引用格式:赵成,苏圣超. 基于海林格距离加权关键主元的流程工业故障检测研究[J]. 北京化工大学学报(自然科学版),2022,49(3):91 - 101.

ZHAO Cheng, SU ShengChao. A fault detection method based on Hellinger distance-weighted key principal components for the process industry[J]. Journal of Beijing University of Chemical Technology (Natural Science), 2022,49(3):91 - 101.

基于海林格距离加权关键主元的流程工业故障检测研究

赵 成 苏圣超*

(上海工程技术大学 电子电气工程学院, 上海 201620)

摘 要: 在采用主成分分析(principal component analysis, PCA)算法进行故障检测时,主元的选取及处理直接影响其故障检测的表现。对此,提出一种基于全变量表达(full variable expression, FVE)和海林格距离(Hellinger distance, HD)的故障检测方法。首先,利用 FVE 得到所有关键主元,即保留所有变量信息;然后考虑到与故障相关主元的重要性,定义基于海林格距离的变化率,用来衡量正常工况下主元与异常工况下主元的差异;对与故障发生更相关的主元进行加权,以突出与故障相关主元对于后续故障检测的影响;最后,考虑到降维后数据通常服从非高斯分布,利用改进的局部离群因子(local outlier factor, LOF)构建统计量,其相应控制限通过核密度估计(kernel density estimation, KDE)确定。数值实例及带钢热连轧实际生产数据验证了所提方法的有效性与优越性。

关键词: 故障检测; 主元分析法; 关键主元; 海林格距离; 局部离群因子

中图分类号: TP277 **DOI**: 10.13543/j.bhxbzr.2022.03.013

引 言

随着现代工业的快速发展,工业生产的安全性与可靠性变得尤为重要。近年来,以故障检测为主导的过程监测技术与方法成为学术界和工业界的研究热点。目前的故障检测方法主要分为以下 3 类:基于数学模型的故障检测方法,基于定性知识的故障检测方法和基于数据驱动的故障检测方法。其中基于数据驱动的故障检测方法与基于数学模型的故障检测方法相比,不需要精准的系统模型,同时随着传感器技术的飞速发展,海量数据的获得也变得更加容易,因此基于数据驱动的故障检测方法扮演着愈加重要的角色[1]。主成分分析(principal component analysis, PCA)作为一种基于数据驱动的故障检测算法,由于其简便的算法流程及对高维数据的高效处理能力而受到广泛的关注与研究[2],相关的拓展算法包括概率 PCA (probability PCA)[3]、核 PCA (kernel PCA)[4]、动态 PCA (dynamic PCA)[5]等。

尽管对于 PCA 算法的各种改进有很多,但针对其中主元的选取及后续的处理问题仍然需要更加深入的研究。传统的主元选取方法有累计方差贡献法(cumulative percent variance, CPV)[6],重构误差方差法(variance of the reconstruction error, VRE)[7]和平均特征值法(average eigenvalue, AE)[8]等。这些方法都是认为较大方差所对应的主元包含更多的信息,而较小的方差所对应的主元则常常被忽视。Jolliffe[9]提出具有较小方差的主元和具有较大方差的主元同等重要,Togkalidou 等[10]指出具有较大方差的主元不一定具有最多的信息量。因此,仅凭方差的贡献度大小来确定主元的方法过于主观机械,有可能造成有用信息的丢失。同时,传统主元选取方法根据正常工况数据进行线下建模,没有考虑故障样本对建模的影响。上述问题都会导致故障检测性能的大幅下降。为此,陶阳等[11]将 ReliefF 算法与

收稿日期: 2021-10-21

基金项目: 国家自然科学基金(61603241);上海工业控制系统安全创新功能型平台开放课题项目(TICPSH202103003-ZC)

第一作者: 男,1996 年生,硕士生

* 通信联系人

E-mail: jnssc@sues.edu.cn

PCA 算法相结合,从故障特征角度出发,避免了主元选取时的主观性。Jiang 等[12]统计单个主元 T^2 统计量的变化率,通过选取与监测敏感主元进行故障检测,但在选取主元阶段仍采用传统的 CPV 方法,导致有用信息丢失,从而影响检测效果。仓文涛等[13]通过构造累计 T^2 统计量的变化率来体现主元的变异程度,但是传统 T^2 统计量要求数据变量服从正态分布,而实际工业流程中采集到的数据显然很难满足此限制条件。同时从几何角度来看,T^2 统计量实质上是一个椭圆形控制边界,误差较大。Song 等[14]提出全变量表达(full variable expression, FVE)的主元选取方法,选取对各个变量解释性最大的主元作为关键主元,保留了所有变量信息,取得了良好的故障检测效果。但该方法仍然是依据 T^2 统计量的形式构造相应的统计量进行检测;其次,上述所有方法都是同等看待被选主元。实际上,故障发生时只有某些主元含有与故障相关的重要信息,这些主元在后续处理中需要加以突出。海林格距离(Hellinger distance, HD)是信息论中的一个概念,在统计学和概率学中用来衡量两个概率密度分布间的差异。在故障检测领域,Jiang 等[15]将 HD 用于变量块的划分,将拥有相似概率密度分布的变量归结到同一个分块进行故障检测;Harrou 等[16]将 HD 作为统计量,用来监测非线性投影产生残差的概率密度分布;Chen 等[17]将 HD 用于高速列车的故障检测,相关实验表明 HD 作为散度的一种,有着比 Kull-back-Leibler(K-L)散度更好的检测效果。

局部离群因子(local outlier factor, LOF)由 Breunig 等[18]提出,用于搜寻数据中的离群点。该方法利用样本点的局部离群因子值的大小来衡量其离群程度,对于样本数据的分布没有要求。故障样本相对于正常样本即可视为离群点,因此该算法已开始被研究者运用到故障检测领域。Lee 等[19]将 LOF 算法与独立元分析法相结合,用于解决数据高斯与非高斯混合分布情况下的故障检测问题,突破了传统独立元分析法对于非高斯分布数据的限制。Deng 等[20]利用双加权策略改进核主元分析法,并将其与 LOF 算法相结合,兼顾了工业过程的非线性特点以及数据的非高斯分布。冯立伟等[21]提出基于时空近邻标准化(time-space nearest neighborhood standardization, TSNS)-LOF 的故障检测方法,提高了对于多模态过程的故障检测能力。由以上文献可以看出,将 LOF 算法融入到故障检测方法中并且构建相应的统计量,有利于提高故障检测能力。

本文在 FVE 方法基础上,利用海林格距离变化率对故障的敏感性提出一种基于全变量表达和海林格距离的故障检测算法(Hellinger distance weighted full variable expression-Mahalanobis distance LOF, HDWFVE-MDLOF)。首先利用 FVE 的主元选取方法有效地保留了全部变量信息,之后提出基于海林格距离变化率的指标,突出故障相关主元;同时,考虑到数据变量的无特定分布以及尺度问题,将加权后的关键主元作为改进 LOF 算法的输入,构造相应统计量进行故障检测。数值仿真实例和带钢热连轧实际生产数据验证了所提方法的有效性与优越性。

1 PCA 算法及海林格距离

1.1 PCA 算法

PCA 是一种用于多变量统计过程监控的基本方法,其具体实施步骤如下。给定一个原始数据矩阵 $\boldsymbol{X}\in\boldsymbol{R}^{n\times m}$,其中 n 为样本数,m 为变量数(传感器个数),其协方差矩阵为

$$\boldsymbol{S}=\frac{1}{n-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} \tag{1}$$

接下来对协方差矩阵进行奇异值分解(SVD)

$$\boldsymbol{S}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{\mathrm{T}} \tag{2}$$

其中,

$$\boldsymbol{P}=[\boldsymbol{P}_{\mathrm{pc}}\quad\boldsymbol{P}_{\mathrm{res}}],\boldsymbol{\Lambda}=\begin{bmatrix}\boldsymbol{\Lambda}_{\mathrm{pc}}&0\\0&\boldsymbol{\Lambda}_{\mathrm{res}}\end{bmatrix} \tag{3}$$

