

基于威布尔分布和支持向量机的 滚动轴承故障诊断方法

姜海燕, 彭涛

(湖南工业大学 电气与信息工程学院, 湖南 株洲 412008)

摘要: 提出了一种基于威布尔分布模型和支持向量机的滚动轴承故障诊断方法。首先对滚动轴承原始振动信号建立威布尔分布模型, 提取其形态参数和尺度参数构建表征轴承运行状态的特征向量, 然后将提取的特征向量输入支持向量机分类器进行故障诊断和识别。分别与基于小波分解和小波包分解特征提取的支持向量机诊断方法进行滚动轴承故障试验仿真比较, 结果表明, 基于威布尔分布模型特征提取的支持向量机诊断方法具有更高的故障识别准确率。

关键词: 威布尔分布模型; 特征提取; 支持向量机; 滚动轴承; 故障诊断

中图分类号: TP206+.3

文献标志码: A

文章编号: 1673-9833(2010)04-0092-06

Fault Diagnosis of Rolling Bearings Based on Weibull Distribution and Support Vector Machine

Jiang Haiyan, Peng Tao

(School of Electrical and Information Engineering, Hunan University of Technology, Zhuzhou Hunan 412008, China)

Abstract: A novel approach to fault diagnosis of rolling bearings based on Weibull distribution model and support vector machine is proposed. Firstly, Weibull distribution model for original vibration signal of rolling bearings is set up, and its shape parameters and scale parameters are extracted. Then the extracted feature vectors are transmitted to the classifier of support vector machine for fault diagnosis and recognition. It is compared to the common feature extraction methods based on wavelet decomposition and wavelet packet decomposition. The experimental simulation results show that the proposed method has the higher accuracy for fault recognition.

Keywords: Weibull distribution model; feature extraction; support vector machine; rolling bearings; fault diagnosis

0 引言

轴承是旋转机械中最常用、也最易发生故障的零部件之一。轴承故障诊断能否成功, 很大程度上取决于故障特征的提取是否有效。常用于轴承故障特征

提取的主要方法有: 基于时域分析及时序建模的特征提取法, 基于常规谱分析的特征提取法, 及基于时-频分析的特征提取法(如小波分解、小波包分解)等^[1-2]。威布尔分布(Weibull distribution)是1939年瑞典人威布尔为描述材料疲劳强度而提出的一种统计模

收稿日期: 2010-03-20

基金项目: 国家自然科学基金资助项目(60774069), 中国博士后科学基金资助项目(20070410462), 湖南省教育厅科技计划基金资助项目(07C005)

通信作者: 姜海燕(1984-), 女, 湖南邵阳人, 湖南工业大学硕士研究生, 主要研究方向为信号处理与模式识别,

E-mail: jianghaiyan1007@163.com

型, 其在产品疲劳寿命和可靠性分析中已获得极其广泛的应用^[3-5], 但在设备状态监测和故障诊断中用于特征提取的研究还不多见。文献[6]通过对感应电机原始振动信号经小波分解的细节系数建立威布尔分布模型, 提取其均值和方差以监测电机运转状态是否正常, 文献[7]则通过对2台风力发电机风速差值建立威布尔分布模型, 提取其形态和尺度参数来监测风速变化, 以达到诊断发电机故障的目的。威布尔分布用于滚动轴承故障特征提取的研究和应用还鲜有报道。

支持向量机 (support vector machine, SVM)^[8]是在统计学习理论上发展起来的一种新的机器学习方法, 它较好地解决了小样本、非线性和高维模式识别、以及传统神经网络模式识别方法中难以解决的网络结构选择和局部极小点等实际问题, 已被成功应用于模式识别与故障诊断^[9-10]。

本文提出一种基于威布尔分布模型的故障特征提取新方法, 在对滚动轴承原始振动信号建立威布尔分布模型基础上, 用威布尔分布的形态参数和尺度参数构建特征向量以表征轴承运行状态, 采用支持向量机分类器对提取的特征向量进行模式分类和故障识别, 并通过试验仿真验证所提方法的有效性。

