

# 基于 MMFD—FCM 的退化状态识别方法及其应用研究

李海涛<sup>1</sup>, 王冰<sup>2</sup>

(1. 商丘师范学院 计算机与信息技术学院, 河南 商丘 476000;

2. 军械工程学院, 石家庄 050003)

**摘要:** 为了有效地对机械设备运行状态进行监测, 进而对其性能退化状态进行识别, 提出一种基于形态多重分形维数 (MMFD) 与模糊 C 均值聚类 (FCM) 的性能退化状态识别方法; 该方法首先计算机械设备振动信号的形态多重分形维数, 以此作为性能退化特征指标; 该特征指标能够有效反映峰值在振动信号中概率分布的不均匀程度, 从而定量描述振动信号的性能退化状态, 并且与多重形态分形维数相比, 利用数学形态学计算的 MMFD 精度更高, 计算速度更快; 在此基础上, 鉴于不同退化状态之间的模糊性, 针对性地采用模糊 C 均值聚类方法对特征指标进行模糊聚类, 从而有效识别性能退化状态; 将该方法应用于滚动轴承全寿命周期振动信号中, 分析结果验证了该方法的有效性。

**关键词:** 机械设备; 数学形态学; 多重分形; 模糊聚类; 退化状态识别; 滚动轴承

## Degraded Status Recognition Method Based on MMFD—FCM and Its Application Research

Li Haitao<sup>1</sup>, Wangbing<sup>2</sup>

(1. Shangqiu Normal University, College of Computer and Information Technology, Shangqiu 476000, China;

2. Ordnance Engineering College, Shijiazhuang 050003, China)

**Abstract:** In order to monitor mechanical equipment operational condition, and then indentify the performance degradation, a performance degradation identification method based on MMFD and FCM is proposed in this paper. For this method, MMFD of mechanical equipment vibrating signal is calculated and chosen as performance degradation characteristic index first. Multi-fractal dimension of vibration signal can reflect the peaks' non uniform probability distribution in the entire signal, and it can quantitative characterize the vibration intensity in the process of performance degradation. Compared with multi-fractal dimension, MMFD calculated based on mathematical morphology has a better performance in precision and computation speed. On this basis, in consideration of the fuzziness among different performance degradation, FCM is introduced into fuzzy clustering for characteristic index, and performance degradation could be recognized effectively. The rolling bearing whole life data verifies the validity of these methods above.

**Keywords:** mechanical equipment; mathematics morphology; multi-fractal dimension; fuzzy clustering; degraded status recognition; rolling bearing

## 0 引言

滚动轴承是机械设备中的典型部件, 其运行状态直接影响着系统是否能够正常工作。据统计, 在齿轮箱中, 19% 的故障由滚动轴承故障引起<sup>[1]</sup>, 在液压泵与电机中, 这一比例更是高达 51%<sup>[2]</sup>, 如何对滚动轴承等机械设备进行状态监测和性能退化状态识别, 进而预防可能产生的故障, 成为当前机械故障诊断与预测领域研究的热点问题。

本文提出一种基于 MMFD—FCM 的退化状态识别方法,

该方法首先计算机械设备振动信号的形态多重分形维数 (MMFD), 以此作为性能退化特征指标, 在此基础上, 结合模糊 C 均值方法对性能退化状态进行模糊聚类, 从而有效识别机械设备的性能退化状态。以机械设备中典型的滚动轴承为分析对象, 分析其全寿命周期振动信号数据, 从而对该方法进行有效性验证。

## 1 基于 MMFD 的退化特征提取

### 1.1 分形维数

分形维数是对一个分形对象的复杂度、不规则度的定量描述。分形维数可分为单重分形维数和多重分形维数两类。单重分形维数的大小不随观测尺度的变化而变化, 其取值与观测尺度大小无关。但实际信号往往在不同的观测尺度上具有非均匀、各向异性的特点, 故引入多重分形维数的概念。

