

基于 ACMPE、ISSL-Isomap 和 GWO-SVM 的行星齿轮箱故障诊断

戚晓利, 王振亚, 吴保林, 叶绪丹, 潘紫微
(安徽工业大学 机械工程学院, 安徽 马鞍山 243032)

摘 要: 针对从行星齿轮箱非线性、非平稳振动信号特征提取困难的问题,提出了一种基于自适应复合多尺度排列熵(ACMPE)、改进监督型自组织增量学习神经网络界标点等度规映射(ISSL-Isomap)和灰狼群优化支持向量机(GWO-SVM)相结合的行星齿轮箱故障诊断方法。利用 ACMPE 从复杂域提取振动信号的故障特征,构建高维故障特征集;采用 ISSL-Isomap 方法对高维故障特征集进行维数约简,提取出低维、敏感故障特征;应用 GWO-SVM 分类器对低维故障特征进行模式识别,判断故障类型。行星齿轮箱故障诊断实验结果表明:与多尺度排列熵(MPE)、复合多尺度排列熵(CMPE)等特征提取方法相比,ACMPE 方法在分类效果和识别精度上更具优势;与局部切空间排列(LTSA)、等度规映射(Isomap)、加权 Isomap(W-Isomap)、监督 Isomap(S-Isomap)和监督型自组织增量学习神经网络界标点 Isomap(SSL-Isomap)等降维方法进行比较,ISSL-Isomap 方法降维效果最佳;所提方法的故障识别率达到 100%,具有一定优越性。

关 键 词: 故障诊断;行星齿轮箱;自适应复合多尺度排列熵(ACMPE);
改进监督型自组织增量学习神经网络界标点等度规映射(ISSL-Isomap);
灰狼群优化支持向量机(GWO-SVM)

中图分类号: V232.8; TH165⁺.3; TN911.7

文献标志码: A

Planetary gearbox fault diagnosis based on ACMPE, ISSL-Isomap and GWO-SVM

QI Xiaoli, WANG Zhenya, WU Baolin, YE Xudan, PAN Ziwei

(School of Mechanical Engineering,
Anhui University of Technology, Maanshan Anhui 243032, China)

Abstract: In view of the difficulty of extracting nonlinear and non-stationary vibration signals from planetary gearboxes, a planetary gearbox fault diagnosis method based on adaptive composite multi-scale permutation entropy (ACMPE), improved supervised self-organizing incremental neural network landmark isometric mapping (ISSL-Isomap) and grey wolf optimizer support vector machine (GWO-SVM) was proposed. Fault features of vibration signals were extracted from the complex domain by using ACMPE, and the high-dimensional fault feature set was constructed. ISSL-Isomap was used to reduce the dimension of the high-

收稿日期:2018-07-24

基金项目:国家自然科学基金(51505002);安徽省自然科学基金(1808085ME152);

安徽省高校自然科学研究重点项目(KJ2017 A053);研究生创新研究基金(2017012)

作者简介:戚晓利(1975—),男,副教授、硕士生导师,博士,主要从事故障诊断和撞击动力学等方面研究。

通讯作者:王振亚(1993—),男,硕士生,主要从事机械系统状态监控和故障诊断等方面研究。E-mail:wzyxaj@163.com

引用格式:戚晓利,王振亚,吴保林,等.基于 ACMPE、ISSL-Isomap 和 GWO-SVM 的行星齿轮箱故障诊断[J].航空动力学报,2019,34(4):744-755. QI Xiaoli, WANG Zhenya, WU Baolin, et al. Planetary gearbox fault diagnosis based on ACMPE, ISSL-Isomap and GWO-SVM[J]. Journal of Aerospace Power, 2019, 34(4): 744-755.

dimensional fault feature set, and the low-dimensional and sensitive fault features were extracted. The low-dimensional fault features were input into a GWO-SVM classifier to recognize fault types. The analysis results of planetary gearbox fault diagnosis show that compared with the feature extraction methods of multi-scale permutation entropy (MPE) and composite MPE (CMPE), ACMPE has more advantages in classification effect and recognition accuracy. ISSL-Isomap has the best dimensionality reduction effect compared with the dimensionality reduction algorithms of local tangent space alignment (LTSA), isometric mapping (Isomap), weighted Isomap (W-Isomap), supervised Isomap (S-Isomap) and supervised self-organizing incremental neural network landmark Isomap (SSL-Isomap). The fault recognition rate of the proposed method reaches 100% with a certain superiority.

Key words: fault diagnosis; planetary gearbox;
adaptive composite multi-scale permutation entropy (ACMPE);
improved supervised self-organizing incremental neural network landmark
isometric mapping (ISSL-Isomap);
grey wolf optimizer support vector machine (GWO-SVM)

直升机动力传动系统作为其最重要的机械结构,直接关系到直升机的飞行寿命、机动性能和飞行安全^[1]。行星齿轮箱具有体积小、传动比大、传动效率高与承载能力强等诸多优点,因此已被广泛应用于直升机传动系统中^[2]。但由于行星齿轮箱结构复杂,不仅承受重载负荷,而且运行工况复杂多变,使其极易发生裂纹、磨损乃至断齿、缺齿等失效情况,影响直升机动力传动系统寿命,甚至导致飞行故障的发生,故对其进行故障诊断有着重要的经济效益和军事意义^[3]。

行星齿轮箱故障诊断的关键在于如何从非线性、非平稳振动信号中提取出能反应其运行状态的故障特征信息^[4]。近年来,基于熵的非线性分析方法已被应用于故障检测与诊断中,如样本熵、排列熵(permutation entropy, PE)、模糊熵和多尺度熵等^[5]。其中,PE作为一种检测时间序列随机性和动力学突变行为的非线性分析方法,具有计算速度快、抗噪能力强、适合在线监测等优点,因此在故障诊断领域应用较为广泛^[6-8]。如刘永斌等^[6]研究了不同工作状态下轴承振动信号的PE,实验结果表明PE方法可以有效地检测出机械设备状态变化。YAN Ruqiang等^[7]将PE应用于旋转机械系统振动信号的特征提取过程。丁闯等^[8]提出了一种PE与局部均值分解相结合的故障特征提取方法。此外,Aziz^[9]针对PE方法单一尺度分析的不足,提出了一种多尺度排列熵(multi-scale permutation entropy, MPE)的概念。相关研究表明,与PE方法相比,MPE能够

更加全面地提取机械系统振动信号的特征信息^[10]。但MPE方法粗粒化方式定义的多尺度计算过程依赖于时间序列的长度,导致MPE值误差会随尺度因子的增大而增加;此外MPE的粗粒化过程本质仍为一种线性和平稳的平滑方法,对于非平稳、非线性行星齿轮箱振动信号而言,直接采用该方法不可避免存在一定的局限性。因此,本文在MPE的基础上提出了一种基于变分模态分解(variational mode decomposition, VMD)的自适应复合多尺度排列熵(ACMPE)方法,以克服MPE粗粒化过程的不足,并将该方法应用于行星齿轮箱故障特征提取过程。

考虑到ACMPE提取的故障特征维数较高,并且可能存在部分冗余和混叠信息,因此需要采用降维方法提取出利于识别的低维、敏感特征。近年来,基于几何思想驱动的流形学习方法为实现高维、非线性的机械故障特征维数约简和高精度的模式分类提供了更为理想的解决思路^[11-13]。如孙斌等^[11]将等距特征映射方法应用于转子高维故障特征的维数约简过程。CHEN Fafa等^[12]利用局部切空间排列(LTSA)方法对滚动轴承多域故障特征进行有效二次特征提取。黄宏臣等^[13]提出了一种基于拉普拉斯特征映射的滚动轴承故障识别方法。但上述流形学习方法均为无监督降维方法,未充分利用面向分类的类判别特征,不太适合行星齿轮箱故障信号这类存在诸多奇异数据点的情况^[14]。针对这一问题,本文提出一种改进监督型自组织增量学习神经网络界标点等度规映射(ISSL-Isomap)方法。一方面将余弦相

