

多切面分类改进独立成份与支持向量机 集成故障诊断方法

薄翠梅¹, 柏杨进¹, 杨海荣², 张广明¹

(1. 南京工业大学 自动化与电气工程学院, 江苏 南京 210009; 2. 上海宝信软件股份有限公司, 上海 201203)

摘要: 本文采用多切面分类方法改进独立成份(ICA)与支持向量机(SVM)集成诊断方法. 在高维独立成份特征空间中采用多切面分类方法在不同切面上分别建立SVM故障分类模型. 对不同切面的分类情况进行故障识别, 改善ICA-SVM集成故障诊断性能. 将ICA-MSVM集成故障诊断方法对动态执行器基准平台(DAMADICS)的19种阀门故障模式进行仿真验证, 结果表明改进的ICA-MSVM方法有效地提高了故障诊断精度.

关键词: 多切面分类; 独立成分分析; 支持向量机; 故障辨识; 执行器基准平台

中图分类号: TP273 文献标识码: A

Multi-section classification improving integrated fault diagnosis method based on independent component analysis and support-vector-machines

BO Cui-mei¹, BAI Yang-jin¹, YANG Hai-rong², ZHANG Guang-ming¹

(1. School of Automation & Electrical Engineering, Nanjing University of Technology, Nanjing Jiangsu 210009, China;
2. Shanghai Baosight Software Limited Liability Company, Shanghai 201203, China)

Abstract: The integrated diagnosis method of independent component analysis (ICA) and support-vector-machines (SVM) is improved by multi-section classification. Fault classification model of SVM is designed for each section in the high dimensional characteristic space. By diagnosing the fault type in different section, we improve the ICA-SVM fault diagnosis performance. This method has been applied to diagnose 19 types of valve failures on the dynamic actuator reference platform (DAMADICS). Simulation results show that the ICA-MSVM fault diagnosis method based on multisection classification effectively improves the accuracy of fault diagnosis.

Key words: multisession classification; independent component analysis (ICA); support-vector-machine (SVM); fault diagnosis; actuator reference platform (DAMADICS)

1 引言(Introduction)

多元统计过程控制在工业过程监控中得到了广泛应用, 且在动态性、非线性、多尺度性等方面被不断改进, 并与多种智能方法相结合, 提高过程监控性能^[1]. 独立成分分析方法(independent component analysis, ICA)在过程数据分析处理上, 能更有效地利用非高斯信息和高阶统计特性, 提取相互独立的过程数据特征信息^[2]. 目前在过程监控方面已取得较好的效果^[3-5], 且在故障识别与隔离方法方面也取得一些新成果^[6-7]. 采用ICA-SVM集成故障诊断方法, 由于SVM分类器的分类精度随着独立成份维数增加导致分类精度下降问题, 本文在文献[7]的基础上, 采用多投影切面分类方法将高维ICA特征空间中的多维分类问题转化为多投影切面的二维分类问

题, 使得SVM分类精度增加, 从而故障诊断的精度被明显提高.

2 ICA过程监测算法(Process monitoring based on ICA arithmetic)

2.1 ICA算法(ICA arithmetic)

独立成分分析的基本思想是假设过程测量信号是由一些互相独立的信号源、过程噪声及干扰混合叠加而成的, 按照信息论准则从过程测量信息中分离出尽可能互相独立的特征信号的过程^[8]. 设数据混合矩阵 $\mathbf{X} = [\mathbf{x}(1) \ \mathbf{x}(2) \ \cdots \ \mathbf{x}(l)] \in \mathbb{R}^{l \times n}$, 分别为 m ($m < l$) 个独立源信号 $\mathbf{S} = [\mathbf{s}(1) \ \mathbf{s}(2) \ \cdots \ \mathbf{s}(m)] \in \mathbb{R}^{m \times n}$ 线性组合, 即

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{S} + \mathbf{E}, \quad (1)$$

式中: $\mathbf{A} = [a_1 \ a_2 \ \cdots \ a_m] \in \mathbb{R}^{l \times m}$ 为混叠系数矩阵, $\mathbf{E} \in \mathbb{R}^{l \times n}$ 为残差矩阵. ICA算法是在独立源 \mathbf{S} 和混叠矩阵 \mathbf{A} 都未知情况下, 找到分离矩阵 \mathbf{W} 从观测信号 \mathbf{X} 中分离源信号 \mathbf{S} .

$$\hat{\mathbf{S}} = \mathbf{W} \cdot \mathbf{X} = \mathbf{W} \cdot \mathbf{A} \cdot \mathbf{S} \approx \mathbf{S}, \quad (2)$$

$\hat{\mathbf{S}}$ 为源信号的估计值. 式(2)中 \mathbf{W} 未知, 可通过计算变量微分熵的大小确定独立成份的个数. 本文采用快速ICA算法, 其具体算法请参见文献[9].