收稿日期: 2014-04-29; 修回日期: 2014-06-09。

基金项目: 河南省基础与前沿技术研究计划项目 (132300410385)。

作者简介: 李海涛 (1978-), 男, 河南长垣人, 讲师, 硕士研究生, 主要从事嵌入式系统, 图像视频处理, 计算机网格应用方向的研究。

多重分形维数随着观测尺度的变化发生相应的变化, 其值与观测尺度的选取相关, 包含了分形理论涉及的大部分分形维数, 能够更全面地反映信号的分形特征<sup>[3]</sup>。相比单重分形维数, 多重分形维数可以反映信号局部的奇异性分布情况, 对信号的分形特征有更完整的反映。一般情况下, 多重分形维数的计算方法如下:

假设  $F$  是分形空间上的分形集, 将其分割为  $N$  个尺度为  $\epsilon$  的超立方体, 若尺度  $\epsilon$  足够小, 则可认为该区域是均匀的。设每个尺度为  $\epsilon$  的超立方体的测度为  $u_i(\epsilon)(i = 1, 2, \dots, N)$ , 则可定义奇异性指数  $\alpha$  如下:

$$u_i(\epsilon) = \epsilon^{\alpha_i} \quad (1)$$

$\alpha_i$  表示测度在  $i$  单元的奇异性指数。

为计算多重分形维数, 给定参数  $q$ , 定义配分函数为:

$$\varphi_q = \sum_{i=1}^N u_i(\epsilon)^q \quad (2)$$

则多重分形维数被定义为:

$$D_q = \frac{1}{q-1} \lim_{\epsilon \rightarrow 0} \frac{\ln \varphi_q(\epsilon)}{\ln \epsilon} \quad (3)$$

当  $q$  分别取 0, 1, 2 时, 多重分形维数  $D_q$  分别为盒维数、信息维数及关联维数<sup>[4]</sup>。

### 1.2 基于 MMFD 的滚动轴承性能退化特征提取

传统的多重分形维数多采用盒计数法, 该方法以规则的网格对不规则的信号进行覆盖, 计算过程中会产生一定的误差<sup>[4]</sup>。基于数学形态学的多重分形维数计算方法以数学形态学理论为基础<sup>[5]</sup>, 以数学形态覆盖代替盒计数法中的网格覆盖, 可以有效提高多重分形维数的计算精度和速度, 本文即采用该方法计算振动信号的多重分形维数, 计算过程如下:

在不同尺度  $\epsilon$  下对信号  $f$  的膨胀和腐蚀分别为  $f \oplus \epsilon g(n)$  和  $f \ominus \epsilon g(n)$ , 定义一个反应局部度量的分布函数  $u_i(\epsilon)$ :

$$u_i(\epsilon) = \frac{f \oplus \epsilon g(n) - f \ominus \epsilon g(n)}{\sum_{n=1}^N [f \oplus \epsilon g(n) - f \ominus \epsilon g(n)]} \quad (4)$$

其中:  $f \oplus \epsilon g(n) - f \ominus \epsilon g(n)$  表示信号在尺度  $\epsilon$  下的覆盖面积, 定义信号在尺度  $\epsilon$  下的  $q$  阶度量如下:

$$K_q(\epsilon) = \alpha \cdot \frac{\ln \sum_{i=1}^N [u_i(\epsilon)]^q}{q-1} \quad (5)$$

式中,  $K_q(\epsilon)$  根据广义信息熵推广构造得到, 系数  $\alpha = \frac{\ln(\frac{A_q(\epsilon)}{\epsilon^2})}{\ln(N(\epsilon))}$ 。

$K_q(\epsilon)$  与尺度  $\epsilon$  之间需满足指数关系如下:

$$K_q(\epsilon) \propto \epsilon^{-D_q} \quad (6)$$

则形态多重分形维数可以用下式进行计算:

$$D_q = \lim_{\epsilon \rightarrow 0} \frac{\ln K_q(\epsilon)}{\ln \epsilon} \quad (7)$$