似度与欧氏距离相结合应用于距离矩阵的构建,并利用样本标签构造监督型成对约束邻域图;另一方面采用自组织增量学习神经网络(self-organizing incremental neural network, SOINN)^[15]进行界标点的选取,从而增强数据的分类能力。

故障诊断的本质在于模式识别,支持向量机(SVM)能较好地解决小样本、非线性和局部极小等问题,因此被广泛应用于模式识别、机械设备状态识别等领域^[16]。但使用 SVM 对机械设备状态进行故障预测时,其预测性能与 SVM 参数选取密切相关。针对这一难题,本文引入一种收敛速度快、全局寻优能力强的新型元启发式优化方法——灰狼群优化(GWO)方法^[17]进行 SVM 参数寻优。

依据上述分析,本文提出了一种基于 ACMPE、ISSL-Isomap 和 GWO-SVM 相结合的行星齿轮箱故障诊断方法。实验结果分析表明,该方法集成了 ACMPE 方法在特征提取、ISSL-Isomap 方法在维数约简和 GWO-SVM 方法在模式识别上的优势,能有效、精准地诊断出行星齿轮箱各故障类型。

1 ACMPE

1.1 MPE 方法

MPE 方法具体过程如下:

1) 对时间序列 $\{x(i), i=1, 2, \dots, N\}$ 进行粗粒化处理后得到粗粒序列 $y^\tau(j)$

$$y^\tau(j) = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i \quad 1 \leq j \leq \frac{N}{\tau} \quad (1)$$

式中 τ 为尺度因子。

2) 计算不同尺度因子下的粗粒化序列 PE 值,对应表达式为

$$E_{\text{mp}}(x, m, \lambda, \tau) = E_p(y^\tau, m, \lambda) \quad (2)$$

式中 m 为嵌入维数, λ 为时延, E_{mp} 为 MPE 值, $E_p(\cdot)$ 为 PE 值。据式(2)可知,当 $\tau=1$ 时,即为单尺度 PE。

1.2 ACMPE 方法流程

MPE 方法衡量了时间序列在不同尺度下的复杂性和随机性,克服了 PE 单尺度分析的不足,但将该方法应用于机械振动信号特征提取过程存在以下两点不足:①MPE 方法(式(1))粗粒化计算过程中, MPE 值估计误差会随粗粒化尺度因子增大而增大。②MPE 方法粗粒化过程本质仍为一种线性、平稳的方法,这导致该方法仅仅捕获了

粗尺度上的低频成分,损失了细尺度上的高频成分。并且在处理非平稳、非线性振动信号时,缺乏自适应性。

针对上述两点不足,本文作了如下改进:①考虑同一尺度下的多个粗粒化时间序列的 PE 值,因此有效抑制由粗粒化时间序列变短而导致的 PE 值突变,使得到的 PE 值更为准确、客观。②VMD 方法是在传统维纳滤波的基础上提出的一种非递归自适应信号分解新方法^[18]。本文利用 VMD 方法将非平稳、非线性复杂信号自适应地分解成不同尺度下的本征模态分量(intrinsic mode function, IMF, 记为 $F_{\text{im},i}$),并通过从细到粗对原始信号进行自适应尺度的重构以解决 MPE 方法线性、平稳和缺乏自适应性等不足。上述分析即为 ACMPE 过程,具体步骤如下:

1) 利用 VMD 方法将原始振动信号自适应地分解为 K 个不同尺度的 IMF 之和,其中 VMD 方法具体过程参见文献^[18]。

2) 通过依次叠加高频 IMF 分量,得到自适应粗粒化序列 $y(s)$,即

$$y(s) = \sum_{i=1}^s F_{\text{im},i}, \quad 1 \leq s \leq K \quad (3)$$

式中 s 为自适应尺度因子, K 为 VMD 方法分解后的 IMF 分量总数,即模态数。

3) 对 $y(s)$ 采用式(4)重新定义粗粒化序列 $y_h^\tau = \{y_{h,1}^\tau, y_{h,2}^\tau, \dots, y_{h,\tau}^\tau\}$ 。

$$y_{h,j}^\tau = \frac{1}{\tau} \sum_{i=(j-1)\tau+h}^{j\tau+h-1} y(i), \quad (1 \leq j \leq \frac{N}{\tau}, 1 \leq h \leq \tau) \quad (4)$$

式中 N 为 $y(s)$ 数据长度。

4) 对于尺度因子 τ , 分别计算出该尺度因子下每个粗粒化序列 y_h^τ 的 PE, 再通过平均 PE 值得到 τ 尺度下的复合多尺度排列熵(CMPE, E_{cmp}) 值,对应表达式为

$$E_{\text{cmp}}(y_h^\tau, m, \lambda, \tau) = \frac{1}{\tau} \sum_{h=1}^{\tau} E_p(y_h^\tau, m, \lambda) \quad (5)$$

5) 计算 $\text{ACMPE}(E_{\text{acmp}})$ 值:

$$E_{\text{acmp}} = E_{\text{cmp}}(y(s), \tau, m, \lambda) \quad (6)$$

2 ISSL-Isomap 降维方法

2.1 SOINN L-Isomap(SL-Isomap)方法

界标点等度规映射(landmark Isomap, L-Isomap)^[19]通过引入界标点多维尺度(landmark multidimensional scaling, LMDS)解决了 Isomap

计算复杂度高的问题。但该方法预先需要在流形表面选取稀疏分布的界标点用以代表流形拓扑结构。SOINN 可在无先验知识的情况下,在线输出数据集拓扑结构,因此,可将其应用于界标点的选取过程。对于样本集 $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_N\}$, $\mathbf{x}_i \in \mathbf{R}^D$, SL-Isomap 方法的具体步骤如下。

1) 界标点的选取

利用 SOINN 方法在线生成神经元集合 $\mathbf{A} = \{\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n\}$ 。为确保 $\mathbf{A}$ 是输入数据集 $\mathbf{X}$ 的子集,对 $\mathbf{A}$ 进行调整得到界标点集合 $\mathbf{L} = \{\mathbf{x}_{l_1}, \mathbf{x}_{l_2}, \dots, \mathbf{x}_{l_n}\}$, 其中 $\mathbf{x}_{l_i}$ 为

$$d_g(\mathbf{x}_{l_i}, \mathbf{x}_j) = \begin{cases} d_g(\mathbf{x}_{l_i}, \mathbf{x}_j) & \mathbf{x}_{l_i} \text{ 与 } \mathbf{x}_j \text{ 为近邻点} \\ \min\{d_g(\mathbf{x}_{l_i}, \mathbf{x}_j), d_g(\mathbf{x}_{l_i}, \mathbf{x}_k) + d_g(\mathbf{x}_k, \mathbf{x}_j)\} & \text{其他} \end{cases} \quad (8)$$

4) 界标点降维映射

从 $\mathbf{D}_g$ 中提取界标点间的测地距离矩阵 $\mathbf{D}_n(n \times n)$, 利用多维尺度(multidimensional scaling, MDS)方法计算出界标点映射矩阵 $\mathbf{M}$

$$\mathbf{M} = \begin{bmatrix} \sqrt{\sigma_1} \mathbf{v}_1^T \\ \sqrt{\sigma_2} \mathbf{v}_2^T \\ \vdots \\ \sqrt{\sigma_d} \mathbf{v}_d^T \end{bmatrix} \quad (9)$$

式中 σ_i 表示第 i 个特征值, $\mathbf{v}_i^T$ 表示 σ_i 对应的特征向量。

5) 基于界标点对数据集降维映射

抽取 $\mathbf{D}_g$ 中界标点与剩余样本点间的测地距离矩阵 $\mathbf{D}_{n, N-n}$, 应用 LMDS 方法计算数据点 $\mathbf{x}_i$ 的低维坐标 $\mathbf{y}_i$

$$\mathbf{y}_i = \frac{1}{2} \mathbf{M}^* (\overline{\Delta_n} - \Delta_{x_i}) \quad (10)$$