式中,$\boldsymbol{\Lambda}_{\mathrm{pc}}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_l)$,$\boldsymbol{\Lambda}_{\mathrm{res}}=\mathrm{diag}(\lambda_{l+1},\lambda_{l+2},\cdots,\lambda_m)$,$\boldsymbol{P}$ 为负载矩阵,$\boldsymbol{\Lambda}$ 为特征值矩阵;l 作为主元个数将负载矩阵划分成两部分,即主元空间的 $\boldsymbol{P}_{\mathrm{pc}}$ 和残差空间的 $\boldsymbol{P}_{\mathrm{res}}$。$l$ 由累计方差贡献率(CPV)确定

$$\frac{\sum_{i=1}^{l}\lambda_i}{\sum_{i=1}^{m}\lambda_i}\times 100\%\geqslant 85\% \tag{4}$$

根据以上描述,样本 $\boldsymbol{X}\in\boldsymbol{R}^{n\times m}$ 可以被分解为

$$\boldsymbol{X}=\hat{\boldsymbol{X}}+\tilde{\boldsymbol{X}}=\boldsymbol{T}\boldsymbol{P}_{\mathrm{pc}}+\tilde{\boldsymbol{X}} \tag{5}$$

式中,$\hat{\boldsymbol{X}}\in\boldsymbol{R}^{n\times m}$ 为主成分矩阵,$\tilde{\boldsymbol{X}}\in\boldsymbol{R}^{n\times m}$ 为残差矩阵,$\boldsymbol{T}\in\boldsymbol{R}^{n\times l}$ 为得分矩阵。

在 PCA 分解后的两个空间里分别建立 T^2 统计量与平方预测误差(squared prediction error, SPE)统计量

$$T^2 = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{P}_{\mathrm{pc}}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}_{\mathrm{pc}}^{\mathrm{T}}\boldsymbol{X} \tag{6}$$

$$E_{\mathrm{SPE}} = \boldsymbol{X}(\boldsymbol{I} - \boldsymbol{P}_{\mathrm{pc}}\boldsymbol{P}_{\mathrm{pc}}^{\mathrm{T}})\boldsymbol{X}^{\mathrm{T}} \tag{7}$$

T^2 统计量反映的是主元空间的变化，而 SPE 统计量反映的是残差空间的变化。关于 T^2 和 SPE 这两个统计量的具体描述以及相应控制限的设定可以参考文献[22]，两个统计量中只要有任意一个超过其对应控制限，即可认定有故障发生。

1.2　海林格距离

HD 作为一种度量两个概率密度分布差异的工具，已得到广泛应用[23-25]。假设有两个连续的概率密度函数 $f(x)$ 和 $g(x)$，那么 $f(x)$ 和 $g(x)$ 之间的海林格距离可以由式(8)描述。

$$D_{\mathrm{HD}}(f,g) = \sqrt{\frac{1}{2}\int\left(\sqrt{f(x)} - \sqrt{g(x)}\right)^2 \mathrm{d}x} \tag{8}$$

$f(x)$ 和 $g(x)$ 的差异越大，HD 的数值就越大。因此 HD 能够用来衡量正常指标与异常指标之间的差异。同时，海林格距离是一个对称有界的距离，即 $0 \leqslant D_{\mathrm{HD}}(f,g) = D_{\mathrm{HD}}(g,f) \leqslant 1$，根据勒贝格测度[26]可以得到式(8)的平方变形形式

$$D_{\mathrm{HD}}^2(f,g) = 1 - \int \sqrt{f(x)g(x)}\,\mathrm{d}x \tag{9}$$

2　基于马氏距离的局部离群因子

对于任意样本 $\boldsymbol{x}_i(i=1,2,\cdots,n)$，在样本数据集 $\boldsymbol{X}$ 中以欧式距离为度量找寻 $\boldsymbol{x}_i$ 的 k 个近邻点组成局部近邻集 $N(\boldsymbol{x}_i) = \{x_i^1, \cdots x_i^f, \cdots x_i^k\}$，$f=1,2,\cdots,k$，并通过式(10)计算样本 $\boldsymbol{x}_i$ 和它的局部近邻集成员 x_i^f 的马氏距离

$$d_{\mathrm{m}}(\boldsymbol{x}_i, x_i^f) = \sqrt{\{\boldsymbol{x}_i - \mathbf{mean}[N(x_i^f)]\}\boldsymbol{\Sigma}^{-1}\{\boldsymbol{x}_i - \mathbf{mean}[N(x_i^f)]\}^{\mathrm{T}}} \tag{10}$$

式中，$\boldsymbol{\Sigma}$ 为 x_i^f 的局部近邻集 $N(x_i^f)$ 的协方差矩阵，$\mathbf{mean}[N(x_i^f)]$ 为其均值向量。式(10)中利用 $\mathbf{mean}[N(x_i^f)]$ 替代 x_i^f 是为了在正常工况下减少关键主元可能受到的污染的影响[27]。接着对每个样本的近邻按照式(10)进行排序并得到每个样本的领域半径 $k\_\mathrm{distance}_{\mathrm{m}}(\boldsymbol{x}_i)$。

$$d_{\mathrm{m}}(\boldsymbol{x}_i, x_i^1) \leqslant \cdots \leqslant d_{\mathrm{m}}(\boldsymbol{x}_i, x_i^f) \cdots \leqslant d_{\mathrm{m}}(\boldsymbol{x}_i, x_i^k)$$

$$k\_\mathrm{distance}_{\mathrm{m}}(\boldsymbol{x}_i) = d_{\mathrm{m}}(\boldsymbol{x}_i, x_i^k) \tag{11}$$

样本 $\boldsymbol{x}_i$ 与其近邻点 x_i^f 之间的可达距离表示为

$$\mathrm{reach\_}d(\boldsymbol{x}_i, x_i^f) = \max\{k\_\mathrm{distance}_{\mathrm{m}}(\boldsymbol{x}_i), d_{\mathrm{m}}(\boldsymbol{x}_i, x_i^f)\} \tag{12}$$

同时，定义样本 $\boldsymbol{x}_i$ 的局部可达密度为

$$D_{\mathrm{Ird}}(\boldsymbol{x}_i) = \frac{k}{\sum_{f=1}^{k}\mathrm{reach\_}d(\boldsymbol{x}_i, x_i^f)} \tag{13}$$

样本 $\boldsymbol{x}_i$ 的局部离群因子表示为 $\boldsymbol{x}_i$ 所有近邻的平均局部可达密度与样本点 $\boldsymbol{x}_i$ 的局部可达密度的比值

$$R_{\mathrm{LOF}}(\boldsymbol{x}_i) = \frac{1}{k}\sum_{f=1}^{k}\frac{D_{\mathrm{Ird}}(x_i^f)}{D_{\mathrm{Ird}}(\boldsymbol{x}_i)} \tag{14}$$

3　基于 HDWFVE-MDLOF 的故障检测方法

3.1　HDWFVE-MDLOF 算法

$\boldsymbol{X} \in \boldsymbol{R}^{n\times m}$ 经过 PCA 分解后得到得分矩阵 $\boldsymbol{T} = [\boldsymbol{t}_1, \boldsymbol{t}_2, \cdots, \boldsymbol{t}_m] \in \boldsymbol{R}^{n\times m}$，传统 PCA 算法根据正常工况下的样本建立相应的模型，如式(15)所示，主元可由所有变量线性表示

$$\boldsymbol{t}_j = p_{j1}\boldsymbol{x}_1 + p_{j2}\boldsymbol{x}_2 + \cdots + p_{jm}\boldsymbol{x}_m \tag{15}$$

由式(15)可知，负载系数 p_{ji} 反映了主元 $\boldsymbol{t}_j$ 与变量 $\boldsymbol{x}_i$ 的相关性，因此可以定义相关性指标 p_{ji}^2，该指标值越大，说明 $\boldsymbol{t}_j$ 与 $\boldsymbol{x}_i$ 的相关性越大。若 p_{ji}^2 在 $[p_{1i}^2, p_{2i}^2, \cdots, p_{mi}^2]$ 中最大，那么主元 $\boldsymbol{t}_j$ 就被选为关键主元。对于每一个变量，都可以通过上述方法判断获取其对应的关键主元。统计量 T^2 的表达式为

$$T^2 = \boldsymbol{x}^{\mathrm{T}}\overline{\boldsymbol{P}}\boldsymbol{\Lambda}^{-1}\overline{\boldsymbol{P}}^{\mathrm{T}}\boldsymbol{x} \tag{16}$$

