

# 基于 LFPCDA 的转子故障数据集降维算法<sup>\*</sup>

陈芳军<sup>1</sup>, 赵荣珍<sup>1,2</sup>, 邓林峰<sup>1</sup>

(1. 兰州理工大学机电工程学院 兰州, 730050)

(2. 广州理工学院智能制造与电气工程学院 广东, 510540)

**摘要** 针对旋转机械高维故障数据集中特征属性冗余导致的故障分类困难问题,提出一种基于局部 Fisher 主成分判别分析(local Fisher principal component discriminant analysis, 简称 LFPCDA)的故障数据集降维算法。首先,利用 Laplacian 得分算法过滤高维故障特征集中冗余特征,并将主成分计算融入局部 Fisher 判别分析(local Fisher discriminant analysis, 简称 LFDA)中,自适应地选取最能反映故障本质的主成分来构成投影矩阵,得到低维特征子集;其次,将低维特征子集输入 K 近邻(K-nearest neighbor, 简称 KNN)分类器中进行故障模式辨识;最后,使用双跨转子实验台模拟的转子故障数据集对所提算法进行验证,并与其他几种典型降维算法进行对比。结果表明:所提算法可剔除高维故障数据集中冗余信息,且保留特征的主要成分,使得故障类别之间的差异性更加突出,从而达到提高故障模式识别准确率的效果。所提算法可为转子故障智能决策技术提供数据降维处理的理论参考依据。

**关键词** 故障诊断;局部 Fisher 主成分判别分析;主成分计算;降维

**中图分类号** TP18;TH165

## 引言

随着计算机、大数据和工业互联网等技术的快速发展,工业智能化已成为我国流程工业、装备制造业发展的重要趋势<sup>[1-2]</sup>。机械设备在运行过程中的状态监测数据量与日俱增,促使相关智能故障诊断技术必须加快迈入“工业大数据”时代<sup>[3-5]</sup>。因此,如何从机械设备在生产过程中产生的低价值密度的海量监测信息中挖掘出有价值的故障状态特征和决策知识,并开发和利用这种大数据资源,以及实现智能决策技术,已成为当前基于数据驱动的智能旋转机械故障诊断研究领域关注的前沿热点问题。作为旋转机械关键部件的转子系统,其对保障石化行业动力机组的安全高效运行发挥着显著作用,因而对转子系统开展监测与诊断研究具有重要科学意义。

近年来,各种维数约简算法相继涌现,主要分为传统的特征降维算法与基于深度神经网络技术的数据降维算法。一般来说,后者在训练网络时需要大量样本数据,这增加了网络训练时间和难度。在特定领域与样本数据量较少时,采用传统降维算法可达到同样效果,且能降低成本。因此,对传统降维算法的研究仍然具有工程价值。故障特征属性的高维性和冗余性依旧是困扰数据分类与辨识技术向前发展的关键科学问题<sup>[6]</sup>。目前,在典型的降维算法中,

线性降维算法主要有主成分分析(principal component analysis, 简称 PCA)<sup>[7]</sup>和线性判别分析(linear discriminant analysis, 简称 LDA)<sup>[8]</sup>等。PCA 是通过选择贡献率趋大的几个主成分来获得最佳投影矩阵,又因该类算法无监督的特性,使得有价值的故障状态信息无法最大化保留,导致降维效果不理想<sup>[9-10]</sup>。

LDA 是一种有监督的降维算法,其目标函数为类间散度与类内散度比值的形式,可通过最大化目标函数的算法求得最佳投影矩阵,从而使得低维特征保持较高的可分性。然而,在 LDA 的降维过程中,因高维故障数据集中通常含有大量的相关信息和冗余特征,导致构建的投影矩阵仍存在一定误差,直接影响着算法的降维效果。此外,该算法确定的目标维数为经验所得,即类别数减 1,因而具有一定的盲目性。主成分计算是 PCA 中控制前几个主成分占有所有成分比例的一种算法<sup>[11]</sup>,可在满足选定贡献率的情况下剔除特征向量中影响降维效果的非主要成分,从而达到提高故障辨识准确率的效果。

