

基于 LCD 和排列熵的滚动轴承故障诊断*

郑近德, 程军圣, 杨宇

(湖南大学汽车车身先进设计制造国家重点实验室 长沙, 410082)

摘要 排列熵(permutation entropy, 简称 PE)是最近提出的一种检测时间序列随机性和动力学突变行为的方法, 可以考虑将其应用于故障诊断。由于机械系统的复杂性, 振动信号的随机性和动力学突变行为表现在不同尺度上, 因此需要对振动信号进行多尺度的排列熵分析。基于此, 提出了基于局部特征尺度分解(local characteristic-scale decomposition, 简称 LCD)和排列熵的滚动轴承故障诊断方法。首先, 采用 LCD 方法对振动信号进行自适应分解, 得到不同尺度的本征尺度分量(intrinsic scale component, 简称 ISC); 其次, 计算前几个包含主要故障信息的 ISC 分量的排列熵; 最后, 将熵值作为特征向量, 输入基于神经网络集成建立的分类器。将该方法应用于滚动轴承实验数据, 分析结果表明, 此方法可有效实现滚动轴承的故障诊断。

关键词 局部特征尺度分解; 排列熵; 滚动轴承; 故障诊断; 神经网络集成

中图分类号 TN911.7; TH165.3

引言

由于机械设备振动信号大部分是非平稳、非线性信号, 因此机械设备故障诊断的关键是如何从非平稳、非线性信号中提取故障特征信息。近年来, 随着非线性科学的发展, 许多非线性分析方法被应用于机械设备的故障诊断。Logan 等^[1]研究了关联维数在滚动轴承中的应用^[1]。彭志科等^[2]将小波和多重分形应用于机械系统中振动信号几何特征的提取。侯荣涛等^[3]将混沌和 Lyapunov 指数等非线性理论应用于汽轮发动机的故障诊断。Yan 等^[4]将近似熵应用于机械系统健康状态监测^[4]。然而, 分形维数的计算依赖数据长度, 计算耗时, 不适合在线监测。Lyapunov 指数计算速度较慢, 精度不高, 易受噪声干扰。近似熵作为一种时间序列复杂性的方法, 其计算也依赖于数据长度, 相对一致性较差^[5]。

排列熵是由 Bandt 等^[6]提出的一种检测时间序列随机性和动力学突变行为的方法, 具有计算简单、快速, 抗噪能力强, 适合在线监测等优点。Yan 等^[7]将其应用于旋转机械振动信号的特征提取, 并将排列熵与近似熵和 Lempel-Ziv 复杂度进行了对比。结果表明, 排列熵能够有效检测和放大振动信号的

动态变化, 表征滚动轴承在不同状态下的工况特征。由于机械系统比较复杂, 振动信号不仅在单一尺度上包含有重要信息, 在其他尺度上也包含有重要信息, 因此有必要对振动信号进行多尺度分析, 而对振动信号进行多尺度分解是一种常用的手段。

文献[8]提出了一种新的信号分解方法——局部特征尺度分解, 该方法能自适应地将一个多分量信号分解为若干个内禀尺度分量之和。与经验模态分解^[9-10](empirical mode decomposition, 简称 EMD)方法相比, LCD 仅采用一次三次样条拟合, 计算所需时间缩短, 拟合误差减小。因此, 将 LCD 应用于滚动轴承振动信号的自适应多尺度分解。在对振动信号进行 LCD 分解并提取各个 ISC 分量的排列熵作为故障特征向量之后, 进一步地需要选择合适的模式识别方法来实现故障的自动分类。神经网络集成(neural networks ensemble, 简称 NNE)^[11]是用有限的几个神经网络对同一个问题进行学习, 集成在某输入示例下的输出由构成集合的各神经网络在这一示例下的输出共同决定。该方法通过训练多个简单的神经网络, 将结果集成作为模型输出, 显著提高了网络的泛化能力, 且该模型易于使用、具有较好的并行计算、自学习以及鲁棒性等优点^[12]。

笔者提出了基于 LCD-排列熵的滚动轴承故障

* 国家自然科学基金资助项目(51075131); 湖南省自然科学基金资助项目(11JJ2026); 中央高校基本科研业务费专项基金资助项目(531107040301)

