

基于 EMD 样本熵-LLTSA 的故障特征提取方法

向 丹^{1,2}, 葛 爽²

(1. 广东技术师范学院 自动化学院, 广州 510635;
2. 华南理工大学 机械与汽车工程学院, 广州 510640)

摘 要: 针对振动信号的非线性、非平稳性以及微弱故障特征难以提取的问题,提出了一种基于经验模态分解(EMD)、样本熵和流形学习的故障特征提取方法.该方法将 EMD、样本熵和流形学习相结合.首先,利用 EMD 的自适应多分辨率的特点计算分解得到的 IMF(固有模态函数)信号的样本熵,初步提取滚动轴承状态特征值;然后利用流形学习方法对初步的提取的滚动轴承状态特征进行进一步的提取;最后利用支持向量机(SVM)对该特征提取方法进行分类评估,并将该方法运用在滚动轴承故障诊断实验中,实验证明该特征提取方法与基于小波包样本熵的故障诊断方法相比具有很好的聚类性能,且对于 SVM 的分类结果可达 100%,在降低了特征数据的复杂度的同时,增强了故障模式识别的分类性能,具有一定的优越性.

关 键 词: 经验模态分解;样本熵;流形学习;特征提取;支持向量机

中图分类号: V263.6; TH165.3

文献标志码: A

Method of fault feature extraction based on EMD sample entropy and LLTSA

XIANG Dan^{1,2}, GE Shuang²

(1. School of Automation, Guangdong Polytechnic Normal University,
Guangzhou 510635, China;

2. School of Mechanical and Automotive Engineering,
South China University of Technology, Guangzhou 510640, China)

Abstract: A fault feature extraction method based on the empirical mode decomposition (EMD), sample entropy and manifold learning was presented to account for a range of issues of the vibration signal, e. g. Nonlinearities, non-stationary and weak fault features hard to extract. The proposed method combined the EMD, sample entropy and manifold learning techniques. Firstly, on the basis of the property of adaptive multi-resolution for the EMD technique, the sample entropy of the IMF (intrinsic mode function) signal reconstructed by using the EMD was calculated, and the state features of the rolling bearing were preliminarily extracted. Secondly, the extraction performance of the state features was further implemented by using the manifold learning technique. Finally, the SVM (support vector machine) was employed to classify and to evaluate the feature extraction method. Moreover, the proposed method was applied to the experiment of the rolling bearing fault diagnosis. The experimental results show that the proposed fault feature extraction method has more

robust clustering performance than the fault diagnosis method based on the sample entropy of wavelet packets. Furthermore, a relatively high precision, namely, 100% of classification result for the SVM, can be obtained. The proposed method not only decreases the complexity of the feature data, but also enhances the classification performance of fault diagnosis and pattern recognition, thus bringing about certain superiority.

Key words: EMD (empirical mode decomposition); sample entropy; manifold learning; feature extraction; SVM (support vector machine)

对机械状态进行检测和诊断的关键在于信号特征的提取.而在实际监测过程中,由于测试环境的影响,提取的信号往往带有噪声干扰,使得一些微弱的表征机械故障状态的信号特征完全淹没在噪声中,而由于负载、间隙、刚度、摩擦等非线性因素的影响,使得利用传统振动信号的分析方法诸如时域和频域分析对机械的工作状态进行精确诊断是比较困难的^[1].

经验模态分解是一种新型的、自适应的信号处理方法,非常适合非线性和非平稳信号的分析^[2].熵作为信号不确定程度的一种度量,近年来开始在信号处理和特征提取领域得到了广泛的应用. Pincus^[3]首次提出了基于近似熵的改进算法,并将样本熵广泛用于生理信号处理上. YAN Ru-qiang 等^[4]将近似熵概念运用到轴承故障诊断上,取得了很好的诊断效果. 赵志宏等^[5]在样本熵的基础上提出了一种基于 EMD 与样本熵的轴承故障诊断方法. 流形学习作为一种非线性的数据维数约简方法,与传统方法相比能有效地发现非线性高维数据的本质结构,利于进行维数约简和非线性数据分析. 夏鲁瑞等^[6]运用流形学习算法对涡轮泵测试海量数据进行故障特征提取,以可视化的方式直观地识别出涡轮泵数据中的异常状态. 李峰等^[7]将流形学习用在滚动轴承故障诊断中,提出了基于线性局部切空间排列维数约简的故障诊断方法,并验证了该方法的有效性.

在理论分析的基础上,本文提出了一种基于 EMD、样本熵和流形学习的特征提取方法,利用 EMD 自适应分解信号和样本熵的容错性强的特点,初步提取振动信号的故障特征,利用 LLTSA (线性局部切空间排列)非线性特征提取能力的优越性对信号进行特征提取,得到一个维数低、敏感度高且分类错误率小的故障特征,最后利用支持向量机对该方法特征提取效果进行评估,并通过实验验证该方法的有效性.