基于上述分析,本研究拟在 LFDA 中融入主成分计算算法,提出一种能达到主成分自适应选取的 LFPCDA 算法,为剔除高维故障数据集中冗余特征属性,减小降维过程的盲目性以及优化目标维数的参数选择提供理论参考依据。

<sup>\*</sup> 国家自然科学基金资助项目(62241308, 51675253);甘肃省优秀研究生“创新之星”资助项目(2022CXZX-415)

收稿日期:2022-11-28;修回日期:2023-02-18

## 1 相关原理简介

### 1.1 Laplacian得分算法

Laplacian得分算法<sup>[12]</sup>根据特征的局部保持能力来衡量特征的重要程度。假设在 $n$ 个样本中,样本 $x_i$ 是样本 $x_j$ 的 $k$ 个最近邻之一,第 $r$ 个特征 $f_r$ 的Laplacian得分定义为

$$\text{Laplacian}(f_r) = \frac{\sum_{i,j} (f_{r_i} - f_{r_j})^2 \omega_{ij}}{\text{Var}(f_r)} \quad (1)$$

其中: $f_{r_i}$ 为样本 $i$ 在特征 $f_r$ 下的取值; $\omega_{ij}$ 为样本 $x_i$ 与样本 $x_j$ 的相似度; $\text{Var}(f_r)$ 为特征 $f_r$ 的方差。

式(1)中, $\omega_{ij}$ 用于刻画数据局部结构,若两近邻样本点距离较近,则其对应权重较大;若两近邻样本点距离较远,则其对应权重较小。具体为

$$\omega_{ij} = \begin{cases} 1 \text{ 或 } \exp(-\frac{\|x_i - x_j\|^2}{t}) & (x_j \text{ 为 } x_i \text{ 的近邻点}) \\ 0 & (\text{其他}) \end{cases} \quad (2)$$

其中: $t$ 为权重参数。

对于每个特征, $\text{Laplacian}(f_r)$ 表示其保持局部结构不变。由式(1)可知,Laplacian得分越低,表示该特征保持局部结构的能力越强。

### 1.2 LFDA算法

LFDA算法的目标之一是减小类内数据的分散程度。在分离不同类时,若两个类包含多个不同聚类时,Fisher判别向量可能不是最优方向。由于同一类的数据不会遵循多元高斯分布,但会表现出类间多模态性,这使得常规的Fisher判别分析(Fisher discriminant analysis,简称FDA)算法的类间可分性效果不理想。FDA算法的不足之处在于其忽略了数据的局部特征,而只专注于全局特征。为克服FDA算法的不足,文献[13]提出了一种LFDA算法,该算法可使类间分离的同时又能保持类内结构。

局部类间离散矩阵 $\tilde{S}_b$ 及类内离散矩阵 $\tilde{S}_w$ 定义为

$$\tilde{S}_b = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \tilde{W}_b (\mathbf{Q}_i - \mathbf{Q}_j)(\mathbf{Q}_i - \mathbf{Q}_j)^T \quad (3)$$

$$\tilde{S}_w = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \tilde{W}_w (\mathbf{Q}_i - \mathbf{Q}_j)(\mathbf{Q}_i - \mathbf{Q}_j)^T \quad (4)$$

其中: $\mathbf{Q}_i = [q_1, q_2, \dots, q_m]$ ,为 $m$ 维空间 $R^m$ 中的 $n$ 个训练样本, $i=1, 2, \dots, n$ 。

运用局部类间散布矩阵和类内散布矩阵进行广义特征值分析求得投影矩阵,即

$$\tilde{S}_b \tilde{\mathbf{P}} = \tilde{\lambda} \tilde{S}_w \tilde{\mathbf{P}} \quad (5)$$

