

研究论文

基于测地线距离统计量的多工况间歇过程监测

郭小萍, 李婷, 李元

(沈阳化工大学信息工程学院, 辽宁 沈阳 110142)

摘要: 针对间歇过程数据具有非线性和多工况的特点, 提出一种基于测地线距离统计量 (geodesic distance statistic, GDS) 的监测方法。首先, 对多工况间歇过程数据按批次方向展开及标准化, 利用主元分析 (principal component analysis, PCA) 方法进行降维; 然后, 在降维空间获得赋权邻接矩阵, 提出采用改进的 Dijkstra (improved Dijkstra, IDijkstra) 算法使 Dijkstra 算法更易于实现, 计算各批次之间的测地线距离, 用以表征非线性多工况数据之间的实际最短距离, 更好地体现批次数据之间的局部近邻关系。通过构造测地线距离 α 次方统计量 D^α 进行过程监测, 与欧氏距离平方和 D^2 相比将减小边缘训练数据距离的偏离程度。最后, 通过在数值仿真和工业仿真实例中的应用, 验证所提算法的有效性。

关键词: 多工况; 非线性; 间歇过程; 测地线距离; 算法

DOI: 10.11949/j.issn.0438-1157.20141439

中图分类号: TP 277

文献标志码: A

文章编号: 0438—1157 (2015) 01—0291—08

Multimode batch process monitoring based on geodesic distance statistic

GUO Xiaoping, LI Ting, LI Yuan

(Institute of Information and Engineering, Shenyang University of Chemical and Technology, Shenyang 110142, Liaoning, China)

Abstract: Process monitoring based on Geodesic Distance Statistic (GDS) is proposed in this article for that fault monitoring method based on Euclidean distance of k Nearest Neighbors (kNN) could not fully reflect the complex characteristics between data with multiple conditions. To start, the batch process data is expanded and standardized by the batch direction. Principal Component Analysis (PCA) is utilized for data dimensionality reduction. Next, Get empowered adjacency matrix in the reduced space. Improved Dijkstra (IDijkstra) algorithm is proposed based on Dijkstra algorithm for easier implement. It can better characterize the actual shortest distance of the nonlinear data and reflect the local neighborhood relations between batch data. Meanwhile, statistics D^α based on α power of Geodesic distance which could reduce the deviation of distance from the edge of the training data is structured for fault monitoring compared with D^2 based on quadratic sum of Euclidean distance. Finally, the effectiveness of the proposed algorithm is verified by applying it in numerical simulation and industry examples.

Key words: multiple conditions; nonlinear; batch process; geodesic distance; algorithm

2014-09-24 收到初稿, 2014-10-08 收到修改稿。

联系人: 李元。第一作者: 郭小萍 (1972—), 女, 博士研究生, 副教授。

基金项目: 国家自然科学基金重点项目 (61034006); 国家自然科学基金面上项目 (60774070, 61174119); 辽宁省教育厅科学研究一般项目 (L2013155); 辽宁省博士启动基金项目 (20131089)。

Received date: 2014-09-24.

Corresponding author: LI Yuan, li-yuan@mail.tsinghua.edu.cn

Foundation item: supported by the Key Program of National Natural Science Foundation of China (61034006), the General Program of National Natural Science Foundation of China (60774070, 61174119), the General Research Project of Department of Education of Liaoning Province (L2013155) and the Liaoning Province Doctor Startup Fund (20131089).

引 言

随着现代社会对多品种、多规格和高质量产品更迫切的市场需求,工业生产更加倚重于生产小批量、高附加值产品的间歇过程。在间歇生产过程中,原料组分、生产负荷、产品特性等因素的改变,过程数据呈现多工况特征。目前,多工况间歇过程的性能监测与故障诊断技术正日益受到工业界和学术界的关注和重视^[1-2]。

PCA 方法^[3-4]是目前广泛应用于过程监测的数据驱动方法,MacGregor 等^[5]提出的多向主元分析(multiway principal component analysis, MPCA)是 PCA 在间歇过程监测中的应用,表明了 MPCA 能够很好地处理间歇过程数据,达到过程检测和诊断的目的。但 MPCA 在进行故障检测时假设过程数据服从正态分布且具有线性关系,不适用于多工况非线性过程,这在一定程度上限制了 MPCA 的应用范围^[5]。核主元分析(kernel principal component analysis, KPCA)方法^[6-7]是常用的非线性分析方法,基于核进行非线性数据映射的分析方法,避免了对非线性映射具体形式的求取,只需定义特征空间中的内积即可。但这些方法中核函数及其中参数的选取目前还没有公认有效的方法,限制了这些算法在更大范围的应用。He 等^[8-9]提出基于欧氏距离 k 近邻的 kNN 算法,能够比较有效地实现复杂批次过程故障检测。该算法中局部近邻的选取以最短欧氏距离为指标,数据空间中两点之间的欧氏距离是这两点连线的直线距离,不能准确描述非线性数据的复杂位置关系。测地线^[10-11]是欧氏空间中的直线段在黎曼流形上的推广,它是黎曼流形上的一条光滑曲线,在局部范围内测地线总是最短线,能够很好地反映两个样本点在该流形上的几何特性,因此在物理学、工程技术学、大地测量学和统计学等领域中越来越受到重视并得以广泛研究和应用^[12-14]。在间