式中，$\overline{\boldsymbol{P}} = [\overline{\boldsymbol{p}}_1, \overline{\boldsymbol{p}}_2, \cdots, \overline{\boldsymbol{p}}_f] \in \boldsymbol{R}^{m\times f}$ 为所选关键主元对应的负载向量构成的负载矩阵，$\boldsymbol{\Lambda} = \mathrm{diag}[\lambda_1, \lambda_2, \cdots, \lambda_f] \in \boldsymbol{R}^{f\times f}$ 为关键主元对应的特征值构成的对角矩阵，f 为关键主元的个数。相应的控制限参照 PCA 算法用于故障检测时 T^2 的控制限

$$T_\alpha^2 = \frac{f(n^2-1)}{n(n-f)}F_{f,n-f;\alpha} \tag{17}$$

式中，n 为样本个数，$F_{f,n-f;\alpha}$ 为置信水平 $1-\alpha$ 下、自由度 f 和 $n-f$ 的 F 分布临界值。关键主元的具体描述可以参考文献[14]。

一般情况下，当故障发生时，不同得分向量与故障之间的相关性是存在差异的。得分向量较正常工况变化幅度越大，说明其包含越多的信息量[12]，与故障的发生也越紧密相关。因此，须根据得分向量与故障的相关性给予其不同的权重值。与故障相关性越大的得分向量，对后续故障检测的贡献度也越大，需赋予较大的权重。本文利用基于海林格距离的得分向量变化率对权重进行定义，并且给出了该

指标对于故障敏感性的证明。从故障对得分向量产生的影响角度出发,文献[28]对加性故障和乘性故障有如下定义:加性故障指故障的发生只影响得分向量的均值,乘性故障指故障的发生只影响得分向量的方差。由文献[28]可知,传统 T^2 统计量对于乘性故障的检测效果较差,这是由于 T^2 统计量本身的定义是从加性故障的角度出发的。接下来,给出基于海林格距离变化率的统计量。

假设正常工况下的得分向量为 $\boldsymbol{t}_j$, $j=1,2,\cdots,m$,故障情况下的得分向量为 $\hat{\boldsymbol{t}}_j$,$j=1,2,\cdots,m$,则基于海林格距离的统计量为

$$hd_j^2=1-\int\sqrt{f(\boldsymbol{t}_j)f(\hat{\boldsymbol{t}}_j)}\,\mathrm{d}t \tag{18}$$

式中,$f(\boldsymbol{t}_j)$ 和 $f(\hat{\boldsymbol{t}}_j)$ 为对应的概率密度函数,可以通过核密度估计获得,核函数采用最常用的高斯核函数形式

$$f(t)=\frac{1}{nh\sqrt{2\pi}}\sum_{i=1}^{n}\exp\left(-\frac{(\boldsymbol{t}_j-t_{ij})^2}{2h^2}\right)$$

$$f(\hat{t})=\frac{1}{nh\sqrt{2\pi}}\sum_{i=1}^{n}\exp\left(-\frac{(\hat{\boldsymbol{t}}_j-\hat{t}_{ij})^2}{2h^2}\right) \tag{19}$$

式中,n 为样本数目,h 为平滑参数,t_{ij} 为得分向量 $\boldsymbol{t}_j$ 的第 i 个元素,$\hat{t}_{ij}$ 为得分向量 $\hat{\boldsymbol{t}}_j$ 的第 i 个元素。

若 $f(\boldsymbol{t}_j)\sim(\mu_j,\lambda_j)$,$f(\hat{\boldsymbol{t}}_j)\sim(\widetilde{\mu}_j,\widetilde{\lambda}_j)$,那么由式(18)可得

$$hd_j^2=1-\left(\frac{2\sqrt{\lambda_j\widetilde{\lambda}_j}}{\lambda_j+\widetilde{\lambda}_j}\right)\exp\left(-\frac{(\widetilde{\mu}_j-\mu_j)^2}{4(\lambda_j+\widetilde{\lambda}_j)}\right) \tag{20}$$

文献[17]证明了当故障引起得分向量的均值和方差发生变化后,hd_j^2 与均值和方差的变化量正相关,即 hd_j^2 对故障的发生十分敏感。参考文献[29]中的表示形式,海林格距离变化率可以表示成单个得分向量与所有得分向量平均值的海林格距离二次型与该统计量的平均值之比

$$Rhd_j^2=hd_j^2/\overline{hd_j^2} \tag{21}$$

式中,hd_j^2 为第 j 个得分向量的海林格距离的二次型统计量,$\overline{hd_j^2}$ 为所有得分向量海林格距离的二次型统计量的平均值,其中 $j=1,2,\cdots,m$。多变量系统中的故障往往是由 PCA 模型中个别得分向量的异常引起的,因此,对于一个数据集中的海林格距离平均值,其受异常得分向量的影响很小,故可以把 $\overline{hd_j^2}$ 视为常数 c。为了体现所提基于海林格距离变化率的统计量对于故障发生的敏感性以及对于噪声的鲁棒性,进行如下证明。

假设故障发生时导致得分向量均值的变化为 $\Delta\mu_j$,方差变化为 $\Delta\lambda_j$,当传感器发生加性故障,即故障的发生只改变了得分向量的均值时,有

$$Rhd_j^2=\frac{hd_j^2}{\overline{hd_j^2}}=\frac{1}{c}\times\left(1-\left(\frac{2\sqrt{\lambda_j\widetilde{\lambda}_j}}{\lambda_j+\widetilde{\lambda}_j}\right)^{\frac{1}{2}}\cdot\right.$$

$$\left.\exp\left(-\frac{(\widetilde{\mu}_j-\mu_j)^2}{4(\lambda_j+\widetilde{\lambda}_j)}\right)\right)=\frac{1}{c}\times\left(1-\exp\left(-\frac{\Delta\mu_j^2}{8\lambda_j}\right)\right) \tag{22}$$

将式(22)中的 Rhd_j^2 对 $\Delta\mu_j$ 求导得

$$\frac{\partial Rhd_j^2}{\partial\Delta\mu_j}=\frac{1}{4c\lambda_j}\times\mathrm{e}^{\frac{\Delta\mu_j^2}{8\lambda}}\times\Delta\mu_j \tag{23}$$

式中,当 $\Delta\mu_j>0$ 时,$\frac{\partial Rhd_j^2}{\partial\Delta\mu_j}>0$,即统计量 $Rhd_j^2(\Delta\mu_j)$ 单调递增。

当传感器发生乘性故障,即故障的发生只改变得分向量的方差时,有

$$Rhd_j^2=\frac{hd_j^2}{\overline{hd_j^2}}=\frac{1}{c}\times\left(1-\left(\frac{2\sqrt{\lambda_j(\lambda_j+\Delta\lambda_j)}}{2\lambda_j+\Delta\lambda_j}\right)^{\frac{1}{2}}\right) \tag{24}$$

将式(24)中的 Rhd_j^2 对 $\Delta\lambda_j$ 求导得

$$\frac{\partial Rhd_j^2}{\partial\Delta\lambda_j}=-\frac{1}{c}\times$$

$$\frac{\lambda_j(\lambda_j^2+\lambda_j\Delta\lambda_j)^{-\frac{1}{2}}(2\lambda_j+\Delta\lambda_j)-2(\lambda_j^2+\Delta\lambda_j\lambda_j)^{\frac{1}{2}}}{(2\lambda_j+\Delta\lambda_j)^2} \tag{25}$$

式中,当 $\Delta\lambda_j>0$ 时,$\frac{\partial Rhd_j^2}{\partial\Delta\lambda_j}>0$,即统计量 $Rhd_j^2(\Delta\lambda_j)$ 单调递增。

