻并构造权重向量实现,提取数据的近邻信息使得近邻点之间一定距离的样本得以保留。

局部目标函数为

$$\begin{aligned} J_l(A) &= \min \sum_i \left\| y_i - \sum_{j=1}^k w_{ij} y_j \right\|^2 = \\ &= \min A^T X (I - W)^T (I - W) X^T A = \\ &= \min A^T X M X^T A, \end{aligned} \quad (9)$$

其中 $M = (I - W)^T(I - W)$ 。

2.3 全局-局部的目标函数

单纯地考虑局部或者全局势必会造成一部分信息无法得到利用,因此,根据映射关系 $Y = A^T X$, 构建综合全局和局部特征提取的优化目标函数如下式所示:

$$J_{g-l}(A) = \max \frac{J_g(A)}{J_l(A)} =$$

$$\begin{aligned} \max \frac{(A^T(x_i - \bar{x})(x_i - \bar{x})^T A)S_{ij}^p}{A^T X(I - W)^T(I - W)X^T A} = \\ \max \frac{A^T S A}{A^T R A}, \end{aligned} \quad (10)$$

其中 $R = X(I - W)^T(I - W)X^T$.

式(10)的优化可通过拉普拉斯方法转换为

$$SA = \lambda RA. \quad (11)$$

通过式(11)求得最终的投影矩阵 A . 得到的新的目标函数既能考虑局部邻近间的信息,又能最大化保持全局结构,实现利用信息最大化,达到有效提高故障检测率的结果.

3 基于CEGLNPE的间歇过程故障检测

3.1 数据预处理

连续过程数据是由时间和变量构成,而间歇过程数据比连续过程多一个“批次”维,构成三维矩阵 $X(I \times J \times K)$. 其中: I 是批次, J 是变量, K 是时间. 然而,在建模时无法对三维数据直接利用,因此,需要通过如图1所示的批次-变量(AT)法将其展开成二维数据. 首先,沿批次方向展开 $X(KJ \times I)$; 然后对其标准化,再沿变量方向排列为 $X(KI \times J)$; 最后,利用 $X(KI \times J)$ 进行建模研究.

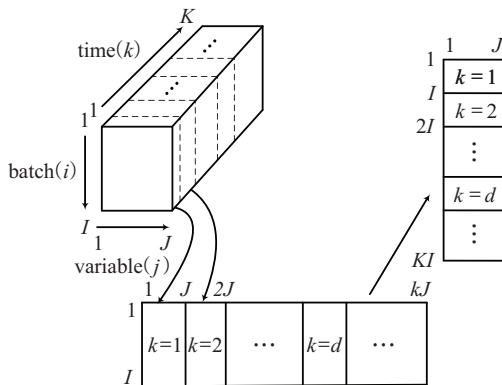


图1 间歇过程数据展开方式

在实际的生产过程中,数据动态性问题往往对故障检测结果造成不良的影响,通过对数据添加滑动窗的方法可以有效消除变量之间的自相关,解决动态性造成的问题.

对于训练数据 $X \in R^{m \times n}$, 添加窗宽为 d 的滑动窗; 而对于 $X_d \in R^{(m-d+1) \times nd}$, 利用 CEGLNPE 算法降维, 第 i 批次的的数据如下所示:

$$X_i(d \times J/K) = \begin{bmatrix} x_{k-d+1,1}^i & x_{k-d+1,2}^i & \cdots & x_{k-d+1,J}^i \\ x_{k-d+2,1}^i & x_{k-d+2,2}^i & \cdots & x_{k-d+2,J}^i \\ \vdots & \vdots & \ddots & \vdots \\ x_{k,1}^i & x_{k,1}^i & \cdots & x_{k,1}^i \end{bmatrix}.$$

本文选取 $d = 2 \sim 7$ 做了实验对比^[15], 通过检测

的速度和精度确定窗宽 $d = 5$.

3.2 监控统计量

SPE、 T^2 统计量分别从样本与残差和主元空间之间的偏离程度描述对数据的欠拟合程度, 定义如下:

$$SPE = e_i e_i^T \leq \delta_\gamma^2, \quad (12)$$

$$T_{(x)}^2 = x^T P \Lambda^{-1} P^T x \leq T_\gamma^2. \quad (13)$$

其中: e_i 表示 x_i 的残差, δ_γ^2 和 T_γ^2 分别是 SPE、 T^2 的监测控制限, P 为投影矩阵, $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_m\}$ 为特征值矩阵. 对于 SPE、 T^2 , 当分别满足 $SPE \leq \delta_\gamma^2$ 和 $T^2 \leq T_\gamma^2$ 时, 为正常状态; 否则, 表示故障发生.