收稿日期: 2012-09-27; 修回日期: 2012-12-14

神经网络集成诊断方法。首先,采用 LCD 方法将振动信号分解为不同尺度的 ISC 分量;然后,对包含主要故障特征信息的分量提取其排列熵;最后,将熵值作为特征向量输入神经网络集成分类器,从而实现滚动轴承故障类别的诊断。

与传统的特征提取方法相比,笔者提出的方法不是简单地基于信号的时域统计特征,而是提取振动信号不同尺度的随机性和动力学突变特征;因此相较于传统方法而言,本研究方法更能反映信号故障的本质特征。笔者将提出的方法应用于轴承实验数据,结果表明,该方法能够有效区分滚动轴承的故障类型,是一种有效的故障诊断方法。

1 LCD 方法概述

LCD 方法^[8]假设任意一个非平稳信号可被分解为有限个 ISC 分量之和,任何两个 ISC 分量之间相互独立,其中 ISC 需满足以下两个条件:

1) 整个数据段内,极大值为正,极小值为负,且任意两个相邻的极大值与极小值之间呈现单调性。

2) 整个数据段内,设其所有极值点为 (τ_k, X_k) ($k=1,2,\dots,M,M$ 为极值点个数)。由任意两个相邻的极大(或小)值点 (τ_k, X_k) 、 (τ_{k+2}, X_{k+2}) 确定的直线 $l_k(y = \frac{X_{k+2} - X_k}{\tau_{k+2} - \tau_k}(t - \tau_k) + X_k)$ 在二者之间的极值点 X_{k+1} 所对应的时刻 τ_{k+1} 处的函数值(记为 A_{k+1})与 X_{k+1} 的关系满足

$$\frac{A_{k+1}}{X_{k+1}} = -1 \quad \text{或} \quad A_{k+1} + X_{k+1} = 0 \quad (1)$$

其中

$$A_{k+1} = X_k + \frac{\tau_{k+1} - \tau_k}{\tau_{k+2} - \tau_k}(X_{k+2} - X_k) \quad (k=1,2,\dots,M-2) \quad (2)$$

对于实信号 $X_i(t \geq 0)$,LCD 方法的分解步骤^[8]如下。

1) 确定信号 X_i 的极值点 (τ_k, X_k) ($k=1,2,\dots,M$),并计算

$$L_k = \frac{X_k + A_k}{2} \quad (k=2,\dots,M-1) \quad (3)$$

其中: A_k 如式(2)所定义, $k=2,\dots,M-1$ 。

2) 由于 L_k 的值是从 2 到 $M-1$,需要估计端点 L_1, L_M 的值,因此需要对序列进行延拓。通过延拓,得到左右两端极值点 (τ_0, X_0) 、 (τ_{M+1}, X_{M+1}) 。令 k 分别等于 0 和 $M-1$,按式(3)求出 L_1 与 L_M 的值,再用三次样条函数拟合所有的 L_k ($k=1,2,\dots,M$),得到均值曲线 $BL_1(t)$ 。

3) 将均值曲线从原信号中分离出来,即

$$h_1(t) = X_i - BL_1(t) \quad (4)$$

若 $h_1(t)$ 是一个 ISC 分量,输出 $h_1(t)$ 并令 $ISC_1 = h_1(t)$ 。否则将 $h_1(t)$ 作为原始数据,继续重复上述步骤,直到 $h_{1k}(t)$ 是一个 ISC 分量,记 $ISC_1 = h_{1k}(t)$ 。

4) 将 ISC_1 分量从原始数据中分量出来

$$u_1(t) = X_i - ISC_1 \quad (5)$$

由此得到一个新的剩余信号 $u_1(t)$ 。

5) 再将 $u_1(t)$ 视为原始数据,重复上述步骤 1~4,重复循环 $n-1$ 次,直到剩余信号 $u_n(t)$ 为一单调或常函数,于是

$$X_i = \sum_{i=1}^{n-1} ISC_i + u_n(t) \quad (6)$$

LCD 方法中均值曲线的定义只采用一次三次样条拟合,与 EMD 方法中采用两次三次样条插值相比,LCD 的拟合误差相对较小且计算时间也比 EMD 少。关于 LCD 和 EMD 的详细比较参见文献[8]。

