

基于 MFE—GASVM 的机载燃油泵故障诊断

戴邵武¹, 陈强强¹, 戴浩然², 邢璐¹

(1. 海军航空大学 岸防兵学院, 山东 烟台 264000; 2. 中国人民解放军空军 95596 部队, 河南 商丘 476000)

摘要: 机载燃油泵的健康状态是保障飞行任务完成的先决条件, 实现机载燃油泵故障诊断的关键是敏感故障特征的提取; 由于机械系统的复杂性, 机载燃油泵振动信号的随机性表现在不同尺度上, 因此需要对振动信号进行多尺度分析; 基于此, 在模糊熵 (fuzzy entropy, FE) 的基础上引入尺度因子, 对振动信号在不同尺度下进行复杂性度量; 然后将多尺度模糊熵 (multiscale fuzzy entropy, MFE) 特征量作为支持向量机 (support vector machine, SVM) 的输入参数以识别故障状态, 并采用遗传算法 (genetic algorithm, GA) 对支持向量机的核函数参数及惩罚参数进行优化; 实验结果表明, 该方法可有效提取故障特征, 实现机载燃油泵的故障诊断。

关键词: 燃油泵; 故障诊断; 多尺度模糊熵; 支持向量机; 遗传算法

Airborne Fuel Pumps Fault Diagnosis Based on MFE—GASVM

Dai Shaowu¹, Chen Qiangqiang¹, Dai Haoran², Xing Lu¹

(1. College of Coastal Defense, Naval Aviation University, Yantai 264000, China;

2. Air force Unit 95596, Shangqiu 476000, China)

Abstract: The health state of airborne fuel pumps is important for the safety of flight mission, so the fault state feature extraction and diagnosis for the pumps become an important factor. Due to the complexity of mechanical systems, the randomness of the vibration signal behave on different scales, making it necessary to analyze the vibration signal in a multi-scale way. Multi-scale fuzzy entropy (MFE) is based on fuzzy entropy is defined to measure the complexity of time series in different scale factors. The MFE characteristic parameters are input to support vector machine for fault classification. Genetic algorithm is applied into adaptive selection of the best penalty parameter and kernel function parameter. By analysis experimental data, the results show that the proposed method can differentiate the fault categories of fuel pumps effectively.

Keywords: fuel pumps; fault diagnosis; multi-scale fuzzy entropy; support vector machine; genetic algorithm

0 引言

随着武器装备制造工艺及设计理念的不断更新换代, 武器装备的作战效能及作战时间都得到了很大程度的提高, 同时现代战场对武器装备的可靠性及维修性也提出了更高的要求^[1]。机载燃油泵系统主要为航空发动机提供指定的流量及压力状态下的燃油, 为航空发动机源源不断的输送“血液”, 是提高航空发动机作战能力并确定航空发动机作战效能稳定发挥的重要保障之一。机载燃油泵系统运行状态的优良与否, 直接程度的关系着飞行任务及飞行安全^[2]。在现代空战领域, 机载燃油泵系统的故障可能导致航空发动机性能的不稳定, 严重时甚至导致机毁人亡, 造成不可估计的损失。因此, 对机载燃油泵系统实现准确的故障诊断, 对于提高机载燃油泵的可靠性具有重要意义。

对机载燃油泵系统中振动传感器所采集到的振动信号进行准确有效的故障特征提取, 是实现机载燃油泵故障诊断的关键^[3]。经典的非线性分析算法如混沌 (Chaos) 理

论、Lyapunov 算法等在线性信号特征提取方面具有广泛应用^[4]。李晓娟等^[5]采用关联维数对飞参系统所记录的航空发动机状态参数进行分析, 提取出反映航空发动机状态参数的特征信息, 为航空发动机状态检测及视情维修提供了理论基础。经典的非线性算法在故障诊断领域具有重要作用, 但在使用过程中存在着依赖数据长度、计算速度较慢、精度不高等不足^[6]。而且随着机械系统复杂程度的不断提高, 机载燃油泵等航空装备也逐步趋向于智能化、精密化、复杂化, 所对应的振动信号也更加复杂, 具有显著的非线性、非稳定性及非高斯型, 为机载燃油泵的故障特征提取增加了难度。

