文章编号：1674-8085(2017)02-0066-08

一种基于经验模态分解和流形学习的滚动轴承故障诊断方法

蔡江林，*戚晓利，叶绪丹，郑近德，潘紫微，张兴权

（安徽工业大学机械工程学院，安徽，马鞍山 243002)

摘　要：提出一种基于经验模态分解（EMD）和流形学习（LTSA）的滚动轴承故障诊断方法。首先，利用 EMD 对滚动轴承振动信号进行自适应分解，计算 IMF 分量的协方差矩阵特征值，组成滚动轴承状态原始特征集；然后利用 LTSA 对原始特征集进一步的融合提取；将所得新特征输入到 K-means 分类器中进行故障识别与聚类。实验分析结果表明：该方法可以有效地对滚动轴承的工作状态和故障类型进行识别。

关键词：滚动轴承；经验模态分解；流形学习；局部切空间排列算法；K-means 分类器

中图分类号：TH165.3　　**文献标识码**：A　　DOI:10.3969/j.issn.1674-8085.2017.02.013

A ROLLER BEARING FAULT DIAGNOSIS METHOD BASED ON THE EMPIRICAL MODE DECOMPOSITION (EMD) AND MANIFOLD LEARNING (LTSA)

CAI Jiang-lin, *QI Xiao-li, YE Xu-dan, ZHENG Jin-de, PAN Zi-wei, ZHANG Xing-quan

(School of Mechanical Engineering, Anhui University of Technology, Ma'anshan 243002, China)

Abstract: A roller bearing fault diagnosis method based on the empirical mode decomposition (EMD) and manifold learning (LTSA) was presented. After the adaptive decomposition of the roller bearing vibration signal by the EMD technique, its original state feature set of the rolling bearing was acquired by calculating eigenvalues of IMF's covariance matrix. The extraction performance of the original feature set was further fusion implemented by using the LTSA. The new features obtained were input into a K-means classifier, and the output of the K-means classifier was clustering results. Finally, the experiment results show that the proposed method can effectively identify work status and fault type of roller bearing.

Key words: roller bearing; EMD(empirical mode decomposition); manifold learning; LTSA(local tangent space alignment algorithm); K-means classifier

0 引言

目前提出的滚动轴承故障诊断方法，基本上是通过深沟球轴承振动实验来验证的，且轴承模拟故障大多数是由激光切割或电火花加工的，故障形状规则狭小。因此在轴承运转过程中产生较为明显的

收稿日期：2016-12-13；修改日期：2017-03-04
基金项目：国家自然科学基金项目(51505002，51375013)
作者简介：蔡江林(1991-)，男，安徽安庆人，硕士生，主要从事信号处理与故障诊断研究(E-mail:18855539706@163.com);
*戚晓利(1975-)，男，安徽马鞍山人，副教授，博士，主要从事故障诊断与撞击动力学研究(E-mail:qxlcheer@sina.com);
叶绪丹(1991-)，女，安徽安庆人，硕士生，主要从事信号处理与故障诊断研究(E-mail:1421866984@qq.com);
郑近德(1986-)，男，安徽临泉人，讲师，博士，主要从事非线性动力学，动态信号处理和机械设备故障诊断等方面的研究(E-mail:jdzheng@ahut.edu.cn);
潘紫微(1956-)，男，江苏南京人，教授，主要从事设备运行性能与状态研究，设备力学行为数值分析与应用研究(E-mail:pzw586@mail.ahwhptt.net.cn);
张兴权(1970-)，男，安徽庐江人，教授，博士，主要从事激光冲击波加工等方面的研究(E-mail:zhxq@ahut.edu.cn).

