

Recognition and Analysis of Mechanical Vibration Signal Based on Support Vector Machine

Chunhong Zhang¹, Yong Hu²

¹Beijing Zhongguancun Yongfeng Industrial Base Development Co., Ltd., Beijing

²Beijing Huanyushengwang Intelligent Technology Co., Ltd., Beijing

Email: 382700893@qq.com, yonghu_iacas@163.com

Received: Jun. 14th, 2019; accepted: Jul. 1st, 2019; published: Jul. 8th, 2019

Abstract

With the continuous development of modern science and technology and industrialization, mechanical equipment has become more and more complex, and the degree of integration and automation has been constantly improved, and the links between components have become more closely. Therefore, when the equipment is in operation, if a small part of the fault occurs, these changes will cause the paralysis of the entire equipment, which will not only cause huge economic losses, but also may cause casualties. If we can make use of the health diagnosis technology of machinery and equipment and understand its health status as soon as possible, we can reduce the incidence of accidents, so that enterprises can obtain greater benefits. In this paper, we collect signals from different mechanical vibration states, extract the frequency domain characteristics of the signals, and make principal component analysis of the characteristics. Finally, we use support vector machine to recognize various behavior signals, so as to realize the fast and intelligent analysis of vibration signals. The experimental results show that four channels (CH1 - CH4) and six kinds of vibration signals can be basically distinguished by extracting frequency domain features and machine learning algorithm.

Keywords

Vibration Signal, Pattern Recognition, Support Vector Machine, Principal Component Analysis, Fourier Transform

基于支持向量机的机械振动信号识别与分析

张春红¹, 胡 勇²

¹北京中关村永丰产业基地发展有限公司, 北京

²北京寰宇声望智能科技有限公司, 北京

Email: 382700893@qq.com, yonghu_iacas@163.com

收稿日期: 2019年6月14日; 录用日期: 2019年7月1日; 发布日期: 2019年7月8日

摘要

随着现代科学技术和工业化的不断发展, 机械设备变得越来越复杂, 且综合程度和自动化程度得到不断的提高, 各部件之间的联系则变得更加紧密。所以当设备在运行的过程中, 如若某一部分发生微小的故障, 这些变化会引起整台设备的瘫痪, 这不仅会造成巨大的经济损失, 还可能造成人员的伤亡。如能利用对机械设备的健康诊断技术, 尽早的了解其健康状况, 就可以降低事故的发生率, 从而使得企业能够获得较大的收益。本文通过对不同机械振动状态进行信号采集, 通过提取信号频域特征, 并对该特征进行主成分分析, 最后利用支持向量机对各种行为信号进行识别, 实现振动信号的快速、智能分析。实验结果表明, 通过提取频域特征结合机器学习算法可以将采集的四个通道(CH1~CH4)、6类振动信号基本区分开。

关键词

振动信号, 模式识别, 支持向量机, 主成分分析, 傅立叶变换

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着现代工业和科学技术的发展以及自动化程度的进一步提高, 机械正朝着大型化、高速化、连续化、集中化、自动化方向发展, 生产系统中各设备之间联系也越来越紧密。由于各种随机因素的影响, 这些设备不可避免地会出现一些故障现象, 即降低或失去一定的功能; 而设备一旦出现故障[1]就可能引起连锁反应, 导致整个生产过程无法正常工作, 造成巨大的经济损失, 甚至还会引起严重的灾难性人员伤亡事故。本文对故障信号识别进行研究, 提取故障信号的各种特征, 构建数据库与分类模型, 对故障信号进行分类识别。对实验中振动行为信号样本采用监督模式识别方法进行分类: 利用一定数量的样本作为训练集, 首先进行特征提取, 选择合适分类器模型, 训练出一个分类超平面。对预测样本, 经过相同特征提取算法, 利用已经训练好的分类超平面判断行为类别, 实现目标行为的智能、自动分析。

2. 振动信号目标特征提取及识别算法

对振动信号进行特征提取是实现故障诊断的前提条件, 也是实现故障诊断的关键技术。实际工程中的机械结构往往较复杂, 且零部件比较繁多, 所以实际中采集到的动态信号是整个系统各个部件振动的共同结果, 且在传输的过程中, 也会受到传递途径的影响, 这些使得采集到信号更加复杂, 包含很多的无用信息。故在诊断过程中, 需要对采集到的信号进行分析, 提取相关的信号特征, 然后分析其是否有相关的故障征兆, 最后利用该征兆来进行故障诊断。

针对机械振动信号特征提取问题, 国内学者提出了基于独立分量分析方法[2]、基于混沌振子的方法[3]、这些方法考虑单故障状态下局部零件的早期故障诊断问题。目前, 最常用的信号特征提取方法是从