式(5)中,与特征值 $\tilde{\lambda}$ 对应的特征向量 $\tilde{\mathbf{P}}$ 是LFDA的方向。通过基于距离的加权矩阵,LFDA可有效保留每类的局部多模态性,同时通过广义特征值分解使类间可分离性趋大。

### 1.3 主成分计算

主成分计算算法是PCA中确定主成分占有成分比例的一种算法,其根据累积特征贡献率 $\rho$ 的大小,选择特征向量中的主要成分。 $\rho$ 越大,表明所选主要成分的个数越多,从而保留的故障特征信息量越多。 $\rho$ 可定义为

$$\rho = \frac{\sum_{i=1}^s \lambda_i}{\sum_{i=1}^r \lambda_i} \quad (s=1, 2, \dots, r) \quad (6)$$

其中: $\lambda_i$ 为特征值; $s$ 为满足贡献率要求的特征值个数; $r$ 为所有特征值数目。

上述3种算法是数据降维研究中的经典算法,因其特点不同,在数据降维应用中的效果亦不同。Laplacian得分算法能有效剔除高维特征数据集中的冗余特征属性,但经其筛选的特征组成的敏感特征集仍然具有较高维数,不利于故障分类,需进一步降维处理;LFDA通过基于距离的加权矩阵可有效保留每类的局部多模态性,同时通过广义特征值分解可最大化类间可分离性;主成分计算可根据累积特征贡献率的大小选择构成低维特征子集的主要成分,可避免LFDA算法在选择主成分时的盲目性。

针对上述3种算法的特点,本研究拟将3种算法进行结合,即采用Laplacian得分算法对原始高维故障数据集进行过滤,再将主成分计算加入LFDA算法中,剔除冗余信息,并通过设定贡献率阈值自适应地得到目标维数,避免在降维过程中选择目标维数的盲目性。

## 2 提出的LFPCDA算法

### 2.1 算法原理

针对高维故障数据集的相关信息和冗余特征导致的故障分类难问题,本研究所提LFPCDA算法的数学原理,即利用Laplacian得分算法对原始高维故障数据集进行第1次降维处理,可有效剔除原始高维故障数据集的冗余特征属性。设特征 $F_r = [f_{r_1}, f_{r_2}, \dots, f_{r_i}]^T$ ,其中: $f_{r_i}$ 为第 $r$ 个特征在第 $i$ 个样本

上的取值。计算Laplacian得分为

$$L_r = \frac{\tilde{F}_r^T L \tilde{F}_r}{\tilde{F}_r^T D \tilde{F}_r} \quad (7)$$

其中: $\tilde{F}_r$ 为去中心化后的特征向量; $L$ 为拉普拉斯矩阵<sup>[14]</sup>; $D$ 为度矩阵。

假设原始高维故障数据集 $X=[x_1, x_2, \dots, x_k]$ ,Laplacian得分为 $L_r=[L_1, L_2, \dots, L_r]$ ,根据设定的得分阈值 $S$ 筛选的特征构成维数约简后的故障数据集,记为 $X_1$ ,经验证后,本研究中Laplacian得分阈值为0.8。以上是对原始高维故障数据集的第1次降维处理。

第2次降维处理的关键点是求得最佳投影矩阵,其求解过程为:定义类间散度矩阵 $S_b$ 与类内散度矩阵 $S_w$ 为

$$S_b = X_1^T (D_b - \tilde{W}_b) X_1 \quad (8)$$

$$S_w = X_1^T (D_w - \tilde{W}_w) X_1 \quad (9)$$

其中: $D$ 为对角矩阵; $\tilde{W}$ 为权重矩阵。

定义权重矩阵为

$$\tilde{W}_b = \begin{cases} A_{ij}(1/h - 1/h_c) & (\text{同类}) \\ 0 & (\text{不同类}) \end{cases} \quad (10)$$

$$\tilde{W}_w = \begin{cases} A_{ij}/h_c & (\text{同类}) \\ 0 & (\text{不同类}) \end{cases} \quad (11)$$