根据上式, 在计算时, 只需对  $\ln(K_q(\epsilon))$  和  $\ln \epsilon$  进行最小二乘线性拟合即可得到形态多重分形维数 (Morphological Multi-fractal Dimension, MMFD)。

形态多重分形维数能够反映振动信号的峰值在整个信号中

概率分布的不均匀程度, 从而定量表征振动信号的形态特征和振动的剧烈程度。基于此, 提出基于 MMFD 的退化特征提取方法, 以 MMFD 作为机械设备性能退化过程的特征指标, 定量描述其性能退化状态。在实际计算过程中, 本文选取其中较敏感的部分。

## 2 基于 MMFD-FCM 退化状态识别

### 2.1 模糊 C 均值聚类

模糊 C 均值聚类算法 (FCM) 是一种非监督动态聚类方法, 其主要思想是将经典划分的定义模糊化, 用隶属度来确定样本的聚类程度<sup>[6]</sup>。在 FCM 中,  $C$  是指将有限个样本集  $X = \{x_1, x_2, \dots, x_n\}$  划分为  $C$  类, 每个样本以一定的隶属度属于  $C$  个不同域。一般用  $\mu_{ij}$  表示第  $j$  ( $j = 1, 2, \dots, n$ ) 个样本属于第  $i$  ( $i = 1, 2, \dots, n$ ) 类的隶属度, 且  $\mu_{ij}$  的值在区间  $[0, 1]$  内, 每个样本的隶属度之和为 1。

模糊 C 均值聚类算法以类内加权平方误差和为目标函数, 其表达式如下:

$$J(U, v_1, \dots, v_c) = \sum_{i=1}^c J_i = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m d_{ik}^2 \quad (8)$$

式中,  $m$  为加权指数,  $m \in (1, +\infty)$ ;  $d_{ij}(x_j, z_i) = \|x_j - z_i\|$  为样本点  $x_j$  与第  $i$  类样本聚类中心  $Z_i$  的欧式距离, 聚类中心  $Z = (z_1, z_2, \dots, z_c)$ 。

FCM 算法的实质就是找出一组中心矢量, 使每个样本到这个矢量的加权距离平方误差和达到最小。利用拉格朗日数乘法对式 (8) 进行求解, 得到下式:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}}\right)^{\frac{2}{m-1}}}, j = 1, 2, \dots, n \quad (9)$$

$$z_i = \frac{\sum_{j=1}^n (\mu_{ij})^m x_j}{\sum_{j=1}^n (\mu_{ij})^m}, i = 1, 2, \dots, C \quad (10)$$

当数据集  $X$ 、划分类别数  $C$  以及加权指数  $m$  已知时, 就能根据式 (9) 和 (10) 确定最佳的模糊分类矩阵及聚类中心。

### 2.2 基于 MMFD-FCM 的退化状态识别流程

基于前文所述, 提出一种基于 MMFD-FCM 的退化状态识别方法, 具体流程如图 1 所示。首先, 通过试验采集机械设备全寿命周期振动信号, 确定并选择全寿命数据训练样本集与待检测样本集, 分别计算每个样本的特征向量, 然后基于 FCM 算法对训练样本进行模糊聚类, 得到不同性能退化状态的聚类中心, 最后根据隶属度矩阵对待检测样本进行模糊分类, 根据隶属度最大原则, 识别待检测样本的退化状态。

## 3 实例分析

为验证本文提出的方法的有效性, 以机械设备中典型的滚动轴承为例进行实例分析。试验数据采用来自 Cincinnati 大学<sup>[7]</sup>轴承疲劳寿命强化试验台上采集的全寿命数据。轴承型号为 Rexnord ZA-115, 转速保持 2 000 rpm, 径向载荷为 26.7 kN, 采样间隔为 10 min, 采样频率为 25.6 kHz, 共采集数据 960 组, 最终失效形式为保持架断裂。采用隔二取一的方