1 威布尔分布

1.1 威布尔分布模型

对于数据序列 $\{x\}$, 其具有2个参数的Weibull分布函数定义为:

$$F(x) = 1 - \exp\left[-(x/\eta)^\beta\right], \quad (1)$$

式中: $\beta > 0$, $\eta > 0$ 分别为形态参数和尺度参数;
 $x > 0$ 为分布变量。

该模型的分布密度函数为:

$$f(x) = (\beta/\eta)(x/\eta)^{\beta-1} \exp\left[-(x/\eta)^\beta\right]. \quad (2)$$

设威布尔变换为:

$$y = \ln(-\ln R(x)), t = \ln x, \quad (3)$$

式中: $R(x) = \exp\left[-(x/\eta)^\beta\right]$ 。

则威布尔分布函数可写成:

$$y = y(t) = \beta(t - \ln \eta), \quad (4)$$

式(4)称为通过威布尔变换后的威布尔概率纸 (Weibull plotting paper)^[3], 即在普通坐标纸上绘制函数 $y(t)$, 若 $y(t)$ 大致为一条直线, 则认为数据序列 $\{x\}$ 符合威布尔分布模型。

1.2 威布尔分布模型的参数估计

形态参数 β 和尺度参数 η 的估计方法有: 图估计法、最佳线性无偏估计法、最佳线性不变估计法、最

大似然估计法以及最小二乘法等。本文采用最大似然估计法, 设一组随机数据序列为 $\{x_1, x_2, \dots, x_n\}$, 且符合威布尔分布模型, 则其似然函数为:

$$L(x_1, x_2, \dots, x_n; \beta, \eta) = \prod_{i=1}^n (\beta/\eta) x_i^{\beta-1} \exp\left[-(x_i/\eta)^\beta\right], \quad (5)$$

由极大似然估计法可得:

$$\begin{cases} \frac{\partial \ln L}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^n \ln x_i - \frac{1}{\eta} \sum_{i=1}^n x_i^\beta \ln x_i = 0, \\ \frac{\partial \ln L}{\partial \eta} = -\frac{n}{\eta} + \frac{1}{\eta^2} \sum_{i=1}^n x_i^\beta = 0, \end{cases} \quad (6)$$

解方程组(6)可得:

$$\sum_{i=1}^n x_i^\beta \ln x_i / \sum_{i=1}^n x_i^\beta - \frac{1}{\beta} = \frac{1}{n} \sum_{i=1}^n \ln x_i, \quad (7)$$

$$\eta = \frac{1}{n} \sum_{i=1}^n x_i^{\hat{\beta}}, \quad (8)$$

式(7)为超越方程, 可用牛顿迭代法^[11]求得估计值 $\hat{\beta}$, 将 $\hat{\beta}$ 带入式(8)中, 即可得尺度参数的估计值 $\hat{\eta}$ 。

3 支持向量机

对给定的2类训练数据样本集:

$$S = \{z_i, y_i\}_{i=1}^n, \quad (9)$$

式中: $z_i \in \mathbf{R}^N$, N 为输入维数; $y_i \in \{-1, 1\}$; n 为训练样本数。

支持向量机的目的是找到一个最优超平面:

$$\begin{cases} \mathbf{w}^\top z_i + \mathbf{b} \geq 1, & \text{对 } y_i = +1; \\ \mathbf{w}^\top z_i + \mathbf{b} \leq -1, & \text{对 } y_i = -1. \end{cases} \quad (10)$$

式中: \mathbf{w} 为最优超平面的权系数向量; \mathbf{b} 为偏置向量。

为了找到全局最优超平面, 可利用拉格朗日函数

将问题(10)转化为求解在约束条件 $\sum_{i=1}^n a_i y_i = 0$ (a_i 为拉格朗日乘子, $a_i \geq 0$)下的函数最大值问题:

$$W(\mathbf{a}) = \sum_{i=1}^n a_i - \frac{1}{2} \sum_{j=1}^n \sum_{i=1}^n a_i a_j y_i y_j z_i^\top z_j. \quad (11)$$