其中: $h$ 为故障数据集 $X_1$ 的维数; $h_c$ 为类别 $c$ 中的样本数; $A_{i,j}$ 为对称矩阵,其编译了数据的局部邻域关系。

设置 $A_{i,j}$ 的第 $(i,j)$ 个元素值的标准方式为

$$A_{i,j} = \exp\left(-\frac{\|X_i - X_j\|^2}{\delta^2}\right) \quad (12)$$

其中: $\delta$ 为可调参数。

目标函数定义为

$$J = \operatorname{argmax}_{W \in (R)^h} \left\{ \frac{W^T S_b W}{W^T S_w W} \right\} \quad (13)$$

用拉格朗日法求解,将 $J$ 转化为广义特征值问题,即

$$S_w^{-1} S_b P = \lambda P \quad (14)$$

计算出特征值 $\lambda_i$ 对应的特征向量 $P_i$ 后,将特征值按由大到小排序,设定累积特征贡献率 $\rho$ ,计算主成分,计算式为

$$s = \operatorname{argmin}_s \left( \frac{\sum_{i=1}^s \lambda_i}{\sum_{i=1}^d \lambda_i} \right) \quad (15)$$

其中: $d$ 为特征值个数; $s$ 为排序后的前 $s$ 个特征值的

个数; $\lambda_i$ 为第 $i$ 个特征值。

最后可得最佳投影矩阵 $W=[P_1, P_2, \dots, P_s]$ 。

## 2.2 算法流程

输入:Laplacian得分阈值 $S$ ,原始高维故障数据集 $X=[x_1, x_2, \dots, x_k]$ ,期望累积特征贡献率 $\rho$ 。

输出:投影矩阵 $W$ ,低维特征子集 $Y$ 。

与式(7)~(15)对应的LFPCDA算法应用步骤如下:

- 1) 在转子实验台上采集6种状态下的振动信号,并进行混合滤波算法消噪处理;
- 2) 计算不同维度特征的值,构建原始高维故障数据集 $X$ ,并进行归一化处理;
- 3) 根据式(7)计算原始高维特征数据集中每个特征的Laplacian得分 $S(k)$ ,设定Laplacian得分阈值 $S$ ,保留 $S(k) \leq S$ 的特征,记为 $X_1$ ;
- 4) 根据式(8)、(9)计算 $X_1$ 的 $S_b$ 和 $S_w$ ;
- 5) 建立如式(13)所示的目标函数,记为 $J_{\text{LFPCDA}}$ ,求解目标函数得到如式(14)所示的广义特征向量 $P$ 与特征值 $\lambda$ ;
- 6) 对特征值进行从大到小的排序,设定期望累积贡献率阈值 $\rho$ ,根据式(15)计算满足条件的前 $s$ 个特征值,并选择与特征值对应的特征向量构成投影矩阵 $W$ ,低维特征子集即为 $Y=X_1 W$ 。