2.2 特征提取与过程监控(Feature extracting and process monitoring)

如图1所示, 选取正常工况下的测量数据 $\mathbf{X} \in \mathbb{R}^{l \times n}$, 并对其进行了标准化和白化处理, 采用快速ICA算法提取正常样本特征, 确定式(2)中独立矩阵成分 \mathbf{S} 和分离矩阵 \mathbf{W} . 依据负熵的最小化原则选择独立成分维数 d , 选取前 d 个独立成分作为特征向量构成正常工况的ICA模型:

$$\mathbf{S}_d = \mathbf{W}_d \cdot \mathbf{X}, \quad \mathbf{S}_e = \mathbf{W}_e \cdot \mathbf{X}, \quad (3)$$

\mathbf{W}_d 表示 \mathbf{W} 矩阵前 d 行向量; \mathbf{W}_e 表示 \mathbf{W} 矩阵剩余部分. 分别建立 I^2 , SPE统计量, 并通过核密度估计的方法确定其对应的置信限, 对过程进行监测. 定义 k 时刻 I^2 统计量可表示为

$$I^2(k) = \hat{\mathbf{S}}_d^T(k) \hat{\mathbf{S}}_d(k), \quad (4)$$

k 时刻SPE统计量为

$$\text{SPE}(k) = (\mathbf{X}(k) - \hat{\mathbf{X}}(k))^T (\mathbf{X}(k) - \hat{\mathbf{X}}(k)), \quad (5)$$

采用核密度函数估计计算统计量界限值表示为^[5]

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left\{\frac{x-x_i}{h}\right\}, \quad (6)$$

x 为控制界限值, x_i 为从样本数据集观测值, h 为平滑参数, K 为核函数, n 为样本数.

3 ICA-MSVM集成故障诊断算法(ICA-MSVM fault integration diagnosis methods)

3.1 高维特征空间的多投影切面原理(Multi-section of high dimensional space)

复杂工业故障传播由于受工业设备的限制, 在故障传播中主要表现在与故障源关联紧密的少数几个独立变量上, 对于其他独立变量的影响比较微弱, 具有“局部性”. 对于不同类型的故障, 其对不同独立变量的影响程度不同, 甚至有些比较微小的故障, 只影响少数的1-2个独立变量, 因此对于不同种类型故障分别在不同投影切面, 建立多个SVM故障分类模型, 从局部反应故障特征更加明显, 用多个局部特征共同识别多种工业故障. 本文对文献[7]中提出的ICA-SVM集成故障诊断方法采用高维投影空间多切面分类方法进行改进研究.

高维投影空间多切面分类改进的ICA-MSVM集

成故障诊断方法包括离线训练(特征提取与建模部分, 如图1所示)和在线故障诊断(如图2所示)两部分. 首先通过独立成分算法建立正常工况数据的独立成份高维特征空间. 分别将正常样本和故障样本进行独立向量提取, 将获得特征独立向量分别在ICA高维特征空间的多切面上进行投影; 然后通过对比分析各故障在不同投影切面上的故障特征投影数据分布, 分别采用二叉树SVM多分类方法建立每个切面上多种故障分类模型. 在线故障诊断应用时, 当故障一旦发生, 通过空间转换矩阵可以将故障数据投影到ICA高维特征空间和不同切面分布特征, 可以得出投影点是否在分类范围内, 根据支持向量机多分类模型输出, 采用专家决策定位和分离故障源, 提高SVM在高维特征空间多分类精度.