对于线性不可分问题, 通常采取先经某种内积函数 $K(z_i, z_j)$ 进行适当的非线性变换后, 再进行线性分类的方式来解决。此时, 式(11)在约束条件 $\sum_{i=1}^n a_i y_i = 0$, $0 \leq a_i \leq C$ (C 为惩罚因子)下变为:

$$W(\mathbf{a}) = \sum_{i=1}^n a_i - \frac{1}{2} \sum_{j=1}^n \sum_{i=1}^n a_i a_j y_i y_j K(z_i, z_j), \quad (12)$$

按式(12)求出最优拉格朗日乘子 a_i^* 后, 支持向量机分类函数可表示为:

$$d(z) = \sum_{i=1}^n y_i a_i^* K(z, z_i) + b^* \quad (13)$$

式中: b^* 为最优偏置向量;

$K(z, z_i) = \exp\left(-\frac{(z - z_i)^2}{\sigma^2}\right)$ 为径向基核函数, 其中 σ 为核函数宽度。

4 故障诊断实例

4.1 试验

实验数据来自美国华盛顿天主教大学 (Case Western Reserve University, 即凯斯西储大学) 电气工程实验室^[12]。试验装置如图 1 所示, 试验中通过对功率计的控制获得期望的扭矩载荷, 振动加速度传感器垂直固定在感应电机输出轴的机壳上, 测试轴承型号为 6205-2RS JEM SKF, 采样频率为 48 kHz。在不同电机负载/转速工况条件 (0, 1, 2 和 3 马力)、4 种不同故障类型 (正常状态、内环故障、滚动体故障、外环故障) 和 3 种不同故障状态 (故障点直径分别为 0.177 8 mm, 0.355 6 mm, 0.533 4 mm) 下, 记录振动加速度传感器的信号数据, 共获得 16 种运行状态的数据, 将每种运行状态数据以 1 024 点作为一个样本进行分析。

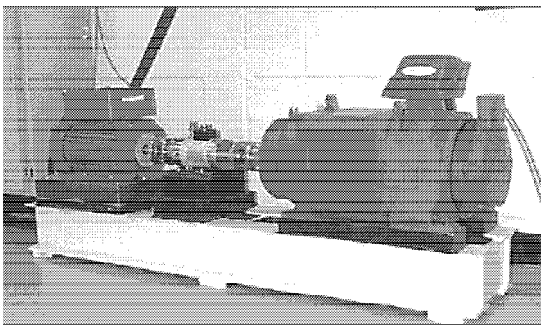


图 1 实验装置

Fig. 1 Experimental device

4.2 基于威布尔分布模型的特征提取

先对采集的滚动轴承原始振动信号进行去噪等预处理, 对预处理后的信号建立威布尔分布模型, 用极大似然法估计威布尔分布模型的形态和尺度参数, 再提取这 2 个参数构建表征轴承运行状态的特征向量。

基于威布尔分布模型的故障特征提取步骤如下:

1) 建模。由威布尔分布模型可知, 数据序列必须满足 $x_i > 0$, 一般情况下获得的滚动轴承振动信号是在 $x=0$ 附近波动的, 因而须对原始信号进行如下处理:

$$x'_i = x_i - \min(x_i) + \delta, \quad (14)$$

式中: δ 为一经验值, 一般取 $\delta \leq 0.1 \min(|x_i|)$ 。

4 种不同故障类型的原始信号和经预处理后信号的时域波形如图 2 和 3 所示。

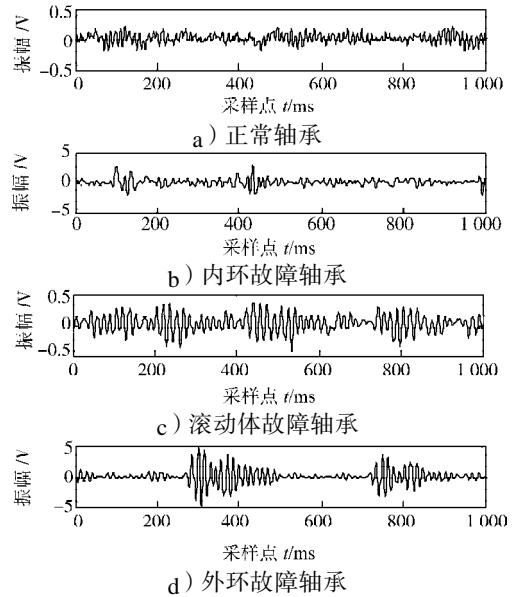


图 2 4 种不同故障原始信号的时域波形

Fig. 2 The time-domain waveform of the primary four fault signals

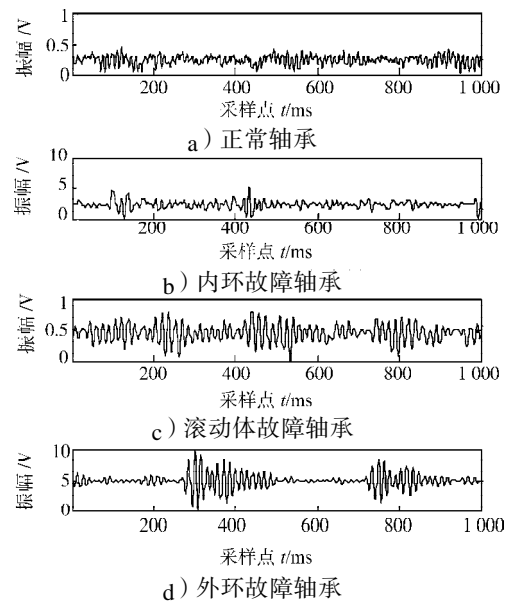
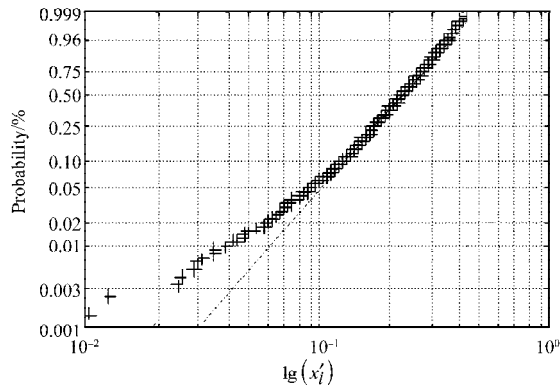


图 3 4 种不同故障预处理后信号的时域波形

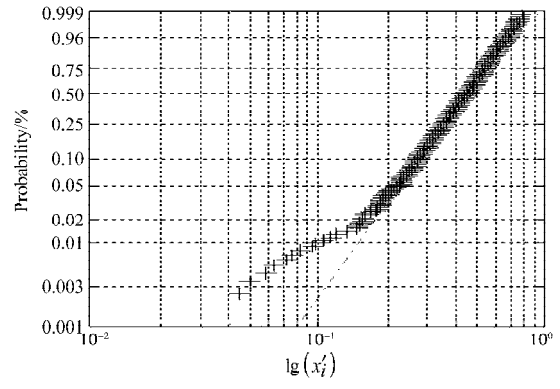
Fig. 3 The time-domain waveform of the pre-processed four fault signals

据预处理后的信号建立威布尔分布模型, 据式 (3) 和 (4), 画出对应 4 种不同故障类型滚动轴承信号的威布尔概率纸图, 如图 4。可见, 各图形大致呈一条直线, 即样本数据基本服从威布尔分布模型。