周期性冲击，使得振动信号中包含一个周期性的故障信号，便于经验模态分解[1]（EMD）、局部特征尺度分解[2]（LCD）以及变分模态分解[3]（VMD）等时频分析方法提取振动信号的故障特征频率从而进行诊断。而实际中的轴承故障振动信号更加复杂，直接使用上述时频分析方法可能效果不明显，但是故障信息仍然包含在信号中，因此可以提取这些故障特征进行故障识别和分类。流形学习[4]是典型的非线性降维方法，能够对高维的故障特征进行约简，获取低维本质。例如文献[5]提出了一种基于主流形识别的非线性时间序列降噪方法，成功提取出齿轮振动信号的冲击特征。栗茂林等[6]利用流形学习成功提取出轴承振动信号时域指标和小波频带能量组成的原始特征空间中的低维本质特征。赵洪杰等[7]提出了将相空间重构技术与非线性流形相结合的滚动轴承故障诊断方法。文献[8]利用样本熵初步提取滚动轴承状态特征值，再利用流形学习进一步降维提取，获得更加精确的低维特征量，并验证了该方法的有效性。但是上述方法在高维特征构造上都存在一些缺点：文献[6]中轴承振动信号时域、频域特征多种多样，从中提取何种特征及选择哪些特征组合进行故障诊断需要不断的尝试，目前没有统一标准；文献[7]中相空间重构的时间延迟和嵌入维数的选取对诊断效果也有很大影响；文献[8]中样本熵的有效性也取决于嵌入维数和阈值的正确选取。

在理论分析的基础上，本文提出了一种基于EMD 分解和流形学习的滚动轴承故障诊断方法。利用 EMD 将振动信号自适应分解成一系列内禀模态分量（IMF），再利用协方差矩阵求取特征值，利用流形学习对这些特征值进行融合提取，得到维数低、敏感度高且分类错误率小的故障特征，并输入到 K-means 分类器进行故障的识别与聚类。最后将该方法应用于调心球轴承故障诊断实验，验证了其有效性。

1 理论基础

1.1 经验模态分解

经验模态分解是依据自身的时间尺度特征来进行信号分解，无须预先设定任何基函数，在理论上可以应用于任何类型信号的分解，因而在处理非平稳、非线性信号中，具有非常明显的优势。其本质是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内禀模态分量。IMF 分量必须满足两个条件：极值点数和零点数相同或最多相差一个；由极大值和极小值生成的上下包络线关于时间轴局部对称。

经验模态分解的基本过程[9]如下所示：

1）确定信号 $f(t)$的所有局部极值点，再用三次样条曲线分别将所有局部极大值点和极小值点分别连接起来形成上下包络线。求出上下包络线的平均值 a_1，计算：

$$f(t)-a_1=h_1 \tag{1}$$

当 h_1 满足 IMF 的两个条件时，则它就是 $f(t)$ 的第一个 IMF 分量。

2）当 h_1 不满足 IMF 的两个条件时，则 h_1 把作为原始数据，重复上述第一步 k 次，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 $f(t)$ 的第一个满足 IMF 条件的分量。

3）将 c_1 从 $f(t)$ 中分离出来，由如下公式计算：

$$f(t)-c_1=r_1 \tag{2}$$

将 r_1 作为原始数据重复上述第一、二步，得到 $f(t)$ 的第二个满足 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 条件的分量 $c_1,c_2,\cdots,c_n$，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：

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

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是一种自适应的信号分析方法。

1.2 流形学习方法

目前已存在的流形学习主要算法可分为两类：一类是全局方法，即计算高维空间中每个数据点和其他所有数据点的关系，如等距映射算法、主曲线算法、流形坐标卡等。另一类是局部方法，这类方法只考虑高维空间中每个数据点与它邻域内点的关系，典型的有拉普拉斯特征映射算法、局部线性嵌入算法、海赛局部线性嵌入算法以及局部切空间排列算法等。通过对比分析，本文选用的是局部切空间排列算法（LTSA）。LTSA 是由浙江大学的张振跃等人[10]于 2004 年首先提出的，其主要的思想是：寻找高维数据集中每个样本点的邻域，利用样本点邻域的切空间来近似表示原始数据局部的几何性质，最后通过对这些局部切空间的仿射变换得到整体的低维嵌入坐标。