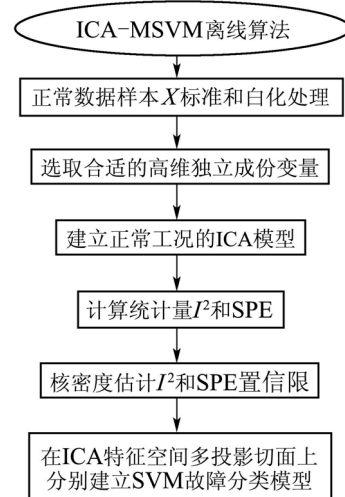


图1 ICA-MSVM离线算法

Fig. 1 Off-line arithmetic of ICA-MSVM

3.2 支持向量机多分类算法(Multi-classification arithmetic based on SVM)

故障诊断问题可以转化为对数据的多类别分类问题. 在实际应用中, 某些重要样本正确分类的要求高, 而某些样本正确分类的要求低, 对每个采样点数据通过采用不同的惩罚系数, 得到更准确的分类, 因此常采用加权支持向量机进行故障分类研究^[10]. 本运用SVM对于多个故障类型模式进行分类, 必须构造多元分类器才能对多个故障进行故障诊断. 本文采用二叉树多类分类算法, 具体二叉树SVM分类算法参见文献[7].

独立成份分析特征信号提取方法可根据微分熵提取数据样本的独立成份, 构成ICA特征空间. 若提取的独立成份维数过多时, 在高维特征子空间运用支持向量机分类算法分类时会遇到两个问题: 一是独立成份维数的增加会增加支持向量机的复杂度, 二是独立成份维数增加将有些主要的故障特征被“淹没”在高维投影空间中, 使得支持向量机分类

精度下降. 本文改进SVM分类模型的建立过程, 根据不同故障类型只对少数几个独立变量影响较大, 由影响较大的独立变量组成投影切面采用SVM算法建立故障分类模型, 可有效提取故障特征, 缩短在线故障诊断时间.

3.3 ICA-MSVM在线故障诊断算法(Fault online diagnosis method based on ICA-MSVM)

多切面ICA-MSVM集成在线故障诊断算法流程如图2所示. 获取新测量数据 $\mathbf{X}_{new}(k) = [x_1 \ x_2 \ \dots \ x_l]^T \in \mathbb{R}^l$, 并进行标准化和白化处理后, 通过混合矩阵 \mathbf{W}_d 和 \mathbf{W}_e 可计算新独立元矢量的估计值:

$$\begin{cases} \hat{\mathbf{S}}_{dnew}(k) = \mathbf{W}_d \mathbf{X}_{new}(k), \\ \hat{\mathbf{S}}_{enew}(k) = \mathbf{W}_e \mathbf{X}_{new}(k), \end{cases} \quad (7)$$

分别利用式(4)和(5)计算新的监控测量数据统计量 I^2 , SPE值, 监控两个统计量的变化是否超过给定的控制置信限, 若超过控制置信限则说明检测出故障.

如检测到故障, 分别将故障ICA独立元矢量在高维ICA特征空间上计算不同切面故障分类SVM效果, 采用专家决策表或举手表决法对多个SVM分类模型输出之进行统计分析, 对上述可能的故障源进行逐一故障分类诊断, 直至诊断出故障源; 输出并显示故障.

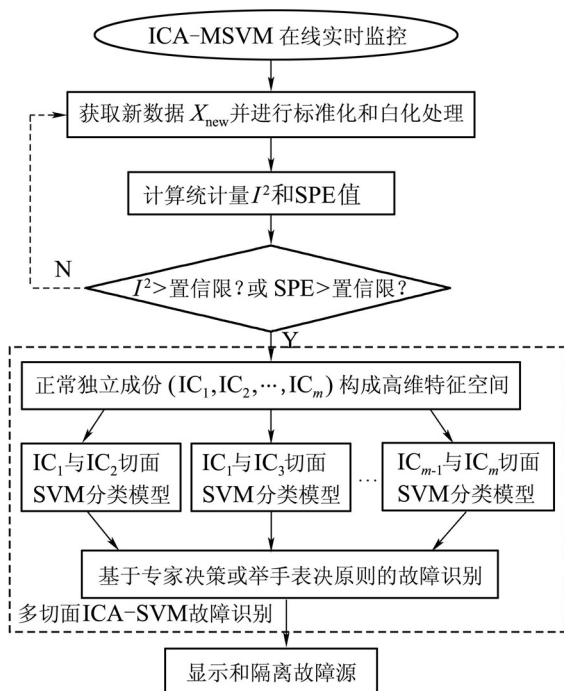


图 2 多切面ICA-MSVM在线故障诊断流程图

Fig. 2 Fault online diagnosis flow chart based on ICA-MSVM

4 仿真(Simulation)

本节研究对象在MATLAB/SIMULINK环境下运行的DAMADICS执行器故障基准仿真平台^[11], 如

图3所示. 故障基准仿真平台分别在调节机构、执行机构、定位器和阀门外部分别设置了19种不同类型故障模式, 如表1所示. 故障的强度可以任意设置, 故障触发模式可以为突变故障或者为缓变故障^[12].