歇过程故障监测中,采用测地线计算批次样本之间的最短距离,能够准确地体现具有非线性特点的多工况数据在空间上的复杂位置关系,为局部近邻的选取提供准确的距离指标。

本文提出一种基于 GDS 的多工况间歇过程监测方法,通过构造邻接图来表示多工况数据的局部几何结构,在一定假设条件下确定数据之间的相似性关系,构建全局模型进行故障监测。首先,将按批次展开和标准化的多工况过程数据采用 PCA 方法进行降维;然后,提出采用 IDijkstra 算法代替 Dijkstra 算法^[15],获得赋权邻接矩阵,计算各批次之间的最短测地线距离,通过构建统计量 D^a 进行故障监测;最后,在人工三臂螺旋数值仿真和半导体工业实例中验证所提算法的有效性。

1 基于测地线距离统计量的间歇过程监测

1.1 基于 IDijkstra 算法的测地线距离

测地线距离是黎曼流形上的一条光滑曲线,对于具有非线性特点的批次过程数据,欧氏距离把两数据点之间的直线距离作为两点之间的最短距离,然而它们沿着非线性流形的距离远远大于它们的输入距离(图 1)^[16]。对于具有强非线性流形结构的 swiss-roll 数据集,从图 1(a)可以看出,高维空间中实线表示的两点之间最短测地线距离远大于虚线表示的欧氏距离;从图 1(c)可以看出,高维样本在低维空间投影后,红线表示的是最短测地线距离与蓝线表示的欧氏距离相比,能够更准确地体现数据间的非线性位置关系。GDS 的主要思想就是首先使用最近邻图上的最短路径得到近似的测地线距离,代替不能表示数据内部非线性流形结构的欧氏距离。测地线距离的计算一般采用 Dijkstra 算法。

Dijkstra 算法^[17]在图论中得到广泛应用,可以

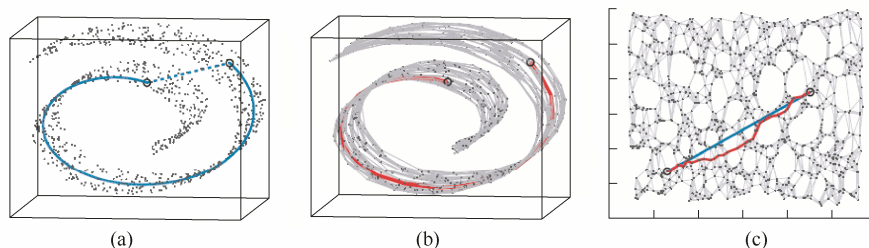


图 1 测地线距离

Fig.1 Schematic diagram of geodesic distance

用来计算加权图里的两个指定顶点 u_0 与 v_0 之间的最短距离, 也可计算 u_0 到所有顶点之间的最短距离。已知图 $G(V, E)$ 及每条边的权 $w(e)$, 对于任意指定的二点 $u_0, v_0 \in V(G)$, 寻找路 $P(u_0, v_0)$, 使得

$$w(P(u_0, v_0)) = \min_{P \in \Omega} \{w(P)\} \quad (1)$$

其中, Ω 是从 u_0 到 v_0 所有路的集合, $w(P)$ 是路 P 上的各边权之和。这样的路 $P(u_0, v_0)$ 称为从 u_0 到 v_0 的最短路, $w(P(u_0, v_0))$ 称为从 u_0 到 v_0 的最短距离。主要方法步骤如下。

① 置 $l(u_0)=0$; 对 $v \neq v_0$, 置 $l(v)=\infty$; $S_0=\{u_0\}$; $i=0$ 。

② 对每个 $v \notin S_i$, 用 $\min\{l(v), l(u_i)+w(u_i, v)\}$ 代替 $l(v)$ 。计算 $\min_{v \notin S_i} \{l(v)\}$, 并把达到这个最小值的一个顶点记为 u_{i+1} , 置 $S_{i+1} = S_i \cup \{u_{i+1}\}$ 。