LFPCDA算法流程如图1所示。

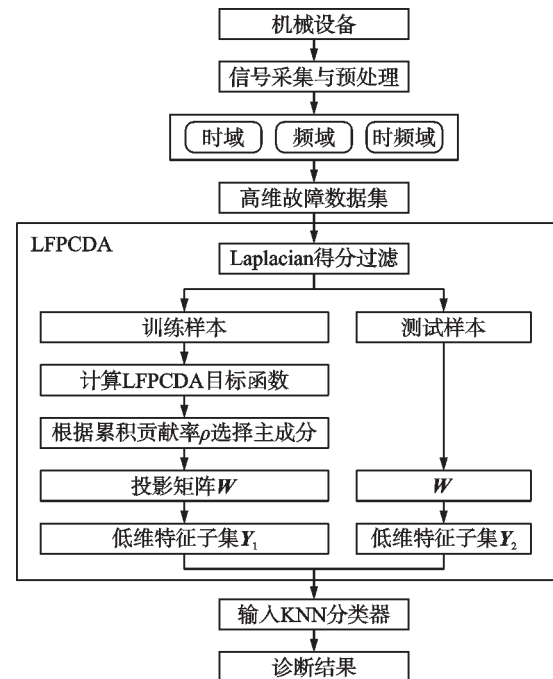


图1 LFPCDA算法流程

Fig.1 Process of LFPCDA algorithm

### 3 实验验证

#### 3.1 转子故障模拟实验

本研究中,双跨转子实验台如图2所示,为无锡厚德仪表公司的HZXT-DS-001型实验台。

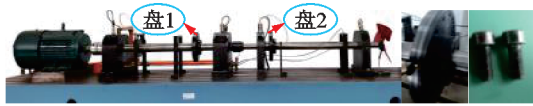


图2 双跨转子实验台

Fig.2 Double span rotor test bench

采集4个轴承座附近截面上水平与垂直方向共8个通道的振动信号。采样频率为20 kHz,在3 600 r/min的工况下采集6种质量不平衡状态的振动信号,从每个通道的振动信号中提取出38个特征,共得 $8 \times 38 = 304$ 个特征。特征参数如表1所示。

表1 特征参数

Tab.1 Characteristic parameters

序号	特征	序号	特征	序号	特征
1	最大值	9	方根幅值	17	均值频率
2	最小值	10	绝对均值	18	中心频率
3	峰峰值	11	波形指标	19	均方根频率
4	方差	12	峰值指标	20	频谱二阶矩
5	均值	13	脉冲指标	21	标准偏差频率
6	歪度	14	裕度指标	22	峭度频率
7	峭度	15	歪度指标	23~38	4层小波包
8	均方根值	16	峭度指标		分解能量特征

特征1~16为时域特征;17~22为频域特征;23~38为时频域小波包能量特征

各质量不平衡状态通过在2个负载盘上分配不同数目的质量块进行模拟,每个负载盘上的质量块数目有0、2、3共3组,保证2个负载盘上添加不同数目质量块便可模拟6种质量不平衡状态,质量块的具体分布情况可参考文献[15]。每种故障状态采集100组样本,其中40组为训练样本、60组为测试样本。

#### 3.2 参数设置

本研究2.1节中设计算法需要确定2个参数,即Laplacian得分阈值 $S$ 和主成分计算中的累积特征贡献率 $\rho$ 。固定一个参数,选取适当步长改变另一个参数,并记录故障分类准确率。故障分类准确率随 $S$ 与 $\rho$ 的变化如表2所示。算法调试结果表明,当

$\rho = 0.9$ 、 $S = 0.8$ 时,故障分类的准确率最高可达99.67%。

表2 故障分类准确率随 $S$ 与 $\rho$ 的变化

Tab.2 Fault classification accuracy varies with  $S$  and  $\rho$

阈值 $S$	贡献率 $\rho$		
	0.7	0.8	0.9
0.75	78.67	78.67	89.00
0.80	91.33	91.67	99.67
0.85	90.00	91.33	91.67

#### 3.3 降维效果的衡量指标

为衡量所提算法的降维效果,本研究引入可分性指标<sup>[16]</sup>来评价低维特征子集的聚类与分类效果, $d_b$ 用于描述不同类样本之间的分散程度, $d_w$ 用于描述同类样本之间的紧密程度。定义如下

$$d_b = \frac{2}{C(C-1)} \sum_{\theta=1}^{C-1} \sum_{l=\theta+1}^C \|\bar{y}_\theta - \bar{y}_l\| \quad (16)$$

$$d_w = \frac{1}{C} \frac{1}{nj} \sum_{\theta=1}^C \sum_{i=1}^{nj} \|\bar{y}_{\theta i} - \bar{y}_\theta\| \quad (17)$$