1 经验模态分解和样本熵

1.1 经验模态分解

EMD 方法是由 Huang 等在 1998 年提出的一种时频分析方法^[8].从本质上讲,EMD 方法将信号中不同尺度的波动或趋势逐步分解出来,是对时间序列进行平稳化处理过程,产生一系列具有不同特征尺度的数据序列,每一个序列称为一个固有模态函数. IMF 分量应满足两个条件:一是其极值点数和零点数相同或最多相差 1 个;二是其上下包络线关于时间轴局部对称.

运用 IMF 可以把任何信号 $f(t)$ 按如下步骤分解^[9]:

步骤 1 确定信号的所有局部极值点,再用 3 次样条线分别将所有局部极大值和极小值点连接起来形成上下包络线.上下包络线的平均值记为 a_1 ,求出

$$f(t) - a_1 = h_1 \quad (1)$$

理想地,如果 h_1 满足 IMF 两个条件,则它就是 $f(t)$ 的第 1 个 IMF 分量.

步骤 2 如果 h_1 不满足 IMF 的条件,则把 h_1 作为原始数据,重复 k 次步骤 1,得到 $h_{1(k-1)} - a_{1k} = h_{1k}$,使得 h_{1k} 满足 IMF 条件.记 $c_1 = h_{1k}$,则 c_1 为信号 $f(t)$ 的第 1 个满足 IMF 条件的分量.

步骤 3 将 c_1 从 $f(t)$ 中分离出来,得到

$$r_1 = f(t) - c_1 \quad (2)$$

将 r_1 作为原始数据重复步骤 1 和步骤 2,得到 $f(t)$ 的第 2 个满足 IMF 条件的分量 c_2 ,重复循环 n 次,得到信号 $f(t)$ 的 n 个满足 IMF 的分量 $c_1, c_2, \dots, c_n$,当 c_n 成为一个单调函数不能再从中提取满足 IMF 的分量时,结束循环,得到

$$f(t) = \sum_{j=1}^n c_j + r_n \quad (3)$$

式中余项 r_n 代表信号的平均趋势,IMF 分量 c_j 包含了信号不同频率段的成分且非等带宽.所以说 EMD 方法是一种自适应的信号分析方法.

1.2 样本熵

样本熵是由 Richman 等^[10]提出的一种时间序列复杂性测度方法.具体算法如下:

1) 将 N 个数据组成的时间序列 $x(n)$ 按序号组成一个维数为 m 的向量.

2) 定义 $\mathbf{X}_m(i)$ 与 $\mathbf{X}_m(j)$ 间的距离 $d[\mathbf{X}_m(i), \mathbf{X}_m(j)]$ 是两个对应元素中最大差值的绝对值,即

$$d[\mathbf{X}_m(i), \mathbf{X}_m(j)] = \max[|x(i+k) - x(j+k)|] \quad (4)$$

3) 给定阈值 r , 对每个 i 值统计 $d[\mathbf{X}_m(i), \mathbf{X}_m(j)]$ 小于 r 的数目与距离总数 $N-m$ 的比值, 记作 $B_i^m(r)$, 即

$$B_i^m(r) = \frac{1}{N-m-1} \{d[\mathbf{X}_m(i), \mathbf{X}_m(j) < r]\} \\ i = 1, 2, \dots, N-m+1 \quad i \neq j$$

4) 则

$$B^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} B_i^m(r) \quad (5)$$

5) 增加维数到 $m+1$, 计算 Q 值, 即 $\mathbf{X}_{m+1}(i)$ 与 $\mathbf{X}_{m+1}(j)$ 小于等于 r 的个数, 记为 A_i 定义

$$A_i^m(r) = \frac{1}{N-m-1} A_i \\ A^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^m(r) \quad (6)$$

6) 样本熵定义为

$$S_n(m, r) = \lim_{N \rightarrow \infty} \left\{ -\ln \left[\frac{A^m(r)}{B^m(r)} \right] \right\} \quad (7)$$

样本熵主要与 m, r 两个参数的取值有关.不同的嵌入维数 m 和阈值 r 对应的样本熵不同.根据文献^[11]的研究成果, 当 $m=1$ 或 2 , r 在 $(0.10 \sim 0.25)E_{\text{std}}$ (E_{std} 为原始数据的标准差) 时计算所得到的样本熵具有较合理的统计特征.因此本研究取 $m=2, r=0.20E_{\text{std}}$.

