

多尺度熵在转子故障诊断中的应用*

郑近德¹, 程军圣¹, 胡思宇²

(1. 湖南大学汽车车身先进设计制造国家重点实验室 长沙, 410082) (2. 湖南大学电气与信息工程学院 长沙, 410082)

摘要 为解决从信号中提取故障特征难的问题, 介绍了一种新的信号故障特征提取方法——多尺度熵(multi-scale entropy, 简称MSE), 并将样本熵和多尺度熵分别应用于转子故障信号复杂性的度量, 以提取故障特征。试验数据分析表明, 与样本熵相比, 多尺度熵更能有效地实现转子故障类型的诊断。

关键词 样本熵; 多尺度熵; 转子; 故障诊断

中图分类号 TN911.7; TH165.3

引言

故障诊断的关键是如何从振动信号中提取故障特征信息。近年来, 随着非线性科学理论的发展, 很多非线性分析的方法, 如小波分析、分形维数、神经网络等已被广泛应用于机械设备故障诊断领域, 丰富了故障诊断的技术和手段^[1-2]。

Pincus^[3-4]提出了一种度量序列的复杂性和统计量化的方法——近似熵(approximate entropy, 简称ApEn), 并成功应用于生理时间序列分析。近似熵算法比较的是数据和其自身, 即包含自匹配, 由于熵是新信息产生率的测度, 所以比较数据和其自身毫无意义。Richman等^[5]提出了经过改进的复杂度测试方法——样本熵(sample entropy, 简称SampEn)。与Lyapunov指数、信息熵、关联维数和K熵等非线性动力学方法相比, 样本熵具有得到稳定估计值所需的数据短、抗噪声和干扰能力强、在参数大取值范围内一致性好等特点^[6]。但是, 样本熵是衡量时间序列在单一尺度上的复杂性, Costa等^[7-8]在样本熵的基础上, 提出了另一种时间序列复杂度的衡量方法——多尺度熵(multiscale entropy, 简称MSE), 用以衡量时间序列在不同尺度上的复杂性, 极大地丰富了熵的含义。

多尺度熵和近似熵、样本熵一样, 都是用来衡量时间序列的复杂性和维数变化时产生新模式的概率大小的方法。产生新模式的概率越大, 序列的复杂性越大, 其熵值越大。对机械振动信号而言, 不同的故

障类型信号的复杂性不同, 产生新模式的概率不同, 因而其熵值也不同。某些故障会在一定的特定频段, 当发生故障时该频段内的信号会发生较大的变化, 其复杂性也会发生变化^[9]。多尺度熵是衡量时间序列在不同尺度因子下的复杂性程度, 其值可以用来作为判断的指标和特征参数, 表征不同故障类型信号的复杂性。胥永刚等^[1]探讨了近似熵在机械设备状态监测和故障诊断领域中的工程应用, 并与分形维数进行了比较, 指出它们在表征振动信号复杂性方面各具特点, 但近似熵包含的信息更多。苏文胜等^[10]将样本熵和小波包相结合应用于滚动轴承故障诊断特征的提取, 取得了很好的效果。基于此, 笔者介绍了样本熵和多尺度熵的概念, 将多尺度熵引入到故障诊断领域, 并将其应用于转子系统信号的故障特征的提取, 通过分析转子径向位移信号的特点, 实现故障类型的诊断。

1 样本熵与多尺度熵的概念

1.1 样本熵的计算

Richman提出的样本熵是一种与近似熵^[3]类似但精度更好的复杂性度量方法, 其计算步骤^[5-6]如下。

1) 设原始数据为 $\mathbf{X}_i = \{x_1, x_2, \dots, x_N\}$, 长度为 N , 预先给定嵌入维数 m 和相似容限 r , 考虑 m 维向量 $\mathbf{x}(i) = [x_i, x_{i+1}, \dots, x_{i+m-1}]$ ($i = 1, 2, \dots, N - m$)

(1)

* 国家自然科学基金资助项目(51075131); 中央高校基本科研业务费专项基金资助项目; 教育部长江学者与创新团队发展计划资助项目(531105050037)

收稿日期: 2011-08-02; 修改稿收到日期: 2011-09-14

2) 定义 $\mathbf{x}(i)$ 与 $\mathbf{x}(j)$ 间的距离 $d[\mathbf{x}(i), \mathbf{x}(j)]$ 为两者对应元素差值的最大值,即

$$d[\mathbf{x}(i), \mathbf{x}(j)] = \max_{k=0,1,\dots,m-1} [|x(i+k) - x(j+k)|] \quad (2)$$