时域角度直接分析信号波形进行特征提取[4], 该方法思路简单易于实现, 但是由于信号采集现场环境往往比较复杂, 诸如风、雨、动物等环境因素产生的波形与许多入侵行为波形极为相似, 降低了系统对时域波形的判断。傅里叶分析[5]是信号处理的经典方法, 对机械振动信号进行傅里叶变换, 利用频谱或功率谱提取其频域特征, 也是目前较为常见的方法, 该方法在振动信号处理中效果较为理想。天津大学曲志刚等人提出使用小波[6]或小波包变换对光纤振动信号进行多层次分解, 提取其在子频带的能量分布特征, 该方法在石油、天然气管道安全检测系统信号处理中取得了较好的识别效果。

机器学习作为多元交叉学科, 综合了统计学、逼近论、算法时间复杂度等多门学科, 常用的机器学习算法有 K 最近邻(K-Nearest Neighbor, KNN)、朴素贝叶斯、支持向量机(Support Vector Machine, SVM)等。湖南大学程军圣提取轴承信号时域特征和能量参数作为神经网络[7]的输入, 有效识别故障类型。东南大学万遂人等则借鉴了语音信号识别的处理方法, 将系统采集到的振动信号当成声音信号进行处理, 利用隐马尔可夫模型(HMM)进行特征识别[8]。

2.1. Fourier 分析

利用 Fourier 分析[5]实现振动信号时域和频域之间的相互转换, 可以将原时域信号的研究转换为频域上的 Fourier 系数的研究。在信号处理领域, Fourier 变换发挥了重要的作用, 具有里程碑式的意义, 它被看作是信号时域与频域之间的桥梁。对于信号 $x(t)$, 其连续 Fourier 变换为:

$$F(\omega) = \int_{-\infty}^{+\infty} x(t)e^{-i\omega t} dt \quad (1)$$

其逆变换为:

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega)e^{i\omega t} d\omega \quad (2)$$

在实际应用时, 信号在时域和频域中往往是离散的, 因此常用到的是离散 Fourier 变换(DFT), 考虑到运算速度与系统消耗方面, 快速 Fourier 变换(FFT)被广泛应用于信号的频域分析中。

2.2. 小波分析

小波分析[6]最早是由法国科学家 Grossman 和 Morlet 于 80 年代提出的, 它与傅里叶分析相比, 有突破性的进展, 它具有时频域局部化的特点, 能够使信号的细节部分特到显示, 从而能够有效地提取特征信息。目前, 小波分析在很多学科领域有着广泛的应用, 如在故障诊断, 图像处理、模式识别等。基于小波分析的故障诊断方法是基于信号处理故障诊断方法的一种, 它不需要建立相应的数学模型, 也不需要先验知识, 而且对输入信号的要求也比较低。同时, 这种方法的计算量比较小, 具有较强的克服噪声能力, 而且灵敏度也很高, 具有较大的应用空间。其基本理论如下:

小波函数 $\psi(t)$, 它局部地存在于时域(空域)轴上, 其波形可在特定的位置上发生变化: 当对低频信号分析时, 小波函数可以放大 a 倍; 当对高频信号分析时, 小波函数可以缩小 a 倍。小波函数 $\psi_a(t)$ 是具有震荡特性、能迅速衰减到零的一类函数, 可表示为:

$$\psi_{ar}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-\tau}{a}\right) \quad (3)$$

其函数图像如图 1 所示:

设 $x(t) \in L^2(\mathbb{R})$ 是平方可积函数, $\psi(t)$ 是基小波或“母小波函数”, 则

$$WT_x(a, \tau) = \frac{1}{\sqrt{a}} \int x(t) \psi^*\left(\frac{t-\tau}{a}\right) dt \quad (4)$$

称为 $x(t)$ 的连续小波变换(CWT), 其内积形式为 $WT_x(a, \tau) = \langle x(t), \psi_{a\tau}(t) \rangle$ 。其中 $a (> 0)$ 称为尺度因子, τ 反映了小波函数在变换中的位移。常用的小波函数有 Haar 小波、Daubechies (dbN) 小波系、Morlet 小波、Mexican Hat 小波、Meyer 小波等。目前关于小波基函数的选择还没有统一标准, 但是基于小波光滑性、正则性、对称性等特点本实验选取 db3 系列小波对原始时域信号进行 k 级多尺度分解。

