

基于核主成分空间支持向量机的过程监视方法

郭金玉,李 涛,李 元

(沈阳化工大学 信息工程学院,辽宁 沈阳 110142)

摘 要:为了有效改进支持向量机(SVM)在工业过程中的故障检测性能,提出一种基于核主成分空间支持向量机的过程监视方法。首先,运用核主成分分析方法(KPCA)处理数据,获取数据的核主成分,在核主成分空间计算 Hotelling's T^2 统计量;然后,在 T^2 统计量基础上加入时滞特性和时差特性,并将三者组合成增广矩阵,运用正常数据和故障数据的增广矩阵训练 SVM 模型;最后,运用 SVM 模型对测试数据进行分类,实现对故障的有效检测。将本研究方法应用于非线性数值例子和田纳西-伊斯曼工业过程中,与 PCA、KPCA、传统的 SVM 和 KPCA-LSSVM 方法作比较,进一步验证了该方法的有效性。

关键词:支持向量机;核主成分分析;过程监视;非线性过程

中图分类号:TP277

文献标志码:A

Process monitoring method based on support vector machine in kernel principal component space

GUO Jinyu, LI Tao, LI Yuan

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

Abstract: In order to effectively improve the fault detection performance of support vector machine (SVM) in industrial process, a process monitoring method based on SVM in kernel principal component space was proposed in this paper. First, the kernel principal component analysis method (KPCA) was used to process the data and obtain the kernel principal components. The Hotelling's T^2 statistic was calculated in the kernel principal component space. Then, the time delay and time difference characteristics were added to the T^2 statistic and an augmented matrix was formed by combining them. The augmented matrix of normal and fault data was used to train the SVM model. Finally, the SVM model was used to classify the test data to achieve effective fault detection. The proposed method was applied to a nonlinear numerical example and the Tennessee-Eastman industrial process, compared with the PCA, KPCA, traditional SVM and KPCA-LSSVM methods, the experimental results further verified its effectiveness.

Key words: support vector machine; kernel principal component analysis; process monitoring; nonlinear process

随着现代工业生产过程日益复杂,有效提高控制系统的故障检测性能变得尤为重要。在工业生产过程中,基于数据驱动的控制系统的故障检测方法快速发展。近年来,以主成分分析(principal component analysis, PCA)为代表的多元统计方法发展迅速,并衍生出了众多故障检测方法^[1-5]。PCA 是一种线性降维方法,在数据满足单一分布的基本假设下,解决相关过程变量引起的线性问题,对数据的信息提取和压缩也取得很好的效果。但实际工业生产过程中存在众多非线性、多模态和非高斯特征的数据,不满足单一分布的假设,PCA 方法难以得到令人满意的检测效果^[6]。核主成分分析(kernel principal component analysis, KP-

收稿日期:2020-10-14

基金项目:国家自然科学基金重大项目(61490701),国家自然科学基金项目(61673279),辽宁省教育厅项目(LJ2019007)

作者简介:郭金玉(1975—),女,山东高唐人,副教授,博士,主要研究方向为工业过程的故障检测与诊断、生物特征识别算法及应用。

李 元(1964—),女,辽宁沈阳人,教授,博士,主要研究方向为基于数据驱动的工业过程故障诊断,本文通信作者。

E-mail: li-yuan@mail.tsinghua.edu.cn

CA)^[7-9]将低维空间中的非线性数据映射到高维特征空间,在高维特征空间中建立 PCA 模型进行故障检测,在一定程度上扩大了 PCA 方法的使用范围,提高了对非线性数据的处理能力。为了解决工业生产过程中的数据非线性和动态性等问题,王亚君等^[10]提出一种多动态核主成分分析的故障检测方法,构造了适合批量生产过程的多模态、非线性和动态模型建立方法。针对复杂工业系统中的动态非线性故障检测方法精度低和计算量大的问题,翟坤等^[11]提出一种改进的动态核主成分分析的故障检测方法,提高了 KPCA 的故障检测性能,同时对微小故障的检测也更为敏感。

支持向量机(support vector machine, SVM)被提出后,因强大的泛化能力和在解决分类问题方面的优势,受到国内外学者^[12-14]的广泛关注,成为机器学习中的经典方法。SVM 的分类思想是找到最大分离超平面,进而将样本有效分类。针对非线性过程中 SVM 的检测性能不足问题,Zhang 等^[15]结合核独立成分分析(kernel independent component analysis, KICA)、KPCA 和 SVM 的优点,开发了一种非线性动态检测方法。Yu^[16]提出基于支持向量聚类的概率方法,通过支持向量建立超球面,将故障数据和正常数据分开,实现故障检测。为了进一步提高 SVM 的故障检测性能,Hsu 等^[17]将独立成分分析(independent component analysis, ICA)和 SVM 方法相结合,提高了 SVM 在多变量过程中的监测能力。王东等^[18]提出基于核主成分分析和最小二乘支持向量机(kernel principal component analysis and least squares support vector machine, KPCA-LSSVM)的故障识别方法,比传统的分类方法故障识别速度快、分类准确率高。后续的研究与应用多集中在提高 SVM 的故障检测性能和丰富其应用场景上。