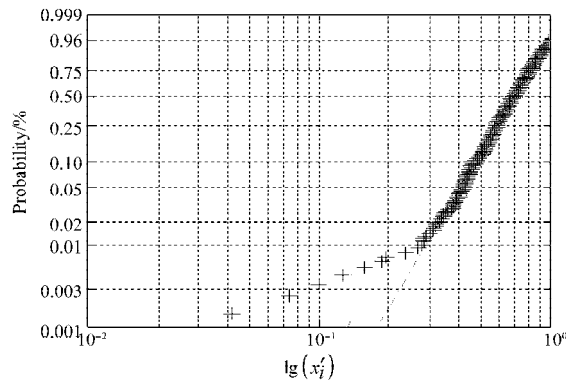
如对式 (14) 中 δ 选取不同值, 可获得与图 4 中某一类型故障轴承信号相对应的威布尔概率纸图, 图 5 为滚动体故障轴承在 δ 分别取 0.001, 0.05, 0.1, 0.2 时的威布尔概率纸图。可见, 无论 δ 取值如何, 经式 (14) 处理后的数据均基本服从威布尔分布模型。



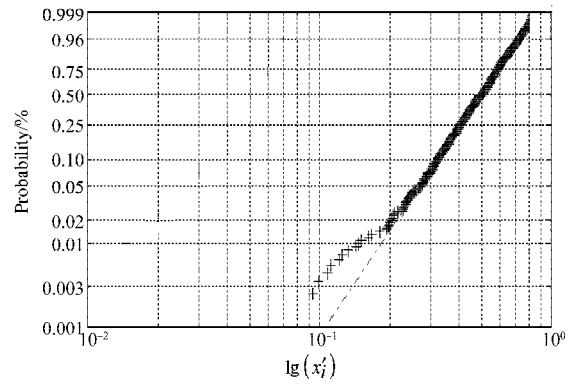
a) 正常轴承



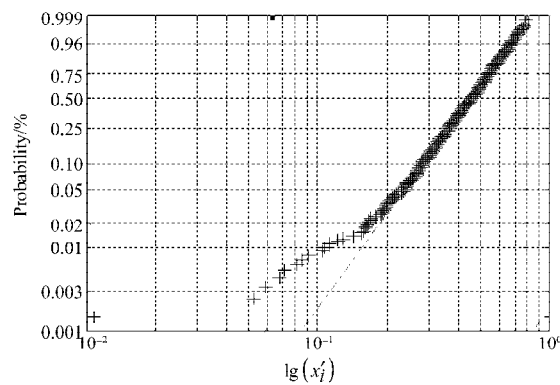
a) $\delta = 0.001$



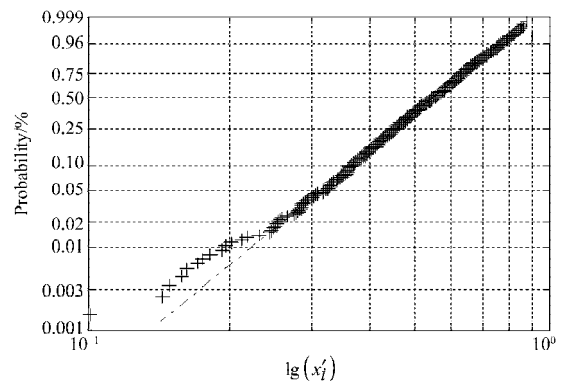
b) 内环故障轴承



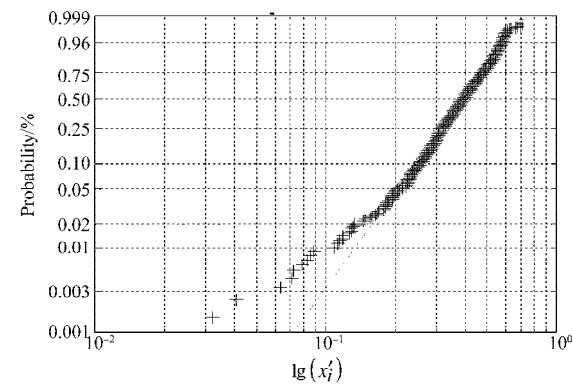
b) $\delta = 0.05$



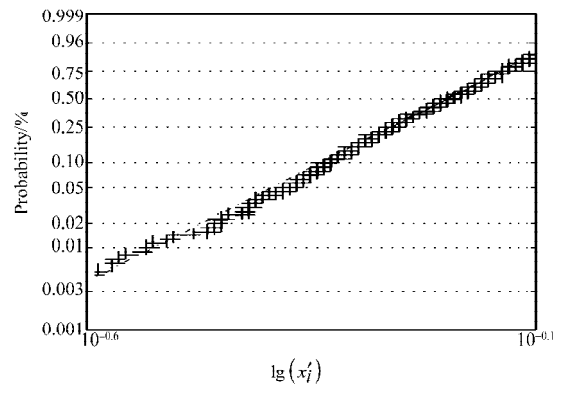
c) 滚动体故障轴承



c) $\delta = 0.1$



d) 外环故障轴承



d) $\delta = 0.2$

图4 4种不同故障信号的威布尔概率纸图

