

本文全文线上首发由
上海电机系统节能工程技术
研究中心有限公司
支持

汤占军,孙润发.
基于多尺度模糊熵和STOA-SVM的风机
轴承故障诊断.
电机与控制应用,2021,48(12):66.



SEARI **NCMS**
上電科 上科電機

www.nems.com.cn

掌握电机行业制造核心技术,
辐射智能制造产业发展



· 业务范围 ·



异步电机装配线

上海电机系统节能工程技术研究中心有限公司 www.nems.com.cn

地址：上海市武宁路505号 手机：18918357509 电话：021-62574990-436

传真：021-62163904

邮箱：wangcj@seari.com.cn

基于多尺度模糊熵和 STOA-SVM 的风机 轴承故障诊断

汤占军，孙润发

(昆明理工大学 信息工程与自动化学院, 云南 昆明 650504)

摘要：针对风机轴承振动信号故障特征提取困难的问题, 提出了一种基于多尺度模糊熵(MFE)特征提取, 并结合乌燕鸥优化算法(STOA)优化支持向量机(SVM)的风机轴承故障诊断方法。首先采集原始振动信号并计算其多层次模糊熵, 其次构造故障特征向量集合作为 SVM 的输入, 最后采用 STOA 优化 SVM 对轴承故障进行分类诊断。通过凯斯西储大学轴承振动数据进行仿真, 结果显示轴承故障诊断准确率达到了 99.3%, 证明了所提方法具有较高的准确度和有效性。

关键词：风机轴承; 多尺度模糊熵; 乌燕鸥优化算法; 支持向量机; 故障诊断

中图分类号: TM 315 文献标志码: A 文章编号: 1673-6540(2021)12-0066-05

doi: 10.12177/emca.2021.139

Fan Bearing Fault Diagnosis Based on Multi-Scale Fuzzy Entropy and STOA-SVM

TANG Zhanjun, SUN Runfa

(Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650504, China)

Abstract: Aiming at the difficulty of fault feature extraction of fan bearing vibration signals, a fault diagnosis method of the fan bearing based on the multi-scale fuzzy entropy (MFE) feature extraction and combined with the sooty tern optimization algorithm (STOA) optimized support vector machine (SVM) is proposed. Firstly, the original vibration signals are collected and the multi-level fuzzy entropy is calculated. Secondly, the fault feature vector set is constructed as the input of SVM. Finally, the STOA is used to optimize SVM for classification and diagnosis of bearing faults. Simulation based on the bearing vibration data from Case Western Reserve University shows that the bearing fault diagnosis accuracy reaches 99.3%, which proves that the proposed method has high accuracy and effectiveness.

Key words: fan bearing; multi-scale fuzzy entropy (MFE); sooty tern optimization algorithm (STOA); support vector machine (SVM); fault diagnosis

0 引言

滚动轴承是风力发电机的关键部件, 及时对其故障进行诊断, 能够确保风电机安全高效运行^[1]。当轴承发生故障时, 安装在风机上的传感

器采集到的振动信号具有非线性、非平稳的特性^[2]。随着非线性理论的发展, 针对从这种特性的振动信号中提取特征困难的问题, 越来越多的非线性理论和方法, 如小波包分解、变分模态分解、样本熵等广泛应用于故障诊断领域^[3]。

收稿日期: 2021-10-30; 收到修改稿日期: 2021-11-18

作者简介: 汤占军(1969—), 男, 教授, 硕士生导师, 研究方向为风力发电。

孙润发(1998—), 男, 硕士研究生, 研究方向为风力发电。

Jaouher 等^[4] 使用经验模态分解(EMD)成功地从轴承信号中提取出了故障特征,但该方法在分解过程中容易出现模态混叠等问题。Dragomiretskiy 等^[5] 提出了变分模态分解(VMD)的方法,相比于 EMD 在分解振动信号方面更具优越性。Chen 等^[6] 在样本熵和近似熵的基础上提出了模糊熵(FE),采用指数函数代替阶跃函数,使熵值具有良好的连续性。郑近德等^[7] 结合模糊熵的概念,提出了多尺度模糊熵(MFE),并将其应用于滚动轴承的故障诊断。