式中 Δ_{x_i} 为数据点 $\mathbf{x}_i$ 与 n 个基准点的距离向量; Δ_n 为 $\mathbf{D}_n$ 的每个值求平方所得矩阵; $\overline{\Delta_n}$ 为 Δ_n 的列平均矩阵; $\mathbf{M}^*$ 为

$$\mathbf{M}^* = \begin{bmatrix} \mathbf{v}_1^T / \sqrt{\sigma_1} \\ \mathbf{v}_2^T / \sqrt{\sigma_2} \\ \vdots \\ \mathbf{v}_d^T / \sqrt{\sigma_d} \end{bmatrix} \quad (11)$$

2.2 ISSL-Isomap 方法

SOINN 界标点等度规映射(SOINN landmark Isomap, SL-Isomap)方法属于无监督降维方法,未将类别信息融入维数约简过程来提高低维特征集的可辨识能力。在高维样本空间中,类别信息相同的样本具有相同的流形结构,类别信息不同的样本流形结构往往不同,因此有效利用

$$\mathbf{x}_{l_i} = \arg \min_{\mathbf{x}_c \in \mathbf{X}} \|\mathbf{L}_i - \mathbf{x}_c\| \quad (7)$$

2) 构造邻域图 G

计算全数据集的欧氏距离矩阵 $\mathbf{D}(N \times N)$ 。采用 k 近邻法(k -nearest neighbors, k -NN)构造邻域图 G , 若样本点 $\mathbf{x}_j$ 为 $\mathbf{x}_i$ 的近邻点, 两点之间有边连接, 边长为 $\mathbf{D}(\mathbf{x}_i, \mathbf{x}_j)$; 否则不连接。

3) 计算测地距离

利用 Dijkstra 方法寻找最短路径, 构造出近似测地距离矩阵 $\mathbf{D}_g = \{d_g(\mathbf{x}_{l_i}, \mathbf{x}_j)\}_{n \times N}$

类别信息能使方法更充分地提取高维空间中的本质流形。

2.2.1 改进距离度量

L-Isomap 方法在邻域图构建过程中仅考虑了样本点之间的欧式距离, 但欧氏距离只能反映样本点之间的空间位置关系, 不能反映流形结构关系。因此, 本文采用余弦相似度与欧氏距离相结合的距离度量方式, 即

$$d_{ij}(\mathbf{x}_i, \mathbf{x}_j) = \left[\frac{1 - d_{ij}^c(\mathbf{x}_i, \mathbf{x}_j)}{2} \right] \times d_{ij}^m(\mathbf{x}_i, \mathbf{x}_j), \quad (i, j = 1, 2, \dots, N) \quad (12)$$

式中 $d_{ij}^c(\mathbf{x}_i, \mathbf{x}_j)$ 表示 $\mathbf{x}_i, \mathbf{x}_j$ 的余弦相似度, $d_{ij}^m(\mathbf{x}_i, \mathbf{x}_j)$ 表示 $\mathbf{x}_i, \mathbf{x}_j$ 的欧氏距离。

2.2.2 构建监督型成对约束邻域图

本文将监督学习思想引入至 SL-Isomap 方法, 具体过程如下。

1) 监督型成对约束邻域图构建

利用 k 近邻法构建无向邻域图 $G(V, E)$ 。其中, 边权重 $e(\mathbf{x}_i, \mathbf{x}_j) \in \{0, 1\}$ 用以表示近邻点的弱连接和强连接两种类型。当标签样本 $\mathbf{x}_i$ 与 $\mathbf{x}_j$ 具有相同类别标签时, 这两点属于强连接类型, $e(\mathbf{x}_i, \mathbf{x}_j) = 1$; 当标签样本 $\mathbf{x}_i$ 与 $\mathbf{x}_j$ 具有不同类别标签时, 这两点属于弱连接类型, $e(\mathbf{x}_i, \mathbf{x}_j) = 0$ 。

对于强连接类型样本, 建立约束集 S_{ml} ; 对于弱连接类型样本, 建立约束集 S_{ll} :

$$\begin{aligned} S_{ml} &= \{(\mathbf{x}_i, \mathbf{x}_j) = 1, l(\mathbf{x}_i) = l(\mathbf{x}_j)\} \\ S_{ll} &= \{(\mathbf{x}_i, \mathbf{x}_j) = 0, l(\mathbf{x}_i) \neq l(\mathbf{x}_j)\} \end{aligned} \quad (13)$$

式中 $l(\mathbf{x}_i)$ 表示样本点 $\mathbf{x}_i$ ($i = 1, 2, \dots, N$) 类别信息。

基于上述定义, 即可对相同标签样本构建强连接型约束邻域图 $G_{ml}(V, E)$; 对不同标签样本构

建弱连接型约束邻域图 $G_{ll}(V, E)$ 。

2) 有监督距离矩阵 D_s 构建

强连接型约束邻域图 $G_{ml}(V, E)$ 上两点 x_i 与 x_j , 其约束关系 $(x_i, x_j) \in S_{ml}$, 对应边长为

$$d_{ml}(x_i, x_j) = \sqrt{1 - \exp\left[\frac{-d^2(x_i, x_j)}{\mu}\right]},$$
$$l(x_i) = l(x_j) \quad (14)$$

式中 $d(x_i, x_j)$ 表示数据点 x_i 和 x_j 之间的改进度量距离, μ 为调节因子 1, 用来抑制类间距离过快增长, 定义为所有数据之间的平均欧氏距离。

弱连接型约束邻域图 $G_{ll}(V, E)$ 上两点 x_i 、 x_j , 其约束关系 $(x_i, x_j) \in S_{ll}$, 对应边长

$$d_{ll}(x_i, x_j) = \sqrt{\exp\left[\frac{d^2(x_i, x_j)}{\mu}\right]} - \varphi,$$
$$l(x_i) \neq l(x_j) \quad (15)$$

式中 φ 为调节因子 2, 用来控制类间距离的最小值及其和类内距离的相近性。

基于上述分析即可得到有监督距离矩阵 $D_s = \{d_s(x_i, x_j)\}$, 定义为

$$d_s(x_i, x_j) = \begin{cases} d_{ml}(x_i, x_j), e(x_i, x_j) = 1 \\ d_{ll}(x_i, x_j), e(x_i, x_j) = 0 \end{cases} \quad (16)$$

2.2.3 改进监督型 ISSL-Isomap 方法流程

对于高维空间样本集 $X = \{x_1, \dots, x_i, \dots,$

$x_N\}$, $x_i \in R^D$, ISSL-Isomap 方法流程如下:

1) 按照 SL-Isomap 方法步骤 1) 所示, 确定界标点集合 $L = \{x_{l_1}, x_{l_2}, \dots, x_{l_n}\}$;

2) 根据式 (12) 计算改进距离矩阵 $D_i = \{d_{ij}(x_i, x_j)\}$;

3) 根据式 (13) 构建监督型成对约束邻域图 $G_{ml}(V, E)$ 、 $G_{ll}(V, E)$;

4) 根据式 (16) 构建监督型距离矩阵 $D_s = \{d_s(x_i, x_j)\}$, 按照 SL-Isomap 步骤 3) ~ 步骤 5) 所示, 计算出低维映射结果。

2.3 仿真实验

图 1(a) 为瑞士卷 (swiss roll) 数据集的三维分布, 4 种形状分别代表 4 类样本, 每类样本拥有 200 个数据点, 共计 800 个数据点, 该数据集包含 1% 的噪声量, 图中 X、Y、Z 分别为数据点的坐标位置。为验证 ISSL-Isomap 方法有效性, 将其与 Isomap 方法、局部线性嵌入 (LLE) 方法、线性判别分析 (LDA) 方法和监督 Isomap (S-Isomap) [20] 方法进行对比, 5 种方法二维嵌入结果如图 1(b) ~ 图 1(f) 所示。本文通过多次实验获取各方法最优参数, 具体设置如下: Isomap 方法近邻参数 $k=18$; LLE 方法近邻参数 $k=15$; S-Isomap 方法近邻参数 $k=17$, 参数 $a=0.38$; ISSL-Isomap 方法近邻参数 $k=20$, 调节因子 2 为 $\varphi=0.5$ 。据图 1 可知,