通过核密度估计(KDE)确定95%为控制限^[27].

3.3 算法实现

3.3.1 离线建模

1) 选择一定批次数据 $X(I \times J \times K)$ 组成训练样本集, 利用批次-变量(AT)法将其展开为二维数据 $X(KI \times J)$;

2) 设滑动窗窗宽 $d = 5$ 进行预处理并标准化;

3) 对参数(如迭代次数等)进行初始化;

4) 通过式(10)进行全局-局部信息保持, 求解式

(11)得到投影矩阵 A ;

5) 根据式(12)、(13)计算 SPE、 T^2 统计;

6) 得到 SPE、 T^2 统计量的控制限.

3.3.2 在线监测

1) 对获取的在线故障数据 X_{new} 所对应的窗口数据标准化;

2) 将窗口内的故障数据投影到低维空间, 得到新的投影矩阵 A ;

3) 根据式(12)、(13)计算故障数据的 SPE、 T^2 统计并得到控制限;

4) 若统计量超过控制限则判定发生故障, 否则视为正常工况.

4 仿真与结果分析

4.1 人工数据集(Swiss-Roll)

Swiss-Roll 瑞士卷作为验证不同算法嵌入结果的标准数据集应用十分广泛, 其在三维欧氏空间中的 Swiss-Roll 曲面上随机采样, 并利用不同颜色区分样本的分布, 近邻由相同颜色标识. 本文利用 Swiss-Roll 数据集, 设本征维数为 2, $N = 2000$, 通过降维将复杂的非线性流形展开, 以观察降维过程中数据结构的保持性. 将本文提出的 CEGLNPE 算法与 CE、NPE 进行对比, 以验证所提出算法在全局、局部的保持性.

图2由左至右分别是 Swiss-Roll 曲面、采样数据

和算法二维嵌入结果. 由图2(a)和图2(b)可知: CE算法在降维嵌入后仍对“瑞士卷”数据有效保持了全局流形特征; 而NPE对Swiss-Roll数据集降维嵌入时是从一个方向将“瑞士卷”“压扁”, 呈现一种“长条状”, 扭曲了原有流形的整体结构, 从而印证了NPE更多保留的是局部特性的结论. 从图2(c)可以看出, CEGLNPE算法能得到更好的结果, 其对全局和局部都有兼顾, 既保留了“瑞士卷”的整体结构, 又使样本点之间的距离减小, 体现了保持局部信息的能力, 更好地利用了数据全局和局部的信息, 减少了数据的损失率.

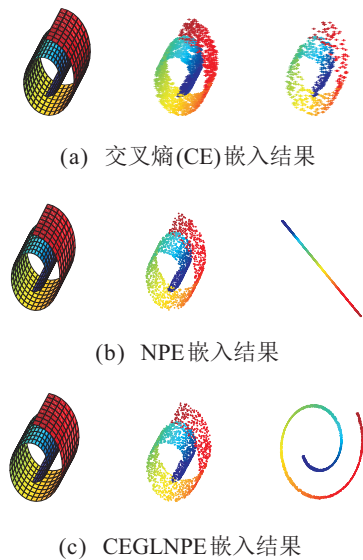


图2 不同算法对Swiss-Roll数据集降维后的二维映射曲线

4.2 青霉素发酵间歇过程

青霉素发酵过程是典型的间歇过程, 具有多阶段、强非线性、多模态以及高斯非高斯混合分布的复杂特性. 本文选择美国Illinois州立理工学院设计的仿真青霉素发酵过程的仿真平台Pensim2.0来模拟间歇过程^[28], 从该平台得到的间歇过程数据已经得到国际上的普遍认可. 除了可以提供正常多批次训练数据, Pensim2.0还可以通过设置故障类型与故障开始和终止的时间得到故障数据集, 而故障类型包括通风率故障、搅拌功率故障和底物流加速率故障, 产生故障的方式为添加阶跃、斜坡信号. 本文从该平台提供的18个变量中选取10个作为监控变量, 选取变量如表1所示.

设定每批次反应时间为400h, 采样时间0.5h, 即采样点为800, 共采集30批次训练样本, 得到 $X(30 \times 800 \times 10)$ 的正常样本; 引入故障类型2(搅拌功率故障), 在采样时间200~300h添加+10%的阶跃信号作为在线检测的样本. 通过正常、故障数据, 分别对KPCA^[29]、NPE算法^[30]和本文提出的CEGLNPE算

法进行间歇过程监测, 结果如图3~图8所示.