本文采用 MFE 对轴承振动信号进行特征提取,并结合乌燕鸥优化算法支持向量机(STOA-SVM)模型对提取到的特征进行分类。首先对原始的振动信号进行采样,然后对采样到的信号求 MFE 来提取故障特征,其次通过 STOA 对 SVM 的惩罚因子 c 和核参数 g 进行寻优,最后将提取到的特征输入到优化后的 SVM 中进行分类,达到了较好的分类效果。

1 MFE

1.1 FE

FE 衡量的是新模式产生的概率大小,测度值越大,新模式产生的概率越大,即序列复杂度越大。

(1) 对于一个 M 点采样序列定义:

$$\{u(j) : 1 \leq j \leq M\} \quad (1)$$

(2) 按照序号的连续顺序重构生成一组 n 维矢量:

$$\begin{aligned} X_j^n = & \{u(j), u(j+1), \dots, u(j+n-1)\} - u_0(j), \\ & j = 1, \dots, (M-n) \end{aligned} \quad (2)$$

式中: $\{u(j), u(j+1), \dots, u(j+n-1)\}$ 为第 j 个点开始连续 n 个 u 的值; $u_0(j)$ 为均值, $u_0(j) = \frac{1}{n} \sum_{i=0}^{n-1} (i+j)$ 。

(3) 定义两个 n 维矢量 X_i^n 和 X_j^n 之间的距离 d_{ij}^n 为两者对应元素中差值最大的一个,即:

$$\begin{aligned} d_{ij}^n = & d[X_i^n, X_j^n] = \\ & \max_{k \in (0, n-1)} \{ |u(j+k) - u_0(j) - [u(i+k) - u_0(i)]| \}, \\ & i, j = 1 \sim (M-n), j \neq i \end{aligned} \quad (3)$$

(4) 用模糊函数 $\mu(d_{ij}^n, m, r)$ 定义两个矢量 X_i^n 和 X_j^n 的相似度 D_{ij}^n ,即:

$$D_{ij}^n = \mu(d_{ij}^n, m, r) = \exp \left[(-d_{ij}^n) \frac{m}{r} \right] \quad (4)$$

式中:函数 $\mu(d_{ij}^n, m, r)$ 为指数函数; m 、 r 分别为指数函数边界的梯度和宽度。

(5) 定义函数:

$$o^n(m, r) = \frac{1}{M-n} \sum_{i=1}^{M-n} \left\{ \frac{1}{M-n-1} \sum_{j=1, j \neq i}^{M-n} D_{ij}^n \right\} \quad (5)$$

(6) 重复(2)~(5)步骤,按照序号顺序重构一组 $n+1$ 维矢量,函数定义如下:

$$o^n(m, r) = \frac{1}{M-n} \sum_{i=1}^{M-n} \left\{ \frac{1}{M-n-1} \sum_{j=1, j \neq i}^{M-n} D_{ij}^{n+1} \right\} \quad (6)$$

(7) FE 定义为

$$\text{FuzzyEn}(n, m, r) = \lim_{M \rightarrow \infty} [\ln o^n(m, r) - \ln o^{n+1}(m, r)] \quad (7)$$

当 M 值为有限值时,按照上述七步得出的序号长度 M 时 FE 的估计为

$$\text{FuzzyEn}(n, m, r, N) = \ln o^n(m, r) - \ln o^{n+1}(m, r) \quad (8)$$

FE 将原始序列映射到高维空间,利用幅值容限内的高维矢量来定义序列复杂度,同时采用模糊函数定义序列相似性,更加符合实际情况。

1.2 MFE

FE 从单一尺度去考虑时间序列的复杂度,很可能丢失轴承振动信号的故障信息,因此提出了从多个尺度去考虑轴承振动时间的序列,这就形成了 MFE。MFE 比 FE 多了尺度因子 τ ,对于长度为 M 的时间序列 $\{x_i\} = \{x_1, x_2, \dots, x_M\}$,对其给定的维数 n 和梯度 r 进行粗粒分割,形成新的粗粒向量 y_i^τ ,其公式如下:

$$y_i^\tau = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{\tau} x_i \quad (9)$$