2 排列熵原理和算法

2.1 排列熵算法

考虑时间序列 $\{x(i), i=1,2,\dots,N\}$,长度为 N ,对其进行相空间重构,得到

$$\begin{cases} X(1) = \{x(1), x(1+\tau), \dots, x(1+(m-1)\tau)\} \\ \vdots \\ X(i) = \{x(i), x(i+\tau), \dots, x(i+(m-1)\tau)\} \\ \vdots \\ X(N-(m-1)\tau) = \{x(N-(m-1)\tau) \\ x(N-(m-2)\tau), \dots, x(N)\} \end{cases} \quad (7)$$

其中: m 为嵌入维数(embedding dimension); τ 为时延(time delay)。

将 $\mathbf{X}(i) = \{x(i), x(i+\tau), \dots, x(i+(m-1)\tau)\}$

按照升序重新排列,即

$$\mathbf{X}(i) = \{x(i+(j_1-1)\tau) \leq x(i+(j_2-1)\tau) \leq \dots \leq x(i+(j_m-1)\tau)\} \quad (8)$$

如果存在 $x(i+(j_{i1}-1)\tau) = x(i+(j_{i2}-1)\tau)$,此时按 j 的值的大小来进行排序,即当 $j_{k1} < j_{k2}$,有 $x(i+(j_{i1}-1)\tau) \leq x(i+(j_{i2}-1)\tau)$ 。所以,对任意一个向量 $\mathbf{X}(i)$ 都可得到一组符号序列

$$S(l) = [j_1, j_2, \dots, j_m] \quad (9)$$

其中: $l=1,2,\dots,k, k \leq m!, m$ 个不同的符号 $[j_1, j_2,$

..., j_m]共有 $m!$ 种不同的排列,对应地共有 $m!$ 种不同的符号序列。

计算每一种符号序列出现的概率: P_1, P_2, \dots, P_k , $\sum_{l=1}^k P_l = 1$, 则时间序列 $\{x(i), i=1, 2, \dots, N\}$ 的排列熵定义为

$$H_p(m) = - \sum_{l=1}^k P_l \ln P_l \quad (10)$$

当 $P_l = 1/m!$ 时, $H_p(m)$ 达到最大值 $\ln(m!)$, 可通过 $\ln(m!)$ 将 $H_p(m)$ 标准化, 即

$$H_p = H_p(m) / \ln(m!) \quad (11)$$

其中: H_p 的取值范围为 $0 \leq H_p \leq 1$ 。

H_p 值的大小表示时间序列的复杂和随机程度。 H_p 越大, 说明时间序列越随机; H_p 越小, 说明时间序列越规则。

2.2 PE 参数的选取

排列熵的计算与参数: 时间序列长度 N , 嵌入维数 m 和时延 τ 的值都有关系。一般地, 嵌入维数 m 取 3~7, 如果 m 等于 1 或 2, 此时重构的向量中包含的状态太少, 算法失去意义和有效性, 不能检测序列的动力学突变; 如果 m 过大, 相空间的重构将会均匀化时间序列, 此时不仅计算比较耗时, 而且也无法反映序列的细微变化^[4], 这里选取 $m=6$ 。时延 τ 对时间序列的计算影响较小^[4], $\tau=1$ 。为研究数据长度 N 对 PE 计算的影响, 分别以长度为 128, 256, 512, 1 024 的随机信号为例, 对应的 PE 分别记为 PE_1, PE_2, PE_3, PE_4 , 如图 1 所示。由图 1 发现, 当 $m=6, \tau=1$ 时, 数据长度为 1 024 与长度 512, 熵值相差 0.057 5, 而与长度 2 048 熵值相差 0.033 8, 因此, 此时选取数据长度大于等于 1 024 较为合适。一般的, 嵌入维数较小时数据长度则要求越小。这里选取 $m=6, \tau=1, N=2 048$ 。