为了弥补经典非线性算法的不足, 熵 (Entropy) 理论逐渐在时间序列分析及故障特征提取中取得广泛的应用。Pincus 等^[7]最早提出近似熵 (approximate entropy, ApEn) 的概念, 并将其用于医学方面的心电信号特征提取。但近似熵在使用过程中存在着自身匹配困难、过于依赖数据长度等不足; 为了弥补近似熵在时间序列分析中的不足, 在近似熵的基础上, Richman 等^[8]提出了样本熵 (sample entropy, SE) 的概念, 用来衡量时间序列的复杂性。样本熵为非线性信号的特征提取提供了思路, 但其不足之处在于所构建的函数基于单位阶跃函数 (0~函数) 定义样本熵值, 无法准确判断样本类别。在此基础上, 陈伟婷等提出

收稿日期: 2019-06-26; 修回日期: 2019-08-02。

基金项目: 山东自然科学基金面上项目 (ZR2017MF036); 国防科技项目基金 (F062102009)。

作者简介: 戴邵武 (1966-), 男, 博士, 教授, 主要从事飞行器综合导航方向的研究。

了模糊熵 (fuzzy entropy, FE) 的概念^[9], 通过采用模糊函数代替阶跃函数, 解决了相似性度量的问题, 并成功应用于机械系统故障诊断^[10]。

机载燃油泵系统工作环境恶劣、制造工艺复杂, 在运行过程中其振动信号具有强烈的复杂程度, 且故障特征信息分布在多个尺度中, 此时单一的 FE 熵值不足以完整概括其故障特征。为了弥补 FE 的不足并提高其特征提取能力, 本文采用多尺度模糊熵 (multi-scale fuzzy entropy, MFE) 对机载燃油泵系统振动信号进行分析; 将 MFE 提取的多尺度熵值作为特征向量输入至由遗传算法 (genetic algorithm, GA) 优化的支持向量机 (support vector machine, SVM) 中完成故障诊断。通过对机载燃油泵系统故障振动信号进行实验分析, 结果表明 MFE-GASVM 算法可以有效的提取故障信息并完成不同故障类型的识别, 是一种有效的故障诊断方法。

1 基于多尺度模糊熵的特征提取

1.1 模糊熵定义

模糊熵通过借用模糊函数概念, 选择指数函数 $e^{-\left(\frac{d}{\tau}\right)^n}$ 作为模糊函数来测量两个向量之间的相似性。指数函数 $e^{-\left(\frac{d}{\tau}\right)^n}$ 具有以下特性^[11]:

- 1) 连续性保证模糊熵的函数值不会产生数据突变;
- 2) 指数函数的凸性质保证了向量本身自相似性值最大。

在理论分析中, 对于长度为 N 的时间序列 $x(i)$, 其模糊熵求解过程定义如下。

步骤 1: 按顺序定义 m 维向量。

$$X_i^m = \{x(i), x(i+1), \dots, x(i+m-1)\} - x_0(i) \quad (1)$$

式中, $i = 1, 2, \dots, N-m+1$ 。 X_i^m 定义为从第 i 个点开始的连续 m 个去掉均值 $x_0(i)$ 的 $x(i)$ 值, 其中

$$x_0(i) = \frac{1}{m} \sum_{j=0}^{m-1} x(i+j) \quad (2)$$

步骤 2: 定义 X_i^m 与 X_j^m 之间的距离为两者对应元素差值的最大值, 即

$$d_{ij}^m = d[X_i^m, X_j^m] = \max_{k \in (0, m-1)} \{ |(x(i+k) - x_0(i)) - (x(j+k) - x_0(i))| \} \\ i, j = 1, 2, \dots, N-m, i \neq j \quad (3)$$

步骤 3: 通过定义模糊函数 $\mu(d_{ij}^m, n, r)$, 确定矢量 X_i^m 与 X_j^m 之间的相似度 D_{ij}^m

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

式中, 模糊函数 $\mu(d_{ij}^m, n, r)$ 为指数函数, n, r 分别表示边界梯度和宽度。

步骤 4: 定义函数

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

步骤 5: 对维数进行 $m+1$ 处理, 在此基础上, 重复步骤 1~4, 得:

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

步骤 6: 当 N 为有限长度时, 根据公式 (6), 可定义模糊熵为

$$FE(m, n, r, N) = \ln \varphi^m(n, r) - \ln \varphi^{m+1}(n, r) \quad (7)$$

1.2 多尺度模糊熵

多尺度模糊熵定义为不同尺度因子下的模糊熵值, 通过引入多尺度粗粒化过程, 实现不同尺度下的模糊熵值分析, 多尺度模糊熵的计算方法如下^[12]:

考虑时间序列 $\{x(i), i = 1, 2, \dots, N\}$, 对其进行多尺度粗粒化, 得到多尺度粗粒化序列 $\{y_j^{(\tau)}\}$

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i \quad (8)$$

式中, $j = 1, 2, \dots, \left[\frac{N}{\tau}\right]$, $\left[\frac{N}{\tau}\right]$ 表示对 $\frac{N}{\tau}$ 进行取整运算; τ 为尺度因子, 当 $\tau = 1$ 时, 即不进行多尺度粗粒化过程, 此时得到的多尺度粗粒化序列即为原始序列, 对应的多尺度模糊熵数值形式即为原始模糊熵值。

通过引入多尺度粗粒化过程, 计算每个多尺度粗粒化序列的模糊熵值, 并将其刻画成尺度因子的函数, 所得多个尺度因子下的模糊熵值即为多尺度模糊熵。尺度因子的取值在一定程度上影响着多尺度模糊熵值分析。

2 基于 GASVM 的故障诊断算法

支持向量机作为经典的机器学习算法, 在模式识别、时间序列预测等方面具有广泛的应用。在采用 SVM 进行机载燃油泵系统故障诊断的过程中, 通过将非线性样本作为输入向量, 并将其映射到高维特征空间中, 然后在高维特征空间进行全局搜索以得到最优超平面, 从而实现样本的分类^[13]。在 SVM 实现分类的过程中, 设样本的训练集为 $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)\}$, 其中 $x_i \in \mathbf{R}^n$ 表示特征向量, y_i 表示类别标签 ($i = 1, 2, \dots, l$)。

将 SVM 的寻找最优超平面问题转换为求解二次规划问题, 可设置目标函数及约束条件为:

$$\min_{\omega, b} \frac{1}{2} \omega^2 + C \sum_{i=1}^l \xi_i \\ s. t. y_i (\langle x_i, \omega \rangle + b) \geq 1 - \xi_i \quad (9)$$

式中, $\langle \cdot \rangle$ 为内积形式, ω 为权值向量, b 为偏置; ξ_i 表示松弛变量, 在数值形式中为非负数, 用来衡量数据点的偏离程度; C 为惩罚参数, C 的值越大表示对错误分类的惩罚越大。

根据 Mercer 定理, 引入核函数概念, 通过采用不同的核函数 $K(x, x_i)$, 以实现非线性分类^[14]。引入 Lagrange 算子 a_i , 则最优分类的超平面决策函数可表示为:

$$f(x) = \text{sgn} \left[\sum_{i=1}^l a_i y_i K(x, x_i) + b \right] \quad (10)$$

使用 SVM 对经过多尺度模糊熵算法所提取出的机载燃油泵系统故障特征向量进行故障诊断时, 核函数参数及惩罚参数对诊断结果影响较大, 需要对其进行最优的自适应选择, 以取得较好的故障诊断结果。GA 算法是基于生物进

化思想的全局寻优算法^[15]。能够在搜索进程中自动获取和积累相关搜索空间知识,从而自适应的控制搜索过程以求得最优解。作为智能优化算法的代表,GA 算法在工程应用中取得了广泛的应用。

3 MFE—GASVM 故障诊断算法实现

综上所述,基于 MFE—GASVM 的机载燃油泵系统故障诊断方法为:

步骤 1:对采集到的机载燃油泵系统振动信号进行分析,选择合适比例的训练集及测试集,为后续故障诊断提供数据准备;

步骤 2:对机载燃油泵振动信号进行多尺度模糊熵分析,得到不同尺度因子下的 FE 熵值,将其作为实现机载燃油泵系统故障诊断所需的特征向量,构建合适的特征向量集;

步骤 3:根据步骤 1 中得到的特征向量,构建 SVM 诊断模型所需的训练模型,其训练集的输入输出形式为:

$$Train = [MFE_1, MFE_2, \dots, MFE_n, label] \quad (11)$$

式中, MFE_i 表示尺度为 i 下的模糊熵值;一般情况下,故障诊断所需的特征向量无需过多,较多的特征向量可能会造成特征数据的冗余及冲突,影响着故障诊断精度。本文参照文献 [16],选择前 5 个尺度下的模糊熵值作为特征向量,即 $n = 5$ 。表示 SVM 模型的标签值,不同的标签值对应不同的故障状态(如正常状态的标签值设为 1)。