对粗粒分割得到的向量分别进行相对运算,即可得到相应的多尺度分析结果。

MFE 的参数有维数 n 、模糊函数的边界梯度 m 和宽度 r 、尺度因子 τ 。FE 的维数决定了细节信息的丰富程度,但是维数越多对于计算的复杂度影响越大,因此一般设定 $n=2$ 。模糊函数的梯度 m 决定了序列相似性度量权值,本文取 $m=2$ 。模糊函数的宽度 r 一般与时间序列有关, $r=(0.1 \sim 0.25) \cdot SD$, SD 为序列的标准差,本文将 r 设定为 $0.1 \cdot SD$ 。尺度因子 τ 是对时间序列进行粗粒

分割的数量,在尺度因子等于1时,MFE即为FE,正常轴承振动信号的FE值比故障状态振动信号的FE值小,但是并不能反应出故障振动信号比正常振动信号复杂,相反,进一步说明了多尺度分析的重要性。本文取 $\tau=2\sim4$ 进行准确度对比,以 $\tau=4$ 为例进行后续步骤的详细说明。时间序列粗粒分割的过程如图1所示。

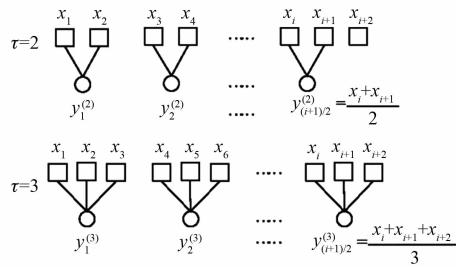


图1 时间序列粗粒化过程

2 STOA

STOA是Dhiman等^[8]于2019年提出的一种新型优化算法,优化灵感来源于乌燕鸥觅食的行为,这种优化算法具有很高的全局搜索能力。其基本原理分为迁徙行为和攻击行为两个部分,分别进行全局探索和局部探索,具有很高的精度^[9]。

2.1 迁徙行为

迁徙行为也就是全局探索,主要分为三个部分:避免碰撞、最优收敛和最优更新。

(1) 避免碰撞。 S_A 用于计算新的搜索乌燕鸥位置,以避免与其相邻搜索乌燕鸥位置发生碰撞,其公式如下:

$$C_{st} = S_A \cdot P_{st} \quad (10)$$

式中: C_{st} 为不与其他乌燕鸥发生冲突的前提下应当处在的位置; S_A 为乌燕鸥在给定空间内的移动; P_{st} 为乌燕鸥的当前位置, Z 表示当前的迭代次数。

S_A 计算式如下:

$$S_A = C_f - [Z \cdot (C_f / \text{MAX}_{\text{iterations}})], \\ Z = 0, 1, 2, \dots, \text{MAX}_{\text{iterations}} \quad (11)$$

这里的 C_f 是一个控制变量,控制 S_A 的线性递减,本文 C_f 的值设置为2,因此 S_A 会从2线性递减为0。 C_f 的取值参考文献[8],通过所涉及算法得到最佳取值,在44个基准测试函数上进行测试,验证了参数的有效性。

- 68 -

(2) 最优收敛。在避免碰撞的前提下,乌燕鸥的位置向最佳位置收敛,其公式如下:

$$M_{st} = C_B^* [\mathbf{P}_{bst}(Z) - \mathbf{P}_{st}(Z)] \quad (12)$$

式中: M_{st} 为乌燕鸥的不同位置; \mathbf{P}_{bst} 为乌燕鸥的最优位置; C_B 为一个使搜索更加全面的随机变量。

随机变量 C_B 为

$$C_B = 0.5 \cdot R_{rand} \in (0, 1) \quad (13)$$

(3) 最佳优更新。乌燕鸥的位置更新为搜索到的最优乌燕鸥位置,其轨迹 D_{st} 的表达式如下:

$$\mathbf{D}_{st} = \mathbf{C}_{st} + M_{st} \quad (14)$$

2.2 攻击行为

乌燕鸥在空中攻击时使用扑翼飞行模式,可以通过翅膀来增加飞行高度,也可以调整速度和角度,转化为数学模型为^[10]

$$\begin{cases} x' = R_{\text{adius}} \cdot \sin(i) \\ y' = R_{\text{adius}} \cdot \cos(i) \\ z' = R_{\text{adius}} \cdot i \\ r = u \cdot e^{kv} \end{cases} \quad (15)$$