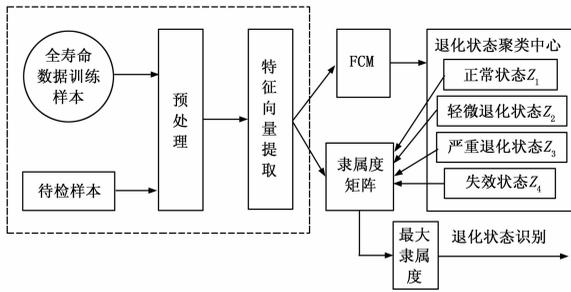


图 1 滚动轴承性能退化识别流程

法，将 960 组数据分为两个 480 组的数据集，分别作为训练样本集和测试样本集。

### 3.1 基于 MMFD 的滚动轴承性能退化特征提取

以前文所述的 MMFD 的计算方法为基础，分别计算训练样本每组数据对应的 MMFD。其中结构元素选择为扁平结构元素  $g = [0 \ 0 \ 0]$ ，尺度范围为 1~64， $q$  取区间  $[-20, 20]$  内的整数。

以第 10 组试验数据为例，其 MMFD 取值与  $q$  的变化关系如图 2 所示。可以看出，随着  $q$  值的不断增大，MMFD 的取值逐渐减小。

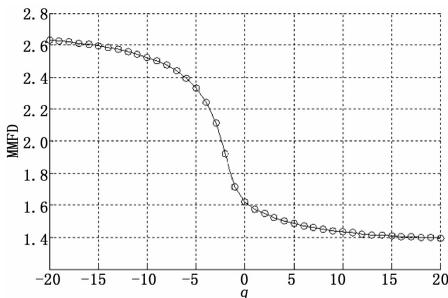


图 2 MMFD 随  $q$  变化趋势图

当  $q$  值一定时，随着组别的升高，MMFD 在不同的  $q$  取值区间，呈现出完全相反的变化趋势。当  $q$  在区间  $[-20, -1]$  取值时，组别越大，性能退化程度越高，MMFD 取值越大，整体上呈现出递减的趋势；当  $q$  在区间  $[0, 20]$  取值时，组别越大，性能退化程度越高，MMFD 取值越小。图 3 显示了  $q$  值分别取 -10 和 10 时的 MMFD-Time 变化曲线。可以看出，图 3 (a) 整体上呈现递增趋势，而图 3 (b) 整体上呈现递减的趋势，相对而言，图 3 (b) 反映的性能退化过程更为精细。在 0~7 000 min 的时间区间内，MMFD 值均保持相对稳定，此时滚动轴承处于正常工作状态；MMFD 在 7 000 min 附近出现小幅下降，之后保持相对稳定直到第 8 000 min。此时滚动轴承处于缓慢退化状态；在 8 000~9 500 min 的时间区间内，MMFD 取值快速的下降，此时，滚动轴承性能快速退化，最后至 9 500 min 滚动轴承完全失效。以上过程反映了滚动轴承正常运行、性能缓慢退化、性能快速退化及完全失效的 4 个阶段，这与文献 [8] 中的描述是一致的。

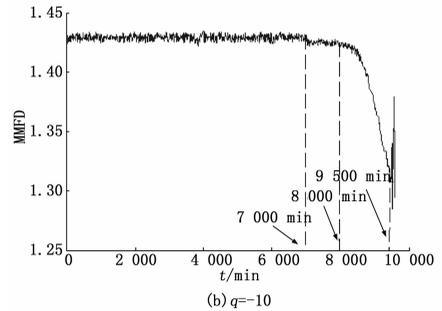
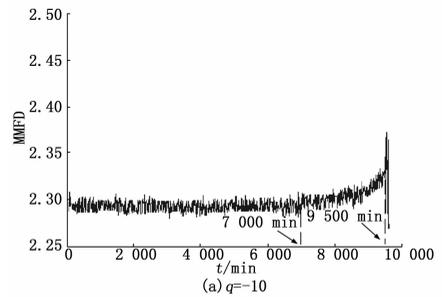