3 方法与试验分析

笔者提出的基于 LCD, PE 和 NNE 的滚动轴承故障诊断方法如下。

首先, 采用 LCD 方法自适应地将滚动轴承振动信号分解为若干个不同尺度的 ISC 分量: $ISC_1, ISC_2, \dots, ISC_n$ 。每个分量都包含了原始信号不同频段和不同尺度的故障信息。

其次, 由于故障信息一般集中在振动信号相对较高的频段, 因选取前 5 个包含主要故障信息的 ISC 分量, 计算它们的排列熵, 作为个体神经网络的特征向量

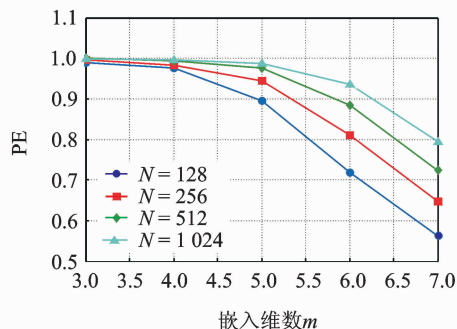


图 1 不同长度的随机噪声的排列熵

Fig. 1 The PE of white Gaussian noises with different lengths

$$T = [PE_1, PE_2, PE_3, PE_4, PE_5] \quad (12)$$

最后, 依据特征向量对多类故障滚动轴承建立神经网络集成分类器。

基于 LCD, PE 和 NNE 的滚动轴承故障诊断方法的流程如图 2 所示。

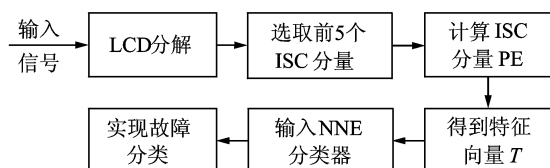


图 2 基于 LCD, PE 和 NNE 的滚动轴承故障诊断方法流程图

Fig. 2 Flow chart of rolling bearing fault diagnosis based on the LCD, PE and NNE

笔者将提出的方法用于滚动轴承试验数据以证明方法的有效性。分析数据采用美国 Case Western Reserve University 电气工程实验室的滚动轴承试验数据。测试轴承为 6205-2RS JEM SKF 深沟球轴承, 电机负载约为 735.5 W, 轴承转速为 1 772 r/min, 试验使用电火花加工技术在轴承上布置单点故障, 故障直径为 0.355 6 mm, 深度为 0.279 4 mm, 在此情况下采集到正常、内圈单点电蚀、外圈单点电蚀和滚动体单点电蚀, 4 种状态的振动信号采样频率为 12 kHz, 每种状态取 30 组数据, 数据样本长度为 2 048。正常 (normal, 简称 NORM)、滚动体故障 (rolling element fault, 简称 REF)、内圈故障 (inner race fault, 简称 IRF) 和外圈故障 (outer race fault, 简称 ORF) 4 种状态轴承的振动加速度信号如图 3 所示。

正常、外圈故障、内圈故障和滚动体故障四种状态轴承的振动信号, 每种状态取 30 组数据, 其中 20 组用来训练, 10 组用来测试。首先, 对 4 种状态的

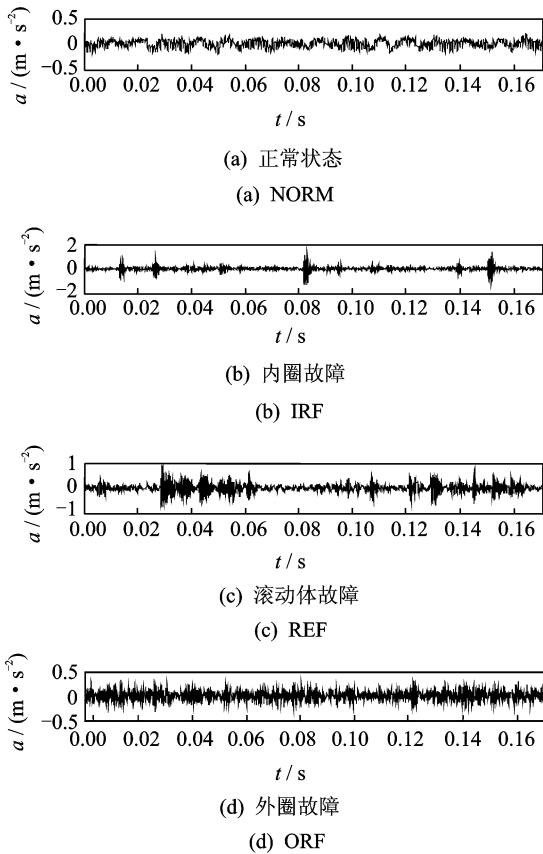


图 3 正常和具有故障轴承的振动信号的时域波形

Fig. 3 The time domain waveforms of the vibration signals from the normal and fault bearings