式中: R_{adius} 为每圈的螺旋半径; $i \in (0, 2\pi)$; u 和 v 为定义螺旋形状的常数。

乌燕鸥的位置更新公式为

$$\mathbf{P}_{st}(Z) = [\mathbf{D}_{st} \cdot (x' + y' + z')] \cdot \mathbf{P}_{bst}(Z) \quad (16)$$

3 STOA-SVM 故障诊断模型

本文提出的STOA-SVM故障诊断模型,采用STOA算法对SVM的惩罚因子 c 和核参数 g 进行优化,以准确率为适应度函数进行寻优^[11]。流程图如图2所示,具体步骤如下:

- (1) 对原始振动信号进行采样;
- (2) 对采样到的时序信号求其MFE;
- (3) 利用MFE构建特征向量集合;
- (4) 对构建的特征向量按照故障类别打标签;
- (5) 将提取的特征向量集合按照一定比例分为训练集和测试集;
- (6) 对STOA算法的各项参数进行初始化;
- (7) 将提取的训练集故障特征信息输入到STOA-SVM模型中进行训练;
- (8) 利用训练好的STOA-SVM模型对测试集进行分类诊断;

(9) 得出其分类的准确率。

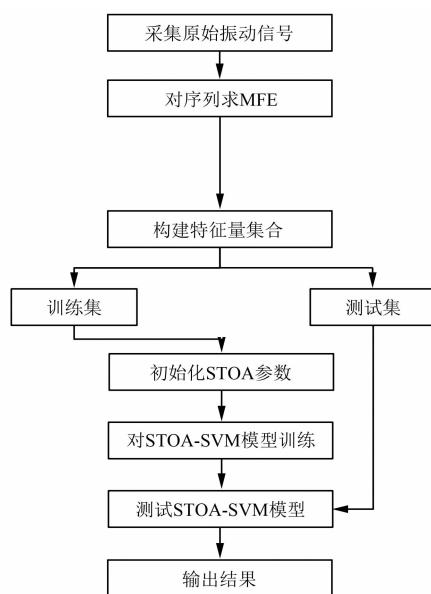


图 2 故障诊断模型流程图

4 轴承故障诊断试验验证

4.1 试验数据

考虑到风机轴承为滚动轴承,其中凯斯西储大学轴承数据^[12]最具代表性,因此为了进一步验证本文所提方法的合理性,采用凯斯西储大学开放轴承数据进行试验。以内圈故障数据为例,数据采集装置参数如表 1 所示。对轴承在相同载荷下 4 种状态的振动信号进行采集,分别为正常状态(NR)、内圈故障(IR)、外圈故障(OR)、滚动体故障(BR),采用控制变量法进行试验,选取 3 种故障状态的故障尺寸均为 0.007 英寸(1 英寸≈2.54 cm)的数据进行试验,每类故障采集 50 组样本数据,共 200 组样本数据,将其中 60 组数据作为训练集样本训练模型,剩余 140 组数据作为测试集样本测试模型准确率。本文分为 3 个尺度进行对比试验,通过准确率的对比说明了本文所取尺度 $\tau=4$ 的优越性。

表 1 数据采集装置参数

故障类型	轴承型号	电机转速/ (r·min ⁻¹)	故障尺寸/ 尺寸	采样频 率/kHz
内圈故障	6205-2RS JEM SKF	1 772	0.007	120 000

4.2 数据仿真

现将本文所提方法应用于上述试验数据,具

体步骤如下:

(1) 计算所有样本的 MFE, 构建 200 组特征向量样本, 分为训练集 60 组和测试集 140 组, 如表 2 所示。

表 2 样本分类数

故障类型	训练样本数	测试样本数	总计
正常状态	15	35	50
内圈故障	15	35	50
外圈故障	15	35	50
滚动体故障	15	35	50

(2) 将训练样本特征集输入到 STOA-SVM 模型中进行训练, 对 SVM 模型的惩罚因子 c 和核参数 g 进行优化。对 STOA 的参数进行初始化设置, 种群大小 20, 最大迭代次数设置为 30 次, 变量个数为 2, 其次对 SVM 的优化参数进行范围设置, 惩罚因子 c 的范围为 [1, 1 000], 核参数 g 的范围为 [1, 100], 最后以准确率为适应度函数开始对 SVM 的参数进行寻优。其适应度函数如图 3 所示, 得到的最优惩罚因子 $c = 877.304$, 核参数 $g = 10.323$ 。