③ 若 $i=|V|-1$, 则停止; 若 $i < |V|-1$, 则用 $i+1$ 代替 i , 并返回步骤②进入下一次迭代。

本文对该方法进行改进提出了 IDijkstra 算法, 通过计算各个批次数据之间的邻接关系权值构造邻接加权图, 邻接加权图上的每个顶点表示多工况数据的每个批次, 得到任意两批次之间的最短距离矩阵, 使其更容易在计算机上进行矩阵运算。具体步骤为: 将已知图 $G(V, E)$ 存储在赋权邻接矩阵 $W(w_{ij})_{n \times n}$ 里, 其中, n 为批次数, x_i 表示各批次对应的数据点

$$w_{ij} = \begin{cases} 0 & \text{当 } i=j \text{ 时} \\ d(x_i, x_j) & \text{当 } i \neq j, \text{ 顶点 } x_i \text{ 与 } x_j \text{ 为 } h \text{ 近邻时} \\ \infty & \text{当 } i \neq j, \text{ 顶点 } x_i \text{ 与 } x_j \text{ 不为 } h \text{ 近邻时} \end{cases} \quad (2)$$

可求出点 x_i ($i=1, 2, 3, \dots, n$) 到其他各点之间的最短距离, 其中 $d(x_i, x_j)$ 为两批次 x_i 与 x_j 之间的欧氏距离; h 为邻接加权图的近邻个数; 将 Dijkstra 算法运用在 W 矩阵的第 t 行进而得到各批次数据点之间的测地线最短距离矩阵 D 。

1.2 多工况批次数据预处理

在进行 GDS 过程监测之前, 需要对多工况间歇过程数据进行预处理, 如图 2 所示。首先, 将每个工况的三维数据 $\hat{X}_c (I \times J \times K)$ 按照批次方向展开成二维数据 $X_c (I \times JK)$, 其中 I 表示批次, J 表示变量, K 表示采样时刻。然后, 将展开后的各工况数据按照 $\bar{X} = [X_1; X_2; \dots; X_C]$ 的方式进行组合, 形成全局的建模数据 $\bar{X} (cN_c \times JK)$ 。用 Z-score 方法对 $\bar{X}$ 进行标准化, 形成零均值方差为 1 的整体建模数据。