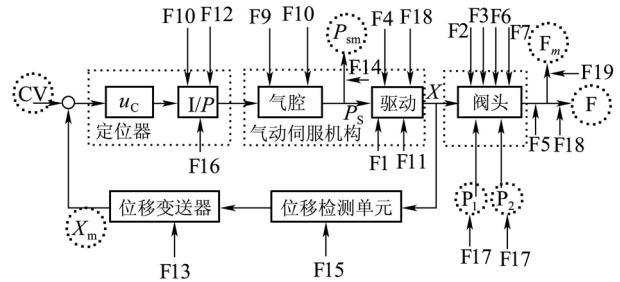


图 3 阀门故障设置系统结构示意图

Fig. 3 The structure sketch map of DAMADICS actuator FDI benchmark

表 1 DAMADICS平台阀门故障描述

Table 1 Fault description of DAMADICS actuator FDI benchmark

故障	故障描述	故障类型
F1	阀门堵塞故障	突变故障
F2	阀塞或阀座下沉故障	衰减故障
F3	阀塞或阀座腐蚀故障	衰减故障
F4	阀或轴承摩擦增加故障	衰减故障
F5	外部泄漏故障	衰减故障
F6	内部泄漏故障	衰减故障
F7	载体过热蒸发故障	突变故障
F8	电机活塞杆扭曲故障	突变故障
F9	机架或端子松动故障	衰减故障
F10	膜头穿孔故障	衰减故障
F11	弹簧弹性故障	衰减故障
F12	电-气转换器故障	突变故障
F13	杆位移传感器故障	衰减故障
F14	压力传感器故障	突变故障
F15	定位器反馈故障	突变故障
F16	压力供给下降故障	突变故障
F17	不可预计压差变化故障	衰减故障
F18	旁路阀故障	突变故障
F19	流量传感器故障	突变故障

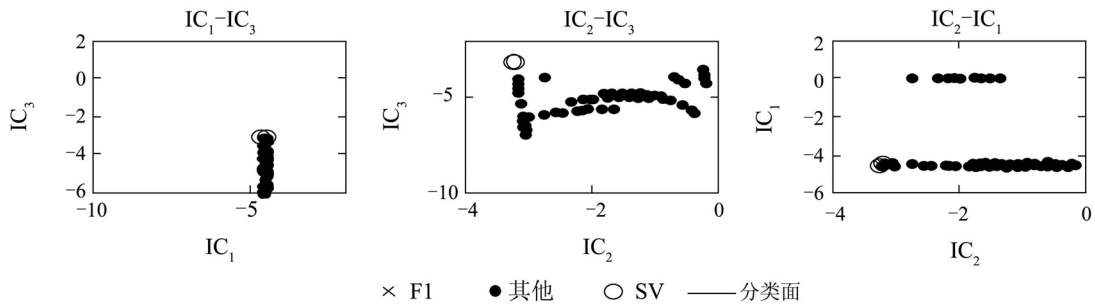
本文将来自Lublin Sugar Factory某一天一小时内实际无故障工业数据导入到DAMADICS模型中进行仿真, 并设置19种不同类型和强度的故障, 运行每种故障得到100个故障样本, 共获取2000个数据样本. 选取控制输出 CV 、输出气压 P_{sm} 、阀杆位移设定值 X 、阀杆位移 X_m 、控制阀输出流量 F 、控制阀输出流量测量值 F_m 、阀前压力 P_1 和阀后压力 P_2 作为监控变量. 将采集正常数据样本线进行标准化和主元分析, 选取 X, X_m, F, P_1 和 P_2 监控变量组成数

据样本进行ICA数据特征提取. 依据微分熵的大小选择3个独立变量.