2 线性局部切空间排列

LLTSA 是在非线性流形学习局部切空间排列(LTSA)的基础上, 借鉴 locality preserving projections 等算法思想, 为解决流形学习算法出现的算法泛化问题而提出的^[12-13].

LLTSA 的主要问题是寻找转换矩阵 $\mathbf{A}$ 将 $\mathbf{R}^p$ 空间中具有 N 个点的含噪声数据集 $\mathbf{X}_{\text{org}}$ (故障样本集) 映射为 $\mathbf{R}^q$ 空间数据集 $\mathbf{Y} = [y_1, y_2, \dots, y_N]$, 即

$$\mathbf{Y} = \mathbf{A}^T \mathbf{X}_{\text{org}} \mathbf{H}_N \quad q < p \quad (8)$$

其中 $\mathbf{H}_N = \mathbf{I} - \mathbf{e}\mathbf{e}^T/N$ 为中心距阵, $\mathbf{I}$ 为单位矩阵, $\mathbf{e}$ 为所有元素为 1 的 N 维列向量, $\mathbf{Y}$ 为 $\mathbf{X}_{\text{org}}$ 潜在的 q 维非线性流形.

LLTSA 具体算法如下:

1) 主成分分析(PCA)投影. 将数据集映射到 PCA 子空间, 转换矩阵设为 $\mathbf{A}_{\text{pca}}$, 为方便描述, 映射后的数据依然用 $\mathbf{X} = [x_1, x_2, \dots, x_N]$.

2) 确定领域. 构建 x_i 点的局部领域图, 可以用 k 近邻法或 ϵ 近邻法.

3) 提取局部信息. 通过计算 $\mathbf{X}_i \mathbf{H}_k$ 的 q 个最大右奇异值对应的奇异向量, 得到局部低维空间矩阵 $\mathbf{V}$, 计算 $\mathbf{W}_i = \mathbf{H}_k(\mathbf{I} - \mathbf{V}_i \mathbf{V}_i^T)$.

4) 构造排列矩阵. 通过局部累加构造排列矩阵 $\mathbf{B}$ 初始化 $\mathbf{B} = \mathbf{0}$.

$$\mathbf{B}(\mathbf{I}_i, \mathbf{I}_i) \leftarrow \mathbf{B}(\mathbf{I}_i, \mathbf{I}_i) + \mathbf{W}_i \mathbf{W}_i^T \quad (9)$$

5) 计算映射. 计算广义特征问题的特征值和特征向量

$$\mathbf{X} \mathbf{H}_N \mathbf{B} \mathbf{H}_N \mathbf{X}^T \boldsymbol{\alpha} = \lambda \mathbf{X} \mathbf{H}_N \mathbf{X}^T \boldsymbol{\alpha} \quad (10)$$

与特征值 $\lambda_1 < \lambda_2 < \dots < \lambda_q$ 对应的特征向量为 $\alpha_1, \alpha_2, \dots, \alpha_q$, 则

$$\mathbf{A}_{\text{llttsa}} = (\alpha_1, \alpha_2, \dots, \alpha_q) \quad (11)$$

因此转换矩阵为 $\mathbf{A} = \mathbf{A}_{\text{pca}} \mathbf{A}_{\text{llttsa}}$, 则原数据的最终低维全局坐标为 $\mathbf{Y} = \mathbf{A}^T \mathbf{X} \mathbf{H}_N$.

3 支持向量机

支持向量机是 Vapnik 首次提出的, 并在 20 世纪 90 年代中期不断发展和成熟, 现已广泛应用于机械故障诊断领域. SVM 在解决小样本、非线性及高维模式识别问题中表现出很大的优势. 其主要思想是建立一个超平面作为决策曲面, 使得正例与反例之间的隔离边缘被最大化^[14-15].

SVM 是从线性可分情况下发展而来的, 设用于分离的超平面方程是

$$\boldsymbol{\omega} \cdot \mathbf{x} + b = 0 \quad (12)$$

式中 $\boldsymbol{\omega}$ 是超平面的法向量, b 是超平面的常数项. 可以对式(12)进行归一化, 使得线性可分的样本集 $(x_i, y_i), i=1, \dots, l, \mathbf{x} \in \mathbf{R}^q, y \in \{+1, -1\}$, 满足

$$y_i[(\boldsymbol{\omega} \cdot x_i) + b] - 1 \geq 0 \quad i=1, \dots, l \quad (13)$$

此时分类间隔为 $2/\|\boldsymbol{\omega}\|$, 要使得分类间隔最大化相当于使 $\|\boldsymbol{\omega}\|^2$ 最小. 满足式(13)且使得 $\|\boldsymbol{\omega}\|^2/2$ 最小的平面就是最优超平面. 而在这个最优超平面上的点称之为支持向量.