1.3 测地线距离统计量的监测方法

基于测地线距离统计量 GDS (geodesic distance statistic) 的故障监测方法由两部分构成: 模型建立和故障监测, 如图 3 所示。

(1) 模型建立主要步骤

① 多工况批次数据预处理, 产生全局建模数据, 并利用 PCA 方法进行降维, 降维后的矩阵为

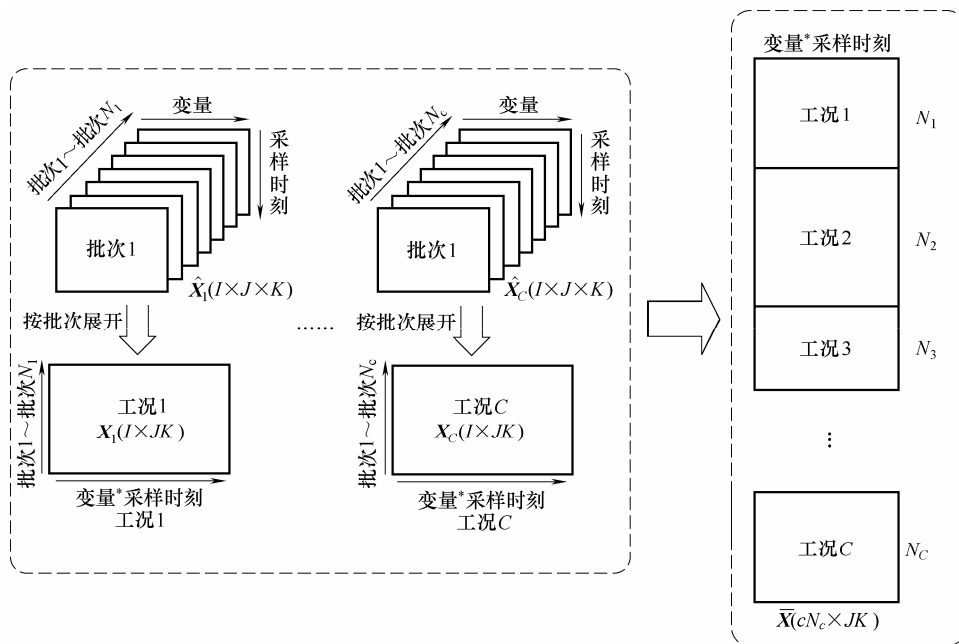


图2 多工况批次过程数据预处理

Fig.2 Pretreatment of multimode batch process data

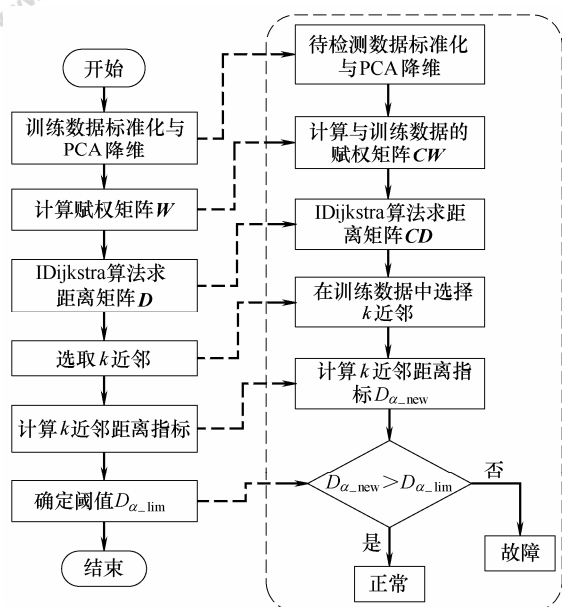


图 3 基于 GDS 的过程监测方法

Fig.3 Process monitoring based on GDS

$X(m \times l)$, 其中, m 为批次数, l 为降维后的维数。

② 计算 X 各数据点之间的欧氏距离 $d(x_i, x_j)$, 利用式 (2) 计算各数据点之间的权重, 求得赋权邻接矩阵 $W(w_{ij})$ 。

③ 利用 IDijkstra 算法计算各批次之间的最短测地线距离矩阵 D 。根据 D 选择各批次数据点的 k 近邻批次。

④ 由 $D_i^\alpha = \sum_{j=1}^k d_{ij}^\alpha$ 计算 k 近邻的 D^α 统计量。其

中, α 为距离的偏离系数。利用核密度估计方法, 计算 D^α 在置信度为 95% 下的控制阈值 D_{α_lim} 。

(2) 故障监测主要步骤

① 对于新来的一批待检测样本, 数据预处理后按批次展开, 利用建模数据的均值和方差进行标准化, 然后用建模的负载矩阵进行 PCA 降维, 降维后的数据用 $x_{new}(1 \times l)$ 表示。

② 计算 x_{new} 与每一批建模数据的欧氏距离 $d_{new}(x_{new}, x_i)$, 组成新的距离矩阵 D_{new} 是 $m+1$ 维的方阵。利用式 (2) 计算 $m+1$ 个点之间的权重, 求得赋权邻接矩阵 CW 。

③ 利用 IDijkstra 算法计算 $m+1$ 个点之间的测地线最短距离矩阵 CD 。根据最短距离 CD 选择各数据点的 k 近邻。

④ 由 $D_{new}^\alpha = \sum_{j=1}^k d_{newj}^\alpha$ 计算待检测批次 k 个近邻的距离指标 D_{α_new} , 并与阈值 D_{α_lim} 比较。当

$D_{\alpha_new} < D_{\alpha_lim}$ 时, 待检测批次是正常的; 否则, 待检测批次是故障的。

2 仿真实例

2.1 数值仿真例子

人工三臂螺旋的每一个臂作为一个工况的数据, 共产生 3 类工况的建模和检测数据。螺旋一个臂表示为

$$\begin{cases} \xi = x \sin(x) + a\varepsilon_e \\ \tau = x \cos(x) + a\varepsilon_e \\ \zeta = \xi\tau \end{cases} \quad (3)$$

其中, ξ 、 τ 、 ζ 为原始数据空间 3 个方向的维度坐标; x 是服从 $\left[\frac{\pi}{8}, \pi\right]$ 的均匀分布; ε_e 是服从 $N(0,1)$ 的噪声; a 是决定噪声幅度的参数, 建模时 $a=0.6$; 由式 (3) 得到人工三臂螺旋一个臂的单工况数据, 然后乘以一个旋转角为 $\left[0, \frac{7\pi}{10}, \frac{7\pi}{5}\right]$ 的旋转矩阵得到另外两个臂的多工况数据。

正常条件下采集螺旋每个臂的 351 个样本作为一个工况的数据, 3 个工况共产生 1053 个采样作为建模数据, 进行全局建模。通过调节噪声幅度 $a=12$ 来产生故障数据, 每个臂产生 21 个故障数据, 3 个工况共产生 63 个采样作为故障监测数据。数据分布如图 4 所示。

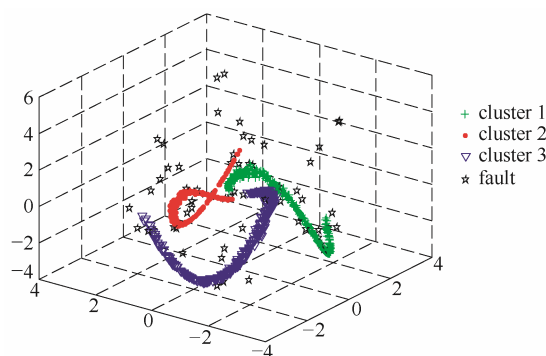


图 4 人工三臂螺旋的三维空间图

Fig.4 Three-dimensional map of artificial three-arm spiral

在该数值例子中, $p=3, h=10, k=3, \alpha=0.7$ 。分别用 KPCA、kNN、PC-kNN^[18] 以及所提的 GDS 算法对数值例子进行检测, 对比效果如图 5 所示, 各种方法检测结果对比如表 1 所示, 3 个工况数据各方法检测结果对比如表 2~表 4 所示。