LTSA 算法的基本步骤如下所示：

1）构造局部邻域。对每个样本点x_i，$x_i \in R^m$，$i=1,2,\cdots,N$，假设$X_i=[x_{i1},x_{i2},\cdots,x_{ik}]$为样本点$x_i$包括自身在内的$k$个邻域点构成的邻域矩阵。

2）局部线性拟合。LTSA 算法求解一个d维的仿射子空间来逼近邻域矩阵X_i中的点，

如下式所示：

$$\min_{x,\Theta,Q}\sum_{j=1}^{k}\left\|x_{i_j}-(x+Q\theta_j)\right\|_2^2=$$

$$\min_{x,\Theta,Q}\sum_{j=1}^{k}\left\|x_i-(xe^T+Q\Theta)\right\|_F^2 \tag{4}$$

Q是$m\times d$阶正交矩阵，$\Theta=[\theta_1,\cdots,\theta_k]$。记$\bar{x}=X_ie_i$为邻域矩阵$X_i$的中心点，$Q_i\Sigma_iV_i^T$为邻域中心矩阵$X_i-\bar{x}e_k^T=[x_{i_1}-\bar{x},\cdots,x_{i_k}-\bar{x}]$的奇异值分解，其中$Q_i\in R^{m\times m}$，$\sum i\in R^{m\times k}$，$V_i\in R^{k\times k}$。求出上式的最优解，从而得到局部坐标系统，如下式所示：

$$\Theta_i=[\theta_1^{(i)},\theta_2^{(i)},\cdots,\theta_k^{(i)}] \tag{5}$$

3）局部坐标全局排列。LTSA 算法认为低维特征空间的全局坐标τ_{i_j}应该能反映由局部坐标系统Θ_i所决定的局部几何结构，即全局坐标τ_{i_j}应满足：

$$\tau_{i_j}=\bar{\tau}_i+L_i\theta_j^{(i)}+\varepsilon_j^{(i)}\text{，}j=1,2,\cdots,k;i=1,2,\cdots,N \tag{6}$$

式中，$\bar{\tau}_i$是τ_{i_j}的均值，L_i是一个待定的局部仿射变换矩阵，$\varepsilon_j^{(i)}$表示局部的重建误差。记$T_i=[\tau_{i_1},\tau_{i_2},\cdots,\tau_{i_k}]$，$E_i=[\varepsilon_1^{(i)},\varepsilon_2^{(i)},\cdots,\varepsilon_k^{(i)}]$，则将上式转换为矩阵形式：

$$T_i=T_ie_ke_k^T\big/k+L_i\Theta_i+E_i \tag{7}$$

从而局部重建误差为：

$$E_i=T_i(I-e_ke_k^T\big/k)-L_i\Theta_i \tag{8}$$

为了在低维特征空间的整体坐标系中尽量保持原始高维数据集局部的低维特征，LTSA 算法将下面的重建误差极小化：

$$E(T)=\sum_i\|E_i\|^2=\sum_i\left\|T_i(I-e_ke_k^T\big/k)-L_i\Theta_i\right\|^2 \tag{9}$$

为了得到唯一的解，LTSA 算法对全局坐标T加以中心化和标准化的约束；$TT^T=I$。上式中的重建误差就可以改写成如下形式：

$$E(T)=\sum_i\left\|T_i(I-e_ke_k^T\big/k)-L_i\Theta_i\right\|^2=tr(T\Phi T^T) \tag{10}$$

其中$\Phi=\sum_{i=1}^{N}S_iW_iW_i^TS_i^T$为排列矩阵，$S_i\in R^{m\times k}$是满足$[x_1,x_2,\cdots,x_N]S_i=[x_{i_1},x_{i_2},\cdots,x_{i_k}]$的 0-1 选择矩阵，且

$$W_i=I-\left[e_k\big/\sqrt{k},V_i\right]\left[e_k\big/\sqrt{k},V_i\right]^T \tag{11}$$