对于非线性分类情况, 用 $\mathbf{X} = \boldsymbol{\phi}(\mathbf{x})$ 将 $\mathbf{x}$ 经变

换映射到高维空间中去,并引入松弛变量 ξ_i 和惩罚因子 C ,则原来的最优函数问题变成

$$\begin{aligned} \min_{w, b, \xi} & \frac{1}{2} \omega^T \omega + C \sum_{i=1}^n \xi_i \\ \text{s. t.} & y_i [\omega^T \phi(x_i) + b] \geq 1 - \xi_i \\ & \xi_i \geq 0 \quad i = 1, 2, \dots, n \end{aligned} \quad (14)$$

4 EMD 样本熵-LLTSA 故障特征提取方法

该特征提取方法利用 EMD 分析的自适应多分辨率和样本熵强抗噪和抗干扰能力的特点,将 EMD 与样本熵结合起来在多个尺度内对信号进行特征的提取,然后利用流形学习的非线性复杂信息处理能力进一步提取高维数据的低维特征,获取维数低、敏感度高且分类错误率小的主要特征向量,并通过支持向量机对该方法提取的特征效果进行评估.该特征提取方法及评估流程图如图 1 所示,具体步骤如下:

1) 对故障信号做 EMD,得到各 IMF 分量.

2) 求出各 IMF 分量的样本熵值.由于不同的振动信号所求 IMF 分量个数不同,但 EMD 也是一种主成分提取方法,分解结果中前面几个 IMF 分量包含了原始信号的主要信息,且本文实验所得原始振动信号是幅值大的非平稳信号,其 EMD 次数均大于 9,因此本文都利用 EMD 的前 9 个 IMF 分量求取样本熵值,记为 S_{ni} .

3) 将求取的 9 个样本熵值构成一个 9 维的特征向量

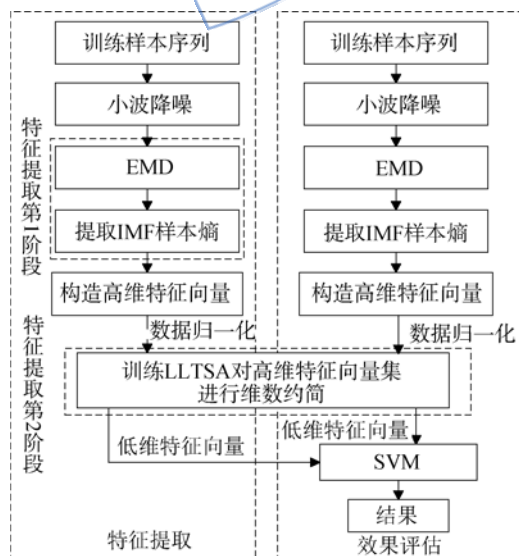


图1 EMD 样本熵-LLTSA 特征提取流程图

Fig. 1 Flow chart of feature extraction based on EMD sample entropy and LLTSA

$$V = [S_{n1}, S_{n2}, \dots, S_{nq}]$$

4) 特征值的归一化处理.

$$S_i = S_{ni} / S \quad S = \sum_{j=1}^9 S_{nj}$$

5) 流形学习方法进一步特征提取:利用 LLTSA 对初提取的 EMD 样本熵特征向量进行进一步的特征提取.

5 实验验证分析

5.1 数据来源

滚动轴承是机械设备中最常见的、最易损坏的器件之一,本文通过对变速箱滚动轴承故障的监测来证明本文提出的基于 EMD 样本熵-LLTSA 故障特征提取方法的有效性.实验数据来自变速箱滚动轴承模拟实验台测试.变速箱型号为 JZQ-250,通过人为制造一些滚动轴承微小故障(故障直径都约为 0.2 mm)来模拟真实故障进行实验和相关数据的采集.该实验平台如图 2 所示.



图2 轴承振动实验台

Fig. 2 Bearing vibration test bench

实验主要采集驱动电动机在 $100 \text{ N} \cdot \text{m}$ 负载情况下 4 种滚动轴承工作状态的数据:轴承外圈故障、轴承内圈故障、轴承滚动体故障和轴承正常状态.数据的采样频率为 12 kHz ,驱动电动机转速为 1750 r/min .

5.2 特征提取分析

图 3 为实验中变速箱滚动轴承滚动体故障的原始振动信号和经过 EMD 后得到的 17 个 IMF 分量和 1 个残余分量 r .由于分解结果中前面几个 IMF 分量已经包含了原始信号的主要信息,限于篇幅,图 3 只列出原始信号和 EMD 的前面 9 个 IMF 分量信号.