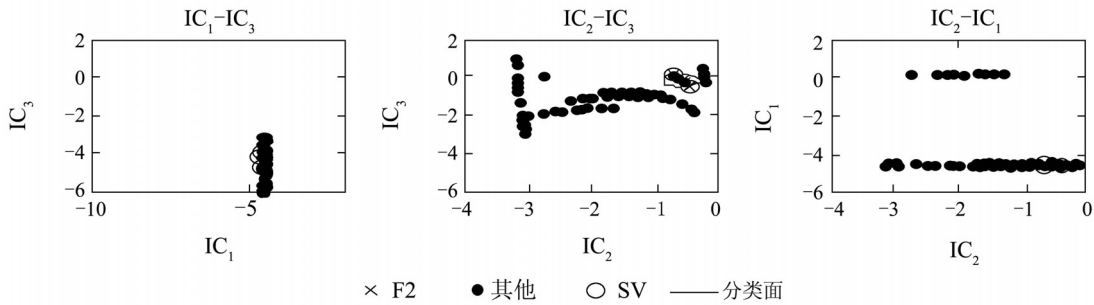
选取可检测的12种故障(F1, F2, F3, F6, F7, F9, F10, F13, F16, F17, F18, F19)作为故障识别的研究对象, 分析故障特征与监测变量的关系有针对性的选取检测数据提取故障特征, 分别在3维特征空间3个不同切面分别采用二叉树SVM多分类方法建立故障SVM分类模型. 本文所选用的核函数为径向基函数 $K(x, x_i) = \exp(-\|x - x_i\|^2/C)$, 以训练确定参数收缩范围, $C \in [10^{-10}, 10^{10}]$, $\sigma \in [0, 1]$,

文中的 $\varepsilon = 0.01$ 不变. 按照一定的步长搜索最优值, 当 $C = 0, \sigma = 0.6$ 时支持向量机的分类效果最好.

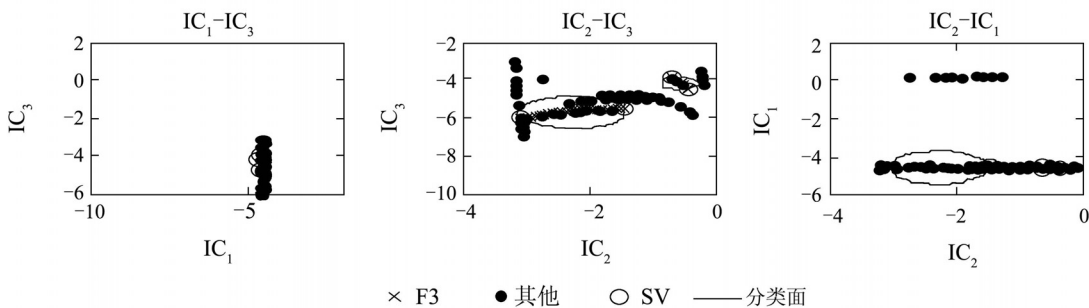
在故障诊断时采用的专家决策表规则即3个SVM分类器若有2个诊断出某一故障就可以诊断是发生了该故障. 支持向量分类器在ICA特征空间3维切面SVM分类效果如图4所示. 对于新的测试样本, 首先通过空间转换矩阵可以将故障数据投影到ICA高维特征空间, 将独立成份分别在不同切面投影, 并导入到训练好的3维空间多切面SVM支持向量分类器中进行故障模式分类, 实现故障的在线诊断.