这样极小化重建误差的最优解就可通过计算矩阵排列矩阵Φ的最小第 2 个到最小第$d+1$个特征值所对应的特征向量$u_1,u_2,\cdots,u_{d+1}$来获得，即全局坐标为$T=[u_1,u_2,\cdots,u_{d+1}]^T$。

1.3 故障分类方法—K-means 分类器

K-means 算法是一种基于划分的聚类算法，因其快速、易行的特点，在遥感图像、机器学习、统计学、数据挖掘等领域得到了广泛的研究和应用。K-means 算法的目标是将样本数据分成K个类别，使每个类别内的数据点相似性尽可能大，而不同的类别间数据点的相似性尽可能小。其算法[11]步骤

如下所示：

1）设有 n 个样本数据 $X=\{x_1,x_2,\cdots,x_n\}\in R^d$（其中 $x_j=(x_{j1},x_{j2},\cdots,x_{jd})^T$ 为 d 维向量）构成待聚类数据集，需要将其划分到 K 个类别中。首先随机选取 K 个样本数据，每个样本数据代表一个初始聚类中心，从而得到初始聚类中心集合 $C=\{c_1,c_2,\cdots,c_k\}^T$。

2）对剩余的每个样本数据，计算各个数据样本到各聚类中心的欧氏距离，然后将各个数据样本划分到具有最小距离的类别中，用 $d(x_j,c_i)$ 表示样本数据 x_j 与聚类中心 c_i 之间的欧氏距离，计算公式如下：

$$d(x_j,c_i)=\left\|x_j-c_i\right\|_2=\left(\sum_{l=1}^{d}\left|x_{jl}-c_{jl}\right|^2\right)^{1/2} \tag{12}$$

3）重新计算 K 个聚类中心的值，计算公式如下：

$$c_i=\frac{1}{n_i}\sum_{x_j\in c_i}x_j \tag{13}$$

式中，$i=1,2,\cdots,k$，n_i 是类 c_i 中样本数据的个数，x_j 代表属于 c_i 类中的样本数据。

4）检验聚类操作是否需要结束，若满足使目标函数公式 14 最小或保持不变，则迭代结束，否则重复步骤二、三、四。最小化目标函数公式如下：

$$J(X,C)=\sum_{i=1}^{k}\sum_{x_j\in S_i}d(x_j,c_i) \tag{14}$$

式中，S_i 代表第 i 个类别中样本的集合。

2 EMD—LTSA—K-means 滚动轴承故障识别方法

该故障识别方法利用 EMD 分析自适应多分辨率的特点，将轴承振动信号分解成一系列多尺度的内禀模态分量 IMF，从而构造 IMF 向量组，然后利用协方差矩阵求取特征值，获得原始高维特征，再利用流形学习的非线性复杂信息处理能力对高维特征进行维数约简，获取维数低、敏感度高且分类错误率小的主要特征向量，最后将这些主要特征量输入 K-means 分类器进行故障类型的识别。该故障识别方法流程图如图 1 所示，具体步骤如下：

1）对滚动轴承振动信号进行 EMD 分解，得到各 IMF 分量。由于 EMD 的自适应性，所以不同振动信号所求的 IMF 分量个数不尽相同，但分解结果中的前几个 IMF 分量包含了原始信号的主要信息，且轴承振动信号具有强烈的非平稳以及非线性特征，故其 EMD 分量个数均大于 10，因此本文选取 EMD 前 10 个 IMF 分量构成向量组。

2）求取上述 EMD 前 10 个 IMF 分量构成的向量组的协方差矩阵，进而求得 10 个协方差矩阵特征值，构成一个 10 维的原始高维特征向量 $V=[S_{n1},S_{n2},\cdots,S_{n10}]$，其中 n 为振动信号的样本个数。