从表 1 可以看出, 对于具有多工况特点的三臂螺旋数值仿真, KPCA 算法检测效果不好, 这与算

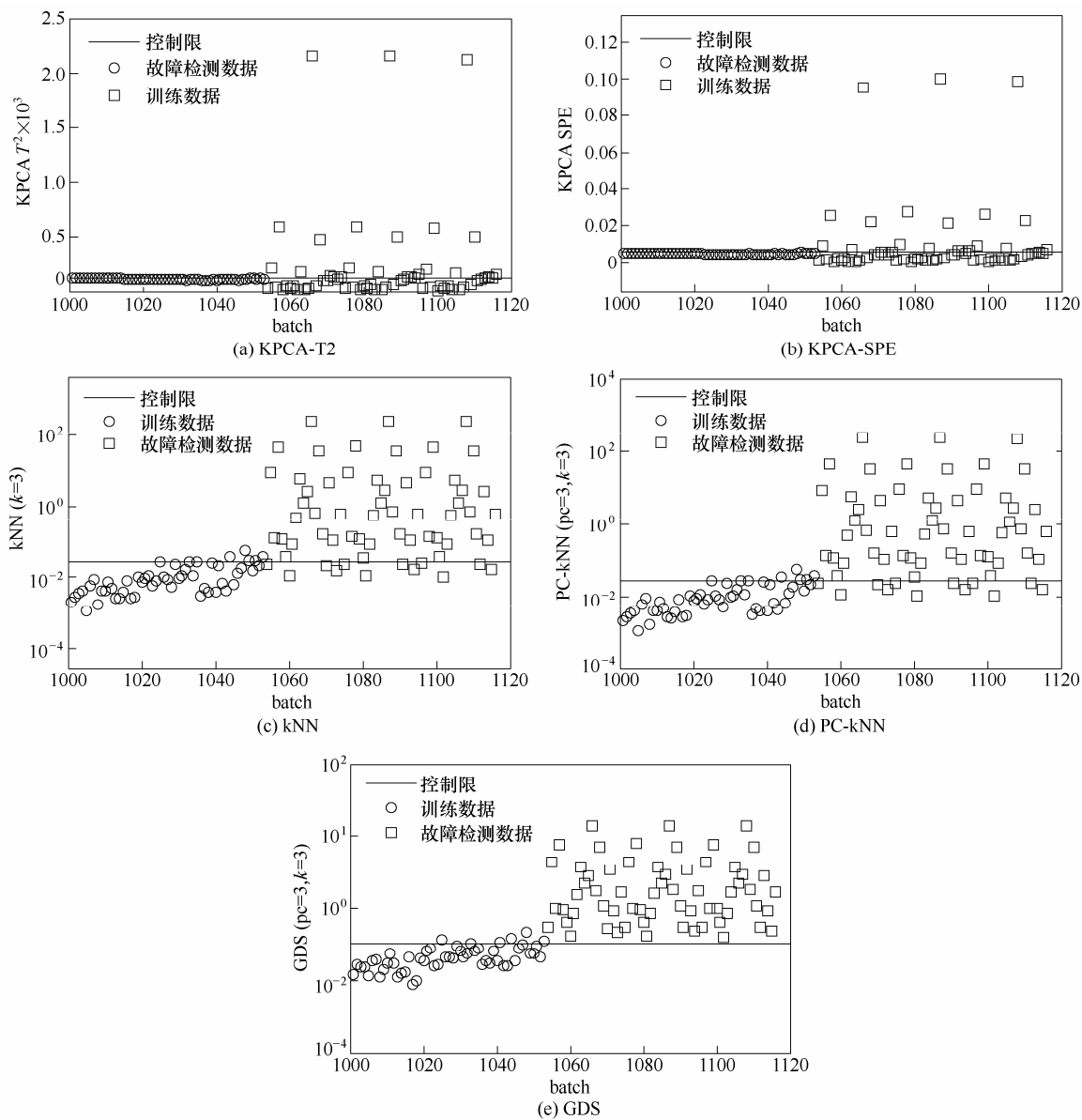


图 5 基于不同方法的过程监测

Fig.5 Process monitoring based on different methods

表 1 数值例子各方法检测对比

Table 1 Detection contrast of numerical example with different methods

方法	被检测出的故障数（共 63 个）	检测率
KPCA	20	32%
kNN	51	81%
PC-kNN	51	81%
GDS	63	100%

表 2 工况 1 检测结果对比

Table 2 Test results contrastion of mode 1

item	检测故障数（共 21 个）	检测率
KPCA T^2	6	28.5%
SPE	7	33.3%
kNN	17	80.9%
PC-kNN	17	80.9%
GDS	21	100%