步骤 4:针对步骤 3 中的 SVM 模型,采用 GA 算法进行核函数参数和惩罚参数的自适应选择,以建立最优的 SVM 训练模型,提高机载燃油泵系统的故障诊断精度。

步骤 5:完成对机载燃油泵系统测试集的故障诊断实验过程,并对故障诊断结果进行分析说明,以论证算法的有效性。

4 实验分析

在某型机载燃油泵故障状态监测系统中,通过搭建机载燃油泵实验平台,在电机外壳上安装振动传感器并获取相应的振动信号。燃油泵的转速为 5 600 r/min;传感器采样频率为 6 kHz;选取样本长度为 2 048;除正常状态(记为标签 1)外,测量得到其他 3 种故障状态的振动信号,分别为扩散管损伤状态(记为标签 2)、扩散管损伤与叶轮损伤并存状态(记为标签 3)、渗漏状态(记为标签 4)。每组状态共得到 30 组样本数据,选择前 15 组数据作为训练集(50%),后 15 组数据作为测试集(50%)进行故障诊断分析。

然后,对样本数据进行多尺度粗粒化分析,选择前 10 个尺度下的模糊熵值进行多尺度因子刻画,为便于画图显示,每组状态选择 2 组样本数据进行展示,所得结果如图 1 所示。

如图 1 所示,横坐标表示不同尺度因子,纵坐标表示在该尺度因子下对应的 FE 熵值。图 1 可表明:1)随着尺度因子的增加,FE 整体呈下降趋势,这是由于多尺度粗粒

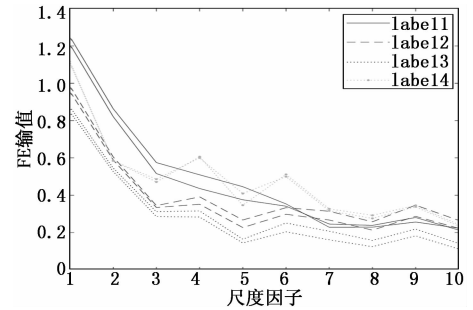


图 1 多尺度模糊熵

化过程降低了原始时间序列的复杂程度,因此对应的 FE 熵值降低。2) label1 状态下的多尺度模糊熵值在大部分尺度下均高于其他状态,这是因为正常状态下,传感器所接受到的轴承振动处于随机状态,所以在大部分尺度因子下振动的无规则程度较高,相应的自相似性较低,从而导致振动信号相对更加复杂。当出现故障状态时,振动信号的规则性增强,自相似性增加,导致模糊熵值降低。3) 尺度因子为 1 时,即为不进行多尺度粗粒化过程,此时所得到的 FE 熵值即为原始时间序列对应的的 FE 熵值。

图 1 中的 10 个尺度因子仅为展示 MFE 随尺度因子变化关系,在进行机载燃油泵故障诊断分析过程中,不需要过多的特征量,参照文献 [16],仅选择前 5 个尺度因子下的 FE 熵值作为故障特征量并进行故障诊断。为了对比验证,首先采用单一尺度下的 FE 熵值作为 GASVM 特征量进行故障诊断,结果如图 2 所示。

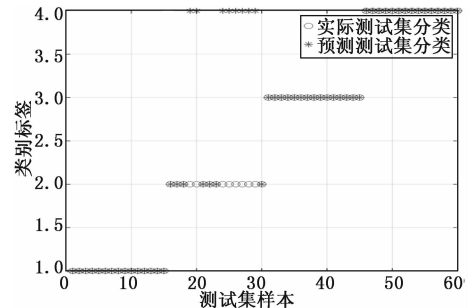


图 2 GASVM 故障诊断结果 (FE)

如图 2 所示,采用单尺度 FE 作为特征量时,故障诊断精度较低,仅为 86.6667% (52/60),此时所错分的 8 个样本将扩散管损伤状态(即标签 2)误判为渗漏状态(即标签 4),误判数目比例达到 53.33%,表明单尺度 FE 算法无法准确实现不同状态下的故障诊断。

采用本文的 MFE—GASVM 方法对相同的测试集进行故障诊断,特征量选择前 5 个尺度因子的 FE 熵值,所得结果如图 3 所示。

如图 3 所示,采用前 5 个尺度下的 FE 熵值作为特征量时,故障诊断精度为 100% (60/60),在图 2 中被误判的 8 个标签 2 样本均得到了准确的分类,实现了所有不同状态的故障识别。