为了提高 SVM 在工业过程中的故障检测性能,提出一种基于核主成分空间支持向量机的过程监视方法。运用 KPCA 对数据进行处理,计算 Hotelling's T^2 统计量;在 T^2 统计量基础上加入时滞特性和时差特性,将三者组合成增广矩阵,作为输入训练 SVM 模型;运用训练的 SVM 模型对测试数据进行故障检测。本研究通过增加模型输入的复杂度在一定程度上去除样本间的自相关性,提高 SVM 的故障检测性能。

1 基于核主成分空间支持向量机的过程监视

1.1 支持向量机

SVM 在处理工业数据时,往往会面临着众多的非线性数据,由于非线性数据无法处理,需要将其投影到高维空间,去除数据非线性,建造最大分离超平面,使数据能够被有效分类。由于分离平面是基于支持向量构造的,所以 SVM 是解决高维问题的一种很好方案。核函数的引入,也避免了高维计算的复杂度。

假设给定的样本训练集为 $H = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_m, y_m)\}$, $y_i \in \{-1, +1\}$, 其中有两类样本,需要找到最大分离超平面将两类样本进行有效分类,这是分类学习最基本的思想。要建立超平面,就要获得权重向量 \mathbf{w} 和位移项 b ,两者决定了超平面的划分,影响样本的分类准确率。找到有最大间隔的超平面,也就是要找到满足条件的参数 \mathbf{w} 和 b ,使得分类间隔 γ 最大,使两类样本最大程度分开,得到更好的分类效果。因此,SVM 基本模型可表示为:

$$\begin{aligned} \min_{\mathbf{w}, b, \xi_i} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^m \xi_i, \xi_i \geq 0, \\ \text{s.t.} \quad & y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i, i = 1, 2, \dots, m. \end{aligned} \quad (1)$$

其中: $C > 0$ 是惩罚参数,其引入能够在一定程度上解决 SVM 对一些样本分类出错的问题; ξ_i 是松弛变量,用于表征样本不满足最大分离超平面的约束程度。

为了求解式(1),需要将其转化成“对偶问题”,使用拉格朗日乘子法求解。该问题的拉格朗日函数表示为:

$$L(\mathbf{w}, b, \alpha, \xi, \mu) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^m \xi_i + \sum_{i=1}^m \alpha_i (1 - \xi_i - y_i(\mathbf{w}^T \mathbf{x}_i + b)) - \sum_{i=1}^m \mu_i \xi_i. \quad (2)$$

其中, $\alpha_i \geq 0$, $\mu_i \geq 0$ 是拉格朗日乘子。在求解过程中,数据往往具有较强的非线性,为了更好地解决这一问题,需要通过非线性映射 φ 投影到高维空间。为了避免高维运算,引入核函数:

$$\mathbf{K}(\mathbf{x}_i, \mathbf{x}_j) = (\boldsymbol{\varphi}(\mathbf{x}_i), \boldsymbol{\varphi}(\mathbf{x}_j)) = \boldsymbol{\varphi}(\mathbf{x}_i)^T \boldsymbol{\varphi}(\mathbf{x}_j). \quad (3)$$

通过核函数能够避免在高维特征空间中内积的计算,对偶问题可表示为:

$$\begin{aligned} \max_a \quad & \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j \mathbf{K}(\mathbf{x}_i, \mathbf{x}_j), \\ \text{s.t.} \quad & \sum_{i=1}^m \alpha_i y_i = 0, 0 \leq \alpha_i \leq C, i = 1, 2, \dots, m. \end{aligned} \quad (4)$$

求解后的支持向量展式为:

$$f(\mathbf{x}) = \text{sign}\left(\sum_{i,j=1}^m \alpha_i y_i \mathbf{K}(\mathbf{x}_i, \mathbf{x}_j) + b\right). \quad (5)$$

因此,SVM在对数据进行分类时,对训练数据进行模型训练,求解式(5)的判别函数。当测试数据输入模型时,判别函数能够对其进行有效分类。无论数据是线性还是非线性,SVM都可以进行有效转化,通过模型训练学习数据变化特征,进而高效准确地对数据进行分类。

1.2 基于核主成分空间的支持向量机方法

假设对于随机变量 \mathbf{x} , 其样本集表示为 $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$, 其中 m 是样本数, $\mathbf{x}_i \in \mathbf{R}^n (i=1, 2, \dots, m)$, 通过非线性映射 φ 将数据投影到高维特征空间中, 得到新矩阵 $\boldsymbol{\varphi}(\mathbf{x}) = [\varphi(\mathbf{x}_1), \varphi(\mathbf{x}_2), \dots, \varphi(\mathbf{x}_m)]^T \in \mathbf{R}^{m \times n}$, 对其进行主成分分析。定义该矩阵的协方差矩阵为:

$$\mathbf{S} = \frac{1}{m} \boldsymbol{\varphi}(\mathbf{x}) \boldsymbol{\varphi}(\mathbf{x})^T. \quad (6)$$