法中核函数及其参数的选择有很大关系；基于欧氏距离 k 近邻的 kNN 和 PC-kNN 算法能够检测出大部分的故障数据，但三臂螺旋正常样本边缘的故障样

本检测效果不好；而本文提出的 GDS 算法能够检测出全部的 63 个故障样本，数据点之间测地线距离的计算更好地体现了数据之间的位置关系，与 D^2

表 3 工况 2 检测结果对比

Table 3 Test results contrastion of mode 2

方法	检测故障数（共 21 个）	检测率	
KPCA	T^2	7	33.3%
	SPE	8	38%
kNN		17	80.9%
PC-kNN		17	80.9%
GDS		21	100%

表 4 工况 3 检测结果对比

Table 4 Test results contrastion of mode 3

方法	检测故障数 (共 21 个)	检测率	
KPCA	T^2	7	33.3%
	SPE	6	28.5%
kNN		17	80.9%
PC-kNN		17	80.9%
GDS		21	100%

统计量相结合增大了建模数据边缘正常样本与故障样本的差异,检测效果明显提高。从表 2~表 4 可以看出,在三臂螺旋数据的多工况数值仿真中,GDS 算法对各个工况的故障敏感度都很高,检测效果非常明显。

2.2 半导体工业过程应用

半导体工业数据来源于在 Lam 9600 上进行的半导体铝蚀反应^[19-20],是包括多个工况的间歇生产过程。本文选取半导体 3 个工况的数据进行多工况故障监测,从 40 个测量变量中选择 17 个变量作为监测变量。采集 107 个批次的正常数据,95 批次用来建模,其中 1~30 批次取自第 1 个工况,31~62 批次取自第 2 个工况,63~95 批次取自第 3 个工况;12 批次用来验证建模的准确性,其中 1~4 批取自第 1 个工况,5~8 批取自第 2 个工况,9~12 批取自第 3 个工况。同时,通过人为改变 TCP 功率、RF 功率等参数分别采集来自 3 个工况 20 批次的故障数据,其中 1~6 批次属于工况一故障,7~13 批次属于工况二故障,14~20 批次属于工况三故障。通过 95 个 3 个工况混合批次的数据进行全局建模,利用 3 个工况 12 批正常数据检验建模准确性,最终验证各工况故障批次能否及时准确地检测出来。

在本例中,PCA 降维的维数 $l=3$,邻接加权图的近邻数 $h=10$; D^α 统计量的近邻数 $k=3$,偏离系数 $\alpha=0.7$ 。将所提 GDS 算法与 KPCA、kNN 以及 PC-kNN 算法对半导体工业过程进行故障监测,对比结果如图 6 所示。各种方法检测结果对比如表 5 所示。

表 5 不同方法对半导体 3 个工况 20 种故障的检测对比

Table 5 Detection contrast of 20 semiconductor faults from 3 modes with different methods

工况	故障	KPCA T^2	KPCA SPE	kNN	PC-kNN	GDS
1	1	×	√	√	√	√
1	2	×	√	√	√	√
1	3	×	×	×	×	√
1	4	√	√	√	√	√
1	5	×	×	×	√	√
1	6	×	×	×	√	√
2	7	√	√	√	√	√
2	8	×	√	√	√	√
2	9	×	×	×	×	√
2	10	×	×	√	√	√
2	11	×	√	√	×	√
2	12	√	√	√	√	√
2	13	√	√	√	√	√
2	14	×	√	√	×	√
3	15	×	√	√	√	√
3	16	×	√	√	√	√
3	17	×	√	√	√	√
3	18	√	√	√	√	√
3	19	√	√	√	√	√
3	20	√	√	√	√	√
检测数		7	15	16	16	20
检测率		35%	75%	80%	80%	100%

从表 5 可以看出,对于具有 3 个工况的半导体间歇过程数据,基于核的 KPCA 算法故障监测效果不理想, T^2 统计量只能检测出 7 个故障批次,出现这种情况的原因是 KPCA 算法中核函数参数的选择没有确定的标准,在多次试验选优的情况下检测效果仍不理想。基于欧氏距离 k 近邻的 kNN 算法和 PC-kNN 算法能检测出 16 个故障批次,对于与边缘建模样本距离很近的故障数据检测效果不佳。而本文提出的 GDS 算法检测出全部 20 个故障,检测率达到 100%。利用测地线距离准确描述了数据之间的非线性位置关系,构造 D^α 统计量减小了边缘建模数据的偏离程度,增大了故障数据与边缘建模数据的差异,提高了控制阈值的准确度,使得其他 3 种方法无法检测出来的第 3 和第 9 个微小故障也能成功检测出来,故障监测效果明显提高。

3 结 论

本文提出了一种基于测地线距离统计量的多工况间歇过程监测 GDS 算法。该算法通过改进的 IDijkstra 算法获得各数据点之间的最短测地线距离,准确描述了多工况数据的非线性位置关系;通