Fig. 4 The Weibull probability paper chart of four fault signals

图5 不同 δ 值时滚动体故障信号的威布尔概率纸图

Fig. 5 The Weibull probability paper chart of rolling elements at different δ

2) 估计参数。由所建威布尔分布模型, 据式(7)和(8)估计出 $\hat{\beta}$ 和 $\hat{\eta}$ 。4种不同故障类型的威布尔分布模型的形态和尺度参数估计值如图6~7所示。

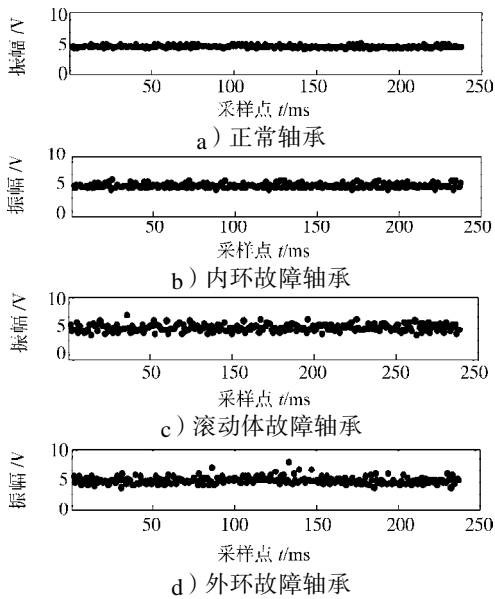


图6 4种不同故障形态参数估计值

Fig. 6 The estimated values for shape parameters of four faults

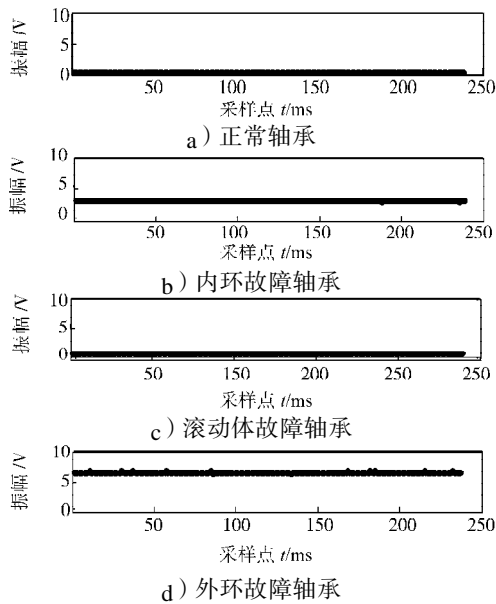


图7 4种不同故障尺度参数估计值

Fig. 7 The estimated values of scale parameters of four faults

3) 构建特征向量。由图6和图7可见, 不同故障类型的威布尔分布模型的形态参数和尺度参数区别较大, 能够较好地刻画轴承运行的状态特性, 因而可以用作反应轴承运行状态的特征向量。