表1 监控变量

变量序号	变量名称
1	通气速率
2	搅拌器速率
3	底物进料速率
4	底物进料温度
5	底物浓度
6	溶解氧(DO)浓度
7	反应器体积
8	二氧化碳浓度
9	pH
10	发酵罐温度

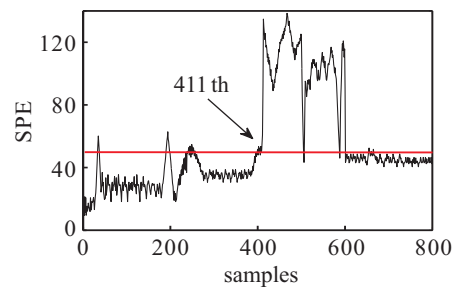


图3 KPCA对故障的SPE监控图

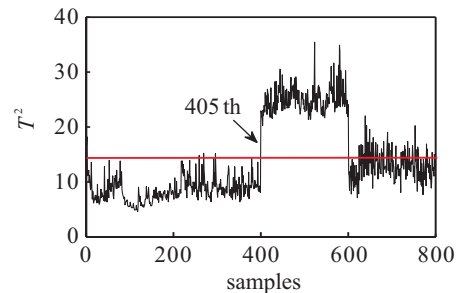


图4 KPCA对故障的T^2监控图

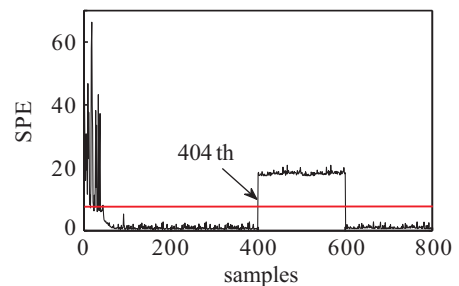


图5 NPE对故障的SPE监控图

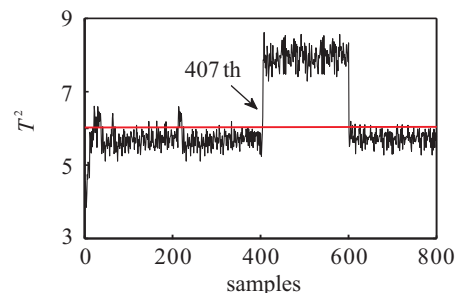


图6 NPE对故障的T^2监控图

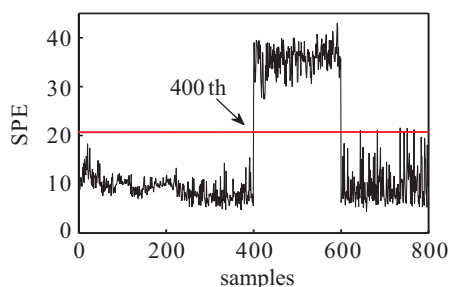


图7 CEGLNPE对故障的SPE监控图

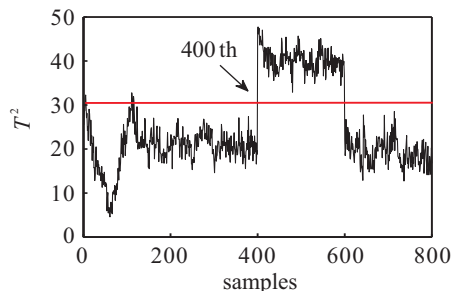
图8 CEGLNPE对故障的 T^2 监控图

图3、图4分别是KPCA算法对所添加的+10%阶跃信号故障2的SPE、 T^2 监控图,可以看出,分别在411采样点和405采样点检测到故障,并且 T^2 监控图在620~780采样点之间有多次误报产生;从图5、图6所示NPE算法对SPE、 T^2 监控图中可以看出,分别在404采样点和407采样点检测出故障,在0~400、600~800采样点均有不同程度的误报;图7、图8为本文所提出算法CEGLNPE的SPE、 T^2 监控图,可以看出, T^2 统计量在0~400采样点未产生误报情况,两种统计量都在400采样点便发生跳升明显超限,说明迅速准确检测到了故障,较其他两种方法检测速度更快,漏、误报率也更低,对故障更为敏感.各方法漏、误报率和检测率见表2.