■ 第1类样本 ● 第2类样本 ▲ 第3类样本 ◆ 第4类样本

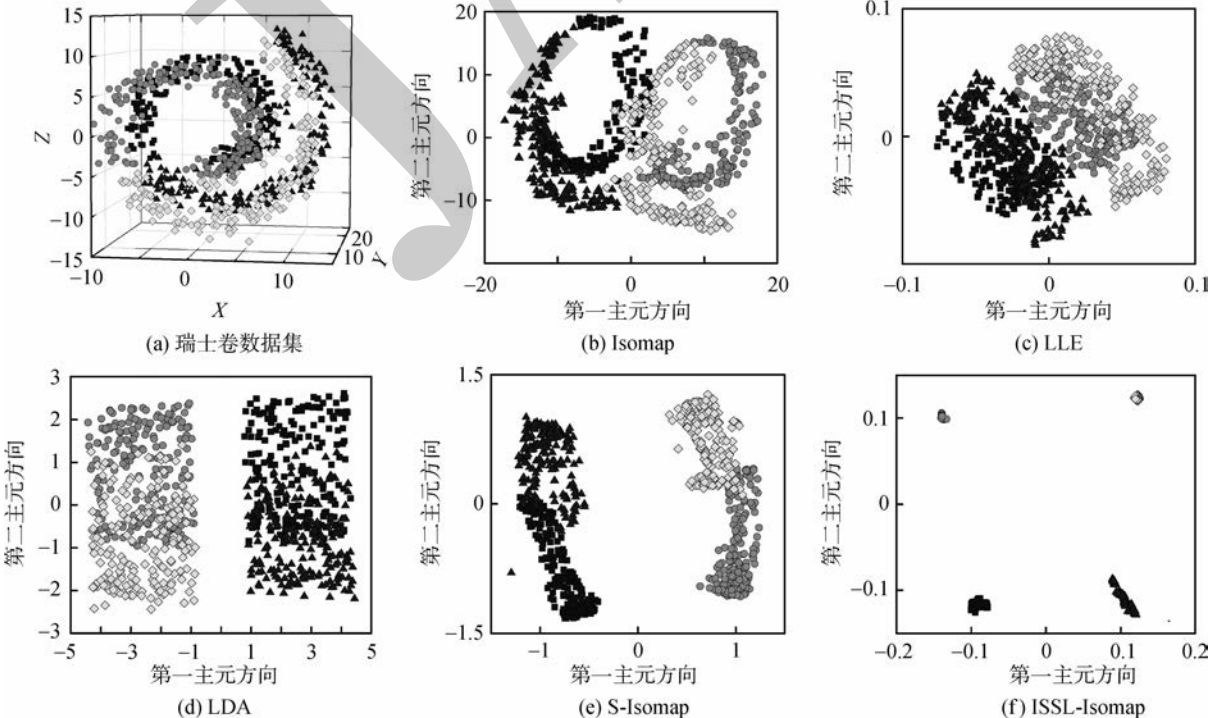


图 1 不同降维方法对瑞士卷数据集特征压缩结果

Fig. 1 Results of different dimensionality reduction methods on the Swiss roll data set

Isomap、LLE、LDA、S-Isomap 等方法样本混叠严重,不能有效将 4 类样本分离;而 ISSL-Isomap 方法可将 4 类样本完全分离,并且各类样本聚集性最好,这验证了本文所提降维方法的可行性。

3 GWO-SVM

3.1 GWO 方法

GWO 方法是一种新型群智能优化方法,通过模拟灰狼群体在捕食过程中的跟踪、包围、追捕、攻击猎物等觅食行为来实现目标优化^[17]。该方法中,狼群按照社会等级从高到低依次标记为 α 、 β 、 δ 、 ω ,下级行为需服从上级领导,并展开群体狩猎行动。GWO 方法的具体步骤如下。

$$\begin{cases} D_\alpha = |C_1 P_\beta - P|, & D_\beta = |C_2 P_\beta - P|, & D_\delta = |C_3 P_\delta - P| \\ P_1 = P_\alpha - H_1 D_\alpha, & P_2 = P_\beta - H_2 D_\beta, & P_3 = P_\delta - H_3 D_\delta \\ P(t+1) = (P_1 + P_2 + P_3)/3 \end{cases} \quad (18)$$

3.2 GWO-SVM 方法流程

SVM 是一种以统计学习理论为基础的机器学习方法,对非线性问题具有较强的拟合能力,但该方法性能与惩罚因子 c 以及核函数参数 g 的选取密切相关。GWO 具有收敛速度快,全局寻优能力强等优势,因此,将 GWO 方法应用于 SVM 模型的参数选取过程,其中 SVM 具体方法参见文献[16],本文选择径向基核函数(radial basis function, RBF)作为 SVM 内核函数。GWO-SVM 流程如图 2 所示,具体步骤如下:

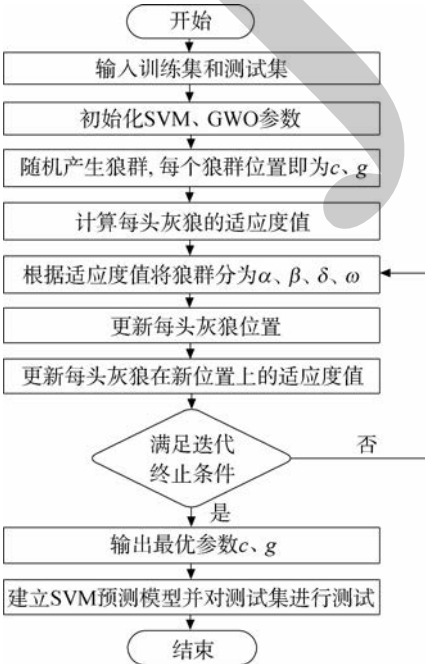


图 2 GWO-SVM 流程

Fig. 2 Flow chart of GWO-SVM

1) 包围猎物。为了捕猎猎物,灰狼首先会采取包围策略。这种包围行为的数学建模为

$$\begin{cases} B = |CP_p(t) - P(t)| \\ P(t+1) = P_p(t) - HD \\ H = 2\vartheta r_1 - \vartheta \\ C = 2r_2 \end{cases} \quad (17)$$

式中 t 为当前迭代次数, H 和 C 为向量系数, P_p 为猎物位置向量, P 为狼群位置向量。在迭代过程中,参数 ϑ 从 2 线性下降至 0, r_1 与 r_2 为区间 $[0, 1]$ 的随机数。

2) 追捕猎物。为模拟灰狼的捕猎过程,将 α 、 β 、 δ 作为主导进行猎物位置的搜索,其他搜索单位 ω 作为辅助,进而更新狼群位置

1) 输入 SVM 的测试集与训练集。

2) 初始化 SVM 与 GWO 参数,利用 GWO 方法自适应搜索全局最优位置,即得到 SVM 中 c 和 g 的最优值。其中,将每头灰狼的位置作为 SVM 的参数进行训练与测试,并将得到的测试样本正确识别率作为该灰狼的适应度值。

3) 将最优参数 c 和 g 用于 SVM 预测模型的构建,并利用训练好的预测模型对测试集进行测试。

4 行星齿轮箱故障诊断方法及应用

4.1 行星齿轮箱诊断方法

基于 ACMPE、ISSL-Isomap 与 GWO-SVM 相结合的行星齿轮箱诊断流程如图 3 所示,具体步骤如下。

步骤 1 在一定转速下以采样频率 f_s 分别采集各运行状态样本 Q 组,其中每种运行状态随机选取 q 组样本作为训练样本,剩余 $Q-q$ 组样本作为测试样本。