3) 对每个 i 值,计算 $\mathbf{x}(i)$ 与其余矢量 $\mathbf{x}(j)$ ($j=1, 2, \dots, N-m, j \neq i$)间的距离 $d[\mathbf{x}(i), \mathbf{x}(j)]$ 。统计 $d[\mathbf{x}(i), \mathbf{x}(j)]$ 小于 r 的数目及此数目与距离总数 $N-m-1$ 的比值,记作 $B_i^m(r)$,即

$$B_i^m(r) = \frac{1}{N-m-1} \{d[\mathbf{x}(i), \mathbf{x}(j)] < r \text{ 的个数} \} \quad (i=1, 2, \dots, N-m, i \neq j) \quad (3)$$

4) 求 $B_i^m(r)$ 的平均值

$$B^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(r) \quad (4)$$

5) 再对维数 $m+1$,即对 $m+1$ 点矢量,重复步骤1~4,得到 $B_i^{m+1}(r)$,进而得到 $B^{m+1}(r)$ 。

6) 理论上此序列的样本熵为

$$\text{SampEn}(m, r) = \lim_{N \rightarrow \infty} \left[-\ln \frac{B^{m+1}(r)}{B^m(r)} \right] \quad (5)$$

当 N 为有限数时,式(5)表示成

$$\text{SampEn}(m, r, N) = \left[-\ln \frac{B^{m+1}(r)}{B^m(r)} \right] = \ln B^m(r) - \ln B^{m+1}(r) \quad (6)$$

SampEn 的值显然与 m, r 的取值有关。不同的嵌入维数 m 和相似容限 r 对应的样本熵也不同。 m, r 的具体取值还没有一个最佳标准,一般取 $m=2, r$ 通常取 $0.1\text{SD} \sim 0.5\text{SD}$ (SD为原始数据的标准差)。本研究取 $m=2, r=0.2\text{SD}$ 。

1.2 多尺度熵

多尺度熵的计算步骤^[7]如下。

1) 设原始数据为 $\mathbf{X}_i = \{x_1, x_2, \dots, x_N\}$,长度为 N ,预先给定嵌入维数 m 和相似容限 r ,依据原始序列建立新的粗粒向量为

$$\mathbf{y}_j(\tau) = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i \quad (1 \leq j \leq N/\tau) \quad (7)$$

其中: $\tau=1, 2, \dots$ 为尺度因子。

一般取 $\tau_{\max} > 10$ 即可,本研究取 $\tau_{\max} = 20$ 。 $\tau=1$ 时, $\mathbf{y}_j(1)$ 就是原序列。对于非零 τ 原始序列, \mathbf{X}_i 被分割成 τ 个每段长为 N/τ 的粗粒序列 $\mathbf{y}_j(\tau)$ 。

2) 对得到的 τ 个粗粒序列分别求其样本熵,并把它画成尺度因子 τ 的函数,称之为多尺度熵分析。

样本熵确定的是时间序列在单一尺度上的复杂度和无规则程度,熵值越小,序列的自相似性越高。熵值越大,序列越复杂。多尺度熵定义为时间序列在不同尺度下的样本熵,多尺度熵曲线反映的是时间

序列在不同尺度下的复杂性程度。如果一个序列的熵值在大部分尺度上都比另一个序列的熵值高,那么就认为前者比后者复杂性更高。如果一个时间序列随着尺度因子递增而熵值单调递减,表明序列结构相对简单,在最小的尺度上包含较多信息。如果一个时间序列随着尺度因子递增熵值也单调递增,表明序列在多个尺度上包含较多信息^[9]。关于多尺度熵的性质详见文献^[7-8]。

2 基于多尺度熵的故障诊断及实验分析

故障诊断的关键是从振动信号中提取故障的特征信息。振动信号一般是非线性和非平稳信号,含有干扰信号和噪声,但样本熵和多尺度熵的计算有很强的抗干扰和抗噪能力;因此,笔者考虑直接求取原始信号的多尺度熵,对其进行MSE分析,进而实现故障类别的诊断。

转子系统常见的故障类型包括不平衡、不对中、碰摩和油膜涡动。笔者从转子振动模拟试验台上分别测得转子不平衡、正常、不对中、碰摩和油膜涡动5种状态下的径向位移振动信号,其采样频率为2 048 Hz,采样点数为1 024,转速为3 kr/min。5种状态的时域波形分别如图1~5所示。

