

# 基于倒谱与 BP 网络的船舶生活垃圾分类方法研究

刘 强<sup>1</sup>, 陈福兰<sup>2</sup>

(1. 南通航运职业技术学院 船舶与海洋工程系, 江苏 南通 226010;

2. 南通航运职业技术学院 机电系, 江苏 南通 226010)

**摘 要:**文章介绍了船舶生活垃圾中不同材质的物体敲击声波的特征提取方法,建立了基于倒谱与神经网络的物体材质的智能识别模型与算法,并通过对玻璃、塑料、金属铝箔等常见物体材质进行分类测试,证实了该分类方法的有效性,为船舶生活垃圾的智能分类提供了新思路。

**关键词:**倒谱;材质分类;神经网络;智能识别

中图分类号:U664.9

文献标识码:A

文章编号:1671-9891(2016)01-0034-05

## 0 引言

船舶规范中对生活垃圾的分类管理有明确要求,但在实际的垃圾分类管理过程中,却面临很多问题。船舶生活垃圾主要有残余食物、饮料瓶、包装盒等垃圾。在垃圾处理过程中,分拣是一项复杂的工作,进行物体材质的智能识别与分类方法研究有助这些问题的解决。针对物体不同材质的相应特点,采用不同的方法进行分离,如食物类垃圾利用切削分离、过筛等方法可与其它材质垃圾进行分离,金属类垃圾可采用电磁感应传感器进行分拣。但对玻璃、塑料类垃圾在分拣时相关传感器的分类效果不是很理想。<sup>[1]</sup>本文结合声波倒谱特征分析与 BP 神经网络分类技术,对垃圾中不同材质物体的智能分类方法进行了研究。

### 1 基于倒谱的声波特征提取方法

#### 1.1 敲击声的时域、频域结构特点

对物体样本的敲击声是一种典型的瞬态声,这类冲击声分类识别的关键在于特征提取。传统的冲击声特征提取方法主要是根据时频分析的谱特征提取,但其效果有限。随着基于人耳听觉原理提取感知特征的深入研究,利用谱质心、谱质心带宽以及不和谐性等特征可实现了对声波的识别。在研究过程中发现,频域特征和时域特征在描述不同材料物体发出的冲击声音时具有互补性。<sup>[2]</sup>

本文的研究对象是在相同条件下,利用金属球对相同大小的玻璃、塑料、铝片试样进行敲击,产生的瞬态声。通过对三种材质的试样进行反复声波信号采集,得到相应的典型时域、频域图如图 1 所示。从图中发现:玻璃、铝板的敲击声持续时间较长,塑料敲击声持续时间较短,衰减速度快;频域上,塑料敲击声频谱主要集中在 3 kHz 以下的低频段,而玻璃板和铝板敲击声的频谱广泛分布于 0-13 kHz 范围内,频带相对较宽。

在分类问题中声波特征提取环节尤为重要。声波属性具有多维性,受声音信号时域结构和频域结构的共同影响。<sup>[3]</sup>本文首先提取了冲击声的相关音色特征用于分类,其中包括时域特征、频域特征及用 Mel 倒谱分析法对声波信号提取的特征,建立了特征库并采用 BP 神经网络对其进行学习、训练,然后对不同材质的物体进行分类并检验分类效果。