当传感器上加性故障与乘性故障并存,即故障的发生既改变了得分向量的均值,又改变了得分向量的方差时,有

$$Rhd_j^2=\frac{hd_j^2}{\overline{hd_j^2}}=\frac{1}{c}\times\left(1-\left(\frac{2\sqrt{\lambda_j(\lambda_j+\Delta\lambda_j)}}{2\lambda_j+\Delta\lambda_j}\right)\right)\cdot$$

$$\exp\left(-\frac{\Delta\mu_j^2}{4(2\lambda_j+\Delta\lambda_j)}\right) \tag{26}$$

将式(26)分别对 $\Delta\lambda_j$ 和 $\Delta\mu_j$ 求偏导,由上述两种情况可知,Rhd_j^2 随着 $\Delta\lambda_j$ 和 $\Delta\mu_j$ 的增大而增大。

综合以上 3 种情况的证明可知，基于海林格距离的变化率对于故障的发生十分敏感。另一方面，噪声对于系统的影响也需要考虑。

假设 $\lambda_j = \lambda_j^* + \nu$，其中 λ_j^* 表示无故障无噪声情况下的第 j 个方差，ν 为噪声，对此有

$$Rhd_j^2 = \frac{hd_j^2}{\overline{hd_j^2}} = \frac{1}{c} \times \left(1 - \left(\frac{2\sqrt{(\lambda_j^* + \nu)\widetilde{\lambda}_j}}{\lambda_j^* + \nu + \widetilde{\lambda}_j}\right)^{\frac{1}{2}}\right) \cdot \exp\left(-\frac{(\widetilde{\mu}_j - \mu_j)^2}{4(\lambda_j^* + \nu + \widetilde{\lambda}_j)}\right) \tag{27}$$

对式(27)关于 ν 求偏导可得其 sgn 函数为

$$\operatorname{sgn}\left(\frac{\partial Rhd_j^2(\nu)}{\partial \nu}\right) = \operatorname{sgn}\left(\widetilde{\lambda}_j^2 - (\lambda_j^*)^2 + (\widetilde{\mu}_j - \mu_j)^2\lambda_j - 2\lambda_j^*\nu - \nu^2\right) \tag{28}$$

当 $\nu = 0$ 时，式(28)的最大值为 0，即

$$\begin{cases} \dfrac{\partial Rhd_j^2(\nu)}{\partial \nu} < 0, & \nu \neq 0 \\ \dfrac{\partial Rhd_j^2(\nu)}{\partial \nu} = 0, & \nu = 0 \end{cases} \tag{29}$$

因此，Rhd_j^2 随着噪声 ν 的增大而减小，即 Rhd_j^2 对噪声具有一定的鲁棒性。

综上可知采用 Rhd_j^2 作为统计量对于故障的检测是有效的，该统计量反映了故障工况的偏离程度，Rhd_j^2 越大，说明第 j 个主元与该故障发生的相关性越大，因此赋予的权重值也越大。令

$$\alpha' = \exp\left(\frac{Rhd_j^2 - 1}{\sigma}\right) \tag{30}$$

式中 $\sigma > 0$ 是可调参数。定义权重 y_j^{new} 如式(31)所示。

$$y_j^{\text{new}} = \begin{cases} 1, & Rhd_j^2 \leqslant \gamma \\ \alpha, & Rhd_j^2 > \gamma \end{cases} \tag{31}$$

式中，γ 为阈值，本文通过 3σ 法则来确定 γ 的数值。假设 Rhd_j^2 的数值服从均值为 M、标准差为 η 的高斯分布，那么阈值 γ 可由式(32)得到。

$$\gamma = M + L\eta \tag{32}$$

式中，L 为预警宽度，一般取值为 3，即 3σ 法则。

那么，新的得分矩阵可表示为

$$\boldsymbol{T}_w = \boldsymbol{T}_{\text{key(new)}}\ \boldsymbol{W} = [\boldsymbol{t}_{1,\text{new}}, \boldsymbol{t}_{2,\text{new}}, \cdots, \boldsymbol{t}_{f,\text{new}}] \cdot \begin{bmatrix} y_1^{\text{new}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & y_f^{\text{new}} \end{bmatrix} = [y_1^{\text{new}}\boldsymbol{t}_{1,\text{new}}, y_2^{\text{new}}\boldsymbol{t}_{2,\text{new}}, \cdots, y_f^{\text{new}}\boldsymbol{t}_{f,\text{new}}] \tag{33}$$

式中，$\boldsymbol{T}_{\text{key(new)}}$ 为故障工况下关键主元构成的得分矩阵，$\boldsymbol{W}$ 为加权矩阵。

传统 PCA 在进行故障检测时，会设定如式(6)、(7)所示的统计量进行监测，这类根据数据间距离所设定的统计量要求采样数据满足正态分布，事实上在实际工业生产过程中，大部分采样数据没有特定的分布规律，这就导致了 PCA 用于故障检测时效果不佳。为此，本文利用基于密度的 LOF 构造了一个单一统计量，相较于 T^2 和 SPE，该统计量集中了故障对过程变量的影响。同时，LOF 从密度的角度出发，依靠数据点的离散程度对故障进行检测，对于数据的分布没有特定的要求，避免了传统 T^2 和 SPE 统计量要求数据服从正态分布的不足。除此之外，基于马氏距离的 LOF 与原始 LOF 算法相比，考虑了加权后的关键主元与其邻域子集之间的马氏距离对过程变量的影响，最终形成一种对数据分布具有鲁棒性的统计量。由于 LOF 统计量的分布情况无法准确获取，因此本文采用核密度估计法确定所构造 LOF 统计量的控制限。令正常工况下获得的 LOF 统计量构成集合 $\Omega = \{R_{\text{LOF},x_1}, R_{\text{LOF},x_2}, \cdots, R_{\text{LOF},x_n}\}$，对于 LOF 统计量进行核密度估计，如式(34)所示。

$$\hat{f}(R_{\text{LOF},x}) = \frac{1}{n\theta}\sum_{i=1}^{n} K\left(\frac{R_{\text{LOF},x} - R_{\text{LOF},x_i}}{\theta}\right) \tag{34}$$

式中，$K(\cdot)$ 为核函数，θ 为核宽，LOF 控制限 $R_{\text{LOF,limit}}$ 在置信水平为 α 的情况下可由式(35)获得

$$\int_{-\infty}^{R_{\text{LOF,limit}}} \hat{f}(R_{\text{LOF},x})\,\mathrm{d}R_{\text{LOF},x} = 1 - \alpha \tag{35}$$

3.2　故障检测步骤

基于 HDWFVE－MDLOF 的故障检测流程如图 1 所示。

离线建模阶段步骤如下。

1)采集正常工况下的样本数据，标准化得到 $\boldsymbol{X} \in \boldsymbol{R}^{n\times m}$。

2)利用 FVE 得到 f 个关键主元 $\boldsymbol{t}_j$ ($j = 1, 2, \cdots, f$) 以及对应的负载向量 $\overline{\boldsymbol{p}}_j$；由关键主元构成的得分矩阵可以表示为 $\boldsymbol{T}_{\text{key}} = [\boldsymbol{t}_1, \boldsymbol{t}_2, \cdots, \boldsymbol{t}_f] \in \boldsymbol{R}^{n\times f}$。

3)将 $\boldsymbol{T}_{\text{key}} = [\boldsymbol{t}_1, \boldsymbol{t}_2, \cdots, \boldsymbol{t}_f] \in \boldsymbol{R}^{n\times f}$ 作为改进 LOF 算法的输入，并利用核密度估计确定其控制限 $R_{\text{LOF,limit}}$。

在线检测阶段步骤如下。

1)采集异常工况下的样本数据，标准化得到 $\boldsymbol{X}_{\text{new}} \in \boldsymbol{R}^{n\times m}$。

2)对于新的数据 $\boldsymbol{x}_{\text{new}} \in \boldsymbol{R}^{m\times 1}$，关键主元由 $\boldsymbol{t}_{j,\text{new}} =$