从图 3 可以看到,由于 EMD 在不同尺度下的分解,滚动轴承的故障特征在不同的时间尺度下的 IMF 分量中显示出来.第 1 个 IMF 分量 c_1 含有的频率成分最高,并随着 IMF 分量逐渐降

低. 从 c_1, c_2 分量中可以明显看出存在周期性冲击信号特征, 而 c_3, c_4 分量中存在冲击信号特征, 但其周期性并不明显.

滚动轴承振动信号经过 EMD 得到一系列 IMF 分量后, 计算各 IMF 分量的样本熵, 表 1 是滚动轴承 4 种不同状态下 2 个样本信号的 EMD 样本熵. 滚动轴承不同状态下的 EMD 样本熵是不同的, 滚动轴承不同状态下的 EMD 样本熵的

大小反映了信号分布的确定性. 由表 1 可以看出, 滚动轴承正常状态下的 EMD 样本熵大于其他故障状态下的样本熵, 这是因为在正常状态下, 振动信号的分布平均和不确定, 当出现故障时, 在相应的频带内出现相应的冲击信号, 使得信号的确定性降低, 从而使样本熵减小, 由于轴承滚动体故障比其他故障程度更严重, 其样本集中更为严重, 所以其 EMD 样本熵最小.

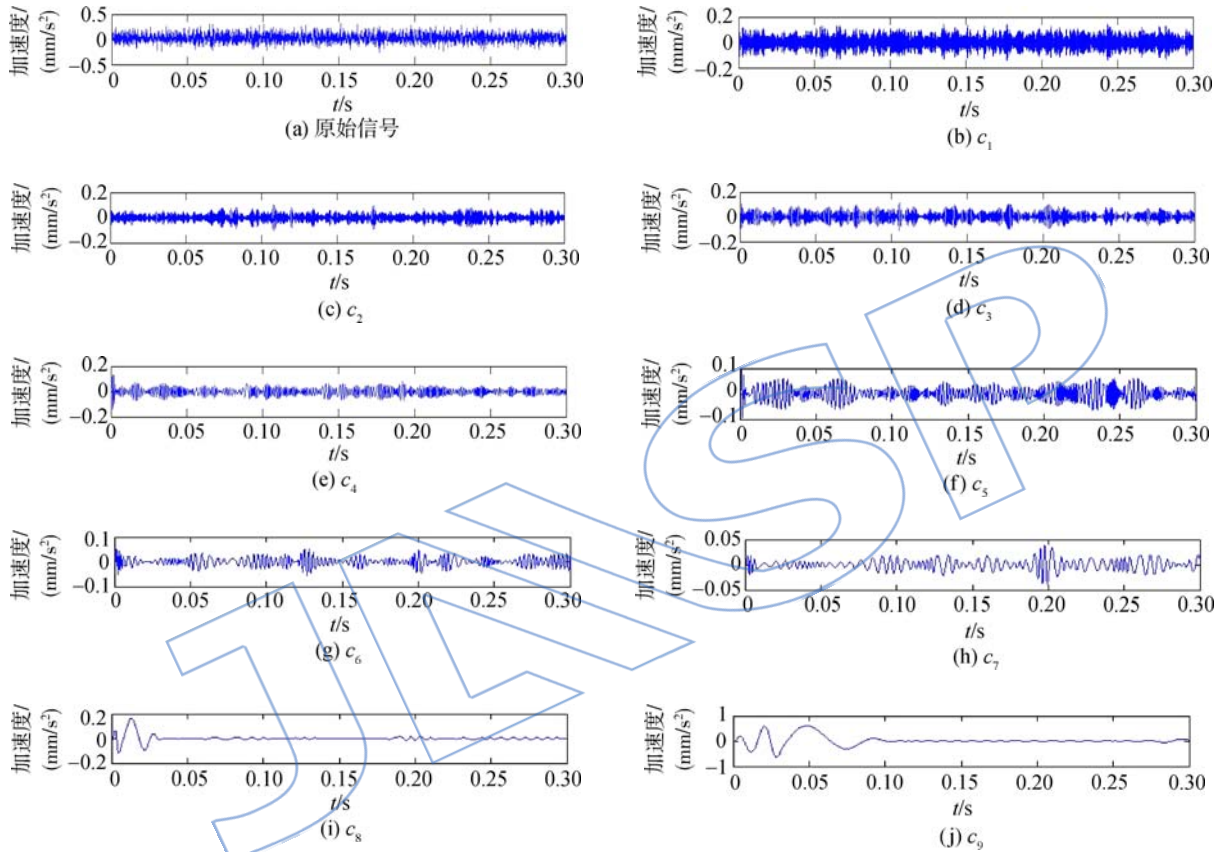


图 3 滚动体故障振动信号与 EMD 结果

Fig. 3 Vibration signal of rolling ball fault and its EMD results

表 1 滚动轴承各状态下的 EMD 样本熵

Table 1 EMD sample entropy of every rolling bearing status

轴承状态	信号序号	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9
正常	1	0.5324	0.5412	0.4913	0.2719	0.2591	0.1911	0.0873	0.0862	0.0795
	2	0.5391	0.5183	0.4872	0.2139	0.2035	0.1732	0.0735	0.0803	0.0709
外圈故障	1	0.5032	0.5423	0.4781	0.2514	0.2231	0.1514	0.0719	0.0783	0.0689
	2	0.5174	0.5419	0.4593	0.2631	0.2383	0.1483	0.0723	0.0694	0.0652
内圈故障	1	0.4753	0.4931	0.4514	0.2619	0.2448	0.1314	0.0690	0.0714	0.0639
	2	0.4832	0.5092	0.4609	0.2594	0.2394	0.1084	0.0645	0.0681	0.0607