对协方差矩阵 \mathbf{S} 进行特征向量分析, 即

$$\frac{1}{m} \boldsymbol{\varphi}(\mathbf{x}) \boldsymbol{\varphi}(\mathbf{x})^T \mathbf{P}_i = \lambda_i \mathbf{P}_i, \quad (7)$$

式中, λ_i 为协方差矩阵的特征值, \mathbf{P}_i 是特征值对应的特征向量, 即

$$\mathbf{P}_i = \sum_{i=1}^m \mathbf{a}_i \boldsymbol{\varphi}(\mathbf{x}_i) = \boldsymbol{\varphi}(\mathbf{x}) \mathbf{a}, \quad (8)$$

其中 $\mathbf{a} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$, \mathbf{a}_i 为核函数矩阵的第 i 个特征向量。根据累计方差贡献率(cumulative percent variance, CPV)^[19]求取核主成分个数 z , 如下式:

$$\frac{\sum_{i=1}^z \lambda_i}{\sum_{i=1}^m \lambda_i} \geq 85\%. \quad (9)$$

式(7)两边同时左乘 $\boldsymbol{\varphi}(\mathbf{x})^T$, 可将其转化为:

$$\frac{1}{m} \boldsymbol{\varphi}(\mathbf{x})^T \boldsymbol{\varphi}(\mathbf{x}) \boldsymbol{\varphi}(\mathbf{x})^T \boldsymbol{\varphi}(\mathbf{x}) \mathbf{a} = \lambda_i \boldsymbol{\varphi}(\mathbf{x})^T \boldsymbol{\varphi}(\mathbf{x}) \mathbf{a}. \quad (10)$$

令 $\mathbf{K} = \boldsymbol{\varphi}(\mathbf{x})^T \boldsymbol{\varphi}(\mathbf{x})$, 如果 \mathbf{K} 有特征值为零的特征向量时, 这些特征向量不会影响主成分投影^[20], 因此将式(10)两边去掉因子 \mathbf{K} , 得到:

$$\mathbf{K} \mathbf{a} = m \lambda_i \mathbf{a}. \quad (11)$$

$\boldsymbol{\varphi}(\mathbf{x})$ 在特征向量 \mathbf{P}_i 上的投影为:

$$\mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})^T \mathbf{P}_i = \boldsymbol{\varphi}(\mathbf{x})^T \sum_{i=1}^m \mathbf{a}_i \boldsymbol{\varphi}(\mathbf{x}_i) = \sum_{i=1}^m \mathbf{a}_i \mathbf{K}(\mathbf{x}, \mathbf{x}_i). \quad (12)$$

故在核主成分空间中 T^2 统计量定义为:

$$T^2 = \mathbf{y}^T \boldsymbol{\Lambda}^{-1} \mathbf{y}. \quad (13)$$

式中 $\boldsymbol{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_z\}$ 。在时刻 t , 建立正常数据的 T^2 统计量 $T_{\text{normal}}^2(t)$, 此时, 构造时滞输入特性 $T_{\text{normal}}^2(t-1)$ 和时差输入特性 $T_{\text{normal}}^2(t) - T_{\text{normal}}^2(t-1)$, 并将其组合成增广矩阵:

$$T_{\text{normal}}^2 = [T_{\text{normal}}^2(t), T_{\text{normal}}^2(t-1), T_{\text{normal}}^2(t) - T_{\text{normal}}^2(t-1)]. \quad (14)$$

在故障操作条件下获取故障数据, 同样在核主成分空间中建立故障数据的增广矩阵:

$$T_{\text{fault}}^2 = [T_{\text{fault}}^2(t), T_{\text{fault}}^2(t-1), T_{\text{fault}}^2(t) - T_{\text{fault}}^2(t-1)]。 \quad (15)$$

将正常数据与故障数据的增广矩阵作为 SVM 模型的训练数据,定义为:

$$T_{\text{train}}^2 = [T_{\text{normal}}^2, T_{\text{fault}}^2]。 \quad (16)$$

将包含正常和故障数据的增广矩阵输入 SVM 模型,进行模型训练,获得判别函数。训练后的 SVM 能学习到正常和故障数据的特性,对数据进行正确分类。将测试数据输入模型,通过超平面划分,正常数据划分成一类,将其定义为标签 0;故障数据划分成另一类,将其定义为标签 1。

1.3 基于核主成分空间支持向量机的过程监视

基于核主成分空间支持向量机的过程监视主要包括离线模型建立和在线故障检测两个部分。

离线模型建立步骤为:

- 1) 收集正常操作条件下的历史数据集 X_1 和故障操作条件下的历史数据集 X_2 ;
- 2) 运用核主成分分析方法在核主成分空间分别建立正常和故障数据的 T^2 统计量;
- 3) 对时刻 t 的 T^2 统计量加入时滞输入特性 $T^2(t-1)$ 和时差输入特性 $T^2(t) - T^2(t-1)$, 根据式 (14)~(15) 分别构造增广矩阵 T_{normal}^2 和 T_{fault}^2 ;

- 4) 构造 SVM 训练数据集 T_{train}^2 , 并对该数据集贴标签, 正常数据定义为标签 0, 故障数据定义为标签 1;

- 5) 对 SVM 模型进行训练, 获得判别函数。

在线故障检测步骤为:

- 1) 对于测试数据集 x_{new} , 利用离线建模数据的均值和方差对其进行标准化处理;
- 2) 求出测试数据对训练数据的矩阵 K , 将数据投影到 KPCA 模型上, 在核主成分空间中计算统计量 T_{new}^2 ;
- 3) 加入时滞输入特性和时差输入特性后生成 SVM 模型测试数据集 T_{test}^2 ;
- 4) 通过 SVM 模型对测试数据进行分类, 实现对故障的检测。分类标签为 0 的数据为正常样本, 而标签为 1 的数据为故障样本。

2 仿真结果与分析

2.1 非线性数值例子

构造由两个变量组成的非线性数值例子, 生成模型如下:

$$\begin{cases} x_1 = t + e_1, \\ x_2 = t^2 + e_2. \end{cases} \quad (17)$$

式中, e_1 和 e_2 分别是服从均值为 0、标准差为 0.01 的正态分布的白噪声, t 服从 $[-2, 2]$ 的均匀分布。

由该模型产生 200 组正常数据, 并分别对变量 1 和变量 2 施加 $0.01 \times (i-100)$ 和 0.5 的阶跃故障, 产生 300 组故障数据。由于 SVM 模型训练需要加入正常和故障数据, 因此, 本例中由 200 组正常数据和 200 组故障数据共同组成训练数据集, 其余 100 组故障数据用于模型测试。PCA、KPCA 和 KPCA-LSSVM 方法的主成分个数由 85% 累计方差贡献率确定, SVM 方法中的惩罚参数 C 和核窗宽参数 g 通过寻优测试分别设置为 0.8 和 0.25。图 1 是样本散点图。图 2 是 5 种方法的故障检测对比图, 从图中可以看出 PCA 和 KPCA 两种方法的故障检测效果并不理想, 在控

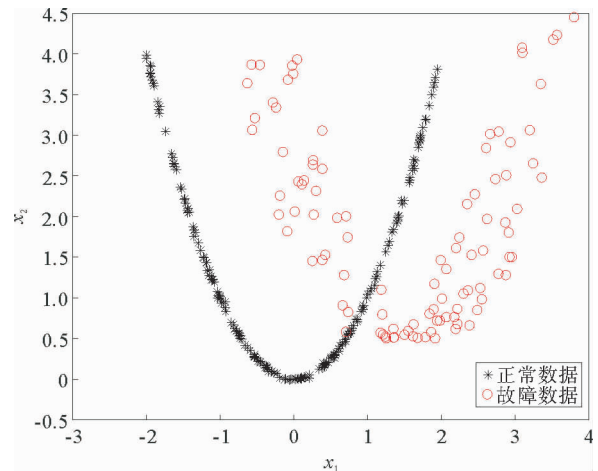
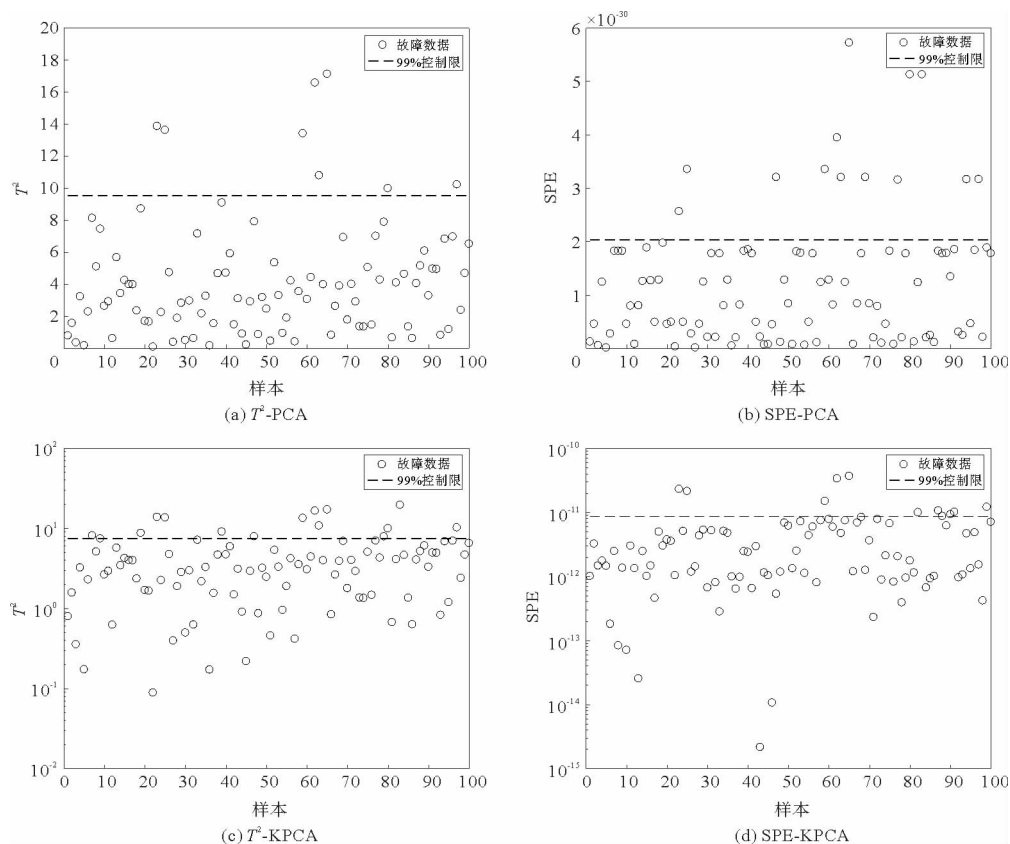