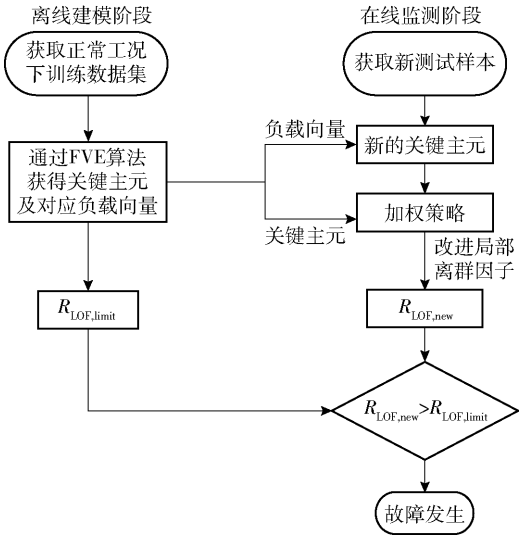

图 1 基于 HDWFVE-MDLOF 的故障检测流程图

Fig. 1 Flowchart of HDWFVE-MLOF based fault detection

$\overline{\boldsymbol{p}}_j^{\mathrm{T}}\boldsymbol{x}_{\mathrm{new}}$ 计算得到。

3)通过式(20)、(21)计算 Rhd_j^2 统计量并利用 3σ 法则计算其控制限 γ,再根据式(30)~(32)赋予得分向量相应权重,根据式(33)形成新的加权后的得分矩阵 $\boldsymbol{T}_w$。

4)将 $\boldsymbol{T}_w=[\boldsymbol{t}_{1,\mathrm{new}},\boldsymbol{t}_{2,\mathrm{new}},\cdots,\boldsymbol{t}_{f,\mathrm{new}}]\in\boldsymbol{R}^{n\times f}$ 作为改进 LOF 算法的输入,求得统计量 $R_{\mathrm{LOF,new}}$,若 $R_{\mathrm{LOF,new}}>R_{\mathrm{LOF,limit}}$,则说明有故障发生。

4 仿真分析

4.1 数值仿真实例

文献[30]中的数值实例被广泛应用于故障检测的仿真实验中,本文采用此数值实例验证所提方法的有效性。该数值例子表示结构如式(36)所示。

$$\begin{aligned}\boldsymbol{r}(i)&=\boldsymbol{A}\times\boldsymbol{r}(i-1)+\boldsymbol{B}\times\boldsymbol{u}(i-1)\\\boldsymbol{u}(i)&=\boldsymbol{C}\times\boldsymbol{u}(i-1)+\boldsymbol{D}\times\boldsymbol{h}(i-1)\\\boldsymbol{g}(i)&=\boldsymbol{r}(i)+\boldsymbol{v}(i)\end{aligned}\tag{36}$$

式中,$\boldsymbol{A}=\begin{bmatrix}0.018&-0.191&0.287\\0.847&0.264&0.943\\-0.333&0.514&-0.217\end{bmatrix}$,$\boldsymbol{B}=\begin{bmatrix}1&2\\3&-4\\-2&1\end{bmatrix}$,$\boldsymbol{C}=\begin{bmatrix}0.811&-0.226\\0.477&0.415\end{bmatrix}$,$\boldsymbol{D}=\begin{bmatrix}0.193&0.689\\-0.320&-0.749\end{bmatrix}$,$\boldsymbol{h}$ 为一个在区间(-2,2)内服从均匀分布的随机向量,$\boldsymbol{v}$ 为一个服从均值为 0、方差为 0.1 的正态分布的随机噪声向量,i 为采样时刻。训练集 X 由 200 个正常采样点构成,$\boldsymbol{x}(i)=[\boldsymbol{g}^{\mathrm{T}}(i)\quad\boldsymbol{u}^{\mathrm{T}}(i)]$,测试集由包含故障的 200 个采样点构成,相关故障的设定及描述如下。

1)故障 1 变量 h_1 从第 51 个样本点开始直到 200 个样本点结束,增加一个幅值为 3 的单位阶跃故障。

2)故障 2 变量 h_2 从第 51 个样本点开始直到 200 个样本点结束,增加一个 $0.06(i-50)$ 的斜坡故障。

在运用 FVE 方法进行主元选取时,特征空间包含了几乎所有原始变量的信息,残差空间中几乎不包含有用信息,因此只需构建反映特征空间的 T^2 统计量进行监测。运用本文所提方法与 CPV-T^2、CPV-SPE、FVE-T^2、FVE-LOF 以及 HDWFVE-LOF 进行仿真对比实验。其中 CPV 的贡献度设定为 85%,控制限采用 95% 的置信度;LOF 中的 k 值设定为 15,核宽设为 $500m$(m 为变量个数)。在 CPV 主元选取方法中,前 3 个主元的累计方差贡献率超过了 85%,因此按照方差从大到小的顺序选取 PC_1、PC_2、PC_3 为主元,而在 FVE 的主元选取方法中,各变量对应的关键主元如表 1 所示,PC_2、PC_3、PC_4、PC_5 被选为主元。表 2 为各个方法对于数值实例中故障 1 和故障 2 的漏报率与误报率。各个方法对故障 1 的检测结果如图 2 所示。图 2(a)、(b)为传统 PCA 算法利用 CPV 主元选取方法所构建的基于 T^2 和 SPE 统计量的故障检测结果,结合表 2 可知,其中 T^2 统计量的漏报率为 74%,检测效果不佳,SPE 统计量

表 1 各变量所对应关键主元

Table 1 Variables and corresponding key PCs

关键主元	变量	关键主元	变量
PC_2	x_1	PC_3	x_4
PC_5	x_2	PC_4	x_5
PC_3	x_3		

表 2 数值实例故障检测结果

Table 2 Fault detection results for numerical cases

算法	故障 1		故障 2	
	漏报率/%	误报率/%	漏报率/%	误报率/%
CPV-T^2	74	0	6	2
CPV-SPE	75	0	57.1	0
FVE-T^2	48.7	0	1.3	4
FVE-LOF	46	0	1.4	0
HDWFVE-LOF	28	0	2	0
HDWFVE-MDLOF	20.8	0	0.7	0