步骤 2 利用 ACMPE 方法提取出训练样本与测试样本的 ACMPE 值,每组样本得到 p 个特征值,其中 $j = K\tau_{\max}$, K 为 VMD 方法中的模态数; $\tau_{\max}$ 为 CMPE 方法中最大尺度因子,一般小于等于 20。

步骤 3 利用 ISSL-Isomap 方法对训练样本与测试样本构成的高维故障特征集进行降维处理,得到 d 维敏感故障特征集。

步骤 4 利用训练样本的 d 维故障特征集构

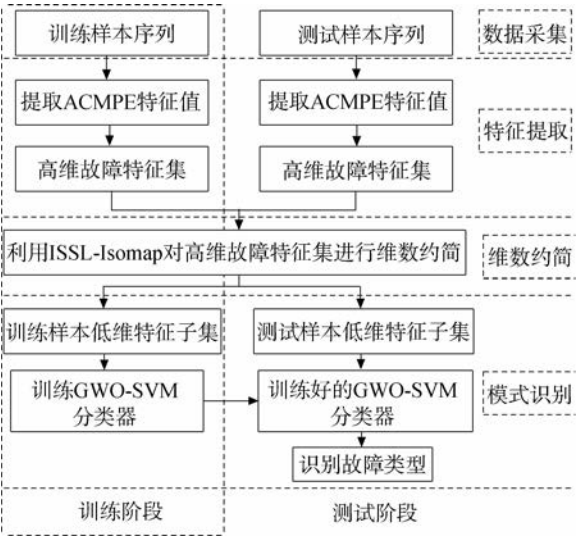


图 3 故障诊断方法流程

Fig. 3 Flow chart of fault diagnosis method

建 GWO-SVM 多类故障预测模型,再利用训练好的模型对测试样本的 d 维故障特征集进行模式识别,根据分类器输出结果确定行星齿轮箱的故障类型。其中,由于 GWO-SVM 是二分类器,因此采用“一对一”的方式建立多故障分类器。

本文所提 ISSL-Isomap 方法在行星齿轮箱故障诊断中起关键性作用,它自动衔接了高维故障特征集和 GWO-SVM 分类器,能够有效提取出便于识别的敏感、低维故障特征。

4.2 行星齿轮箱故障诊断实验

为模拟行星齿轮箱实际工作条件,实验在美国 Spectra Quest 公司开发的动力传动系统故障诊断综合实验台上进行。实验台固定于工作台上,变速驱动电动机提供动力,作为整个系统的输入端,经 2 级行星齿轮箱和 2 级平行齿轮箱,最终到达磁粉制动器。驱动电动机和磁粉制动器由相应调节器来控制,以模拟不同转速和负载等工况环境,实验平台如图 4 所示。本次实验中,行星齿轮箱基本参数如表 1 所示,输入轴转速为 2 083 r/min,负载扭矩为 14 N · m,对第 2 级太阳轮分别施加裂纹、磨损、缺齿以及断齿等故障,利用 1 号加速度传感器分别采集太阳轮正常和具有裂纹、磨损、缺齿、断齿等故障的振动加速度信号,采样频率为 7 680 Hz,通过滑动时间窗口得到各组样本,其中滑动窗口包含 2 048 个采样点,滑动步长为一个窗口。通过加窗分别提取太阳轮 5 种运行状态振动信号各 60 组,共计 300 组样本,行星齿轮箱 5 种状态振动信号的时域波形如图 5 所示。

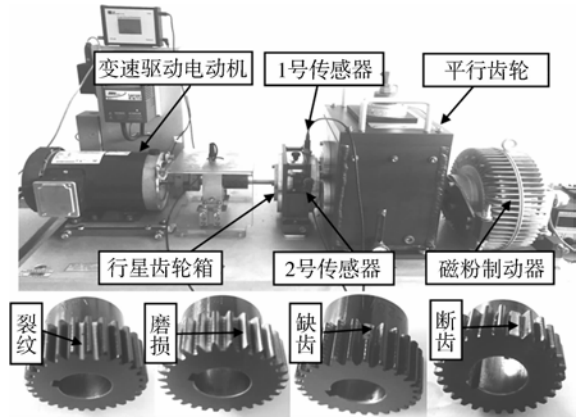


图 4 行星齿轮箱故障诊断平台

Fig. 4 Planetary gearbox fault diagnosis platform

表 1 行星齿轮箱参数

Table 1 Planetary gearbox parameters

级数	齿轮	齿数	级数	齿轮	齿数
1	太阳轮	20	2	太阳轮	28
	行星轮(3个)	40		行星轮(4个)	36
	内齿圈	100		内齿圈	100

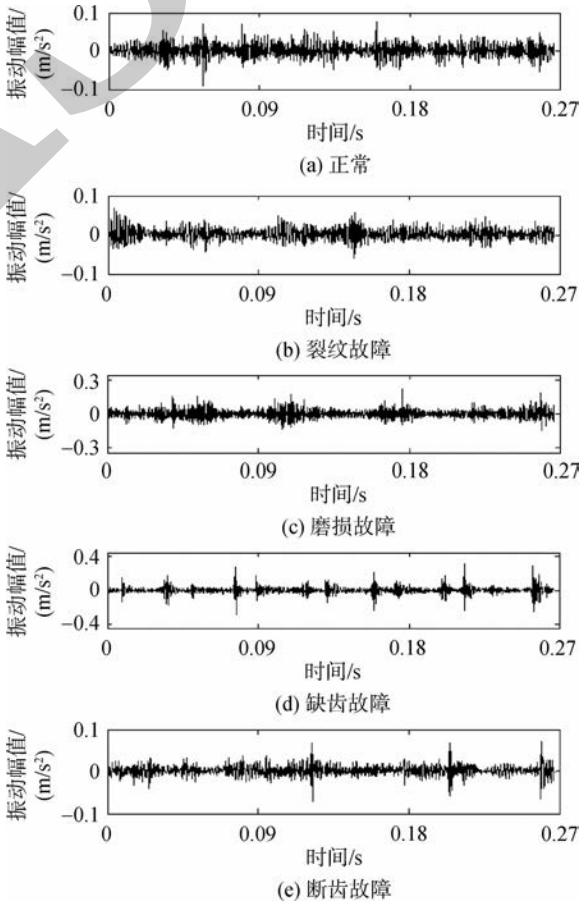


图 5 行星齿轮箱 5 种运行状态振动信号的时域波形
Fig. 5 Time domain waveform of five operating state vibration signals of planetary gearbox

利用所提 ACMPE + ISSL-Isomap + GWO-SVM 方法对齿轮箱 5 种运行状态进行故障诊断之前,对该方法中涉及的相关参数设置如下:①每种运行状态随机选取 20 组样本作为训练样本,剩余 40 组作为测试样本。②利用 ACMPE 进行复杂域特征提取时,设置模态数 5,裂纹故障振动信号的 VMD 分解结果如图 6 所示。此外,设置 ACMPE 嵌入维数 $m=6$,时延 $\lambda=1$,最大尺度因子 $\tau_{\max}=16$ 。③GWO-SVM 多故障分类器设置灰狼群规模为 10,最大迭代次数为 100,核函数为 RBF。

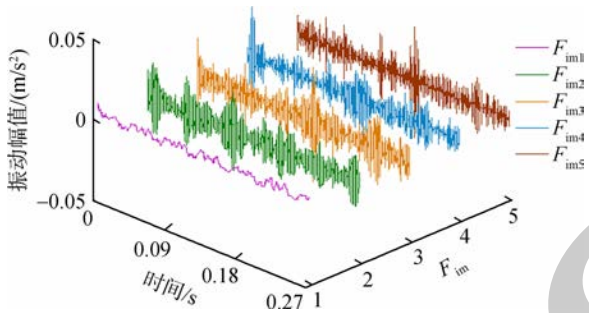


图 6 裂纹故障信号经 VMD 分解的时域波形
Fig. 6 Time domain waveform of the crack fault signal decomposed by VMD