#### 1.2 Mel 倒谱特征系数

倒谱分析技术是一个很有用的方法,在声波信号处理中应用普遍,与时域相比,倒谱系数在倒谱域中变

收稿日期:2015-11-20

基金项目:江苏省交通科学研究计划项目“基于知识的标准船型动力装置智能化设计系统”(项目编号:2011C04-1)。

作者简介:刘强(1970—),男,陕西宝鸡人,南通航运职业技术学院船舶与海洋工程系副教授。

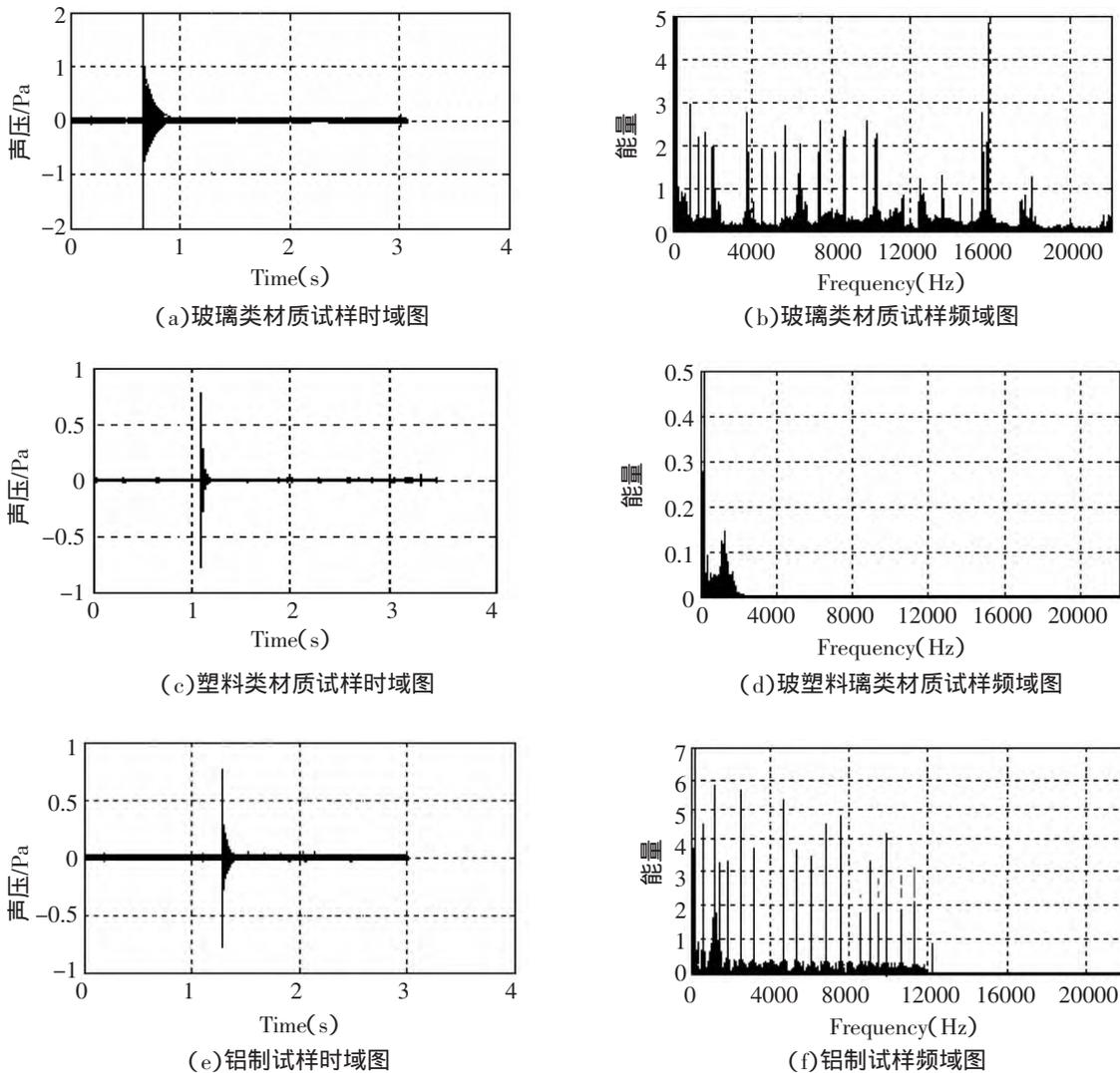


图 1 三种材质试样敲击声的时域与频域图

化很小。由于倒谱是声波信号的对数功率谱的逆傅立叶变换,是一个时域函数,故称该函数为倒谱。<sup>[4]</sup>基于人耳听觉特性的 Mel 频率是一种非常重要的音色特征,与线性频率的关系式如式(1)所示:

$$f_{Mel} = 2.595 \times \lg(1 + f/700) \tag{1}$$

式(1)中,  $f_{Mel}$  为 Mel 频率,  $f$  为线性频率。

Mel 频率倒谱系数(Mel Frequency Cepstrum Coefficient,以下简称“MFCC”)考虑到了人类的听觉特征,先将线性频谱映射到基于听觉感知的 Mel 非线性频谱中,然后转换到倒谱上。大量研究证实,该参数很适用于声音音色识别,MFCC 参数提取过程如图 2 所示。具体计算流程如下:

- (1)对声波信号进行预加重、分帧与加窗处理;
- (2)通过 FFT,得到每一个短时分析窗对应的频谱;
- (3)将以上频谱用 Mel 滤波器组滤波得到 Mel 频谱;
- (4)在 Mel 频谱上面进行倒谱分析,获得 Mel 频率倒谱系数 MFCC(即声波的 MFCC 特征向量)。

### 1.3 其他声波特征值

在声波特征分析中,谱质心、谱通量、谱下降值、振幅因数等谱特征也有利于提高材质分类的准确性。<sup>[5]</sup>谱质心(SC)是在一定频率范围内通过能量加权平均的频率,是频率成分的重心,反映声音的明亮程度,SC 越高,声音越明亮。其计算公式如式(2):



图2 MFCC 参数提取过程

$$SC = \frac{\sum_{n=1}^N f(n)E(n)}{\sum_{n=1}^N E(n)} = \sum_{n=1}^N f(n)P(E(n)) \quad (2)$$

式中  $f(n)$  为信号  $x(n)$  傅里叶变换后对应的频率;  $N$  为 DFT 的长度;  $E(n)$  为离散时域信号  $x(n)$  傅里叶变换后对应的谱能量  $P(E(n))$  为各点能量在总能量上的概率值。

谱通量(SF)是声音信号频谱的包络面积,反映信号中各频带成分的能量之和。本文将整个频带内的谱通量作为识别特征。其计算公式如式(3):

$$SF = \sum_{n=1}^N \Delta f(n)E(n) \quad (3)$$

式中  $\Delta f(n)$  为信号  $x(n)$  傅里叶变换后两点频率之差。作为频谱幅度增量, SF 曾被用于乐音音色识别。

谱下降值(SRO)表示频谱的倾斜程度,反映声音能量开始下降的特定频率点,常用于区分敲击声。SRO 的计算方法如式(4):

$$\sum_{n=1}^R E(n) = C \sum_{n=1}^N E(n) \quad (4)$$

式中  $R$  即为声音能量开始下降的特定频率点,也就是谱下降值;  $C$  是一个经验系数,本文选用  $C=0.75$ 。

振幅因数(CF)是声音信号识别中的重要参数,反映了信号最大振幅在整个包络中所占的重要程度,定义为信号最大振幅与信号的有效值之比。<sup>[6]</sup>计算公式如式(5):

$$CF = \max\{x(n)\} / \sqrt{\frac{1}{N} \sum_{n=1}^N x(n)^2} \quad (5)$$

式中  $x(n)$  为信号振幅。

## 2 基于神经网络的声波分类方法

基于模式识别的声波特征识别过程如下:首先,待识别声波转化为电信号后输入识别系统,经过预处理后用数学方法提取特征信号,提取出的特征信号可以看成该段声波的模式;其次,将该声波模式同已知参考模式相比较,获得最佳匹配的参考模式为该段声波的识别结果。

本文采用 BP 神经网络实现对这三类声波的有效分类。经过声波特征提取后,每段声波特征信号由 16 维声波特征信号组成(其中包括 12 维 MFCC 特征系数、SC、SF、SRO、CF 组成)。结合 BP 多层神经网络的前馈特点(信号前向传递,误差反射传播),声波特征信号分类算法建模主要包括 BP 神经网络构建、训练和 BP 神经网络分类三步,其算法流程如图 3 所示。

### 2.1 BP 神经网络模型建立

首先根据系统输入、输出数据特点确定 BP 神经网络的结构。根据倒谱系数法提取三类声波特征信号,不同的声波信号分别用 1、2、3 标识,提取出的信号分别存储于 data1.mat、data2.mat、data3.mat 数据库文件

中,每组数据为 17 维(第 1 维为类别标识,后 16 维为声波特征信号)。由于声波特征输入信号有 17 维,待分类的语音信号共有 3 类,所以 BP 神经网络的结构为 17-12-3,即输入层有 17 个节点,隐含层有 12 个节点,输出层有 3 个节点。声波特征信号共有 500 组,从中随机选取 350 组作为训练数据,150 组数据作为测试数据。