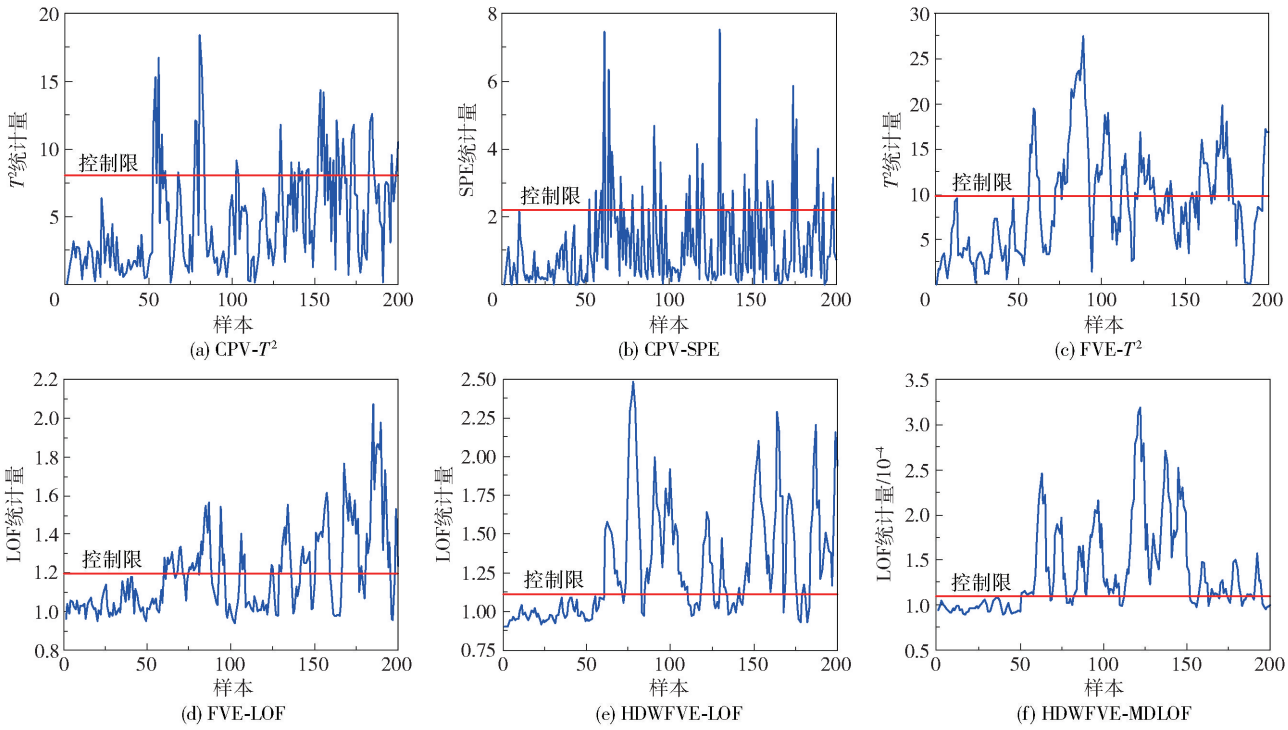

图 2　数值实例中故障 1 的检测结果

Fig. 2　Detection results of fault 1 in a numerical example

的漏报率达到 75%，表明该统计量基本无法检测到故障的发生。图 2(c)为利用 FVE 主元选取方法进行故障检测的结果，结合表 2 可知，T^2 统计量的漏报率降低为 48.7%。图 2(d)为 FVE 主元选取法利用 LOF 作为统计量的故障检测结果，由表 2 知其漏报率为 46%，漏报率的降低是因为该方法从数据密度大小出发来搜寻故障点。图 2(e)给出了 HDWFVE-LOF 的故障检测结果，得益于对故障相关主元的合理加权，由表 2 可知其漏报率为 28%，得到大幅降低。图 2(f)为本文所提方法的故障检测结果，可以看出基于马氏距离的 LOF 算法起到了很好的效果，相比于其他方法，本文所提方法的漏报率最低，为 20.8%。

图 3 给出了不同方法对故障 2 的检测结果。结合表 2 的结果，总体上，除了传统 PCA 利用 CPV 的主元选取法将 SPE 作为统计量的漏报率较高外，其余方法均能够保持较低的故障漏报率。其中，本文所提方法最早检测到故障的发生，并且在检测到故障发生后依然保持着很好的稳定性，统计量曲线始终位于控制限上方，而其他方法在检测到故障发生后，统计量曲线有回落到控制限的现象(图 3(a)、(b)、(c)、(d))，引起漏报。

为了进一步体现 LOF 作为统计量的优越性，图 4(a)、(b) 给出了故障 2 中运用本文方法加权后关键主元 PC_3 和 PC_5 的正态概率分布，其中横坐标为主成分样本点数据，纵坐标为先验概率。数据的分布越接近标准正态分布，相应图上散点的分布就越近似直线。由图 4 可知，图中条形点的头部和尾部较直线还是有着较大的偏移，意味着这两个加权后的主元并不服从正态分布，这也从另一个角度说明了利用 LOF 构建统计量的必要性。综上，数值仿真结果体现了本文所提方法的有效性与优越性。

4.2　带钢热连轧过程验证

带钢热连轧过程(hot strip mill process, HSMP)是工业生产中的一个重要工序，其生产流程如图 5 所示。由于 HSMP 在运行过程中有高温度、高速度、多阶段的特征，导致其变量具有高耦合的特点。对此，本文将该过程用于故障检测的研究。在带钢热连轧现场(鞍钢集团 1 700 mm 带钢热连轧生产线)收集能够反映生产过程的数据，相关的 20 个过程变量的具体描述如表 3 所示。

首先选取 4 000 个正常工况下带钢热连轧过程样本数据进行离线建模，再收集 4 000 个带有故障的样本数据进行在线检测，其中故障产生时间为样本点 2 001 到 3 000，故障原因为第 5 机架的弯辊力测量传感器故障。LOF 算法中近邻数 k 设置为 150。利用 KDE 进行控制限计算时，核宽 θ 设置为

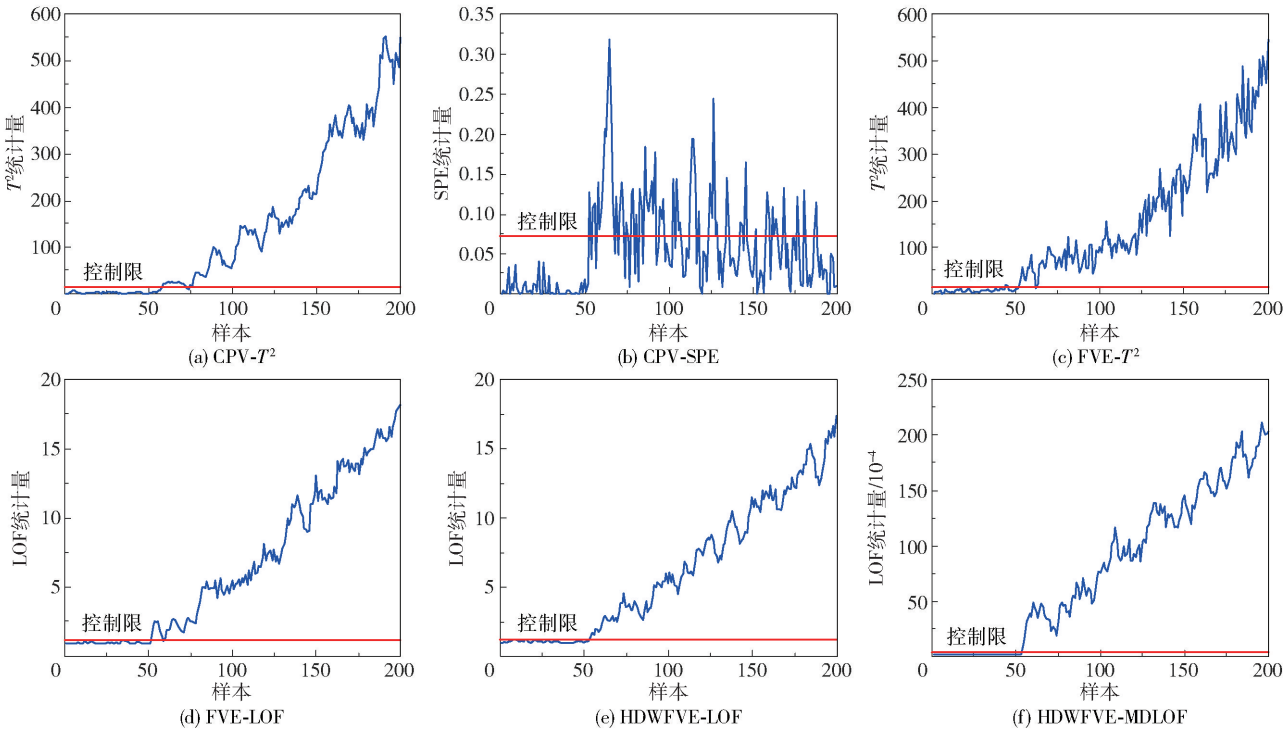

图 3 数值实例中故障 2 的检测结果

Fig. 3 Detection results of fault 2 in a numerical example

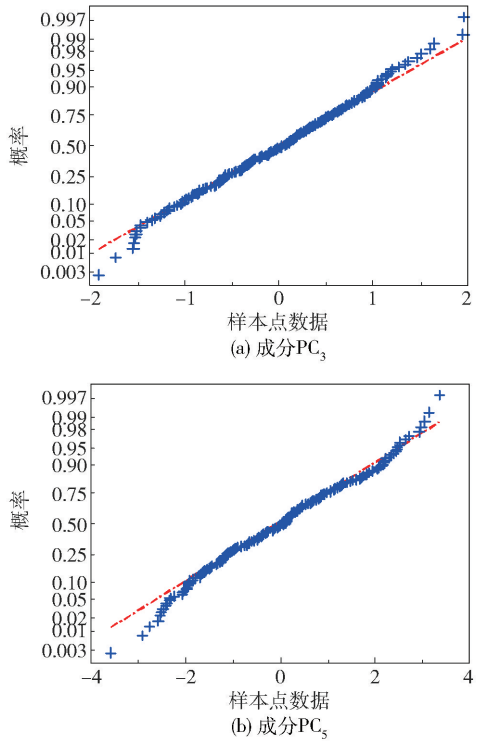

图 4 成分正态概率分布

Fig. 4 Normal probability distribution of components

500m(m 为变量数)。利用 FVE 得到的关键主元与变量的对应关系如表 4 所示。各方法对于该故障的检测结果如图 6 所示,具体结果列于表 5。图 6(a)、(b)为 PCA 利用传统 CPV 方法对于该故障的检测结果,T^2 和 SPE 统计量的漏报率分别为 74.2% 和 90.2%,数值都较高,尤其是 SPE 统计量,基本无法检测到故障的发生。图 6(c)为 FVE 方法选取关键主元后构建 T^2 统计量的故障检测结果,关键主元的选取使得故障漏报率大大降低,且故障误报率没有出现较大的增加。对比图 6(d)、(e)、(f)可以得出,本文所提方法在检测出所有故障的同时,还获得了最低的故障误报率。