120 个样本数据进行 LCD 分解, 每个样本分解得到若干个 ISC 分量; 其次, 对包含主要故障信息的前 5 个 ISC 分量计算其排列熵, 并将熵值组成特征向量, 共得到 120 个特征向量。图 4 为正常和 3 种故障轴承振动信号特征向量的关系。

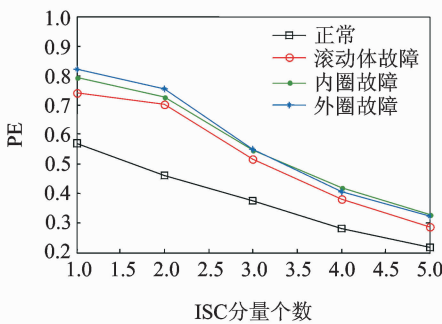


图 4 正常和具有故障轴承的振动信号的特征向量

Fig. 4 Characteristic vectors of vibration signals from the normal and fault bearings

从图 4 可以发现, 正常状态滚动轴承的振动信号各分量的排列熵较小, 当滚动轴承发生故障时, 信

号各分量的排列熵值发生了较大变化。这是因为, 当滚动轴承发生故障时, 信号的随机性发生改变, 振动信号的动力学行为也发生较大突变, 因此熵值变大。然而, 至于外圈故障, 内圈故障还是滚动体故障, 三者的区分仍不明显。因此, 为了实现故障类别的诊断, 论文建立基于神经网络集成的滚动轴承多故障分类器。

采用训练样本分别对每个 BP 神经网络进行训练, 再将测试样本输入到已训练的神经网络分类器中, 通过相对多数投票法集成神经网络的输出, 结果如表 1 所示。

表 1 个体 BP 神经网络和神经网络集成输出识别率比较
Tab. 1 Accuracy rate comparison of individual BP neural network and NNE outputs

BP	隐含层数与动量因子	识别率/%			
		正常	外圈故障	内圈故障	滚动体故障
BP ₁	(12, 0.4)	100	100	100	100
BP ₂	(18, 0.4)	100	100	100	100
BP ₃	(24, 0.4)	100	100	100	80
BP ₄	(12, 0.6)	100	100	100	90
BP ₅	(18, 0.6)	100	100	100	100
BP ₆	(24, 0.6)	100	100	90	100
BP ₇	(12, 0.8)	100	100	90	100
BP ₈	(18, 0.8)	100	100	100	100
BP ₉	(24, 0.8)	100	100	100	100
集成输出		100	100	100	100

表 1 说明, 论文提出的方法有较高的故障识别率, 对试验样本达到 100%。而且表 1 也说明集成后的神经网络的泛化能力确实得到了提高, NNE 输出的准确率比单个输出要高和稳定。

为了说明进行多尺度分析的重要性, 采用原始信号的排列熵作为特征向量, 训练 9 个个体 BP 神经网络, 参数选择与上述选取相同, 同样的方法集成输出。4 种状态的故障识别率分解为: 正常 100%, 外圈故障 70%, 内圈故障 80%, 滚动体故障 100%。这说明直接对原始信号提取排列熵作为特征向量的分类效果并不理想, 这也验证了论文方法中进行 LCD 分解的必要性。