本文将 ACMPE 方法应用于行星齿轮箱故障特征提取过程中,为验证该方法的有效性,将其与 MPE 和 CMPE 特征提取方法进行对比。其中设置 MPE 与 CMPE 的嵌入维数 $m=6$,时延 $\lambda=1$,最大尺度因子 $\tau_{\max}=16$ 。每种运行状态随机选取 3 组样本,MPE、CMPE 与 ACMPE 方法对齿轮箱 5 种运行状态分析结果如图 7 所示。据图 7 可知,对于同一故障的 3 组振动信号而言,CMPE 曲线一致性优于 MPE,这表明 CMPE 方法克服了传统 MPE 方法粗粒化方式的不足(即遗漏了部分尺度信息),能更准确、有效地提出不同尺度熵值。但 CMPE 在不同尺度下的 5 种运行状态振动信号的熵值较为接近,无法有效区分出各故障类型。与 MPE、CMPE 相比,ACMPE 方法在部分尺度上($\tau<20$),对 5 种不同运行样本区分较为明显,这验证了 ACMPE 方法应用于齿轮箱非线性振动信号的有效性。

为量化 MPE、CMPE 和 ACMPE 方法的特征提取效果,采用 GWO-SVM 多故障分类器对上述 3 种方法得到的故障特征集进行训练与测试,GWO 方法的参数优化结果(c, g)_{best}、最佳优化参数下的故障识别率 ξ 以及测试集样本的分类结果如图 8 所示,图 8 中纵坐标 1、2、3、4、5 分别代

表正常、裂纹故障、磨损故障、缺齿故障和断齿故障。据图 8 可知,GWO-SVM 多故障分类器能够在较少的迭代次数下达到最佳识别效果,验证了 GWO 方法的优势。此外,ACMPE 故障特征集测试样本故障识别率($\xi=91.5\%$)高于 MPE 故障特征集的故障识别率($\xi=84.5\%$)和 CMPE 故障特征集的故障识别率($\xi=88.5\%$),这进一步验证了 ACMPE 特征提取的优越性。

本文将 ISSL-Isomap 方法应用于行星齿轮箱高维 ACMPE 故障特征集的维数约简过程。为验证该降维方法的有效性,将其与局部切空间排列(LTSA)、Isomap、S-Isomap、加权 Isomap(weighted Isomap, W-Isomap)^[21]、监督型 SOINN L-Isomap(supervised SOINN L-Isomap, SSL-Isomap)等方法进行对比,6 种方法降维后得到的样本分布如图 9 所示。其中,上述方法涉及参数通过多次实验确定出最优值,具体设置如下:通过关联维数法确定最佳本征维数为 3。LTSA 近邻参数 $k=68$;Isomap 近邻参数 $k=51$;W-Isomap 近邻参数 $k=50$,权重因子 $w=0.3$ ^[21];S-Isomap 近邻参数 $k=50$,参数 $a=0.5$;SSL-Isomap 近邻参数 $k=51$,调节因子 2 为 $\varphi=0.38$;ISSL-Isomap 近邻参数 $k=140$,调节因子 2 为 $\varphi=0.38$ 。据图 9 可知:基于无监督学习的 LTSA 与 Isomap 方法降维后异类样本混叠现象严重,降维效果最差;基于监督学习的 W-Isomap、S-Isomap、SSL-Isomap 等方法虽可将 5 类样本基本分离,但仍存在部分异类样本混叠现象,降维效果次之;与上述 5 种方法相比,ISSL-Isomap 方法可将 5 类样本完全分离,并且各类样本聚集性最好,降维效果最佳。上述分析验证了利用 ISSL-Isomap 方法进行行星齿轮箱故障特征集维数约简的有效性。

为量化上述 6 种降维方法的降维效果,将各方法对应特征压缩结果分别输入至 GWO-SVM 多故障分类器中进行训练与测试,GWO-SVM 多故障分类器对 6 种方法的测试集样本的正确识别率如表 2 所示。据表 2 可知:LTSA、Isomap 无监督方式导致其无法在类别标签信息的指导下进行维数约简,易丢失大量利于分类的信息,出现一定数量样本类别误判现象,降维效果不如监督型降维方法;基于监督学习的 ISSL-Isomap 方法综合考虑了样本点的空间位置和夹角信息,因此,降维效果要优于仅考虑样本点空间位置信息的 W-Isomap、S-Isomap、SSL-Isomap 等监督降维方法。

上述分析进一步验证了 ISSL-Isomap 方法的优越性。

为避免应用单一分类器出现的偶然性,利用粒子群优化 SVM(particle swarm optimization SVM, PSO-SVM)多故障分类器、人工鱼群优化 SVM(artificial fish swarm algorithm SVM, AFSA-SVM)

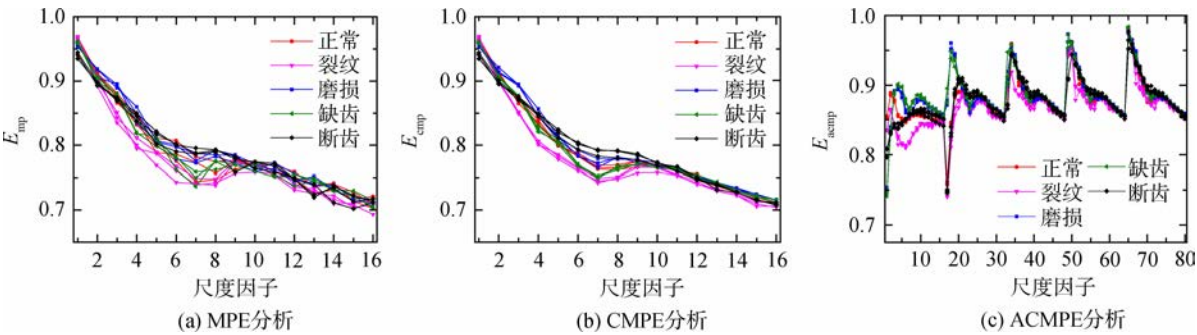


图 7 5 种状态行星齿轮箱振动信号的 MPE、CMPE、ACMPE 分析

Fig. 7 MPE, CMPE and ACMPE analysis of vibration signals of planetary gearbox with five states