表2 不同方法对故障的监控统计 %

统计量	KPCA		NPE		CEGLNPE	
	SPE	T^2	SPE	T^2	SPE	T^2
漏报率	7.50	6.50	4.50	4.00	0.50	1.00
误报率	6.17	9.67	4.83	8.33	1.17	2.17
检测率	92.50	93.50	95.50	96.00	99.50	99.00

由表2可知,本文所提出算法相比于KPCA和NPE算法,漏、误报率均有所降低,检测率也有所提高.虽然KPCA在处理数据非线性上有不错的效果,但由于其只关注数据全局特性,因而造成故障检测率不高的结果.NPE算法只寻求近邻点之间的关系,但忽略了较远的点和全局特性对监测结果的重要性,因此精度也受到了影响.而本文所提出的CEGLNPE算法既对数据进行全局优化以保证其全局结构特征,又保留了数据的局部流形结构,同时还兼顾全局和局部,最

大化利用了数据信息,通过滑动窗考虑了数据的动态性,检测速度和精度都有了提高,仿真结果表明了该方法的有效性.

5 结论

间歇过程中多种特性并存、分布复杂,对其进行有效的监控和故障检测是十分必要的.本文提出的基于交叉熵的全局-局部邻域保持嵌入算法,利用交叉熵和邻域保持嵌入算法分别对数据提取全局和局部的特性,达到对数据流形结构两方面的保持.利用Swiss-Roll人工数据集证明其数据结构的保持性,并通过青霉素仿真过程与KPCA、NPE算法进行了比较验证,所得结果表明,该算法对间歇过程提高了故障检测精度并加快了故障检测速度.

参考文献(References)

- [1] 周东华, 史建涛, 何潇. 动态系统间歇故障诊断技术综述[J]. 自动化学报, 2014, 40(2): 161-171. (Zhou D H, Shi J T, He X. Review of intermittent fault diagnosis techniques for dynamic systems[J]. Acta Automatica Sinica, 2014, 40(2): 161-171.)
- [2] Pieter V, Geert G, Jef V. Dynamic model-based fault diagnosis for (bio)chemical batch processes[J]. Computers & Chemical Engineering, 2012, 40(1): 12-21.
- [3] Wang L M, He X, Zhou D H. Average dwell time-based optimal iterative learning control for multi-phase batch processes[J]. Journal of Process Control, 2016, 40: 1-12.
- [4] Zhang H L, Qi Y S, Wang L, et al. Fault detection and diagnosis of chemical process using enhanced KECA[J]. Chemometrics and Intelligent Laboratory Systems, 2017, 161: 61-69.
- [5] Ben I K, Liama M, Weihs C. Variable window adaptive kernel principal component analysis for nonlinear nonstationary process monitoring[J]. Computers & Industrial Engineering, 2011, 61(3): 437-446.
- [6] 翟坤, 杜文霞, 吕锋, 等. 一种改进的动态核主元分析故障检测方法[J]. 化工学报, 2019, 70(2): 716-722. (Zhai K, Du W X, Lyu F, et al. Fault detect method based on improved dynamic kernel principal component analysis[J]. CIESC Journal, 2019, 70(2): 716-722.)
- [7] Tan R M, Cao Y. Multi-layer contribution propagation analysis for fault diagnosis[J]. International Journal of Automation and Computing, 2019, 16(1): 40-51.
- [8] Fouzi Harrou, Mohamed N Nounou, Hazem N Nounou, et al. PLS-based EWMA fault detection strategy for process monitoring[J]. Journal of Loss Prevention in the Process Industries, 2015, 36: 108-119.
- [9] Hui Y Y, Zhao X Q. Multi-phase batch process monitoring based on multiway weighted global neighborhood preserving embedding method[J]. Journal of Process Control, 2018, 69: 44-57.
- [10] 马萍, 张宏立, 范文慧. 基于局部与全局结构保持算法