用估计出的形态参数和尺度参数构建特征向量 $Y_i = [\hat{\beta}^i, \hat{\eta}^i]$, 其中 $i = 1, 2, \dots, m$, 它表示第 i 个样本,

m 为样本总数。

4.3 基于支持向量机的状态分类与故障识别

从4.2节第三步得到的特征向量中, 随机抽取16种运行状态下各20个样本, 建立支持向量机的训练样本, 其余样本作为测试样本。

基于支持向量机的分类识别步骤如下:

1) 用4.2节方法提取的特征向量 Y_i 构造出训练样本集 $\{(Y_l^{\text{train}}, y_l)\}$ 和测试样本集 $\{(Y_j^{\text{test}}, y_j)\}$, 其中 $l=1, 2, \dots, 20, j=21, 22, \dots, m, y_l, y_j \in \{1, 2, 3, 4\}$ 为分类器目标输出值(1标记正常状态, 2标记内环故障, 3标记滚动体故障, 4标记外环故障)。

2) 用训练样本集的特征向量 Y_l^{train} 训练支持向量机分类器。

由于SVM分类器主要是进行2类分类, 而本实验中轴承的故障类型不止2类, 于是须要构造多个支持向量机, 每2类间训练1个分类器, 因此对于1个 k 类分类问题, 将有 $k(k-1)/2$ 个如式(13)所示的分类函数^[14]。

3) 用训练好的支持向量机多类分类器对测试样本进行故障识别。

为了验证基于威布尔分布模型的故障特征提取方法的有效性, 将所提方法与传统的小波分解和小波包分解特征提取方法进行比较, 都采用支持向量机进行状态分类, 识别结果见表1。

表1中的 A, B, C 分别表示以小波分解后各频带信号能量、小波包分解后各频带信号能量、威布尔分布模型形态参数和尺度参数作为轴承运行状态特征向量时的SVM识别结果。由表1可见, 基于威布尔分布模型特征提取的SVM故障识别方法, 不仅识别正确率比前两者都高, 且计算简单, 容易实现故障识别。

表1中的 $C_1 \sim C_4$ 分别对应式(14)中 δ 取值为0.001, 0.05, 0.1, 0.2时的SVM识别结果。表1结果说明 δ 取值不同, 对SVM识别效果不会产生多大影响, 也进一步验证了数据预处理方法的合理性。

5 结论

1) 将威布尔分布首次应用于滚动轴承故障的特征提取, 以威布尔分布模型的形态参数和尺度参数构建的特征向量能有效地表征滚动轴承运行状态的特性。

2) 试验表明, 以威布尔分布的2个参数作为特征向量的SVM诊断方法, 相比以小波分解和小波包分解后各频带信号能量作为特征向量的SVM诊断方法, 具有更高的识别准确率。

3) 基于威布尔分布模型和SVM的诊断方法能够准确有效地对滚动轴承故障类型进行分类和识别, 为滚动轴承故障诊断提供了一种新的有效方法。

表1 基于3种不同特征提取方法的SVM故障识别结果

Table 1 The results of SVM fault recognition based on three different feature extraction methods

样本号	负载 / 马力	转速 /(r · min ⁻¹)	故障点直径 /mm	故障类型	SVM 分类器的识别率/%							
					不同故障特征提取法			不同 δ 值				
					A	B	C	C ₁	C ₂	C ₃	C ₄	
1	0	1 797		正常	100	100	100	100	100	100	100	100
2	1	1 772		正常	100	100	100	100	100	100	100	100
3	2	1 750		正常	100	100	100	100	100	100	100	100
4	3	1 730		正常	100	100	100	100	100	100	100	100
5				内环	97.71	99.08	100	100	100	100	100	100
6	0	1 797	0.177 8	滚动体	100	90.87	100	100	100	100	100	99.54
7				外环	100	95.85	99.54	99.54	99.54	99.54	99.54	99.54
8				内环	99.34	91.85	100	100	100	99.78	99.78	
9	1	1 772	0.177 8	滚动体	99.34	93.19	99.78	100	100	98.46	98.46	
10				外环	100	97.58	100	100	100	100	100	100
11				内环	92.54	93.64	96.27	96.49	96.27	96.27	96.03	
12	2	1 750	0.355 6	滚动体	78.24	29.89	80.22	80.44	80.44	80.22	80	
13				外环	83.52	81.32	97.14	97.14	97.14	97.14	97.36	
14				内环	92.41	81.18	100	100	100	100	100	
15	3	1 730	0.533 4	滚动体	80.78	81.10	81.46	81.46	81.46	81.46	81.46	
16				外环	85.02	84.25	100	100	100	100	100	