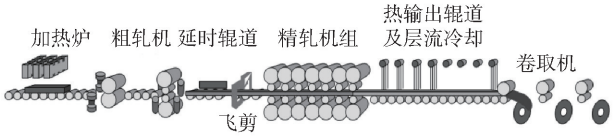

图 5 带钢热连轧生产流程图

Fig. 5 HSMP flow chart

表 3 精轧过程变量

Table 3 Process variables in the finishing mill

变量编号	类别	具体描述
1 ~ 7	测量变量(mm)	第 i 机架的平均辊缝,i = 1,2,…,7
8 ~ 14	测量变量(MN)	第 i 机架的轧制力,i = 1,2,…,7
15 ~ 20	测量变量(MN)	第 i 机架的弯辊力,i = 2,3,…,7

括号内为该变量的单位。

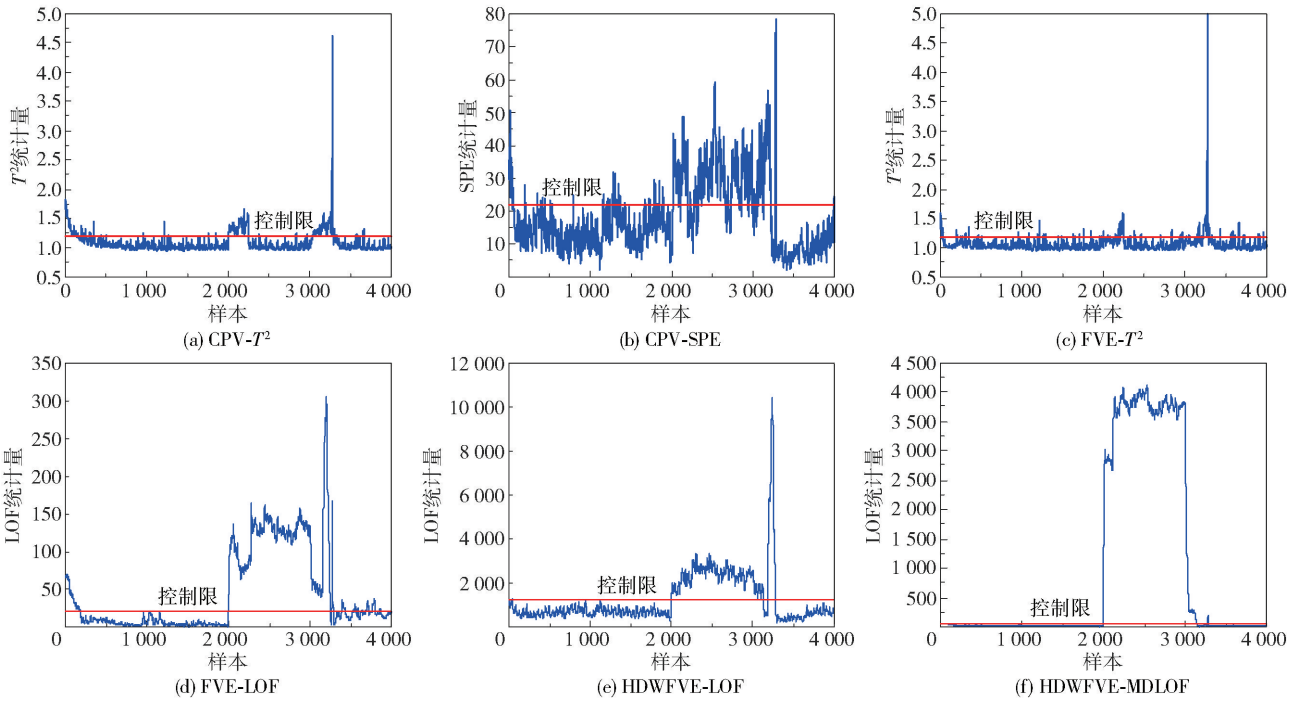

图 6　故障工况下各个方法的检测结果

Fig. 6　Detection results for each method under fault operation

表 4　各变量所对应关键主元

Table 4　Variables and corresponding key PCs

关键主元	变量	关键主元	变量
PC_3	x_{20}	PC_{16}	x_{10}
PC_4	x_{18}	PC_{17}	x_2, x_6, x_9
PC_5	x_3	PC_{18}	x_7, x_{14}
PC_6	x_{12}	PC_{19}	x_{16}, x_{17}
PC_7	x_4, x_{15}	PC_{20}	x_1
PC_8	x_8	PC_{12}	x_{13}, x_{19}
PC_{10}	x_5	PC_{13}	x_{11}

表 5　不同方法的故障检测结果

Table 5　Fault detection results of different methods

算法	漏报率/%	误报率/%
CPV-T^2	74.2	16.8
CPV-SPE	90.2	8.6
FVE-T^2	14.5	15.5
FVE-LOF	0	22.6
HDWFVE-LOF	0	8.0
HDWFVE-MDLOF	0	6.83

表 6 给出了 ReliefF-PCA 算法[11]、敏感主成分分析(SPCA)算法[12]、故障相关-贝叶斯推断主成分分析(FBPCA)算法[31]以及敏感主元多块主成分分析(MBSPCA)算法[32]对于第 5 机架的弯辊力测量传感器故障的检测漏报率和误报率。其中 MBSPCA 算法中 ω 取值 0.2,即敏感主元阈值 $\varepsilon_{\text{limit}}$ 为 0.003 9,所有方法的置信度均为 95%。由表 6 的数据对比可发现,本文所提方法的漏报率与误报率均为最低。综上可说明本文所提 HDWFVE-MDLOF 算法的有效性与优越性。

表 6　几种现有主元选取方法性能比较

Table 6　Performance comparison of some states of principal component selection monitoring methods

算法	统计量	漏报率/%	误报率/%
ReliefF-PCA	T^2	4.88	6.97
FBPCA	BIC_T^2	6.00	8.45
SPCA	T^2	4.71	7.96
MBSPCA	BIC_T^2	4.32	9.56
HDWFVE-MDLOF	LOF	0	6.83

5　结论

本文提出一种基于 HDWFVE-MDLOF 的故障检测算法,在 PCA 框架下的主元选取阶段利用 FVE 方法获取关键主元,保留了全部变量信息,并利用海林格距离变化率对关键主元中的故障相关主元进行加权突出,有效地降低了故障漏报率;利用改进的

LOF 构造统计量,突破了传统统计量对于数据分布的限定。数值实例及带钢热连轧实际生产数据结果表明,与 ReliefF－PCA、SPCA、FBPCA 以及 MBSPCA 等算法相比,本文所提算法的漏报率及误报率均为最低,证明了所提方法的有效性与优越性。

参考文献:

[1] CHEN H T, JIANG B, DING S X, et al. Data-driven fault diagnosis for traction systems in high-speed trains: a survey, challenges, and perspectives[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(3): 1700－1716.

[2] YIN S, DING S X, XIE X C, et al. A review on basic data-driven approaches for industrial process monitoring[J]. IEEE Transactions on Industrial Electronics, 2014, 61(11): 6418－6428.