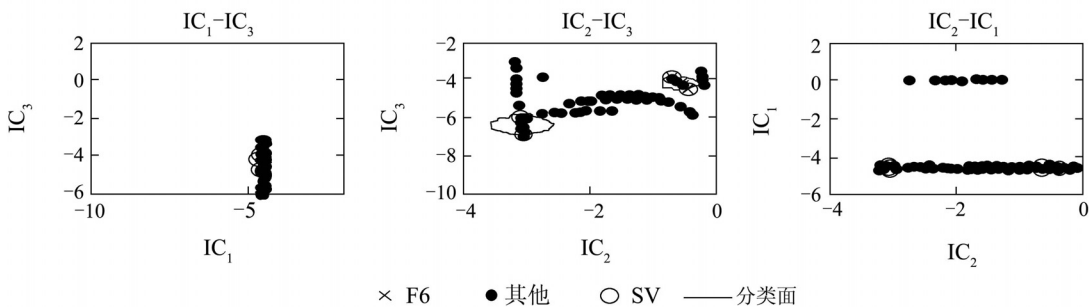
(a) 故障F1多切面分类效果



(b) 故障F2多切面分类效果



(c) 故障F3多切面分类效果



(d) 故障F6多切面分类效果

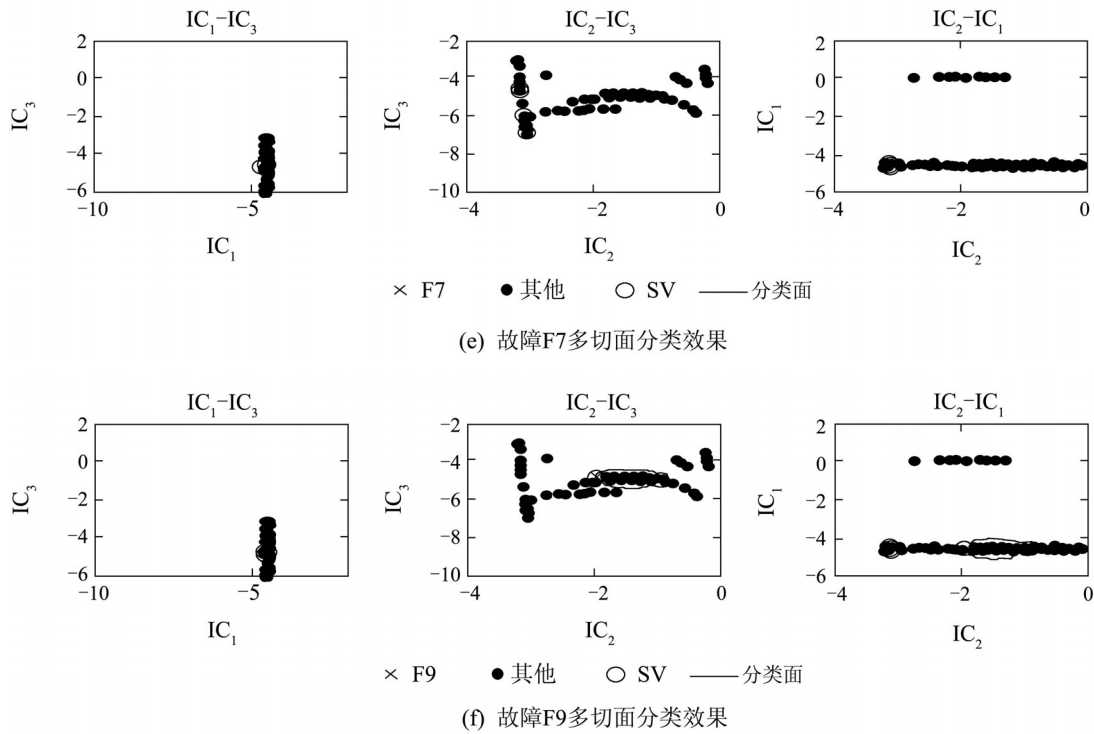


图 4 ICA特征空间多切面ICA-MSVM故障分类效果图

Fig. 4 Fault classification effect of ICA-MSVM in three dimensional ICA feature space

针对上述12种可检测故障, 分别采用本文提出的多切面ICA-MSVM集成故障诊断方法和文[7]提出的ICA-SVM诊断方法进行对比研究, 故障诊断精度(正确诊断率)如表2所示。

表2中, ICA-SVM故障诊断对故障F9和故障

F17诊断效果不好, 由于只提取了2维的独立分量没有提取更多的故障信息, 改进后方法提取了3维的独立变量并在每个2维空间中建立支持向量机模型进行诊断保留了更多的故障信息, 仿真实验表明, 该方法可以较好识别故障。

表 2 ICA-MSVM集成故障诊断算法仿真结果对比分析

Table 2 Comparative analysis of simulation with ICA-MSVM and ICA-SVM

诊断正确率/%	F0	F1	F2	F3	F6	F7	F9	F10	F13	F16	F17	F18	F19
ICA-SVM	97.3	89.2	91.8	80.4	28.1	60.2	5.8	90.6	32.1	60.2	10.4	82.1	86.9
ICA-MSVM	98.6	94.5	92.3	94.1	83.6	81.7	70.1	92.1	76.9	90.1	80.4	91.1	87.3