滚动体故障	1	0.4613	0.4704	0.4398	0.2439	0.2648	0.1149	0.0584	0.0514	0.0501
	2	0.4594	0.4695	0.4417	0.2384	0.2384	0.1091	0.0432	0.0432	0.0418

EMD方法可以将任何一个信号 $x(t)$ 分解成若干个 IMF 分量和一个残余分量之和, 这样可以使信号特征自适应地在不同的分辨率下显示出来, 和小波包的多分辨率相比, 采用 EMD 信号效果更好更稳定, 因为 EMD 不像小波包分解那样需要基函数, 自适应能力强, 更能有效地反映信号的内部特征, 避免了信号能量的扩散和泄漏. 图 4 和图 5 为滚动轴承两种状态下的 EMD 样本熵和小波包样本熵对比图. 实验取 6 个数据样本分别求其 EMD 样本熵和小波包样本熵, 小波包分解所用基函数为 db4 小波基进行 3 层分解.

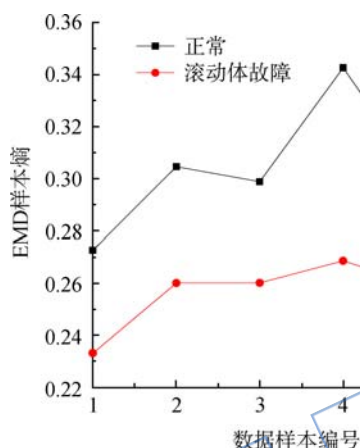


图 4 实测信号 EMD 样本熵对比图

Fig. 4 Comparison chart of actual measured signal of EMD sample entropy

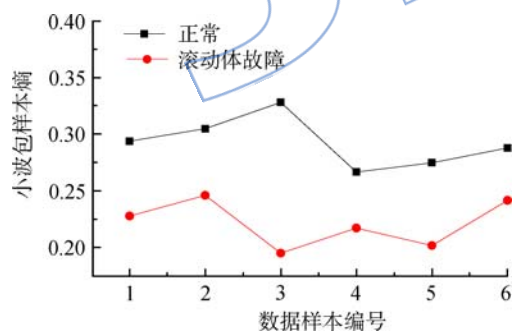


图 5 实测信号小波包样本熵对比图

Fig. 5 Comparison chart of actual measured signal of wavelet packet sample entropy

比较图 4 和图 5 可以看出, 经过 EMD 得到正常和滚动体故障时的 EMD 样本熵均值差为 0.1042, 而经过小波包 3 层分解得到的正常和滚动体故障时的小波包样本熵均值差为 0.0463, 这不仅证明了 EMD 样本熵和小波包样本熵可以作为故障特征用于模式识别, 而且证明了基于 EMD 样本熵提取的故障信息更为有效.

因此, 可以通过 EMD 对振动信号进行多分辨率的分解, 求出不同频带的 IMF 分量, 并求出各 IMF 分量的样本熵, 将之作为故障特征向量 (此为特征提取第 1 阶段), 再利用 LLTSA 进行进一步的故障特征提取 (特征提取第 2 阶段). 针对滚动轴承 4 种不同的工作状态分别选取各状态振动信号 4000 个数据作为一个样本, 对 20 个样本分别进行 EMD, 求出各 IMF 分量的样本熵值, 组成一个 9 维的特征向量, 然后将 9 维的特征向量用 LLTSA 降到 3 维以提取低维的故障特征. 作为比较, 同样求取这 20 个样本的 3 层小波包样本熵. 图 6 和图 7 为 LLTSA 对于两种方法特征提取的效果图, 领域参数 $K=7$, 目标维数为 3.