3）利用流形学习 LTSA 对初步提取的 EMD 特征向量进行维数约简，从而获得低维本质特征。

4）将上述低维本质特征输入 K-means 分类器，从而实现故障类型的识别诊断。

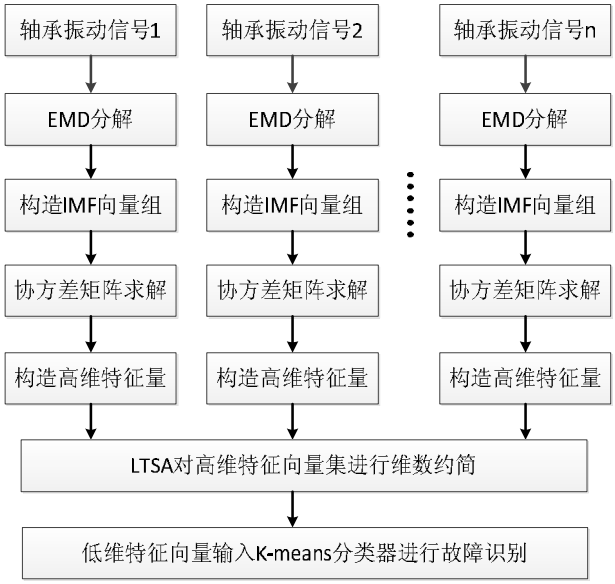


图 1 EMD—LTSA—K-means 滚动轴承故障识别流程图

Fig.1 EMD—LTSA—K-means rolling bearing fault identification flow chart

3 实验验证分析

3.1 实验轴承及故障

本文选取型号为 1210 的调心球轴承为研究对

象，并利用 SG 双金属电刻机对其进行加工，模拟滚动轴承内圈、外圈及滚子的早期微弱点蚀故障。模拟故障类型分别如图 2~4 所示。

图 2　轴承滚子故障

Fig.2 Bearing roller failure

图 3　轴承外圈故障

Fig.3 Bearing outer ring failure

图 4　轴承内圈故障

Fig.4 Bearing inner ring failure

图 5　轴承实际故障

Fig.5　Bearing actual failure

如图 5 所示，通过与轴承在实际工作中所产生的故障相比，除了故障程度较为微弱，本实验所模拟的轴承故障与实际应用中出现的轴承故障基本吻合，因此应用本实验所采集的轴承振动信号来验证本文所提方法的有效性更加具有实际意义。

3.2　实验数据采集

实验设备采用的是 BVT-5 轴承振动测量仪，如图 6 所示。装上实验轴承后，通过径向加载器和轴向加载器固定轴承外圈，保持实验过程中外圈固定不动，内圈随着主轴同步转动。本实验中，轴承径向与轴向施加载荷分别为 300 N 和 100 N，主轴转速为 1800 r/min，对正常、内圈故障、外圈故障及滚子故障四种状态的轴承采集加速度振动信号，采样频率为 5120 Hz，采样点数 3000 个。

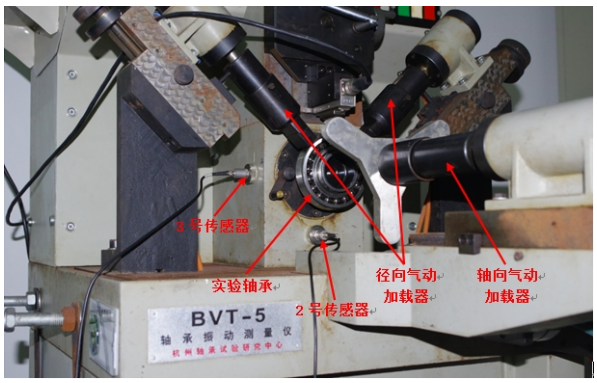


图 6　BVT-5 轴承振动测量仪

Fig.6 BVT-5 bearing vibration measuring instrument

3.3 轴承实验数据分析

1）轴承实验数据的 EMD 分析

利用 MATLAB 软件对上述实验采集的四种状态轴承振动数据进行直接分析，得到 4 种时域波形见图 7。从图中可以看出这四种轴承状态的时域波形图基本相同，不能仅从时域波形图判别轴承的故障状态。