根据图1~5中转子系统5种状态径向位移信号

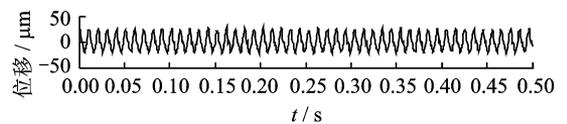


图1 转子不平衡状态的径向位移振动信号

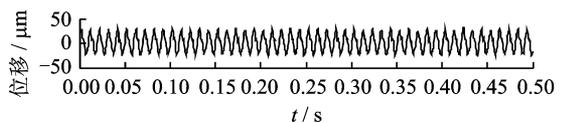


图2 转子正常状态的径向位移振动信号

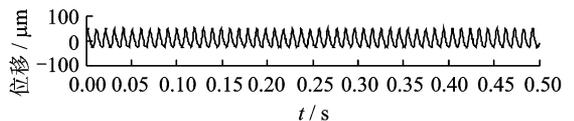


图3 转子不对中状态的径向位移振动信号

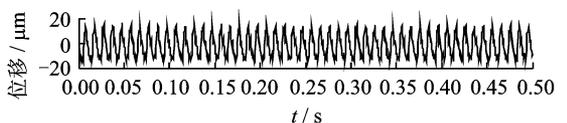


图4 转子碰摩状态的径向位移振动信号

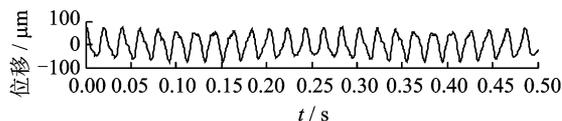


图5 转子油膜涡动状态的径向位移振动信号

的时域波形不易发现它们的区别,因此先考虑它们的样本熵值的关系。以上5种状态,每种状态分别取3组数据,计算其样本熵,并求其平均值,如表1所示。

表1 转子5种状态信号的样本熵值

样本	不平衡	正常	不对中	碰摩	油膜涡动
样本1	0.515 6	0.460 1	0.304 5	0.728 7	0.257 3
样本2	0.502 8	0.462 1	0.294 8	0.673 9	0.258 4
样本3	0.525 3	0.420 2	0.291 9	0.615 2	0.261 8
样本熵均值	0.514 6	0.447 5	0.297 1	0.672 6	0.259 2

设正常、不对中、不平衡、碰摩和油膜涡动故障信号的样本熵分别对应为 SE_1, SE_2, SE_3, SE_4 和 SE_5 。由表1容易看出,不同故障类型转子的径向位移信号的样本熵不同,同种故障类型的位移信号的样本熵在一个固定的值附近波动。故障状态与正常状态的转子位移信号,按熵值大小关系为

$$SE_4 > SE_3 > SE_1 > SE_2 > SE_5$$

碰摩和不平衡的样本熵值相对较大,这说明二者的振动信号的复杂性较高,维数变化时产生新模式的概率也越大。不对中和油膜涡动的样本熵值相对较小,说明二者的振动信号的复杂性较低,序列自相似性较高。由上述可知,样本熵可以实现故障类型的诊断,但各种故障类型的样本熵值相差不大,区分效果不明显。

由表1可知,原始信号在单一尺度上的样本熵值虽然能够区分转子的故障类型,但是区分效果不明显。下面考虑各种转子径向位移信号在不同尺度下的样本熵,即多尺度熵,将多尺度熵应用于转子故障类型的诊断。仍然考虑上述5种状态下的各3组样本,分别求其多尺度熵,并画成尺度因子的函数关系图,对其进行多尺度熵分析,如图6所示。

由图6可以得出如下结论。

1) 由于引入了尺度因子,与样本熵相比,多尺度熵能够更明显、更直观地区分转子的几种故障状态类型。不同的故障类型的转子径向位移信号在不同尺度下的熵值不同。

2) 多尺度熵值和样本熵所得的结果一致。正常、不对中、不平衡、碰摩和油膜涡动转子信号的多尺度熵分别为 $MSE_1, MSE_2, MSE_3, MSE_4$ 和 MSE_5 ,不同故障状态与正常状态的多尺度熵值的大小关系

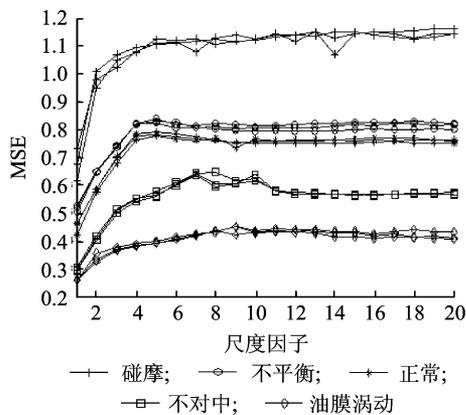


图6 转子5种状态的多尺度熵

是 $MSE_4 > MSE_3 > MSE_1 > MSE_2 > MSE_5$,区分非常明显。图6说明,碰摩状态的转子径向位移信号在大部分尺度上较正常状态的复杂,无规则程度较高,不平衡和正常状态的次之。不对中和油膜涡动状态径向位移信号在大部分尺度下的样本熵值较小,说明不对中和油膜涡动状态径向位移信号自相似性较高,信号较为规则。

3) 5种状态的多尺度熵曲线都是随着尺度因子的增加而渐变地趋向于某一值,这说明多尺度熵不仅反映了时间序列本身的复杂性程度,而且还包含了序列在其他尺度上的信息,这是分形维数和样本熵所无法反映的,说明了多尺度熵的优越性。