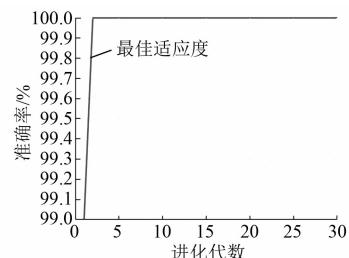


图 3 STOA 进化曲线

(3) 将测试集样本特征集输入到训练好的 STOA-SVM 模型中, 得到整个模型的故障分类诊断结果如图 4 所示, 140 个测试样本的具体情况如表 3 所示。可见本文所提 MFE-STOA-SVM 方

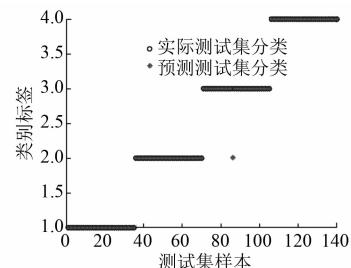


图 4 故障分类结果

法在轴承故障诊断中,准确率达到了 99.3%,并与 SVM 和粒子群算法优化的支持向量机(PSO-SVM)模型进行对比,如表 3 所示,证明了本文所提方法具有很高的准确性和有效性。

表 3 不同算法模型的对比

分类模型	准确度/%
STOA-SVM($\tau=2$)	98.6
STOA-SVM($\tau=3$)	98.6
STOA-SVM($\tau=4$)	99.3
SVM	92.9
PSO-SVM	97.8

5 结语

本文提出了一种基于 MFE 和 STOA-SVM 结合的风机轴承故障诊断方法,运用凯斯西储大学轴承数据集进行验证,轴承故障诊断的准确率达到了 99.3%,并与不同尺度模型以及 SVM 和 PSO-SVM 模型进行对比,证明了本文所提方法的准确性和有效性。

【参考文献】

- [1] 杜小磊,陈志刚,张楠,等.基于同步挤压 S 变换和深度学习的轴承故障诊断[J].组合机床与自动化加工技术,2019(5): 90.
- [2] 李益兵,王磊,江丽.基于 PSO 改进深度置信网络的滚动轴承故障诊断[J].振动与冲击,2020,39(5): 89.
- [3] 刘敏,范红波,张英堂,等.机械振动信号自适应多尺度非线性动力学特征提取方法研究[J].振动与冲击,2020,39(14): 224.

- [4] JAOUHER B A, NADER F, LOTFI S, et al. Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals [J]. Applied Acoustics, 2015, 89: 16.
- [5] DRAGOMIRETSKIY K, ZOSO D. Variational mode decomposition [J]. IEEE Transactions on Signal Processing, 2014, 62(3): 531.
- [6] CHEN W T, WANG Z Z, XIE H B, et al. Characterization of surface EMG signal based on fuzzy entropy [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2007, 15(2): 266.
- [7] 郑近德,陈敏均,程军圣,等.多尺度模糊熵及其在滚动轴承故障诊断中的应用[J].振动工程学报,2014,27(1): 145.
- [8] DHIMAN G, KAUR A. STOA: a bio-inspired based optimization algorithm for industrial engineering problems [J]. Engineering Applications of Artificial Intelligence, 2019, 82: 148.
- [9] 贾鹤鸣,李瑶,孙康健.基于遗传乌燕鸥算法的同步优化特征选择[J].自动化学报,DOI: 10.16383/j.aas.c200322.
- [10] TAMURA K, YASUDA K. The spiral optimization algorithm: convergence conditions and settings [J]. IEEE Transactions on Systems, Man & Cybernetics: System, 2017, 50(1): 360.
- [11] 朱兴统.基于蝙蝠算法优化相关向量机的轴承故障诊断方法[J].自动化与仪表,2021(2): 21.
- [12] SMITH W A, RANDALL R B. Rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study [J]. Mechanical Systems and Signal Processing, 2015, 64-65: 100.

~~~~~

欢迎订阅 2022 年《电机与控制应用》

2021 年在邮局漏订的读者,可联系编辑部补订,具体请咨询编辑部

地址: 上海市武宁路 505 号《电机与控制应用》编辑部

邮编: 200063

电话: 021-62574990-462

邮箱: eec@seari.com.cn

国内邮发代号: 4-199

每册定价: 15.00 元

全年定价: 180.00 元