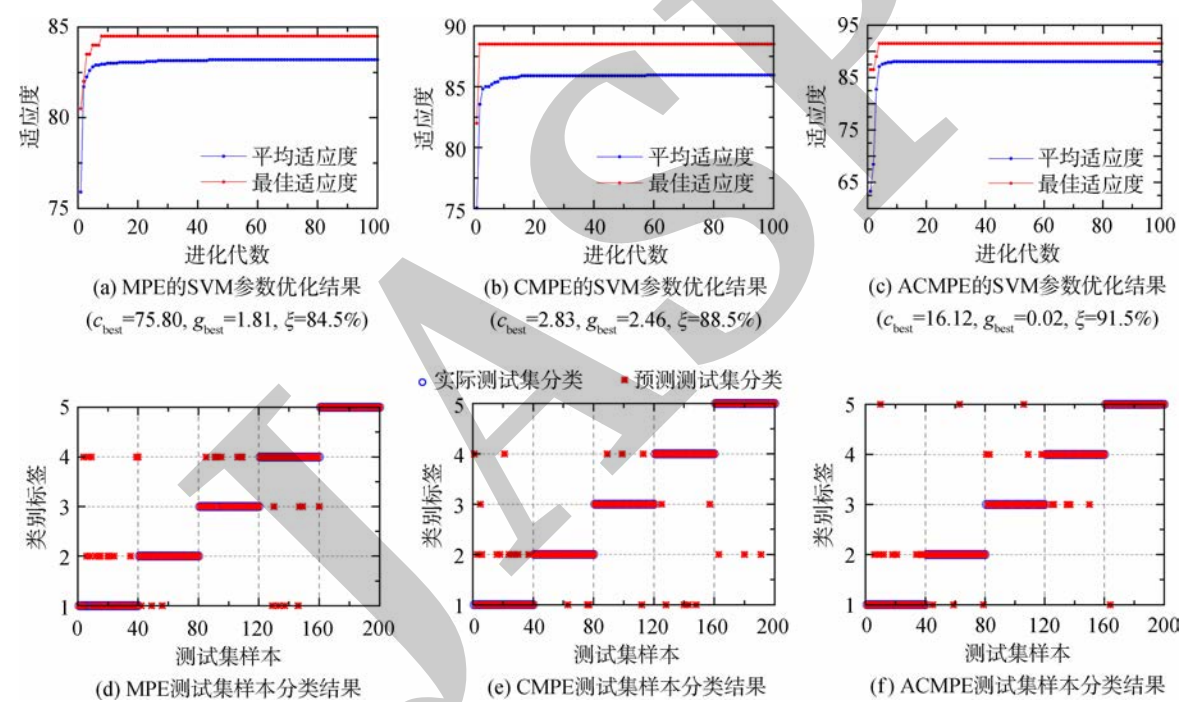
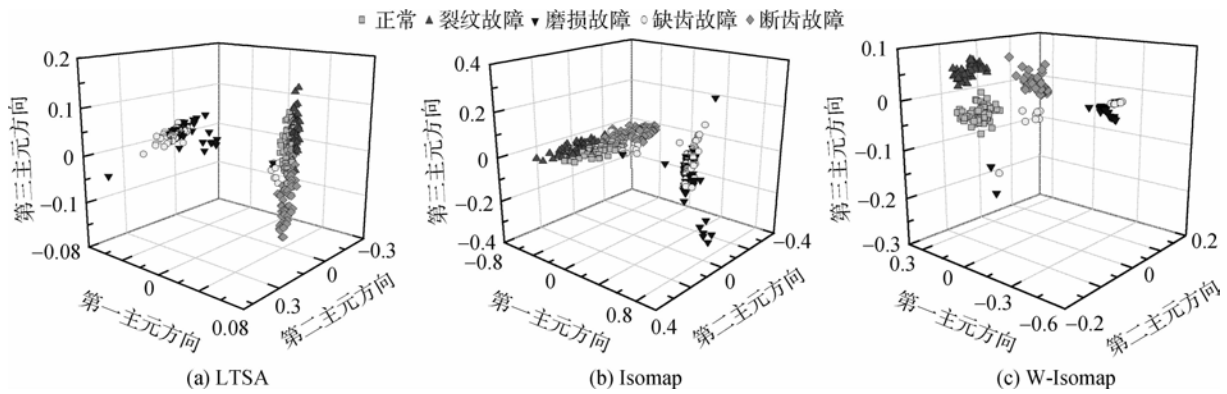


图 8 GWO-SVM 分类器对 MPE、CMPE 和 ACMPE 故障集参数优化及分类结果

Fig. 8 Parameters optimization and classification results of MPE、CMPE、ACMPE fault sets based on GWO-SVM classifier



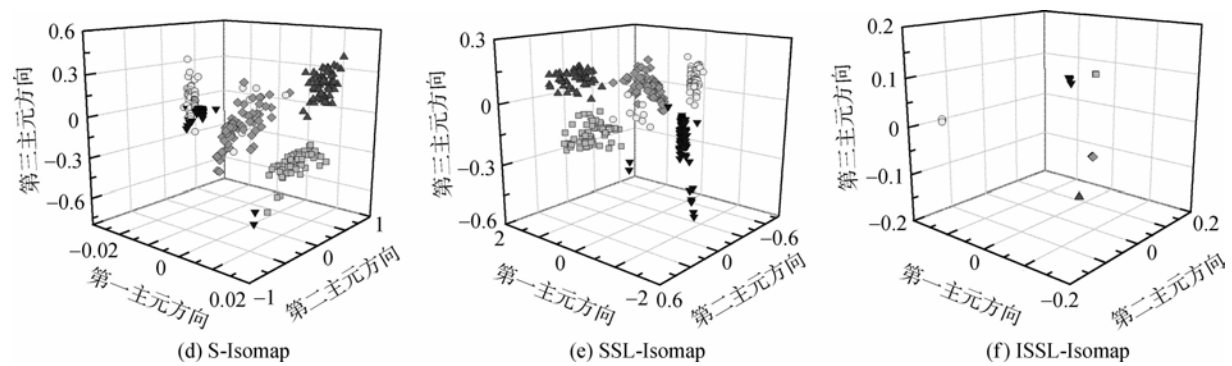


图 9 6 种方法降维结果

Fig. 9 Dimensionality reduction results of six methods

表 2 GWO-SVM 分类器对测试样本的正确识别率

Table 2 Recognition accuracy of testing samples under GWO-SVM classifiers							
降维方法	$(c, g)_{\text{best}}$	GWO-SVM 分类器对测试样本的正确识别率/%					
		正常	裂纹	磨损	缺齿	断齿	平均值
LTSA	(18.27,17.50)	75.0	92.5	80.0	65.0	95.0	81.5
Isomap	(9.98,1.71)	85.0	92.5	90.0	55.0	87.5	82.0
W-Isomap	(38.49,14.74)	100	100	97.5	90.0	97.5	97.0
S-Isomap	(49.38,3.44)	100	100	95.0	92.5	97.5	97.0
SSL-Isomap	(64.12,0.32)	100	100	97.5	90.0	100	97.5
ISSL-Isomap	(22.71,9.76)	100	100	100	100	100	100

多故障分类器、模拟退火 SVM(simulated annealing SVM, SA-SVM)多故障分类器以及 GWO-SVM 多故障分类器对上述 6 种维数约简后的特征集进行训练与测试,测试样本的平均识别率如图 10 所示,各分类器参数设置如表 3 所示。图 10 中横坐标 P1、P2、P3、P4、P5、P6 分别表示 LT-SA、Isomap、W-Isomap、S-Isomap、SSL-Isomap 和 ISSL-Isomap 特征压缩结果。

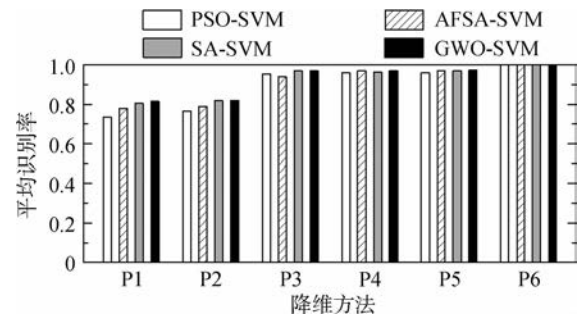


图 10 不同分类器下测试样本平均正确识别率

Fig. 10 Comparison of average recognition accuracy of testing samples under different classifiers

表 3 分类器参数设置

Table 3 Classifiers parameters setting

分类器	参数设置
PSO-SVM	局部搜索能力为 2,全局搜索能力为 2,粒子群数 10,终止迭代为 100
AFSA-SVM	拥挤度因子为 0.618,感知范围为 0.01,最大移动步长为 0.01,最大试探次数为 20,鱼群数为 10,终止迭代为 100
SA-SVM	马可夫链长度为 100,衰减参数为 0.85,步长因子为 0.01,初始温度为 8,终止温度为 3,Boltzmann 为 1
GWO-SVM	灰狼群数为 10,终止迭代为 100

据图 10 可知,相比于 LTSA、Isomap、W-Isomap、S-Isomap 和 SSL-Isomap 等方法,ISSL-Isomap 方法特征压缩结果在不同分类器下均具有最高的测试样本平均正确识别率,排除单一分类器识别率高的偶然性,验证了 ACMPE 与 ISSL-Isomap 相结合的方法能够精确地提取出能反应行星齿轮箱不同运行状态的故障特征;相比于其他分类器,GWO-SVM 多故障分类器对 6 种降维

方法特征压缩结果均具有最高平均正确识别率,验证了 GWO-SVM 多故障分类器在模式识别方面的优越性。

5 结 论

1) 提出一种衡量时间序列复杂性的方法——ACMPE,在一定程度上解决了 MPE 粗粒化过程的不足。实验结果表明,相比于 MPE 和 CMPE 方法,ACMPE 方法更能提取表征行星齿轮箱不同故障类型的特征信息。