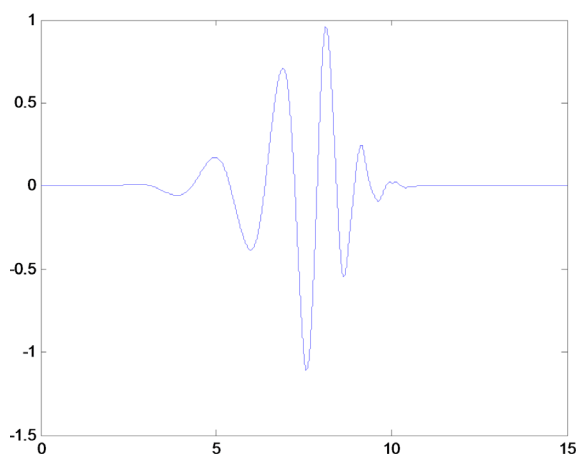


Figure 1. Wavelet function
图 1. 小波函数

2.3. 主成分分析

主成分分析[9] (Principal component analysis: PCA)是一种统计方法。通过正交变换将一组可能存在相关性的变量转换为一组线性不相关的变量, 转换后的这组变量叫主成分。其基本思路是寻找高维数据在低维空间重构误差最小意义下最优的投影子空间, 从而用较少数目的特征向量刻画样本的全局信息, 最终达到降低原始空间维数的目的。设样本矩阵为 $X_{m \times n}$ (m 个样本, n 个特征), 首先对样本矩阵 X 均值中心化, PCA 希望找到一个方向 w , 使得样本投影到方向 w 上方差最大(主方向), 优化目标为:

$$\hat{w} = \arg \min_{\|w\|^2=1} w^T X^T X w \quad (5)$$

提取前 k 个主方向构成载荷矩阵 $W = (w_1, \dots, w_k)$, 则对原始样本特征提取后的结果为: $X_{new} = XW$ 。

2.4. 支持向量机

支持向量机[10] (Support Vector Machine, SVM)是一种常用的定性分析算法。它是 Vapnik 等人在统计学习理论和结构风险最小化(Structural Risk Minimization, SRM)原则的基础上, 建立起来的一种通用的机器学习算法。它根据有限的样本信息, 在模型的复杂性和学习能力之间寻求最佳折衷, 以期获得最好的推广能力。其基本原理如下:

假设训练样本集 $(x_1, y_1), \dots, (x_n, y_n)$, 其中 $x_i \in X$ 为第 i 个样本, y_i 是 x_i 的标记, 即第 i 个样本所对应的类别信息。在二分类问题中将其记做 $y_i \in \{1, -1\}$ 。如果训练样本线性可分或近似线性可分, 则一定可以找到一个最优超平面将两类样本正确分开, 且使得两类样本到超平面的间隔最大。设该最优超平面为:

$$w * x + b = 0 \quad (6)$$

超平面将两类样本正确分开意味着:

$$y_i [w * x_i + b] \geq 1 - \xi_i, i = 1, 2, \dots, n, \xi_i \geq 0 \quad (7)$$

其中 ξ_i 是引入的松弛变量, 有: $\xi_i \geq 0, i = 1, 2, \dots, l$

即求解优化问题

$$\hat{w} = \arg \min_w \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \quad (8)$$

求解最优分类面为:

$$f(x) = \text{sign} \left(\sum_{i=1}^n y_i \alpha_i^* (x_i * x) + b^* \right) \quad (9)$$

3. 振动信号目标识别实验

3.1. 样本说明

实验测试数据包含六种状态: Healthy_{20Hz}、Healthy_{40Hz}、YangCo_{20Hz}、Yangpa_{40Hz}、YangUn_{20Hz}、YangUn_{40Hz}。每一类数据采集 60 秒左右, 采样频率 $F_s = 10K$ Hz, 包含四个通道(CH1、CH2、CH3、CH4), 测得了 10 组数据。四个通道中六类典型样本的时域表示如图 2~图 5。从图中可以看出, 四个通道中六类信号分布杂乱, 从时域上很难将其区分开来。