图 3 MMFD 随滚动轴承性能退化的变化趋势图

通过以上分析，为了涵盖不同  $q$  值下的 MMFD 状态信息，选取描述效果较好的  $[0, 20]$  区间内  $q = 5, 10, 15$  时的 MMFD 作为三维特征向量，以此描述滚动轴承性能退化状态信息。在滚动轴承性能退化状态划分中，沿用文献 [8] 的传统的划分方法，将滚动轴承性能退化状态分为正常状态、轻微退化状态、严重退化状态和失效状态。

### 3.2 基于 MMFD-FCM 的退化状态识别

在分别计算每组滚动轴承全寿命数据的三维退化特征向量的基础上，利用 FCM 方法对训练样本进行模糊聚类，四种退化状态的聚类中心如表 1 所示。

表 1 4 种退化状态的聚类中心

| 聚类中心   | MMFD(5) | MMFD(10) | MMFD(15) |
|--------|---------|----------|----------|
| 正常状态   | 1.479 7 | 1.429 4  | 1.390 9  |
| 轻微退化状态 | 1.472 1 | 1.423 1  | 1.370 2  |
| 严重退化状态 | 1.436 0 | 1.387 6  | 1.343 0  |
| 失效状态   | 1.393 6 | 1.330 0  | 1.311 8  |

从另外 480 组测试样本中随机抽取 5 组样本进行结果分析，利用 FCM 方法求取 5 组样本与 4 个聚类中心的隶属度，结果如表 2 所示，根据最大隶属度原则，隶属度最大的为识别结果。

将表 2 中的识别结果与实际情况进行对比，结果显示 5 组样本的性能退化识别结果全部正确，从而验证了本文提出的方法在滚动轴承性能退化状态识别中的有效性。

## 4 结论

本文提出了一种基于 MMFD-FCM 的性能退化状态识别方法，详细论述了该方法的具体建模过程，将该方法应用于滚动轴承全寿命周期数据中，结果表明，该方法计算得到的形态多重分形维数能较准确地反映滚动轴承的性能退化程度，基于 MMFD-FCM 性能退化状态识别方法能够有效地识别出滚动轴承的性能退化状态。该方法的提出对实现基于状态的维修具

表 2 待识别样本模糊识别结果

| 样本序号 | 正常状态    | 轻微退化状态  | 严重退化状态  | 失效状态    | 识别状态   |
|------|---------|---------|---------|---------|--------|
| 1    | 0.956 5 | 0.040 9 | 0.002 0 | 0.000 6 | 正常状态   |
| 2    | 0.081 2 | 0.879 0 | 0.027 6 | 0.012 2 | 轻微退化状态 |
| 3    | 0.059 6 | 0.935 1 | 0.003 2 | 0.002 1 | 轻微退化状态 |
| 4    | 0.043 2 | 0.067 5 | 0.735 1 | 0.154 2 | 严重退化状态 |
| 5    | 0.016 8 | 0.023 1 | 0.086 4 | 0.873 7 | 失效状态   |

有一定的促进意义。

参考文献:

[1] 钟秉林, 黄仁. 机械故障诊断学 [M]. 北京: 机械工业出版社, 2007.  
 [2] Cheung M, Li W. Probabilistic Fatigue and Fracture Analysis of Steel Bridges [J]. Journal of Structural Safety, 2003, 23: 245-262.  
 [3] 李兵, 张培林, 任国全, 等. 形态学广义分形维数在发动机故障诊断中的应用 [J]. 振动与冲击, 2011, 30 (10): 209-211.  
 [4] Li B, Zhang P L, Wang Z J, etc. Morphological Covering Based Generalized Dimension for Gear Fault Diagnosis [J]. Nonlinear Dynamics, 2012, 67 (4): 2561-2571.