- 的滚动轴承故障诊断[J]. 机械工程学报, 2017, 53(2): 20-25.
(Ma P, Zhang H L, Fan W H. Fault diagnosis of rolling bearings based on local and global preserving embedding algorithm[J]. Chinese Journal of Mechanical Engineering, 2017, 53(2): 20-25.)
- [11] Zhang H Y, Tian X M, Deng X G. A local and global statistics pattern analysis method and its application to process fault identification[J]. Chinese Journal of Chemical Engineering, 2015, 23(11): 1782-1792.
- [12] 张晓涛, 唐力伟, 王平, 等. 基于半监督PCA-LPP流形学习算法的故障降维辨识[J]. 中南大学学报: 自然科学版, 2016, 47(5): 1559-1564.
(Zhang X T, Tang L W, Wang P, et al. Fault identification and dimensionality reduction method based on semi-supervised PCA-LPP manifold learning algorithm[J]. Journal of Central South University: Science and Technology, 2016, 47(5): 1559-1564.)
- [13] Zhou X M. A formula of conditional entropy and some applications[J]. Discrete and Continuous Dynamical Systems, 2016, 36(7): 4063-4075.
- [14] Fei L G, Deng Y. A new method to identify influential nodes based on relative entropy[J]. Chaos Solitons & Fractals, 2017, 104: 257-267.
- [15] 赵化良. 基于传递熵的MPCA间歇过程监测方法[J]. 计算机系统应用, 2016, 25(2): 146-151.
(Zhao H L. MPCA online monitoring based on transfer entropy for batch process[J]. Computer Systems & Applications, 2016, 25(2): 146-151.)
- [16] Rubinstein R Y. Optimization of computer simulation models with rare events[J]. European Journal of Operational Research, 1997, 99(1): 89-112.
- [17] Rubinstein R. The cross-entropy method for combinatorial and continuous optimization[J]. Methodology and Computing in Applied Probability, 1999, 1(2): 127-190.
- [18] 边莉, 张欣欣, 高雪, 等. 交叉熵在油浸式变压器故障诊断中的应用[J]. 黑龙江科技大学学报, 2018, 28(4): 410-414.
(Bian L, Zhang X X, Gao X, et al. Cross entropy application in fault diagnosis of oil immersed transformer[J]. Journal of Heilongjiang Institute of Science and Technology, 2018, 28(4): 410-414.)
- [19] 周军妮, 杨润玲, 王燕妮, 等. 一种结合交叉熵和投影特征的图像匹配算法[J]. 小型微型计算机系统, 2013, 34(2): 405-408.
(Zhou J N, Yang R L, Wang Y N, et al. Image matching algorithm based on cross-entropy and projection features[J]. Journal of Chinese Computer Systems, 2013, 34(2): 405-408.)
- [20] Farid MiarNaeimi, Gholamreza Azizyan, Mohsen Rashki. Multi-level cross entropy optimizer(MCEO): An evolutionary optimization algorithm for engineering problems[J]. Engineering with Computers, 2018, 34(4): 719-739.
- [21] 苏志刚, 赵子轩, 吴仁彪. 基于交叉熵的缓变故障检测技术[J]. 信号处理, 2011, 27(11): 1640-1644.
(Su Z G, Zhao Z X, Wu R B. Technique of detecting the degradation of failure based on cross-entropy[J]. Journal of Signal Processing, 2011, 27(11): 1640-1644.)
- [22] He X F, Cai D, Yan S C, et al. Neighborhood preserving embedding[C]. Proceedings of the 10th IEEE International Conferences on Computer Vision. Beijing, 2005: 1208-1213.
- [23] Jiang Q C, Yan X F. Probabilistic weighted NPE-SVDD for chemical process monitoring[J]. Control Engineering Practice, 2014, 28: 74-89.
- [24] Song B, Tan S, Shi H B. Process monitoring via enhanced neighborhood preserving embedding[J]. Control Engineering Practice, 2016, 50: 48-56.
- [25] Hui Y Y, Zhao X Q. Batch process monitoring based on WGNPE-GSVDD related and independent variables[J]. Chinese Journal of Chemical Engineering, 2018, 26(12): 2549-2561.
- [26] Wang X G, Huang L W, Zhang Y W. Modeling and monitoring of nonlinear multi-mode processes based on similarity measure-KPCA[J]. Journal of Central South University, 2017, 24(3): 665-674.
- [27] 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(5): 531-543.
- [28] Birol G, Ündey C, Çinar A. A modular simulation package for fed-batch fermentation: Penicillin production[J]. Computers and Chemical Engineering, 2002, 26(11): 1553-1565.
- [29] Lee J M, Yoo C, Lee L B. Fault detection of batch processes using multiway kernel principal component analysis[J]. Computers & Chemical Engineering, 2004, 28(9): 1837-1847.
- [30] Miao A M, Ge Z Q, Song Z H, et al. Nonlocal structure constrained neighborhood preserving embedding model and its application for fault detection[J]. Chemometrics and Intelligent Laboratory Systems, 2015, 142: 184-196.

作者简介

赵小强(1969—), 男, 教授, 博士生导师, 从事故障检测与诊断、图像处理等研究, E-mail: xqzhao@lut.cn;

张和慧(1995—), 女, 硕士生, 从事间歇过程故障检测与诊断的研究, E-mail: zhh950620@163.com.

(责任编辑: 李君玲)