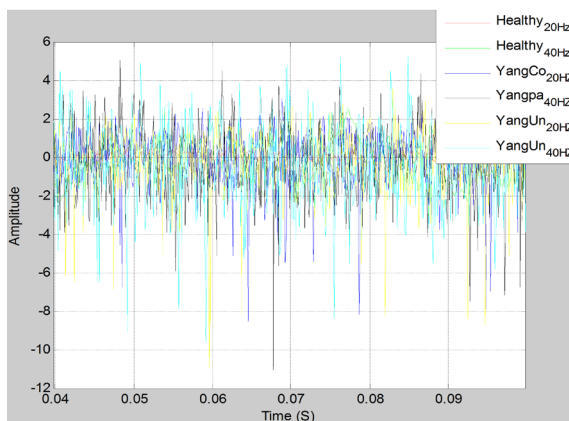


Figure 2. CH1 six kinds of time domain samples

图 2. CH1 六类时域样本

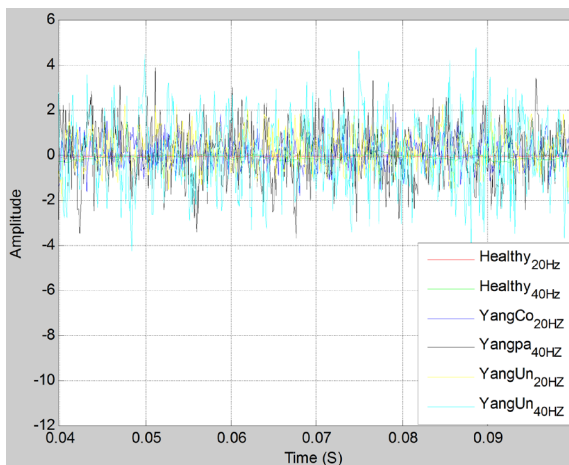


Figure 3. CH2 six kinds of time domain samples

图 3. CH2 六类时域样本

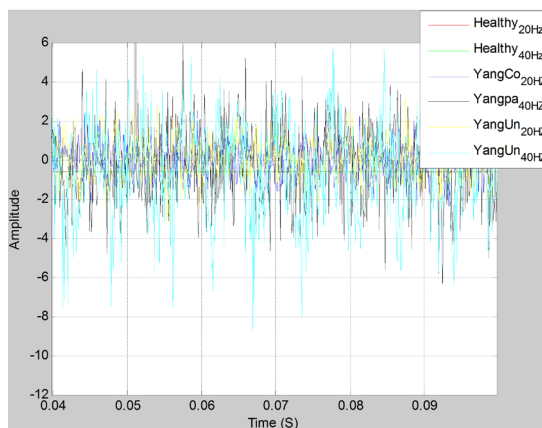


Figure 4. CH3 six kinds of time domain samples

图 4. CH3 六类时域样本

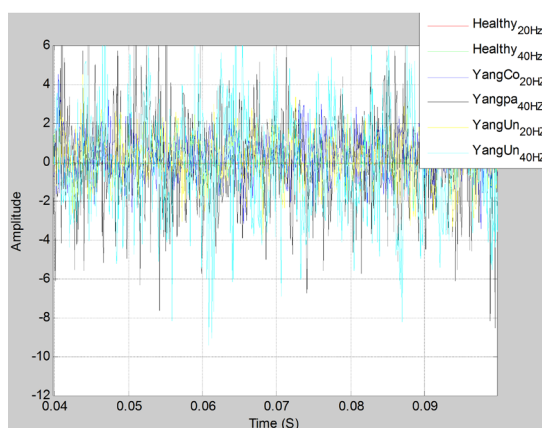


Figure 5. CH4 six kinds of time domain samples

图 5. CH4 六类时域样本

对数据进行连续小波变换(CWT)分析提取特征：基于小波光滑性、正则性、对称性等特点本实验选取 db3 系列小波对原始时域信号进行 32 级多尺度分解，以一通道信号(CH1)为例(图 6~图 11)，从图上分析，同一通道中不同类别信号在小波域上具有一定差异性。