参考文献:

- [1] Samantha B, Al-Balushi K R. Artificial Neural Network Based Fault Diagnostic of Rolling Element Bearing Using Time Domain Features[J]. Mechanical System and Signal Process, 2003, 17(2): 317-328.
- [2] Bojan K, Zdravko K. A Noise Robust Feature Extraction Algorithm Using Joint Wavelet Packet Subband Decomposition and AR Modeling of Speech Signals[J]. Signal Processing, 2007, 87(6): 1202-1223.
- [3] 蒋仁言. 威布尔模型族-特性、参数估计和应用[M]. 北京: 科学出版社, 1998.
Jiang Renyan. Weibull Models-Characteristics, Parameters Estimation and Applications[M]. Beijing: Science Press, 1998.
- [4] 崔卫民, 薛红军, 喻天翔, 等. 试验数据服从 Weibull 分布时可靠性试验最少试件数的确定[J]. 机械工程学报, 2008, 44(1): 51-55.
Cui Weimin, Xue Hongjun, Yu Tianxiang, et al. Determination of Sample Size for Weibull Distribution in Structural Reliability Tests[J]. Chinese Journal of Mechanical Engineering, 2008, 44(1): 51-55.
- [5] Heng A, Zhang S, Tan A C C, et al. Rotating Machinery Prognostics: State of the Art, Challenges and Opportunities [J]. Mechanical Systems and Signal Processing, 2009, 23(3): 724-739.
- [6] Jiménez G A, Muñoz A O, Duarte-Mermoud M A. Fault Detection in Induction Motors Using Hilbert and Wavelet Transforms [J]. Electrical Engineering, 2007, 89(3): 205-220.
- [7] Tai-Her Y, Li W. A Study on Generator Capacity for Wind Turbines under Various Tower Heights and Rated Wind Speeds Using Weibull Distribution[J]. IEEE Transaction on Energy Conversion, 2008, 23(2): 592-601.
- [8] Burges C J C. A Tutorial on Support Vector Machine for Pattern Recognition[J]. Data Mining and Knowledge Discovery, 1998, 2(2): 121-167.
- [9] Guyon I, Weston J, Barnhill S, et al. Gene Selection for Cancer Classification Using Support Vector Machines[J]. Machine Learning, 2002, 46(1/3): 389-422.
- [10] 杨正友, 彭涛. 基于振动信号分析和支持向量机的滚动轴承故障诊断[J]. 湖南工业大学学报, 2009, 23(1): 96-99.
Yang Zhengyou, Peng Tao. Fault Diagnosis of Rolling Element Bearing Based on Vibration Signal Analysis and Support Vector Machine[J]. Journal of Hunan University of Technology, 2009, 23(1): 96-99.
- [11] 李剑, 张群会. 两参数威布尔分布中参数的极大似然估计的迭代求解方法[J]. 机械强度, 1994, 16(3): 67-68.
Li Jian, Zhang Qunhui. A Iterative Method for the Maximum Likelihood Estimate of the Two Parameters in the Two-Parameters Weibull Distribution[J]. Journal of Mechanical Strength, 1994, 16(3): 67-68.
- [12] Case Western Reserve University Bearing Data Center. Bearing Data Center Fault Test Data[EB/OL]. [2009-10-01]. Http://www.eecs.case.edu/laboratory/bearing.

(责任编辑: 李玉珍)