图 1 样本散点图

Fig. 1 Scatter plot of samples

制限上方的故障样本数较少,说明这两种方法对故障样本的检测效果较差。传统的 SVM 方法在训练时加入了故障数据,模型训练时能够有效学习到故障的变化特征,因此检测效果优于 PCA 和 KPCA。但是样本间的时刻相关性会影响传统 SVM 的检测效果,因此传统 SVM 方法不能得到满意的分类效果。传统 SVM 方法把标签为 1 的故障样本识别为标签 0 的正常样本数量多于本方法,即漏报样本较多,因此传统 SVM 方法对故障样本的识别能力不如本方法。在漏报的故障样本数量方面,KPCA-LSSVM 好于传统 SVM 方法,但比本研究方法差。本研究方法通过引入时滞和时差输入特性,在一定程度上增加矩阵的复杂度,去除样本间的自相关性,使数据特征更加显著,能够识别出更多的故障样本,检测效果优于传统的 SVM 和 KPCA-LSSVM 方法。

表 1 是 5 种方法故障检测率的对比。从表 1 中可知,PCA 的 T^2 、SPE 的检测率分别为 9% 和 13%,KPCA 的 T^2 和 SPE 的检测率分别为 14% 和 12%,传统 SVM 的检测率为 68%,KPCA-LSSVM 的检测率为 73%,而本研究方法的故障检测率为 87.88%。原因在于:PCA 缺乏对非线性数据的处理能力,难以得到满意的检测效果;KPCA 寻找方差差异最大化的方向投影,其残差空间和主成分空间在一定程度上会丢失部分数据特征,故不能获得较好的检测效果;而 SVM 具有一定的学习能力,对故障样本的分类准确性高于 PCA 和 KPCA;KPCA-LSSVM 结合两者的优点,检测效果略高于传统 SVM 方法;本研究方法通过引入时滞和时差输入特性,在核主成分空间上建立增广矩阵,增强输入矩阵的复杂度,去除样本间的自相关性,故障检测效果优于传统 SVM 和 KPCA-LSSVM 方法。综上,相较于其他 4 种方法,本研究方法的故障检测率最高。