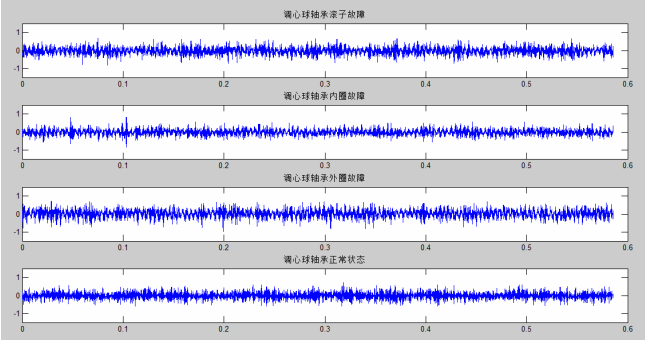
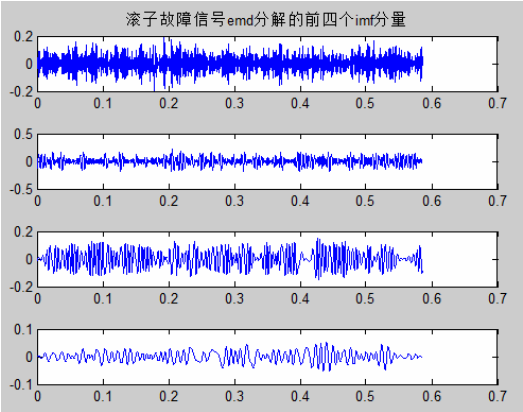


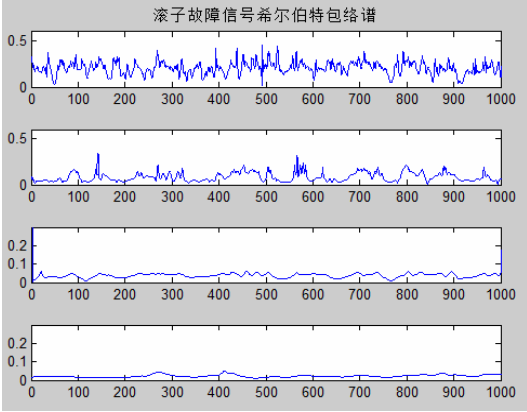
图 7　调心球轴承 4 种状态时域波形图

Fig.7 Self-aligning ball bearings with 4 state-time waveforms

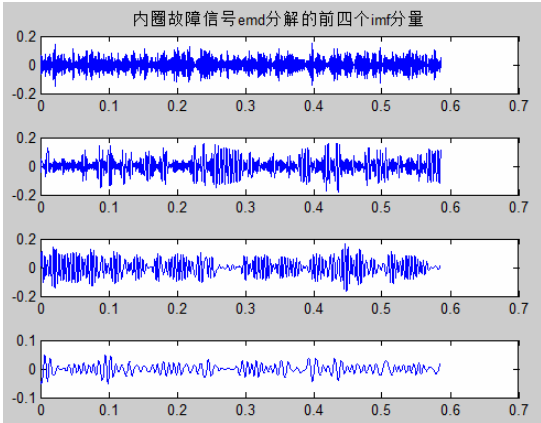
采用 EMD 分别对滚子、内圈和外圈故障时域信号进行分解，对得到的前 4 个 IMF 分量进行希尔伯特变换，获得每个 IMF 的希尔伯特谱，如图 8 所示。从图中可见，IMF 分量的信号波形中无明显的周期性冲击特征，希尔伯特谱中也没有与故障特征频率及倍频相匹配的谱峰，因此，仅仅使用 EMD 分析方法不能判断出实验轴承是否存在故障，更不用说故障类型了。