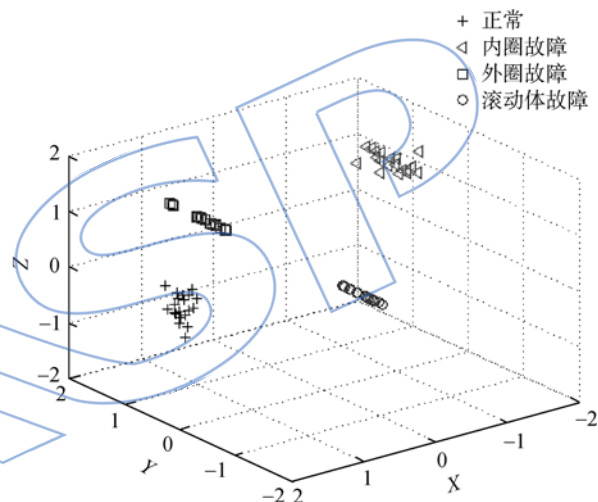


图 6 EMD 样本熵 LLTSA 特征提取效果

Fig. 6 LLTSA features extraction effect of EMD sample entropy

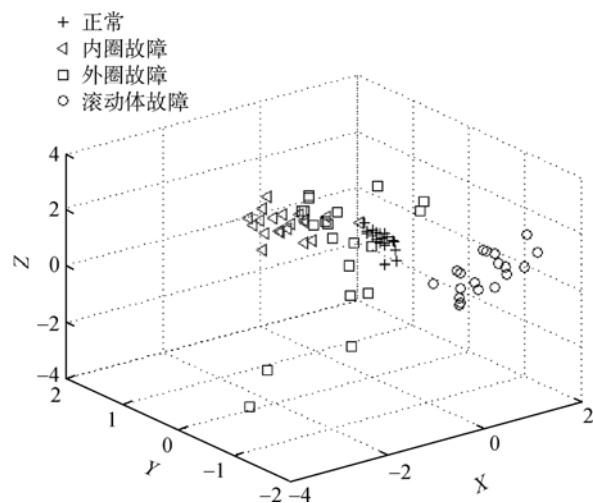


图 7 小波包样本熵 LLTSA 特征提取效果

Fig. 7 LLTSA features extraction effect of wavelet packet sample entropy

由图 6 可以看出,滚动轴承 4 种状态特征已经完全分开,并且提取的特征具有非常好的聚类性能,这证明了基于 EMD 样本熵-LLTSA 故障特征提取方法的有效性.由图 7 可以看出基于小波包样本熵和 LLTSA 的故障特征提取方法只能将滚动轴承的正常、内圈故障和滚动体故障 3 种状态分开,提取的特征聚类性能差,且由于小波基函数和分解层数必须人为选择,分解得到分量的模式比较固定,对信号的适应性差,从而造成经过小波包样本熵提取的效果差异很大.由此可以证明本文所提方法对于特征提取具有一定的优越性.

5.3 故障识别分析

为了验证本文所提特征提取方法对于故障诊断的效果,利用支持向量机对提取的故障特征进行分类和诊断.作为比较,实验选用第 1 阶段特征提取(EMD 样本熵)效果和第 2 阶段特征提取效果进行比较,为了证明本文所提方法对于故障特征提取的优越性,又将该方法与基于小波包样本熵-LLTSA 特征提取方法进行故障识别比较,各方法故障特征识别结果如表 2 所示.实验将 20 个样本数据分成两类,将 10 个样本数据作为训练样本,另外 10 个样本数据作为测试样本.实验中 SVM 核函数选择最常用的高斯核函数,核函数高度控制参数 $\sigma=1$,惩罚因子 $C=10$.

表 2 支持向量机识别结果
Table 2 Support vector machine recognition results

参数	正常	内圈故障	外圈故障	滚动体故障
训练样本/个	10	10	10	10
训练标签	1	2	3	4
测试样本/个	10	10	10	10
第 1 阶段识别率/%	80	80	60	90
小波包样本熵-LLTSA 识别率/%	90	100	80	100
EMD 样本熵-LLTSA 识别率/%	100	100	100	100

从表 2 可以看出,第 1 阶段提取的故障特征虽然能较准确地诊断出滚动轴承的工作状态,但其识别率非常低,而且 SVM 训练时间长达 20.5 s,这说明提取的 EMD 样本熵中含有大量的重叠的故障信息,使得后面分类器训练时间加长,而且

降低了故障特征的分辨率,从而证明有必要进行进一步的特征提取.而基于小波包样本熵-LLTSA 故障特征提取方法虽然能对滚动轴承状态进行准确的判别,且识别率可达 80% 以上,但由于该方法对于小波基和分解层次不好把握,对于不同信号提取的特征差异较大,之所以有 80% 以上的故障分辨率是因为选择的样本非常小,对于大样本的分量效果明显降低.基于 EMD 样本熵-LLTSA 的故障特征提取方法对于其 4 种工作状态的识别率为 100%,且 SVM 训练所费时间短,只有 2.2 s,且聚类性能好.