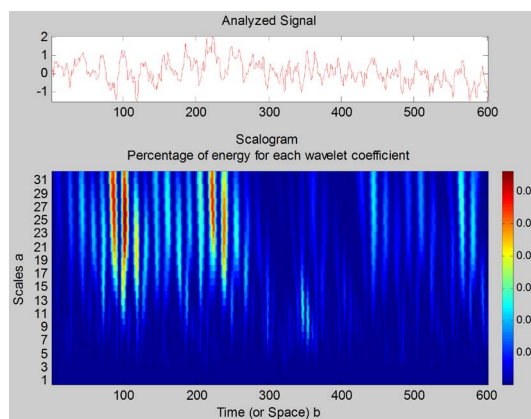


Figure 6. Healthy_{20Hz} CWT chart

图 6. Healthy_{20Hz} CWT 效果图

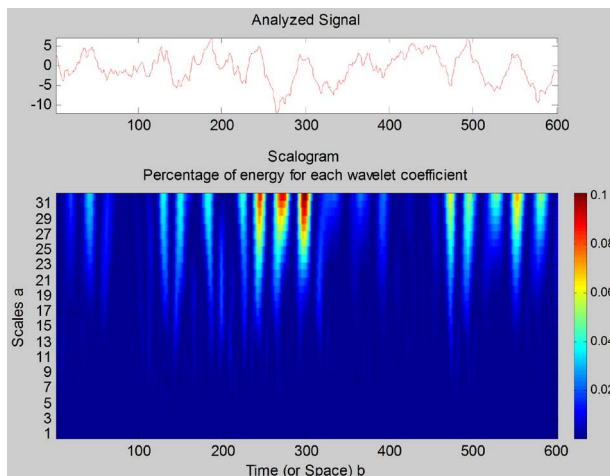


Figure 7. Healthy_{40Hz} CWT chart

图 7. Healthy_{40Hz} CWT 效果图

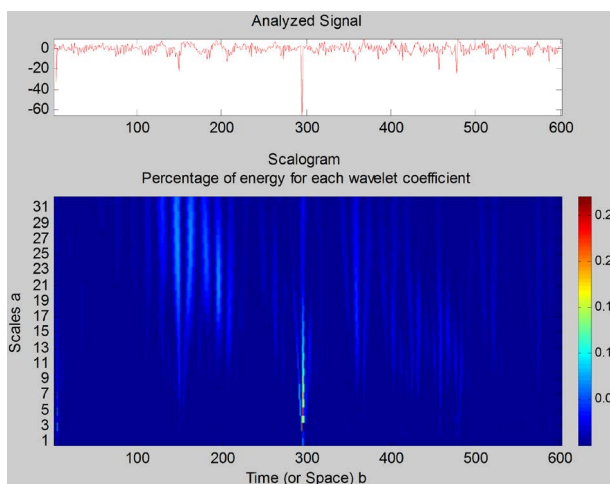


Figure 8. YangCo_{20Hz} CWT chart

图 8. YangCo_{20Hz} CWT 效果图

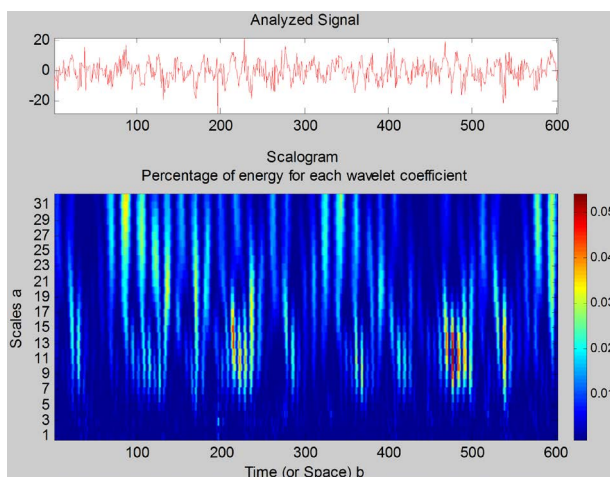
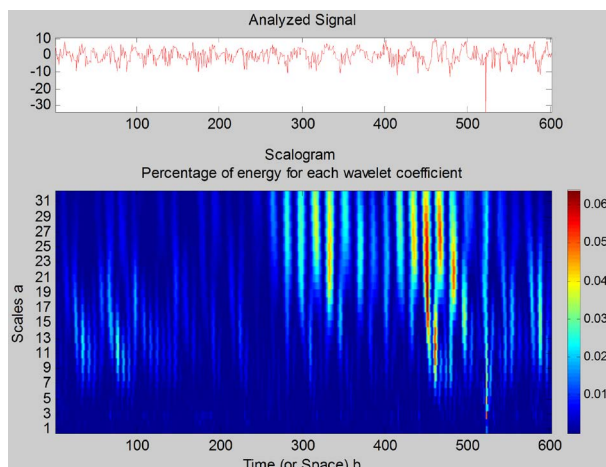
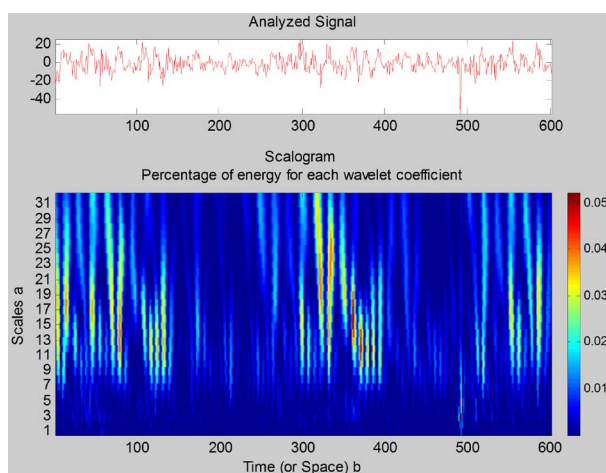


Figure 9. Yangpa_{40Hz} CWT chart

图 9. Yangpa_{40Hz} CWT 效果图

Figure 10. YangUn_{20HZ} CWT chart图 10. YangUn_{20HZ} CWT 效果图Figure 11. YangUn_{40HZ} CWT chart图 11. YangUn_{40HZ} CWT 效果图

对每一段数据(60 秒)进行分割, 取 1 秒数据(一帧)为一个样本(每两个样本重合率为 0.4)。实验样本信息见表 1 所示:

Table 1. Sample information

表 1. 样本信息

样本类别	时域样本(样本数 × 特征数)
Healthy _{20Hz}	1042 × 10,000
Healthy _{40Hz}	1021 × 10,000
YangCo _{20Hz}	1027 × 10,000
Yangpa _{40Hz}	1030 × 10,000
YangUn _{20Hz}	1037 × 10,000
YangUn _{40Hz}	1029 × 10,000

对六类光纤信号提取 FFT 特征(NFFT = 4096、Fs = 10K Hz), 四个通道(CH1、CH2、CH3、CH4) FFT 谱图均值如图 12~图 15 所示: 从该图中看出此六类样本在 FFT 空间具有一定的差异性。

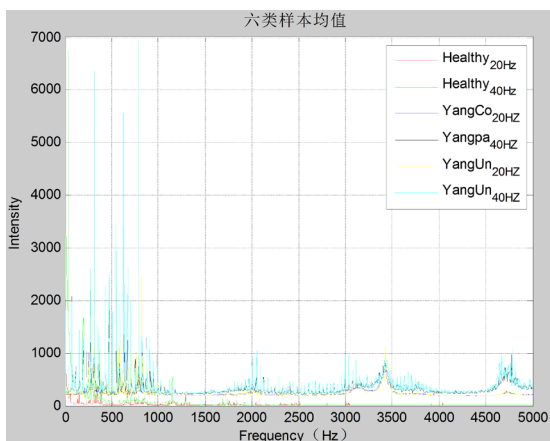


Figure 12. FFT means of six kinds of CH1 samples
图 12. CH1 六类样本 FFT 均值

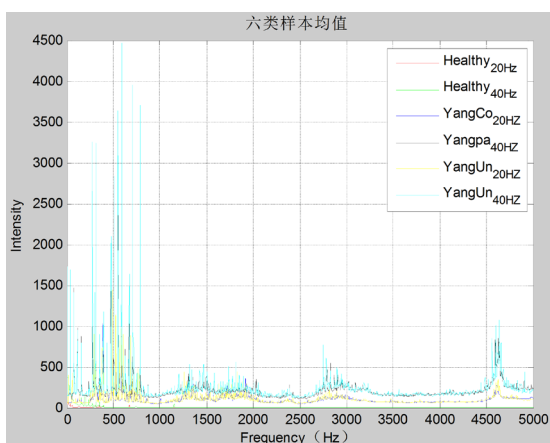


Figure 13. FFT means of six kinds of CH2 samples
图 13. CH2 六类样本 FFT 均值

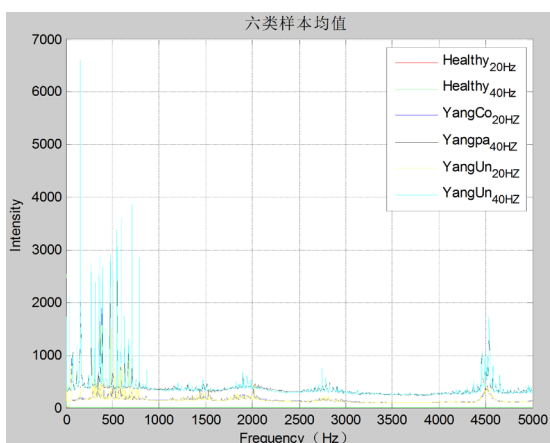


Figure 14. FFT means of six kinds of CH3 samples
图 14. CH3 六类样本 FFT 均值

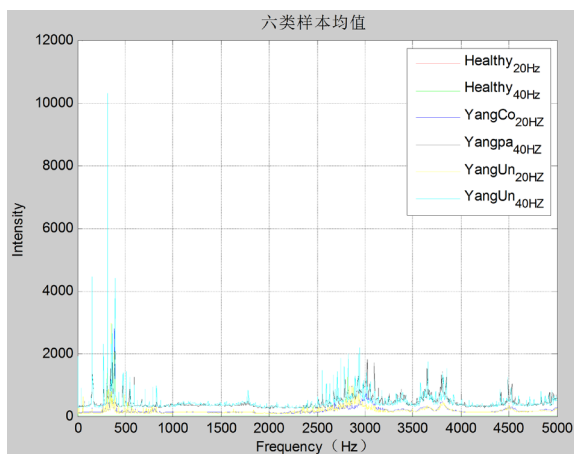


Figure 15. FFT means of six kinds of CH4 samples
图 15. CH4 六类样本 FFT 均值

实验中特征提取及分类算法见表 2 所示:

Table 2. Feature extraction and classifier description
表 2. 特征提取及分类器说明

特征提取方法	说明
FFT	快速傅里叶变换
PCA	主成分分析
WT	小波变换
LibLinSvm	SVM 线性分类器

3.2. 实验结果分析

为了分析样本在空间的分布状况, 对 FFT 特征进一步提取 PCA 分析见图 16~图 19 所示: 从图中看出, 通过 PCA 处理后六类样本在空间部分可分。

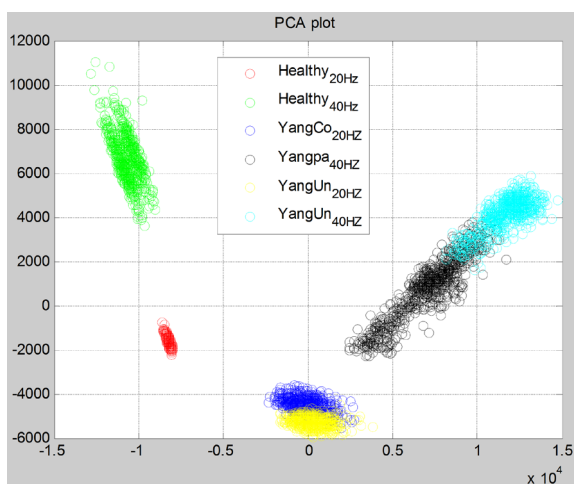


Figure 16. FFT feature of six kinds of CH1 samples for PCA chart
图 16. CH1 中六类样本 FFT 空间 PCA 效果图

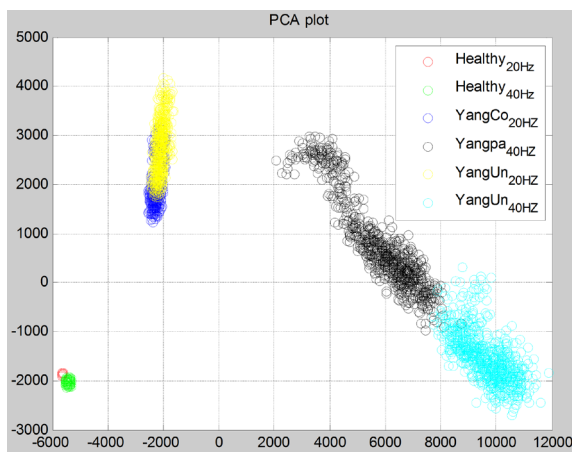


Figure 17. FFT feature of six kinds of CH2 samples for PCA chart
图 17. CH2 中六类样本 FFT 空间 PCA 效果图

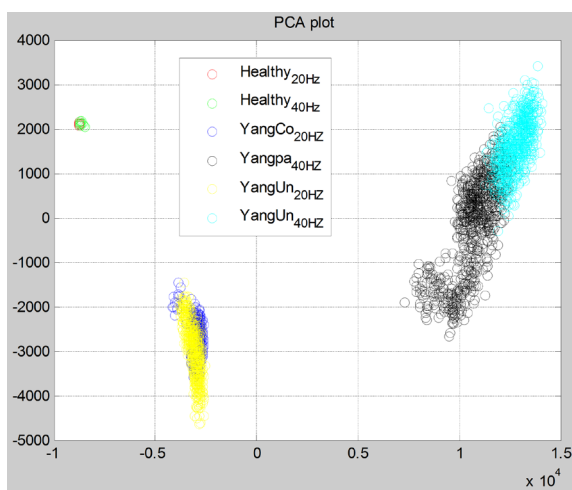


Figure 18. FFT feature of six kinds of CH3 samples for PCA chart
图 18. CH3 中六类样本 FFT 空间 PCA 效果图

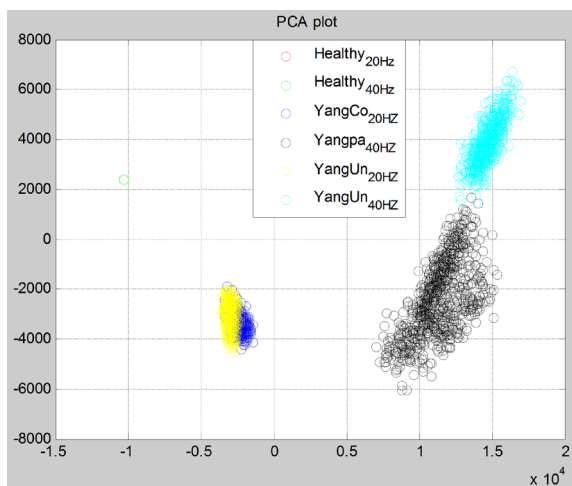


Figure 19. FFT feature of six kinds of CH4 samples for PCA chart
图 19. CH4 中六类样本 FFT 空间 PCA 效果图

为了准确评价 FFT 特征提取及 SVM 分类算法效果, 对每类样本随机选取[11] 0.75 训练, 0.25 测试。该建模过程如图 20 所示:

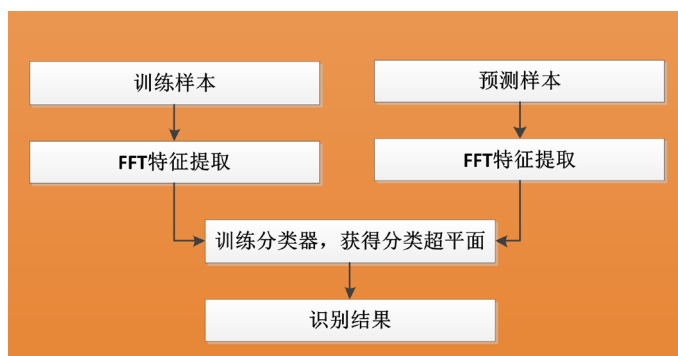


Figure 20. Flow of supervised pattern recognition
图 20. 监督模式识别过程

10 次平均结果, 测试总体分类正确率及每类样本分类正确率如表 3 所示:

Table 3. Comparison of classification results

表 3. 分类结果比较

通道	总体正确率	Healthy20	Healthy40	YangCo20	Yangpa40	YangUn20	YangUn40
CH1	0.998	1	1	0.996	1	0.992	1
CH2	0.995	1	1	0.995	1	0.993	1
CH3	0.988	1	1	0.996	0.961	0.992	0.984
CH4	0.839	1	1	0.853	0.893	0.828	0.812

从上述四个通道(CH1~CH4)实验结果分析, 仅利用 CH1 通道即可将 6 类振动信号基本区分开。

4. 总结

由于 FFT 样本在原始特征空间中具有一定的线性结构, 所以基于 PCA 特征提取方法(线性变换)分类效果显著, 同时利用 SVM 方法也取得了比较满意的分类效果。这是由于 SVM 基于结构风险最小化原则最大化分类间隔, 特别针对小样本学习问题, 具有良好的推广性能, 随着后续训练样本的不断丰富, 该方法也会进一步提高目标识别的正确率。

参考文献

- [1] 盛兆顺. 设备状态监测与故障诊断技术及应用[M]. 北京: 化学工业出版社, 2003.
- [2] 陈仲生, 杨拥民, 沈国际. 独立分量分析在直升机齿轮箱故障早期诊断中的应用[J]. 机械科学与技术, 2004, 23(4): 481-483+500.
- [3] 李强, 王太勇, 胥永刚. 基于混沌和二维近似熵的滚动轴承故障诊断[J]. 振动工程学报, 2007, 20(3): 268-273.
- [4] 王燕花. 新型光纤传感系统的研究与实现[D]: [博士学位论文]. 北京: 北京交通大学, 2009.
- [5] 杨渤, 程化, 陈树琪, 等. 基于傅里叶分析的超表面多维光场调控(特邀综述)[J]. 光学学报, 2019, 39(1): 1-20.
- [6] Mallat, S.G. and Zhang, Z. (1993) Matching Pursuits with Time-Frequency Dictionaries. *IEEE Trans on Signal Processing*, **41**, 3397-3415. <https://doi.org/10.1109/78.258082>
- [7] 虞和济, 陈长征, 张省. 基于神经网络的智能诊断[J]. 振动工程学报, 2000, 13(2): 202-209.

-
- [8] Tokuda, K., Yoshimura, T., Masuko, T. and Kobayashi, T. (2000) Speech Parameter Generation Algorithms for HMM-Based Speech Synthesis. 2000 *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Istanbul, Turkey, 5-9 June 2000, 1315-1318.
- [9] 张鹏. 基于主成分分析的综合评价研究[D]: [硕士学位论文]. 南京: 南京理工大学, 2004.
- [10] 丁世飞, 齐丙娟, 谭红艳. 支持向量机理论与算法研究综述[J]. 电子科技大学学报, 2011, 40(1): 2-10.
- [11] 范永东. 模型选择中的交叉验证方法综述[D]: [硕士学位论文]. 太原: 山西大学, 2013.

知网检索的两种方式:

1. 打开知网首页: <http://cnki.net/>, 点击页面中“外文资源总库 CNKI SCHOLAR”, 跳转至: <http://scholar.cnki.net/new>, 搜索框内直接输入文章标题, 即可查询;
或点击“高级检索”, 下拉列表框选择: [ISSN], 输入期刊 ISSN: 2325-6753, 即可查询。
2. 通过知网首页 <http://cnki.net/>顶部“旧版入口”进入知网旧版: <http://www.cnki.net/old/>, 左侧选择“国际文献总库”进入, 搜索框直接输入文章标题, 即可查询。

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: jjsp@hanspub.org