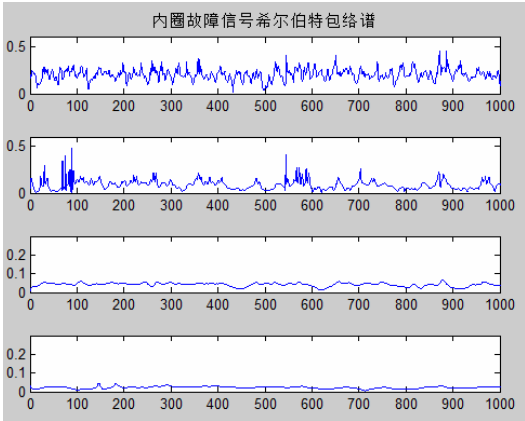
（a）滚子故障信号 IMF 分量



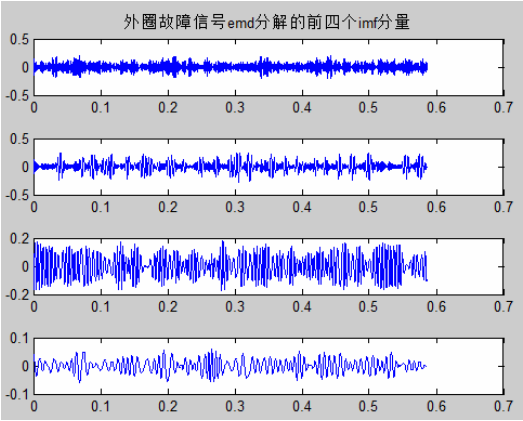
（b）滚子故障信号希尔伯特谱



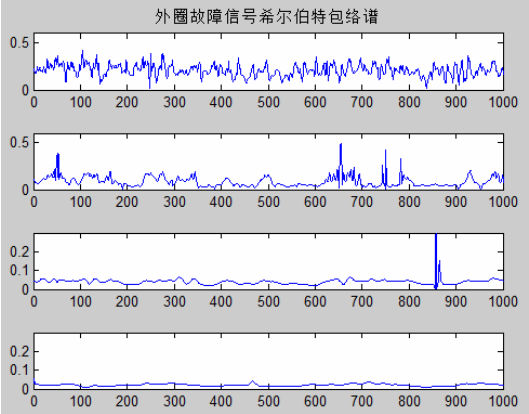
（c）内圈故障信号 IMF 分量



（d）内圈故障信号希尔伯特谱



（e）外圈故障信号 IMF 分量



（f）外圈故障信号希尔伯特谱

图 8　故障轴承振动信号 EMD 分析结果

Fig.8 EMD analysis result of fault bearing vibration signal

2）基于 EMD—LTSA—K-means 方法的滚动轴承故障识别

通过 EMD 对振动信号进行自适应的分解，求出不同尺度频率带的 IMF 分量，构成 IMF 分量向量组，计算其协方差矩阵并求得特征值，作为故障原始高维特征量，再利用 LTSA 进行维数约简后用于故障识别。本实验首先选取调心球轴承滚子、内圈、外圈和正常这 4 种状态下的样本实验振动信号 120 组，每种状态各选 30 组，每组信号含 3000 个采样点。对每组信号分别进行 EMD 分解，提取前 10 个 IMF 分量，从而构成矩阵 $X=R^{300\times10}$，进而提取 X 相应的协方差矩阵的特征值，能得到 10 个特征值，所以 4 种状态下所有样本的原始高维特征集为 $R^{120\times10}$。然后将这些 10 维的特征向量用 LTSA 降到 3 维以提取低维本质特征 $R^{120\times3}$。最后将低维本质特征输入到 K-means 分类器进行故障类型的识别。作为比较，将高维特征集不经过 LTSA 降维而直接输入 K-means 分类器进行故障识别。图 9 和图 10 为是否使用 LTSA 这两种方法的效果图。