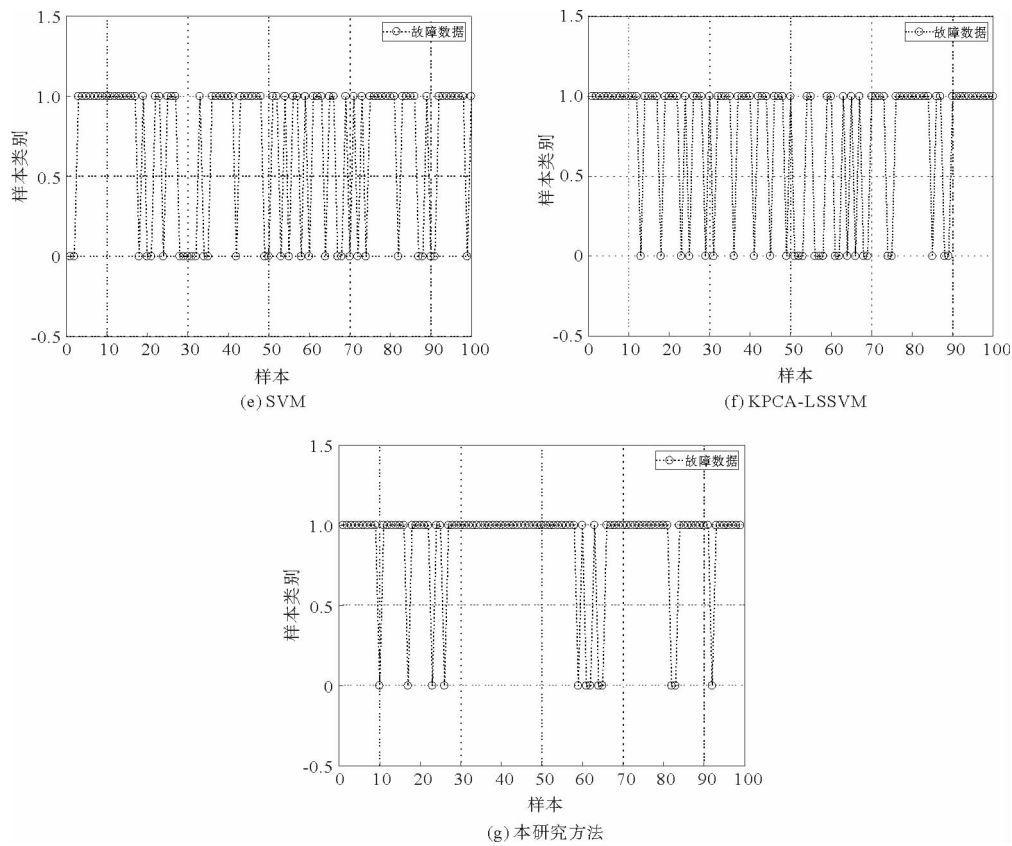


图 2 5 种方法的检测结果

Fig. 2 Detection results of five methods

表 1 5 种方法的故障检测率对比

Tab. 1 Comparison of fault detection rates of five methods

方法	PCA		KPCA		SVM	KPCA-LSSVM	本研究方法
	T^2	SPE	T^2	SPE			
故障检测率/%	9	13	14	12	68	73	87.88

2.2 田纳西-伊斯曼 (Tennessee Esatman, TE) 过程仿真

TE 过程仿真平台已成为国际上通用的工业过程模型仿真平台^[21-23],在故障检测和诊断领域得到广泛使用。TE 过程变量众多,其中 2 个气液放热反应会产生 2 种主产品,与 5 个主要操作单元等共同组成 TE 过程。利用 TE 过程模拟 21 种预编程故障,丰富多样的故障类型可真实反映实际工业过程中的众多问题。

在 TE 过程仿真中,以故障类型 3、8、9、10、12、13、15、16 和 19 为例,分别从每个故障类型中选取 500 组故障数据和 500 组正常数据共同组成训练数据集用于 SVM 模型的训练,另外选取 160 组故障数据作为测试数据。在本例中,PCA、KPCA、SVM、KPCA-LSSVM 和本研究方法的参数设置与非线性数值例子保持一致。图 3 是 4 种方法对故障 9 的检测结果,其中传统 SVM、KPCA-LSSVM 和本研究方法都将正常样本的标签定义为 0,故障样本的标签定义为 1。

由图 3 可知,KPCA-LSSVM 的检测效果优于 PCA、KPCA 和传统 SVM 方法,能够检测出更多的故障样本,这得益于 KPCA 方法对非线性特征的提取,但样本间时刻相关性仍影响故障检测的效果,而本研究方法能够有效地克服这一缺点,检测效果优于其他几种方法。