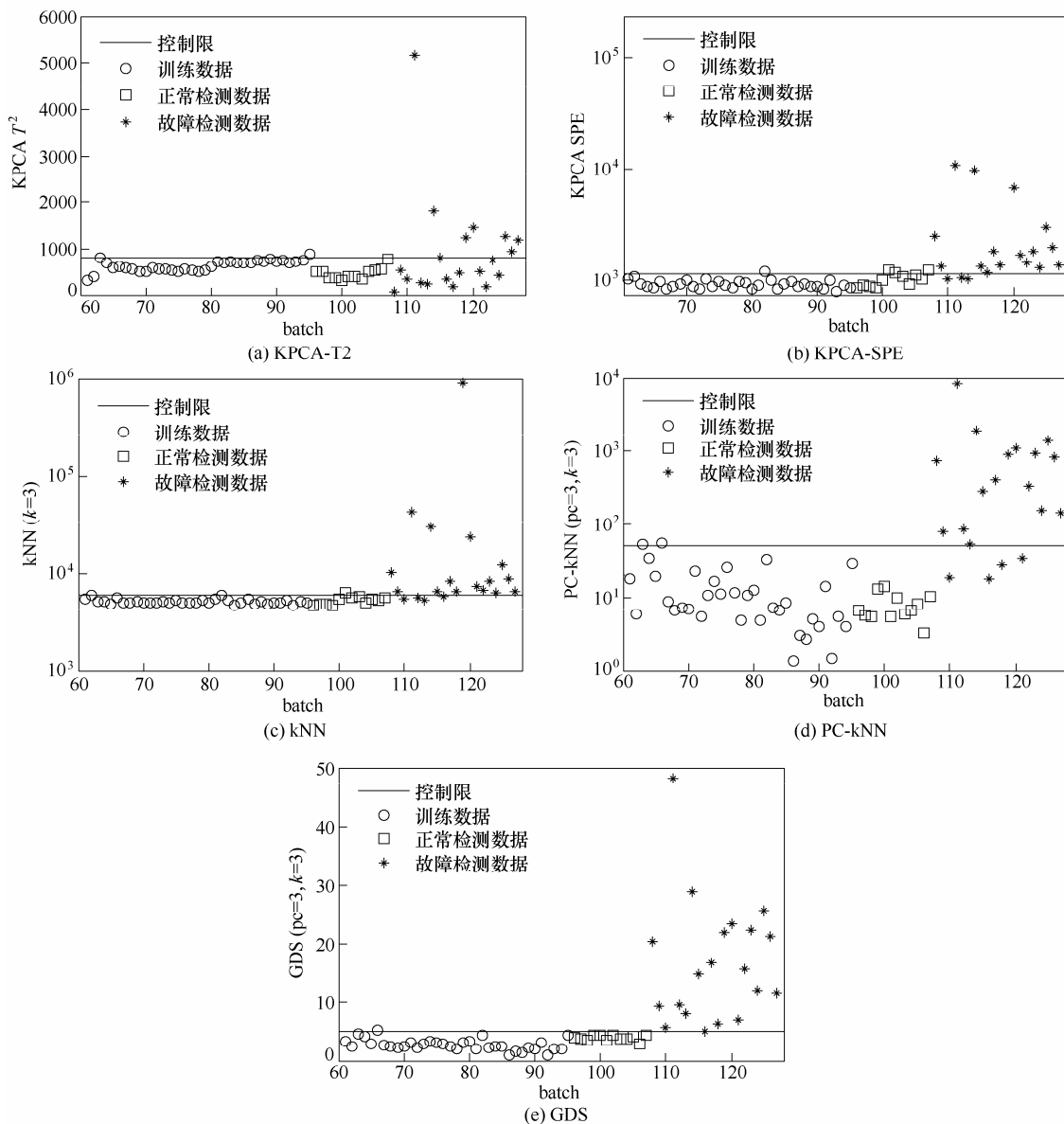


图6 基于不同方法的过程监测

Fig.6 Process monitoring based on different methods

过构造故障监测统计量 D^a 减小了边缘训练数据的偏离程度, 提高了控制阈值的准确度。利用人工三臂螺旋数值仿真和半导体工业过程应用, 验证了所提 GDS 算法在故障检测中的突出优越性。但本文偏离系数的选择对实验结果有一定影响, 选择经验值不具有普遍性, 因此偏离系数的选取原则仍待继续研究。

符号说明

a —— 决定噪声幅度的参数
 D, D^a —— 分别为算法模型的测地线距离矩阵, GDS 统计量

$D_{\alpha_{lim}}$ —— GDS 统计量的控制阈值
 $D_{\alpha_{new}}$ —— 待检测批次数据的统计量
 $d(x_i, x_j)$ —— i 和 j 数据点之间的欧氏距离
 h —— 邻接加权图的近邻数
 k —— D^a 统计量的近邻数
 m, l —— 分别为批次数, 降维后数据维数
 $P(u_0, v_0)$ —— 邻接图中任意两点之间的路径
 p —— 主元个数
 u_0, v_0 —— 邻接图中任意两个顶点号
 $W(w_{ij})_{n \times n}$ —— 赋权邻接矩阵
 $w(e)$ —— 邻接图中各顶点之间的权值
 x —— 服从 $\left[\frac{\pi}{8}, \pi\right]$ 的均匀分布

x_i, x_j ——第 i 和 j 个数据点
 α ——距离的偏离系数
 ε_e ——服从 $N(0,1)$ 的噪声
 ξ, τ, ζ ——原始数据空间 3 个方向的维度坐标