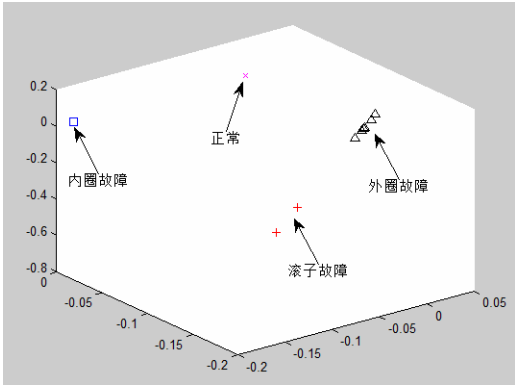


图 9　EMD-LTSA 的故障分类结果

Fig.9 Fault classification results of EMD-LTSA

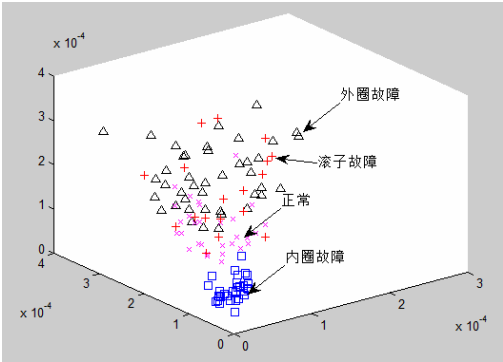


图 10　滚动轴承故障直接分类结果

Fig.10 Rolling bearing fault classification result

从图 9 可以看出，调心球轴承的 4 种状态特征已经完全分开，这说明了经 LTSA 特征融合提取后的低维特征具有非常好的聚类性能，并证明了基于 EMD—LTSA—K-means 的滚动轴承故障诊断方法的有效性。由图 10 则能看出，调心球轴承 4 种状态非常分散，彼此渗透，不能很好的将其识别出来，进一步证明了 LSTA 方法对于故障识别的优势。综上所述，充分证明了本文所提方法对于滚动轴承故障诊断的实际效用。

4　结论

1）本文通过 EMD 方法对调心球轴承振动数据

的分析证明了：当轴承出现故障且故障不规则时，振动信号中的特征频率将不够明显，此时通过提取故障特征频率的方法来进行故障诊断会有一定的局限性。

2）经过 EMD 分解获得的故障特征量直接用于故障分类时，不能将调心球轴承滚子、内圈、外圈和正常这 4 种状态完全分离出来。

3）EMD 对振动信号自适应分解后能够获得包含故障信息的高维特征量，而 LTSA 对高维特征量维数约简后能得到具有内在规律的低维本质特征，将两者结合，可以有效的对调心球轴承进行故障诊断和识别。

综上所述，本文所提出的基于经验模态分解（EMD）和流形学习（LTSA）的滚动轴承故障诊断方法具有一定的实际意义。

参考文献：

[1] 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 A Mathematical Physical & Engineering Sciences, 1998, 454(1971):903-995.

[2] 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.

[3] He G, Ding K, Lin H. Fault feature extraction of rolling element bearings using sparse representation[J]. Journal of Sound & Vibration, 2016, 366:514-527.

[4] Seung H S, Daniel D L. The manifold ways of perception[J]. Science, 2000, (12): 2268-2269.

[5] 阳建宏,徐金梧,杨德斌,等. 基于主流形识别的非线性时间序列降噪方法及其在故障诊断中的应用[J]. 机械工程学报, 2006, 42(8):154-158.

[6] 栗茂林,王孙安,梁霖. 利用非线性流形学习的轴承早期故障特征提取方法[J]. 西安交通大学学报,2010, 44(5):45-49.

[7] 赵洪杰,潘紫微,童靳于,等. 基于相空间重构与非线性流形的滚动轴承复合故障诊断[J]. 振动与冲击, 2013, 32(11):41-45.

[8] 向丹,葛爽. 基于 EMD 样本熵-LLTSA 的故障特征提取方法[J]. 航空动力学报, 2014, 29(7):1535-1542.

[9] Fegeant O. Closed form solutions for the point mobilities of axisymmetrically excited cylindrical shells[J]. Journal of Sound & Vibration, 2001, 243(1):89-115.

[10] Zhang Z Y, Zha H. Principal manifolds and nonlinear dimensionality reduction by local tangent space alignment[J]. SIAM Journal of Scientific Computing, 2004, 26: 313-338.

[11] Reddy D, Jana P K. Initialization for K -means Clustering using Voronoi Diagram[J]. Procedia Technology, 2012, 4(4):395-400.