$$\gamma = \frac{d_b}{d_w} \quad (18)$$

其中: $d_b$ 、 $d_w$ 分别为类间距离与类内距离; $C$ 为类别数目; $y_{\theta i}$ 为第 $\theta$ 类中包含 $n_j$ 个样本点的数据; $\theta = 1, 2, \dots, C$ ;  $i = 1, 2, \dots, n$ ;  $\bar{y}_\theta = \frac{1}{nj} \sum_{i=1}^{nj} y_{\theta i}$ ,为每类样本的聚类中心。

式(18)表明, $\gamma$ 值越大,数据的可分性越好。

#### 3.4 降维效果对比与分析

为验证所提LFPCDA算法的有效性,选取PCA、LFDA算法与其进行对比验证,得到测试样本的前3个主分量效果。测试样本基于不同算法的降维效果如图3所示。其中,图3(a)为原始数据集的完全不可分状况。图3(d)中LFPCDA算法的分类性能显著优于图3(b)、(c)所示的PCA与LFDA算法。LFPCDA算法使不平衡的3类数据子集被分开,使得各类内部和各类之间分别呈现出明显的集聚和明显的区分。由此可知,LFPCDA算法有类间分离,同时又能保持类内结构不变的功能。

为量化描述所提算法的优势,给出如表3所示的3种降维算法及其KNN辨识准确率,表中1、2、3、4、5、6分别对应6种不平衡状态。分析表中数据可知,LFPCDA算法的各故障状态识别准确率均高于其他2种算法,其平均识别准确率可达99.67%。



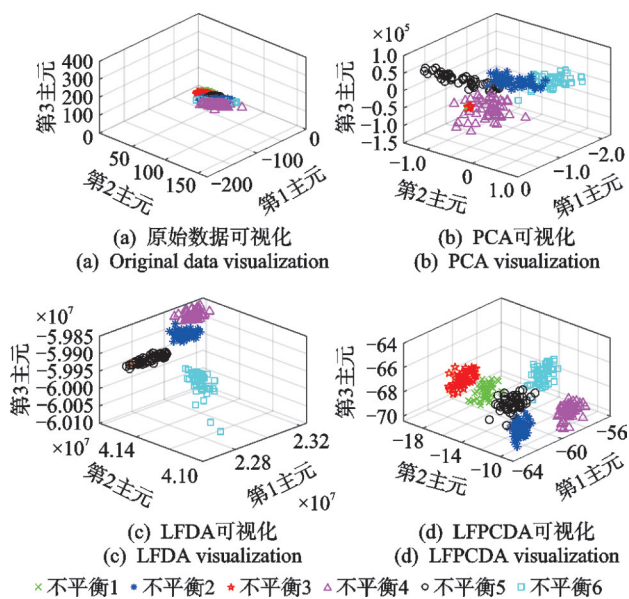


图3 测试样本基于不同算法的降维效果

Fig.3 Dimensionality reduction effect of test samples based on different algorithms

表3 3种降维算法及其KNN辨识准确率

Tab.3 3 dimensionality reduction algorithms and their KNN classification accuracy %

降维 算法	不平衡状态						平均值
	1	2	3	4	5	6	
PCA	60.00	80.00	64.00	88.00	82.00	84.00	76.33
LFDA	70.00	98.00	62.00	96.00	88.00	100.00	85.67
LFPCDA	100.00	100.00	100.00	100.00	100.00	98.00	99.67

为进一步验证所提算法的优势与泛化性,加入LDA与边际Fisher分析(marginal Fisher analysis,简称MFA)算法进行可分性指标与不同转速下故障识别准确率的对比实验。测试样本的可分性指标对比如图4所示。式(18)为类间分散程度与类内紧致程度的量化指标, $\gamma$ 值越大,说明数据的可分性越好。由图4可知,LFPCDA算法的 $\gamma$ 值最大。