2) 提出一种 ISSL-Isomap 降维方法。一方面将余弦相似度与欧氏距离相结合应用于距离矩阵的构建,并利用样本标签构造监督型成对约束邻域图;另一方面采用 SOINN 方法进行界标点的选取,增强了数据的分类能力。实验结果表明,该方法降维效果优于 LTSA、Isomap、W-Isomap、S-Isomap 和 SSL-Isomap 等方法。

3) 通过对行星齿轮箱的故障诊断实验分析表明,本文所提基于 ACMPE、ISSL-Isomap 和 GWO-Isomap 相结合的行星齿轮箱故障诊断方法集成了 ACMPE 方法在特征提取、ISSL-Isomap 方法在维数约简以及 GWO-SVM 方法在模式识别上的优势,能有效、精准地诊断出齿轮箱不同故障类型。

4) 通过将 GWO-SVM 分类器的分类效果与 PSO-SVM、AFSA-SVM、SA-SVM 分类器等方法进行对比,结果表明,在 SVM 参数优化方面,GWO 方法要优于现有的粒子群优化方法、人工鱼群优化方法和模拟退火优化方法。

尽管如此,本文方法也存在不足之处,如 ISSL-Isomap 中个别参数的选取还需人工设定,以及传动系统组合失效的问题该如何识别,目前正在对这些问题进行研究与完善。

参考文献:

[1] 范磊,王少萍,张超,等.直升机行星架疲劳裂纹扩展寿命预测[J].北京航空航天大学学报,2016,42(9):1927-1935.
FAN Lei,WANG Shaoping,ZHANG Chao,et al. Life prediction of helicopter planetary carrier plate fatigue crack propagation[J]. Journal of Beijing University of Aeronautics and Astronautics,2016,42(9):1927-1935. (in Chinese)

[2] 赵磊,郭瑜,伍星.基于振动分离信号构建和同步平均的行星齿轮箱轮齿裂纹故障特征提取[J].振动与冲击,2018,37(5):142-147.
ZHAO Lei, GUO Yu, WU Xing. Fault feature extraction

of gear tooth crack of planetary gearbox based on constructing vibration separation signals and synchronous average[J]. Journal of Vibration and Shock,2018,37(5):142-147. (in Chinese)

[3] LI Guoyan,LI Fangyi,WANG Yifan,et al. Fault diagnosis for a multistage planetary gear set using model-based simulation and experimental investigation[J]. Shock and Vibration,2016,2016:9263298. 1-9263298. 19.

[4] 胥永刚,赵国亮,马朝永,等.双树复小波域 MCA 降噪在齿轮故障诊断中的应用[J].航空动力学报,2016,31(1):219-226.
XU Yonggang,ZHAO Guoliang,MA Chaoyong,et al. Denoising method based on dual-tree complex wavelet transform and MCA and its application in gear fault diagnosis[J]. Journal of Aerospace Power,2016,31(1):219-226. (in Chinese)

[5] TIAN Ye,WANG Zili,LU Chen. Self-adaptive bearing fault diagnosis based on permutation entropy and manifold-based dynamic time warping[J]. Mechanical Systems and Signal Processing,2019,114:658-673.

[6] 刘永斌,龙潜,冯志华,等.一种非平稳、非线性振动信号检测方法的研究[J].振动与冲击,2007,26(12):131-134.
LIU Yongbin, LONG Qian, FENG Zhihua, et al. Detection method for nonlinear and nonstationary signals[J]. Journal of Vibration and Shock,2007,26(12):131-134. (in Chinese)

[7] YAN Ruqiang,LIU Yongbo,GAO R X. Permutation entropy:a nonlinear statistical measure for status characterization of rotary machines[J]. Mechanical Systems and Signal Processing,2012,29:474-484.

[8] 丁闯,张兵志,冯辅周,等.局部均值分解和排列熵在行星齿轮箱故障诊断中的应用[J].振动与冲击,2017,36(17):55-60.
DING Chuang,ZHANG Bingzhi,FENG Fuzhou,et al. Application of local mean decomposition and permutation entropy in fault diagnosis of planetary gearboxes[J]. Journal of Vibration and Shock,2017,36(17):55-60. (in Chinese)

[9] AZIZ W. Multiscale permutation entropy of physiological time series[C]//2005 Pakistan Section Multitopic Conference. Piscataway,US:Institute of Electrical and Electrical Engineering Computer Society,2005:10-23.

[10] ZHAO Liye,WANG Lei,YAN Ruqiang. Rolling bearing fault diagnosis based on wavelet packet decomposition and multi-scale permutation entropy[J]. Entropy,2015,17(9):6447-6461.

[11] 孙斌,薛广鑫.基于等距特征映射和支持矢量机的转子故障诊断方法[J].机械工程学报,2012,48(9):129-135.
SUN Bin,XUE Guangxin. Method of rotor fault diagnosis based on isometric feature mapping and support vector machine[J]. Journal of Mechanical Engineering,2012,48(9):129-135. (in Chinese)

[12] CHEN Fafa,TANG Baoping,SONG Tao,et al. Multi-fault diagnosis study on roller bearing based on multi-kernel

support vector machine with chaotic particle swarm optimization[J]. *Measurement*, 2014, 47(1): 576-590.

[13] 黄宏臣, 韩振南, 张倩倩, 等. 基于拉普拉斯特征映射的滚动轴承故障识别[J]. *振动与冲击*, 2015, 34(5): 128-134.
HUANG Hongchen, HAN Zhennan, ZHANG Qianqian, et al. Method of fault diagnosis for rolling bearings based on Laplacian eigenmap[J]. *Journal of Vibration and Shock*, 2015, 34(5): 128-134. (in Chinese)

[14] 李锋, 王家序, 汤宝平, 等. 有监督不相关局部 Fisher 判别分析故障诊断[J]. *振动工程学报*, 2015, 28(4): 657-665.
LI Feng, WANG Jiaxu, TANG Baoping, et al. Fault diagnosis method based on supervised uncorrelated local Fisher discriminant analysis[J]. *Journal of Vibration Engineering*, 2015, 28(4): 657-665. (in Chinese)

[15] ZHANG Tianyue, XU Baile, SHEN Furao. Fuzzy self-organizing incremental neural network for fuzzy clustering [C]// *International Conference on Neural Information Processing*. Cham, Switzerland: Springer, 2017: 24-32.

[16] 张小龙, 张氢, 秦仙蓉, 等. 基于 ITD 复杂度和 PSO-SVM 的滚动轴承故障诊断[J]. *振动与冲击*, 2016, 35(24): 102-107.
ZHANG Xiaolong, ZHANG Qing, QIN Xianrong, et al. Rolling bearing fault diagnosis based on ITD Lempel-Ziv complexity and PSO-SVM[J]. *Journal of Vibration and Shock*, 2016, 35(24): 102-107. (in Chinese)

[17] MIRJALILI S, MIRJALILI S M, LEWIS A. Grey wolf optimizer[J]. *Advances in Engineering Software*, 2014, 69(3): 46-61.

[18] ZHANG Ming, JIANG Zhinong, FENG Kun. Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump[J]. *Mechanical Systems and Signal Processing*, 2017, 93: 460-493.

[19] SILVA V D, TENENBAUM J B. Sparse multidimensional scaling using landmark points[R]. Palo Alto, US: Stanford University, 2004.

[20] GENG Xing, ZHAN Dechuan, ZHOU Zhihua. Supervised nonlinear dimensionality reduction for visualization and classification[J]. *IEEE Transactions on Systems, Man and Cybernetics*, 2005, 35(6): 1098-1107.

[21] VLACHOS M, DOMENICONI C, GUNOPOULOS D, et al. Non-linear dimensionality reduction techniques for classification and visualization[C]// *Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM, 2002: 645-651.

(编辑: 陈 越)