[5] Sharif M, Bilal M, Khan S, etc. Adaptive Filter and Morphological Operators Using Binary PSO [J]. Lecture Note in Computer Science, 2010, 6377 (1): 525-532.  
 [6] 刘德馨, 李晓理, 周翔, 等. 模糊 C 均值聚类算法在高炉料面分类中的应用 [J]. 北京科技大学学报, 2012, 34 (6): 683-690.  
 [7] Lee J, Qiu H, Yu G. Rexnord Technical Services. "Bearing Data Set" [EB/OL]. IMS, University of Cincinnati, NASA Ames Prognostics Data Repository. <http://ti.arc.nasa.gov/project/prognostic-data-repository>, NASA Ames, Moffett Field, CA, (accessed 09.05.09).  
 [8] 潘玉娜, 陈进, 李兴林. 基于模糊 C-均值的设备性能退化评估方法 [J]. 上海交通大学学报, 2009, 43 (11): 1795-1797.

(上接第 2881 页)

化的情况, 从图 5 可以看出, 文中方法在第 15 个情节处就已经趋于收敛, 最终达到目标所需要的时间步为 99, 而经典的 Sarsa ( $\lambda$ ) 算法和 KNN-Sarsa ( $\lambda$ ) 方法则分别等到情节数为 40 和 24 时才收敛, 且所需要的时间步分别为 103 和 101。

在图 6 中, 虽然障碍物增加且布局更为复杂, 但文中方法仍然能在第 15 个情节处趋于收敛, 收敛对应的时间步为 98, 经典的 Sarsa ( $\lambda$ ) 算法收敛于第 19 个情节, 对应的时间步为 98, KNN-Sarsa ( $\lambda$ ) 方法则收敛于第 25 个情节, 收敛对应的时间步为 98。

综合上述, 文中方法不仅收敛效果好, 同时具有很强的稳定性, 在环境变化, 其收敛的效果不受影响, 因此, 较经典的 Sarsa ( $\lambda$ ) 算法和 KNN-Sarsa ( $\lambda$ ) 方法更优。

5 总结

为了实现爬壁机器人的自适应在线规划, 提出了一种基于模糊 K-Means 算法和 Sarsa ( $\lambda$ ) 的机器人在线规划算法, 首先对爬壁机器人的动力学模型进行了建模和描述, 然后, 设计了改变的模糊 K 均值聚类算法的实现状态的自适应在线聚类, 将聚类中心对应的值函数作为整个聚类所有数据对象的价值函数近似值, 最后, 对基于模糊 K 均值聚类算法和 Sarsa ( $\lambda$ ) 的爬壁机器人在线规划算法进行了

详尽地定义和描述。实验表明文中方法不仅具有收敛效果好, 规划路径短的有点, 而且在环境变化的情况下, 具有很好的稳定性。

参考文献:

[1] 衣正尧, 弓永军, 王祖温, 等. 新型船舶壁面除锈爬壁机器人动力学建模与分析 [J]. 机械工程学报, 2010, 46 (15): 24-29.  
 [2] Viet, Hoang Huu. Simulation-based evaluations of reinforcement learning algorithms for autonomous mobile robot path planning [J]. Springer Science business media B. V. 2012: 467-476.  
 [3] 李志海. 轮足混合驱动爬壁机器人及其关键技术的研究 [D]. 哈尔滨: 哈尔滨工业大学, 2010.  
 [4] Wang M, Chen X Q, Nayyerloo M, et al. A novel wall climbing robot based on Bernoulli effect [J]. Mechatronic and Embedded systems and applications, 2008, 210-215.  
 [5] 李志海, 付宜利. 基于遗传算法的仿生双足爬壁机器人越障运动规划 [J]. 机器人, 2012, 6 (34): 751-757.  
 [6] 禹建丽, 张晓梅, 程思雅, 等. 爬壁机器人路径规划研究 [J]. 系统仿真学报, 2009, 15 (21): 4748-4751.  
 [7] 岳荣刚, 王少萍. 基于混合整数线性规划的爬壁机器人路径规划 [J]. 北京航空航天大学学报, 2013, 6 (39): 792-797.  
 [8] 王金秋, 孙晓松, 秦华. 基于强化学习的爬壁机器人路径规划方法 [J]. 计算机测量与控制, 2013, 21 (11): 3093-3095.