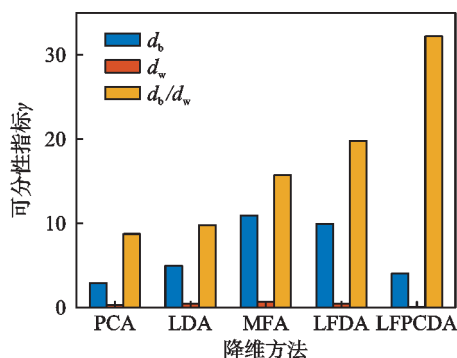


图4 测试样本的可分性指标对比

Fig.4 Comparison of separability indexes of test samples

不同转速下各算法的KNN辨识准确率如图5所示。可以看出,在3种转速下,本研究所提LFPCDA算法识别准确率均高于其他算法,识别准确率接近100%,说明其具有更加广泛的普适性。

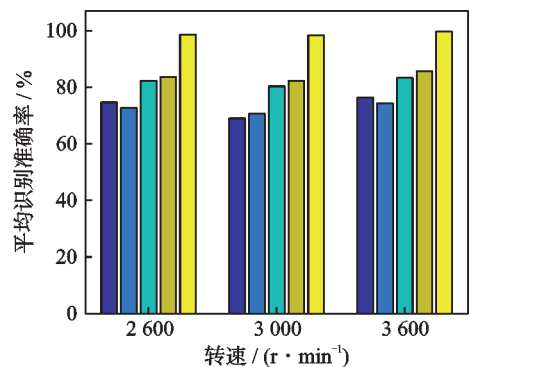


图5 不同转速下各算法的KNN辨识准确率

Fig.5 KNN classification accuracy of different algorithms with different speeds

## 4 结 论

1) 将Laplacian得分算法、PCA算法与LFDA算法进行集成,提出的LFPCDA算法可剔除高维故障数据集中冗余特征属性。同时,保留了最有价值特征成分的性能,并达到了构成投影矩阵的主成分自适应选取效果。

2) 双跨转子实验台模拟数据集的实验表明,与其他降维算法相比,LFPCDA算法的降维能力更强,故障识别精度更高,且具有更好的通用性,可为旋转机械故障诊断提供新的数据降维思路。

## 参 考 文 献

- [1] 李杰,李响,许元铭,等.工业人工智能及应用研究现状及展望[J].自动化学报,2020,46(10):2031-2044.  
LI Jie, LI Xiang, XU Yuanming, et al. Recent advances and prospects in industrial AI and applications[J]. Acta Automatica Sinica, 2020, 46(10): 2031-2044. (in Chinese)
- [2] 柴天佑.工业人工智能发展方向[J].自动化学报,2020,46(10):2003-2012.  
CHAI Tianyou. Development directions of industrial artificial intelligence[J]. Acta Automatica Sinica, 2020, 46(10): 2003-2012. (in Chinese)
- [3] 雷亚国,贾峰,孔德同,等.大数据下机械智能故障诊断的机遇与挑战[J].机械工程学报,2018,54(5):94-104.  
LEI Yaguo, JIA Feng, KONG Detong, et al. Opportu-