3 结束语

多尺度熵、关联维数、近似熵和样本熵等都是非线性动力学的方法,研究表明,多尺度熵比分形维数和样本熵等包含更多的信息。笔者将多尺度熵的概念引入到故障诊断领域,并将其作为诊断转子系统故障的特征参数,提出了一种基于多尺度熵的转子系统的故障诊断的新方法。试验表明,该方法能够有效地识别转子故障的类型,不足之处是多尺度熵的计算量较大,耗时多。

参 考 文 献

- [1] 胥永刚,何正嘉. 分形维数和近似熵用于度量信号复杂性的比较研究[J]. 振动与冲击, 2003, 22(3): 25-27.
Xu Yonggang, He Zhengjia. Research on comparison between approximate entropy and fractal dimension for complexity measure of signals[J]. Journal of Vibration and Shock, 2003, 22(3): 25-27. (in Chinese)

- [2] 徐玉秀,王志强,梅元颖. 叶片振动响应的长度分形故障特征提取与诊断[J]. 振动、测试与诊断, 2011, 31(2):190-192.
Xu Yuxiu, Wang Zhiqiang, Mei Yuanying. Fault feature extraction and diagnosis for wind turbine blades from vibration response based on theory of length fractal dimension[J]. Journal of Vibration, Measurement & Diagnosis, 2011, 31(2):190-192. (in Chinese)
- [3] Pincus S M. Approximate entropy as a measure of system complexity[J]. Proceedings of the National Academy of Sciences, 1991, 88:2297-2301.
- [4] 白冬梅. 脑电信号的特性分析与特征提取[D]. 大连: 大连理工大学, 2006.
- [5] Richman J S, Moorman J R. Physiological time-series analysis using approximate entropy and sample entropy[J]. American Journal of Physiology-Heart Circulatory Physiology, 2000, 278:2039-2049.
- [6] 来凌红,吴虎胜,吕建新,等. 基于EMD和样本熵的滚动轴承故障SVM识别[J]. 煤矿机械, 2011, 32(1): 249-252.
Lai Linghong, Wu Husheng, Lü Jianxin, et al. SVM recognition method based on EMD and sample entropy in roiling bearing fault diagnosis[J]. Coal Mine Machinery, 2011, 32(1):249-252. (in Chinese)
- [7] Costa M, Goldberger A L, Peng C K. Multiscale entropy analysis of biological signals[J]. Physical Review E, 2005, 71:1-18.
- [8] Costa M, Goldberger A L, Peng C K. Multiscale entropy analysis of complex physiologic time series[J]. Physical Review Letters, 2002, 89(6):1-18.
- [9] 齐子元,徐章遂,卢志才. 近似熵在发动机故障诊断中的应用研究[J]. 军械工程学院学报, 2008, 20(2): 39-42.
Qi Ziyuan, Xu Zhangsui, Lu Zhicai. Application of approximate entropy in fault diagnosis of engine[J]. Journal of Ordnance Engineering College, 2008, 20(2):39-42. (in Chinese)
- [10] 苏文胜,王奉涛,朱泓. 基于小波包样本熵的滚动轴承故障特征提取[J]. 振动、测试与诊断, 2011, 31(2): 162-166.
Su Wensheng, Wang Fengtao, Zhu Hong, et al. Rolling bearing fault feature extraction based on wavelet packet and sample entropy[J]. Journal of Vibration, Measurement & Diagnosis, 2011, 31(2): 162-166. (in Chinese)



第一作者简介:郑近德,男,1986年3月生,博士研究生。主要研究方向为机械故障诊断及动态信号处理。

E-mail:lqdlzheng@126.com

欢迎订阅《振动、测试与诊断》

《振动、测试与诊断》由工业和信息化部主管,南京航空航天大学 and 全国高校机械工程测试技术研究会联合主办,是反映振动、动态测试及故障诊断学科领域的科研成果及其应用情况的技术性刊物。主要刊登国内外以振动测试与故障诊断为中心的动态测试理论、方法和手段的研究及应用方面的技术文献,包括实验测试技术、测试仪器的研制、方法和系统组成、信号分析、数据处理、参数识别与故障诊断以及有关装置的设计、使用、控制、标定和校准等,不拘泥于行业和测试项目。

本刊入选EI Compendex 数据库收录期刊和中文核心期刊,为双月刊,每逢双月末出版,每本定价20元,全年120元。欢迎订阅和投稿,欢迎在本刊刊登各类广告和科技信息。

编辑部地址:南京市御道街29号 邮政编码:210016

电话:(025)84893332

传真:(025)84893332 E-mail:qchen@nuaa.edu.cn

网址:http://zdc.nuaa.edu.cn