References

- [1] Wang Shu. A study of batch process fault diagnosis and prediction methods based on data [D]. Shenyang: Northeastern University, 2010
- [2] Chen Yong(陈勇). Fault diagnosis of the production process based on multiway principal components analysis[D]. Hangzhou: Zhejiang University, 2003
- [3] Zhou Donghua(周东华), Li Gang(李钢), Li Yuan(李元). Industrial Process Fault Diagnosis of Data-Driven(数据驱动的工业过程故障诊断技术)[M]. Beijing: Science Press, 2011
- [4] Wang Haiqing(王海清), Song Zhihuan(宋执环), Wang Hui(王慧). Fault detection behavior analysis based on PCA process monitoring [J]. *Journal of Chemical Industry and Engineering (China)*(化工学报), 2002(3): 297-301
- [5] Nomikos P, MacGregor J F. Monitoring batch processes using multiway principal components analysis [J]. *AIChE Journal*, 1994, **40**(8): 1361-1375
- [6] Scholkopf B, Smola A, Muller K R. Nonlinear component analysis as a kernel eigenvalue problem [J]. *Neural Computation*, 1998, **10**(5): 1299-1319
- [7] Guo Xiaoping(郭小萍), Yuan Jie(袁杰), Li Yuan(李元). Nearest neighbors index of kernel features based batch processes monitoring and simulation research [J]. *Journal of System Simulation*(系统仿真学报), 2014, **26**(7): 1424-1429
- [8] He Q P, Wang J. Fault detection using the k-nearest neighbor rule for semiconductor manufacturing processes [J]. *IEEE Transactions on Semiconductor Manufacturing*, 2007, **20**(4): 345-354
- [9] Guo Xiaoping(郭小萍), Yuan Jie(袁杰), Li Yuan(李元). Batch process monitoring based on k-nearest neighbor of spatial [J]. *Journal of Automation*(自动化学报), 2014, **40**(1): 135-142
- [10] Bai Zhengguo(白正国), Shen Yibing(沈一兵), Shui Naixiang(水乃翔), et al. Riemannian Introduction(黎曼几何引论)[M]. Beijing: Higher Education Press, 2004
- [11] Jurgen Jost. Riemannian Geometry and Geometric Analysis[M]. Berlin: Springer-Verlag, 2002
- [12] Quan Yong, Yang Jie. Geodesic distance for support vector machines [J]. *Acta Automatic Sinica*, 2005, **31**(2): 202-208
- [13] Yang Qing(杨庆), Chen Guiming(陈桂明), Liu Jingjie(刘靖洁), He Qingfei(何庆飞). Kernel principal components analysis based on geodesic distance and its application in gear pump fault recognition [J]. *Journal of Shanghai Jiao Tong University*(上海交通大学学报), 2011, **45**(11): 1632-1637
- [14] Wu Xiaoting(吴晓婷). Data dimensionality reduction algorithm research based on manifold learning[D]. Dalian: Liaoning Normal University, 2010
- [15] Zhang Fuhao(张福浩), Liu Jiping(刘记平), Li Qingyuan(李青元). Shortest path optimization algorithm based on Dijkstra algorithm [J]. *Remote Sensing Information*(遥感信息), 2004, **2**: 38-41
- [16] Li Chenbo(李臣波), Liu Rentao(刘润涛). A shortest path algorithm based on Dijkstra [J]. *Journal of Harbin University of Science and Technology*(哈尔滨理工大学学报), 2008, **13**(1): 35-37
- [17] Su Baoli(苏宝丽), Li Ning(李宁). The optimization algorithm of Dijkstra and its application for the best path in GIS system [J]. *Remote Sensing Technology and Application* (遥感技术与应用), 2013, **28** (5): 866-870
- [18] He Q P, Wang J. Principal component based k-nearest-neighbor rule for semiconductor process fault detection//Proceedings of the 2008 American Control Conference[C]. Seattle, WA: IEEE, 2008: 1606-1611
- [19] Yu J. Multiway discrete hidden Markov model-based approach for dynamic batch process monitoring and fault classification [J]. *American Institute of Chemical Engineers Journal*, 2012, **58**(9): 2714-2725
- [20] Lee S P, Chao A K, Tsung F, Wong D S H, Tseng S T, Jang S S. Monitoring batch processes with multiple on-off steps in semiconductor manufacturing [J]. *Journal of Quality Technology*, 2011, **43**(2): 142-157