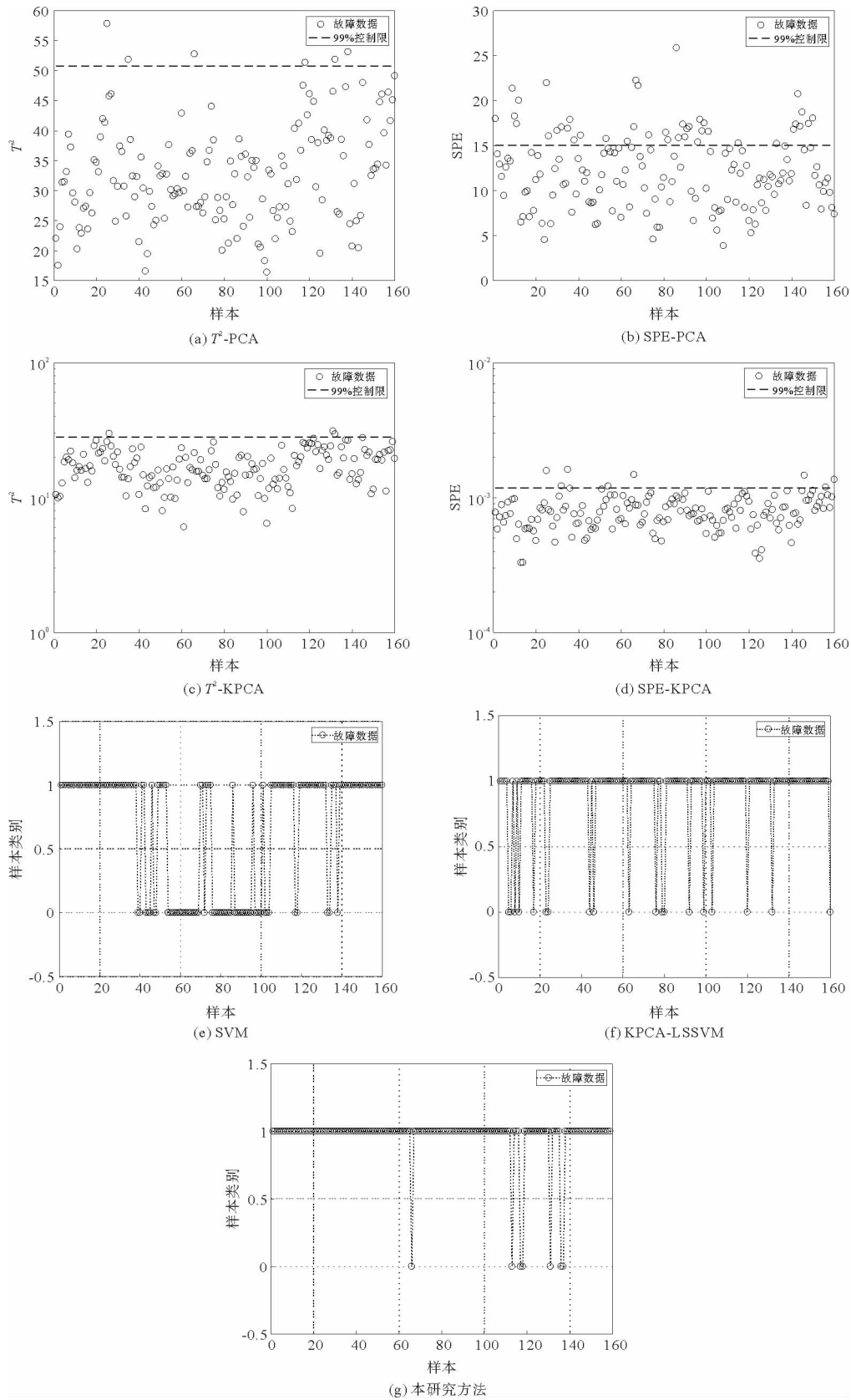


图3 5种方法对故障9的检测效果

Fig. 3 Detection results of five methods for Fault 9

表 2 是 5 种方法对 9 种不同类型故障的检测率对比。由表 2 可知,对于 9 种不同类型的故障,本研究方法的故障检测率比 PCA、KPCA、传统 SVM 和 KPCA-LSSVM 方法均有不同程度的提高,验证了该方法的有效性。

表 2 5 种方法的故障检测率对比

Tab. 2 Comparison of fault detection rates of five methods

故障类型	PCA		KPCA		SVM	KPCA-LSSVM	本研究方法
	T^2	SPE	T^2	SPE			
3	0	23.75	1.25	4.38	37.50	54.37	59.12
8	100	100	100	99.38	69.37	97.50	100
9	4.38	25.62	1.88	5.00	65.63	90.63	95.60
10	91.88	97.50	92.50	61.25	100	100	100
12	100	100	100	100	100	100	100
13	100	100	100	100	100	100	100
15	13.13	42.50	25.62	14.37	83.13	100	100
16	50.63	77.50	47.50	37.50	92.50	100	100
19	5.00	61.25	1.88	30.00	80.00	88.12	94.97

3 结论

本研究提出一种基于核主成分空间支持向量机的过程监视方法。运用 KPCA 方法在低维核主成分空间建立 T^2 统计量,并引入时滞输入特性和时差输入特性作为 SVM 模型的输入,通过增强模型输入矩阵的复杂度能够有效地降低样本间的自相关性,提高了 SVM 的故障检测性能。将本研究方法应用于非线性数值实例和 TE 过程仿真,仿真结果表明,与 PCA、KPCA、传统 SVM 以及 KPCA-LSSVM 方法相比,该方法能够有效地提高故障检测率。

参考文献:

- [1] 周东华,李钢,李元.数据驱动的工业过程故障检测与诊断技术:基于主元分析与偏最小二乘的方法[M].北京:科学出版社,2011:1-76.
- [2] 张成,郭青秀,冯立伟,等.基于局部近邻标准化和动态主元分析的故障检测策略[J].计算机应用,2018,38(9):2730-2734.
ZHANG Cheng, GUO Qingxiu, FENG Liwei, et al. Fault detection strategy based on local neighbor standardization and dynamic principal component analysis[J]. Journal of Computer Applications, 2018, 38(9): 2730-2734.
- [3] HAMADACHE M, LEE D. Principal component analysis based signal-to-noise ratio improvement for inchoate faulty signals; Application to ball bearing fault detection[J]. International Journal of Control Automation & Systems, 2017, 15(2): 1-12.
- [4] 李元,杨东昇,冯立伟,等.基于局部近邻标准化和主多项式算法的故障检测[J].山东科技大学学报(自然科学版),2020,39(3):76-83.
LI Yuan, YANG Dongsheng, FENG Liwei, et al. Fault detection based on local neighborhood standardization and principal polynomial analysis[J]. Journal of Shandong University of Science and Technology (Natural Science), 2020, 39(3): 76-83.
- [5] 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 Cheng, LI Yuan. 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.
- [6] YU J, QIN S J. Multimode process monitoring with Bayesian inference-based finite Gaussian mixture models[J]. IChe Journal, 2008, 54(7): 1811-1829.
- [7] SHIOKAWA Y, DATE Y, KIKUCHI J. Application of kernel principal component analysis and computational machine learning to exploration of metabolites strongly associated with diet[J/OL]. Scientific Reports, 2018, 8(1): 3426. DOI: 10.

1038/s41598-108-20121-w.

- [8] SOH W, KIM H, YUM B J. Application of kernel principal component analysis to multi-characteristic parameter design problems[J]. *Annals of Operations Research*, 2018, 263(1/2): 69-91.
- [9] JIA M, FEI C, WANG F, et al. On-line batch process monitoring using batch dynamic kernel principal component analysis[J]. *Chemometrics & Intelligent Laboratory Systems*, 2010, 101(2): 110-122.
- [10] 王亚君, 周岐. 基于多动态核 PCA 的统计过程监测策略研究[J]. *辽宁工业大学学报(自然科学版)*, 2012, 32(5): 295-298.
WANG Yajun, ZHOU Qi. Research on statistical process monitoring strategy based on multi-dynamic kernel PCA[J]. *Journal of Liaoning University of Technology(Natural Science Edition)*, 2012, 32(5): 295-298.
- [11] 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 Kun, DU Wenxia, LÜ Feng, et al. Fault detect method based on improved dynamic kernel principal component analysis[J]. *CIESC Journal*, 2019, 70(2): 716-722.
- [12] HSU C W, LIN C J. A comparison of methods for multiclass support vector machines[J]. *IEEE Transactions on Neural Networks*, 2002, 13(2): 415-425.
- [13] SHEN L, WANG H, XU L D, et al. Identity management based on PCA and SVM[J]. *Information Systems Frontiers*, 2016, 18(4): 711-716.
- [14] FAN R E, CHANG K W, HSIEH C J, et al. LIBLINEAR: A library for large linear classification[J]. *Journal of Machine Learning Research*, 2008, 9(9): 1871-1874.
- [15] ZHANG Y. Enhanced statistical analysis of nonlinear processes using KPCA, KICA and SVM[J]. *Chemical Engineering Science*, 2009, 64(5): 801-811.
- [16] YU J. A support vector clustering-based probabilistic method for unsupervised fault detection and classification of complex chemical processes using unlabeled data[J]. *AIChE Journal*, 2013, 59(2): 407-419.
- [17] HSU C C, CHEN M C, CHEN L S. Integrating independent component analysis and support vector machine for multivariate process monitoring[J]. *Computers & Industrial Engineering*, 2010, 59(1): 145-156.
- [18] 王东, 王新晴, 闫凤国, 等. 基于 KPCA-LSSVM 的发动机 PT 燃油系统故障诊断[J]. *解放军理工大学学报(自然科学版)*, 2016, 17(5): 499-504.
WANG Dong, WANG Xinqing, YAN Fengguo, et al. Fault diagnosis of PT fuel system based on KPCA-LSSVM[J]. *Journal of PLA University of Science and Technology(Natural Science Edition)*, 2016, 17(5): 499-504.
- [19] MA H H, HU Y, SHI H B. Fault detection and identification based on the neighborhood standardized local outlier factor method[J]. *Industrial & Engineering Chemistry Research*, 2013, 52(6): 2389-2402.
- [20] SCHÖLKOPF B, SMOLA, MJULLER K. Nonlinear component analysis as a kernel eigenvalue problem[J]. *Neural Computation*, 1998, 10(5): 1299-1319.
- [21] DOWNS J J, VOGEL E F. A plant-wide industrial process control problem[J]. *Computers & Chemical Engineering*, 1993, 17(3): 245-255.
- [22] MCAVOY T J, YE N. Base control for the Tennessee Eastman problem[J]. *Computers & Chemical Engineering*, 1994, 18(5): 383-413.
- [23] YIN S, DING S X, HAGHANI A, 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.

(责任编辑: 齐敏华)