6 结 论

本文在基于 EMD、样本熵和线性局部切空间排列的基础上提出了一种基于 EMD 样本熵-LLTSA 故障特征提取方法.结合 EMD 处理非平稳信号的优越性、样本熵的容错性和 LLTSA 对于非线性特征提取能力的特点,应用于滚动轴承故障诊断实验.实验结果表明,基于 EMD 样本熵-LLTSA 故障诊断方法对于滚动轴承 4 种不同工作状态的识别率为 100%,而且提取的特征向量具有很好的聚类性能,这样不仅提高了故障特征的分辨率,而且简化了特征数据,增强了故障模式识别的分类性能.

参考文献:

- [1] 翟旭升,胡金海,谢寿生,等.基于 DSMT 的航空发动机早期振动故障融合诊断方法[J].航空动力学报,2012,27(2):301-304.
ZHAI Xusheng, HU Jinhai, XIE Shousheng, et al. Diagnosis of aero-engine with early vibration fault symptom using DsmT[J]. Journal of Aerospace Power, 2012, 27(2): 301-304. (in Chinese)
- [2] 鞠萍华,秦树人,秦毅,等.多分辨 EMD 方法与频域平均在齿轮早期故障诊断的研究[J].振动与冲击,2009,28(5):97-101.
JU Pinghua, QIN Shuren, QIN Yi, et al. Study of the average in the early fault diagnosis of gear multiresolution and frequency domain EMD method[J]. Journal of Vibration and Shock, 2009, 28(5): 97-101. (in Chinese)
- [3] Pincus S M. Approximate entropy as a complexity measure[J]. Chaos, 1995, 5(1): 110-117.
- [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] 赵志宏,杨绍普.一种基于样本熵的轴承故障诊断方法[J].振动与冲击,2012,31(6):136-140.
ZHAO Zhihong, YANG Shaopu. A bearing fault diagnosis

- method based on sample entropy[J]. *Journal of Vibration and Shock*, 2012, 31(6):136-140. (in Chinese)
- [6] 夏鲁瑞, 胡芑庆, 秦国军, 等. 基于流形学习的涡轮泵海量数据异常识别算法[J]. *航空动力学报*, 2011, 26(3):698-703.
- XIA Lurui, HU Niaoqing, QIN Guojun, et al. Abnormal recognition algorithm based on manifold learning for turbopump mass data[J]. *Journal of Aerospace Power*, 2011, 26(3):698-703. (in Chinese)
- [7] 李峰, 汤宝平, 陈法法. 基于线性局部切空间排列维数化简的故障诊断[J]. *振动与冲击*, 2012, 31(13):36-41.
- LI Feng, TANG Baoping, CHEN Fafa. Fault diagnosis based on dimension reduction using linear local tangent space alignment[J]. *Journal of Vibration and Shock*, 2012, 31(13):36-41. (in Chinese)
- [8] Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. *Proceedings of the Royal Society of London: Series A Mathematical, Physical and Engineering Science*, 1998, 454:903-995.
- [9] Fegeant O. Closed form solutions for the point mobilities of axisymmetrically excited cylindrical shells[J]. *Journal of Sound and Vibration*, 2001, 243(1):89-115.
- [10] Richman J S, Moorman J R. Physiological time series analysis using approximate entropy and sample entropy[J]. *American Journal of Physiology-Heart Circulatory Physical*, 2000, 278(3):2039-2049.
- [11] Pincus S M. Assessing serial irregularity and its implications for health[J]. *Annals of the New York Academy of Sciences*, 2002, 954(1):245-267.
- [12] Zhang T, Yang J, Zhao D, et al. Linear local tangent space alignment and application to face recognition[J]. *Neurocomputing*, 2007, 70(7/8/9):1547-1553.
- [13] 梁霖, 徐光华, 栗茂林, 等. 冲击故障特征提取的非线性流形学习方法[J]. *西安交通大学学报*, 2009, 43(11):95-99.
- LIANG Lin, XU Guanghua, LI Maolin, et al. The nonlinear manifold learning method of feature extraction of impulse fault[J]. *Journal of Xi'an Jiaotong University*, 2009, 43(11):95-99. (in Chinese)
- [14] LI Hua, ZHANG Yongxin. An algorithm of soft fault diagnosis for analog circuit based on the optimized SVM by GA [C] // *The Ninth International Conference on Electronic Measurement and Instruments*. [S. l.]: IEEE, 2009:1023-1024.
- [15] Vapnik V N. *The nature of statistical learning theory*[M]. 2nd ed. New York: Springer, 1995.