[3] QU L, LI L, ZHANG Y, et al. PPCA-based missing data imputation for traffic flow volume: a systematical approach[J]. IEEE Transactions on Intelligent Transportation Systems, 2009, 10(3): 512－522.

[4] LUO K W, LI S L, DENG R, et al. Multivariate statistical kernel PCA for nonlinear process fault diagnosis in military barracks[J]. International Journal of Hybrid Information Technology, 2016, 9(1): 195－206.

[5] KU W F, STORER R H, GEORGAKIS C. Disturbance detection and isolation by dynamic principal component analysis[J]. Chemometrics and Intelligent Laboratory Systems, 1995, 30(1): 179－196.

[6] MALINOWSKI E R. Factor analysis in chemistry[M]. 2nd ed. New York: Wiley & Sons. Inc., 1991.

[7] VALLE S, LI W H, QIN S J. Selection of the number of principal components: the variance of the reconstruction error criterion with a comparison to other methods[J]. Industrial & Engineering Chemistry Research, 1999, 38(11): 4389－4401.

[8] KAISER H F. The application of electronic computers to factor analysis[J]. Educational and Psychological Measurement, 1960, 20(1): 141－151.

[9] JOLLIFFE I T. A note on the use of principal components in regression[J]. Applied Statistics, 1982, 31(3): 300－303.

[10] TOGKALIDOU T, BRAATZ R D, JOHNSON B K, et al. Experimental design and inferential modeling in pharmaceutical crystallization[J]. AIChE Journal, 2001, 47(1): 160－168.

[11] 陶阳, 王帆, 侍洪波, 等. 基于 ReliefF 的主元挑选算法在过程监控中的应用[J]. 化工学报, 2017, 68(4): 1525－1532.
TAO Y, WANG F, SHI H B, et al. Principal component selection algorithm based on ReliefF and its application in process monitoring[J]. CIESC Journal, 2017, 68(4): 1525－1532. (in Chinese)

[12] JIANG Q C, YAN X F, ZHAO W X. Fault detection and diagnosis in chemical processes using sensitive principal component analysis[J]. Industrial & Engineering Chemistry Research, 2013, 52(4): 1635－1644.

[13] 仓文涛, 杨慧中. 基于主元子空间富信息重构的过程监测方法[J]. 化工学报, 2018, 69(3): 1114－1120.
CANG W T, YANG H Z. A process monitoring method based on informative principal component subspace reconstruction[J]. CIESC Journal, 2018, 69(3): 1114－1120. (in Chinese)

[14] SONG B, MA Y Y, SHI H B. Improved performance of process monitoring based on selection of key principal components[J]. Chinese Journal of Chemical Engineering, 2015, 23(12): 1951－1957.

[15] JIANG Q C, WANG B, YAN X F. Multiblock independent component analysis integrated with Hellinger distance and Bayesian inference for non-Gaussian plant-wide process monitoring[J]. Industrial & Engineering Chemistry Research, 2015, 54(9): 2497－2508.

[16] HARROU F, MADAKYARU M, SUN Y. Improved nonlinear fault detection strategy based on the Hellinger distance metric: plug flow reactor monitoring[J]. Energy and Buildings, 2017, 143: 149－161.

[17] CHEN H T, JIANG B, LU N Y. A newly robust fault detection and diagnosis method for high-speed trains[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(6): 2198－2208.

[18] BREUNIG M M, KRIEGEL H P, NG R T, et al. LOF: identifying density-based local outliers[C]// ACM Sigmod International Conference on Management of Data. Dallas, 2000: 93－104.

[19] LEE J, KANG B, KANG S H. Integrating independent component analysis and local outlier factor for plant-wide process monitoring[J]. Journal of Process Control, 2011, 21(7): 1011－1021.

[20] DENG X G, 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.

[21] 冯立伟, 李元, 张成, 等. 基于时空近邻标准化和局部离群因子的复杂过程故障检测[J]. 控制理论与应用, 2020, 37(3): 651－657.
FENG L W, LI Y, ZHANG C, et al. Time-space neighborhood standardization-local outlier factor based fault detection for complex process[J]. Control Theory and Applications, 2020, 37(3): 651－657. (in Chinese)

[22] QIN S J. Statistical process monitoring: basics and beyond[J]. Journal of Chemometrics, 2003, 17: 480－502.

[23] BASSEVILLE M. Divergence measures for statistical data processing-an annotated bibliography[J]. Signal Processing, 2013, 93(4): 621 – 633.

[24] DITZLER G, POLIKAR R. Hellinger distance based drift detection for nonstationary environments[C]//2011 IEEE Symposium on Computational Intelligence in Dynamic and Uncertain Environments (CIDUE). Paris, 2011: 41 – 48.

[25] LI C, HUANG B, QIAN F. Hellinger distance based probability distribution approach to performance monitoring of nonlinear control systems[J]. Chinese Journal of Chemical Engineering, 2015, 23(12): 1945 – 1950.

[26] 童武. 谈谈 Lebesgue 测度概念的建立[J]. 数学通报, 1988(3): 38 – 41.
TONG W. On the establishment of Lebesgue's measure [J]. Bulletin des Sciences Mathematics, 1988 (3): 38 – 41. (in Chinese)

[27] MA H H, HU Y, SHI H B. A novel local neighborhood standardization strategy and its application in fault detection of multimode processe[J]. Chemometrics and Intelligent Laboratory Systems, 2012, 118(1): 287 – 300.

[28] ZHANG K, DING S X, SHARDT Y A W, et al. Assessment of T2- and Q-statistics for detecting additive and multiplicative faults in multivariate statistical process monitoring[J]. Journal of the Franklin Institute, 2017, 354 (2): 668 – 688.

[29] 韩敏, 张占奎. 基于加权核独立成分分析的故障检测方法[J]. 控制与决策, 2016, 31(2): 242 – 248.
HAN M, ZHANG Z K. Fault detection method based on weighted kernel independent component analysis [J]. Control and Decision, 2016, 31 (2): 242 – 248. (in Chinese)

[30] LEE J M, YOO C K, LEE I B. Statistical process monitoring with independent component analysis[J]. Journal of Process Control, 2004, 14(5): 467 – 485.

[31] JIANG Q C, YAN X F, HUANG B. Performance-driven distributed PCA process monitoring based on fault-relevant variable selection and Bayesian inference[J]. IEEE Transactions on Industrial Electronics, 2016, 63 (1): 377 – 386.

[32] 顾炳斌, 熊伟丽, 史旭东. 基于故障敏感主元的多块 PCA 故障监测方法[J]. 高校化学工程学报, 2019, 33 (6): 1499 – 1508.
GU B B, XIONG W L, SHI X D. Multi-block PCA process monitoring based on fault sensitive principal components[J]. Journal of Chemical Engineering of Chinese Universities, 2019, 33(6): 1499 – 1508. (in Chinese)

A fault detection method based on Hellinger distance-weighted key principal components for the process industry

ZHAO Cheng　SU ShengChao*

(School of Electric and Electronic Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

Abstract: When a principal components analysis (PCA) algorithm is used for fault detection, the selection of principal components and how to manage them directly affects the efficacy of fault detection. A new fault detection method based on full variable expression(FVE) and Hellinger distance(HD) is proposed in this work. All the key principal components are first obtained by FVE, and all the variable information is retained. The change rate based on Hellinger distance is then defined, considering the importance of fault-relevant principal components. This is used to measure the difference between the principal components under normal operation and abnormal operation. The principal components which are more relevant to the occurrence of faults are weighted in order to highlight the effect of fault-relevant principal components in subsequent fault detection. Finally, considering that the dimensionality reduction data often follow a non-Gaussian distribution, the improved local outlier factor (LOF) is used to construct the statistics and the corresponding control limit is determined by kernel density estimation(KDE). Finally, a numerical case and the actual data for a hot strip mill process have been used to verify the effectiveness and superiority of the proposed method in fault detection.

Key words: fault detection; principal components analysis; key principal component; Hellinger distance; local outlier factor

(责任编辑:吴万玲)