5 结论(Conclusion)

本文采用高维投影空间多切面分类方法改进的ICA-SVM集成故障诊断的性能。由于支持向量多分类器在二维平面分类效果较好, 且可以直观显示其数据分类效果。文中采用高维投影空间多切面分类方法将高维独立成份投影空间按照不同切面将多维分类问题转化为二维分类问题, 使得SVM分类精度增加, 提高故障诊断精度。针对DAMADICS基准平台19种阀门故障模式, 进行改进前后的ICA-SVM故障诊断方法仿真研究, 仿真结果表明改进的多切面ICA-SVM故障诊断方法有效提高了故障诊断精度。

参考文献(References):

- [1] WANG H, CHAI T Y, DING J L, et al. Data driven fault diagnosis and fault tolerant control: some advances and possible new directions[J]. *Acta Automatica Sinica*, 2009, 35(6): 739 – 747.
- [2] 杨沐青, 李勇, 胡德文. 独立成分分析方法综述[J]. *自动化学报*, 2002, 28(5): 762 – 772.
(YANG Zhuqing, LI Yong, HU Dewen. Independent component analysis: a survey[J]. *Acta Automatica Sinica*, 2002, 28(5): 762 – 772.)
- [3] 张沐光, 宋执环. 一种基于独立元贡献度的子空间故障检测方法[J]. *控制理论与应用*, 2010, 27(3): 296 – 302.
(ZHANG Muguang, SONG Zhihuan. Subspace fault detection method based on independent component contribution[J]. *Control Theory & Applications*, 2010, 27(3): 296 – 302.)
- [4] LEE J M, QIN S J, LEE I B. Fault detection and diagnosis based on modified independent component analysis[J]. *AIChE Journal*, 2006, 52(10): 3501 – 3514.

- [5] GE Z Q, SONG Z H. Process monitoring based on independent component analysis-principal component analysis (ICA-PCA) and similarity factors[J]. *Industrial & Engineering Chemistry Research*, 2007, 46(7): 2054 – 2063.
- [6] 谢磊, 刘雪芹, 张建明, 等. 基于NGPP-SVDD的非高斯过程监控及其应用研究[J]. *自动化学报*, 2009, 35(1): 107 – 112.
(XIE Lei, LIU Xueqin, ZHANG Jianming, et al. Non-gaussian process monitoring based on NGPP-SVDD[J]. *Acta Automatic Sinica*, 2009, 35(1): 107 – 112.)
- [7] BO C M, QIAO X, ZHANG G G, et al. An integrated method of independent component analysis and support vector machines for industry distillation process monitoring[J]. *Journal of Process Control*, 2010, 20(10): 1133 – 1140.
- [8] HYVARINEN A, OJA E. Independent component analysis: algorithms and applications[J]. *Neural Networks*, 2000, 13(4/5): 411 – 430.
- [9] CICHOCKI A, AMARI S. *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*[M]. Chichester, England: John Wiley & Sons, 2003.
- [10] 袁胜发, 褚福磊. 支持向量机及其在机械故障诊断中的应用[J]. *振动与冲击*, 2007, 26(11): 29 – 35.
(YUAN Shengfa, CHU Fulei. Support vector machines and its application in machine fault diagnosis[J]. *Journal of Vibration and Shock*, 2007, 26(11): 29 – 35.)
- [11] BARTYS M, PATTON R, SYFERT M, et al. Introduction to the DAMADICS actuator FDI benchmark study[J]. *Control Engineering Practice*, 2006, 14(6): 577 – 596.
- [12] KOS J M, BARTYS M, RZEPIEJIEWSKIA P. Actuator fault distinguishability study for the damadics benchmark problem control[J]. *Control Engineering Practice*, 2006, 14(6): 645 – 652.

作者简介:

薄翠梅 (1973—), 女, 副教授, 研究方向为复杂过程系统建模、优化控制与过程监控, E-mail: lj.bcm@163.com;

柏杨进 (1985—), 男, 硕士研究生, 研究方向为化工过程系统建模与优化控制, E-mail: byj0911@163.com;

杨海荣 (1982—), 男, 工程师, 研究方向为工业过程建模与过程监控技术, E-mail: yhrnasa@163.com;

张广明 (1965—), 男, 教授, 研究方向为工业过程智能控制与监控, E-mail: zgmchina@163.com.