- nities and challenges of machinery intelligent fault diagnosis in big data era[J]. *Journal of Mechanical Engineering*, 2018, 54(5): 94-104.(in Chinese)
- [4] 金晓航, 王宇, ZHANG Bin. 工业大数据驱动的故障预测与健康管理[J]. *计算机集成制造系统*, 2022, 28(5): 1314-1336.  
JIN Xiaohang, WANG Yu, ZHANG Bin. Industrial big data-driven fault prognostics and health management [J]. *Computer Integrated Manufacturing Systems*, 2022, 28(5): 1314-1336.(in Chinese)
- [5] 王庆锋, 卫炳坤, 刘家赫, 等. 一种数据驱动的旋转机械早期故障检测模型构建和应用研究[J]. *机械工程学报*, 2020, 56(16): 22-32.  
WANG Qingfeng, WEI Bingkun, LIU Jiahe, et al. Research on construction and application of data-driven incipient fault detection model for rotating machinery[J]. *Journal of Mechanical Engineering*, 2020, 56(16): 22-32.(in Chinese)
- [6] 陈晓云, 廖梦真. 基于稀疏和近邻保持的极限学习机降维[J]. *自动化学报*, 2019, 45(2): 325-333.  
CHEN Xiaoyun, LIAO Mengzhen. Dimensionality reduction with extreme learning machine based on sparsity and neighborhood preserving[J]. *Acta Automatica Sinica*, 2019, 45(2): 325-333.(in Chinese)
- [7] WOLD S, ESBENSEN K, GELADI P. Principal component analysis[J]. *Chemometrics and Intelligent Laboratory Systems*, 2002, 2(3): 37-52.
- [8] MARTINEZ A M, KAK A C. PCA versus LDA[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23(2): 228-233.
- [9] JOLLIFFE I T, CADIMA J. Principal component analysis: a review and recent developments[J]. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 2016, 374(2065): 20150202.
- [10] 朱文昌, 罗梦婷, 倪广县, 等. 随机矩阵理论和主成分分析融合的滚动轴承性能退化评估方法[J]. *西安交通大学学报*, 2021, 55(2): 55-63.  
ZHU Wenchang, LUO Mengting, NI Guangxian, et al. An evaluation method for rolling bearing performance degradation combining random matrix theory and principal component analysis[J]. *Journal of Xi'an Jiaotong University*, 2021, 55(2): 55-63.(in Chinese)
- [11] KHEDIRI I B, LIMAM M, WEIHS C. Variable window adaptive kernel principal component analysis for nonlinear nonstationary process monitoring[J]. *Computers & Industrial Engineering*, 2011, 61(3): 437-446.
- [12] HE X, DENG C, NIYOGI P. Laplacian score for feature selection[C]//*Proceedings of the 19th International Conference on Neural Information Processing System*. Cambridge, MA, USA: MIT Press, 2005.
- [13] 郭金玉, 韩建斌, 李元, 等. 基于局部 Fisher 判别分析的复杂化工过程故障诊断[J]. *计算机应用研究*, 2018, 35(4): 1122-1125, 1129.  
GUO Jinyu, Han Jianbin, Li Yuan, et al. Fault diagnosis of complex chemical process based on local Fisher discriminant analysis[J]. *Application Research of Computers*, 2018, 35(4): 1122-1125, 1129.(in Chinese)
- [14] SUGIYAMA M, IDÉ T, NAKAJIMA S N H, et al. Semi-supervised local Fisher discriminant analysis for dimensionality reduction[J]. *Machine Learning*, 2010, 78(1): 35-61.
- [15] YUAN J, ZHAO R, HE T, et al. Fault diagnosis of rotor based on semi-supervised multi-graph Joint Embedding[J]. *ISA Transactions*, 2022, 131: 516-532.
- [16] 李学军, 李平, 蒋玲莉. 类均值核主元分析法及在故障诊断中的应用[J]. *机械工程学报*, 2014, 50(3): 123-129.  
LI Xuejun, LI Ping, JIANG Lingli. Class mean kernel principal component analysis and its application in fault diagnosis[J]. *Journal of Mechanical Engineering*, 2014, 50(3): 123-129.(in Chinese)



**第一作者简介:**陈芳军,男,1997年2月生,硕士生。主要研究方向为旋转机械故障诊断。

E-mail: 1343961699@qq.com

**通信作者简介:**赵荣珍,女,1960年12月生,博士、教授、博士生导师。主要研究方向为旋转机械故障诊断、机械系统动力学。

E-mail: zhaorongzhen@lut.edu.cn