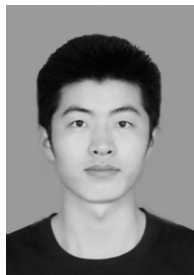
4 结束语

笔者提出了一种基于 LCD-排列熵和神经网络集成的滚动轴承故障诊断方法。排列熵可用于振动

信号随机性和动力学突变行为的检测,但由于机械系统的复杂性,振动信号表现不同尺度的动力学突变,基于此,笔者提出用LCD方法自适应地将振动信号进行多尺度的分解,再检测不同尺度下分量的动力学突变行为,以此作为故障特征信息。将其作为特征向量输入神经网络集成分类器,实现滚动轴承的故障诊断。通过对原始振动信号直接求排列熵与文中方法进行对比,结果表明了进行多尺度分解的必要性。实验数据分析表明,笔者提出的方法能够有效地识别滚动轴承故障类型,为滚动轴承的故障诊断提供了另一种思路和手段。

参 考 文 献

- [1] Logan D, Mathew J. Using the correlation dimension for vibration fault diagnosis of rolling element bearing-I: Basic concepts [J]. *Mechanical Systems and Signal Processing*, 1996, 10(3):241-250.
- [2] 彭志科,何永勇,卢青,等. 小波多重分形及其在振动信号分析中应用的研究[J]. *机械工程学报*, 2002, 38(8):59-63.
Peng Zhike, He Yongyong, Lu Qing, et al. Wavelet multiracial spectrum: application to analysis vibration signals [J]. *Chinese Journal of Mechanical Engineering*, 2002, 38(8):59-63. (in Chinese)
- [3] 侯荣涛,闻邦椿,周飙. 基于现代非线性理论的汽轮发电机组故障诊断技术研究[J]. *机械工程学报*, 2005, 41(2):142-147.
Hou Rongtao, Wen Bangchun, Zhou Biao. Study on fault diagnosis technique to turbo unit based on modern nonlinear theories[J]. *Chinese Journal of Mechanical Engineering*, 2005, 41(2): 142-147. (in Chinese)
- [4] Yan Ruqiang, Gao R X. Approximate entropy as a diagnostic tool for machine health monitoring[J]. *Mechanical Systems and Signal Processing*, 2007, 21(2): 824-839.
- [5] Zhang Long, Xiong Guoling, Liu Hesheng. Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference [J]. *Expert Systems with Applications*, 2010, 37: 6077-6085.
- [6] Bandt C, Pompe B. Permutation entropy: a natural complexity measure for time series [J]. *Physical Review Letters*, 2002, 88(17): 174102(1-4).
- [7] Yan Ruqiang, Liu Yongbin, Gao X. Permutation entropy: a nonlinear statistical measure for status characterization of rotary machines [J]. *Mechanical Systems and Signal Processing*, 2012, 29(5): 474-484.
- [8] 程军圣,郑近德,杨宇. 一种新的非平稳信号分析方法—局部特征尺度分解法[J]. *振动工程学报*, 2012, 25(2):215-220.
Cheng Junsheng, Zheng Jinde, Yang Yu. A new non-stationary signal analysis approach—the local characteristic-scale decomposition method[J]. *Journal of Vibration Engineering*, 2012, 25(2): 215-220. (in Chinese)
- [9] Huang N E, Wu Zhaohua. A review on Hilbert-Huang transform: method and its applications to geophysical studies[J]. *Reviews of Geophysics*, 2008, 46(2):RG2006(1-23). DOI:10.1029/2007RG000228.
- [10] 张德祥,汪萍,吴小培,等. 基于EMD和非线性峭度的齿轮故障诊断[J]. *振动、测试与诊断*, 2012, 32(1): 56-61.
Zhang Dexiang, Wang Ping, Wu Xiaopei, et al. Gear fault diagnosis based on EMD and nonlinear kurtosis [J]. *Journal of Vibration, Measurement & Diagnosis*, 2012, 32(1): 56-61. (in Chinese)
- [11] Hansen L, Salamon P. Neural network ensembles[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, 12(10):993-1001.
- [12] 周志华,陈世福. 神经网络集成[J]. *计算机学报*, 2002, 25(1): 1-8.
Zhou Zhihua, Chen Shifu. Neural network ensemble [J]. *Chinese Journal of Computers*, 2002, 25(1): 1-8. (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设备故障诊断,动态信号处理等。曾发表《基于局部特征尺度分解的经验包络解调方法及其在机械故障诊断中的应用》(《机械工程学报》2012年第48卷第19期)等论文。

E-mail:lqdlzheng@126.com