### 2.2 BP 神经网络的 MATLAB 实现

据 BP 网络理论,在 MATLAB 中编程实现基于 BP 网络的声波特征分类算法,程序流程图如图 4 所示。

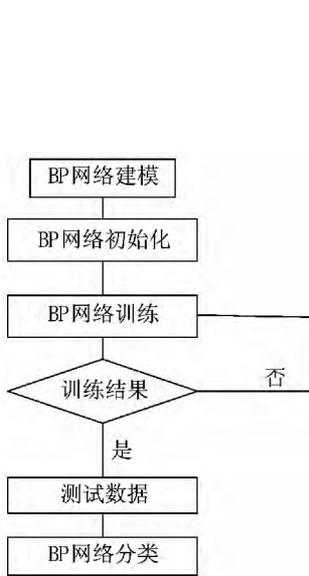


图 3 声波特征信号分类算法流程图

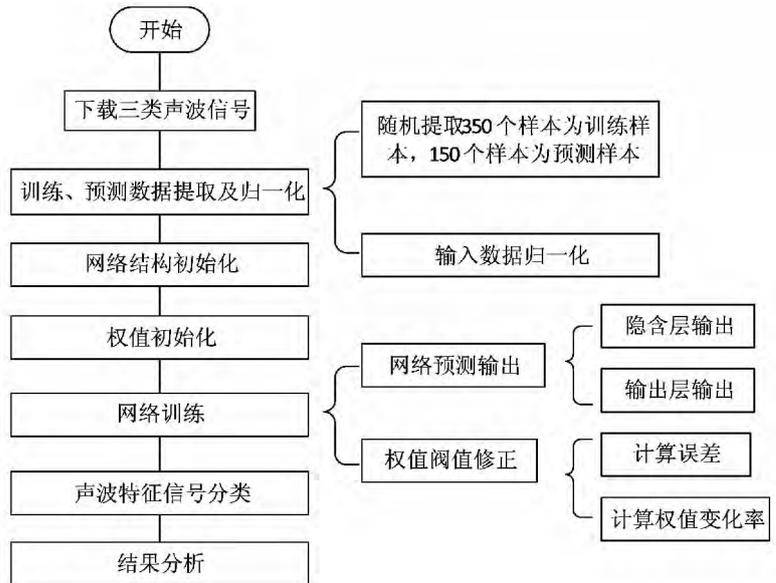


图 4 MATLAB 实现声波特征分类程序流程图

首先用 load 函数从 data1.mat、data2.mat、data3.mat 声波特征数据库中提取声波特征数据,构造一个 500X17 的矩阵,随机从矩阵中提取 350 组样本作为训练样本,150 组作为预测样本,用 mapminmax 函数对训练样本进行数据归一化处理,构造 17-12-3 的 BP 网络结构并初始化权值与学习率,定义学习率为 0.1,误差为 0.01,循环次数为 10 次,用训练样本对网络进行训练,根据误差计算修正权值阈值,训练误差曲线如图 5 所示,直到满足误差要求,对预测样本进行分类,得出分类结果及分类准确率,根据分类结果分析 BP 神经网络分类能力。

### 2.3 结果分析

用训练好的 BP 神经网络分类声波特征信号,并计算测试数据分类正确率。经计算,BP 神经网络分类正确率如表 1 所示。

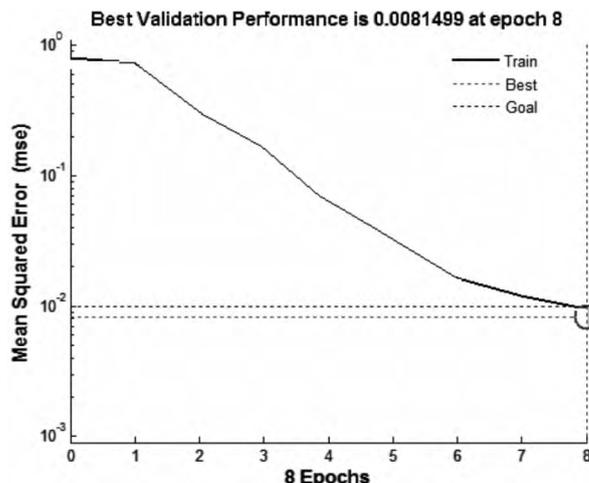


图 5 BP 神经网络训练误差曲线

表1 BP神经网络分类正确率

声波信号类别	识别正确率
塑料类	0.966 8
玻璃类	0.863 7
金属类	0.877 0

从表中可以看出,基于BP神经网络的声波信号分类算法具有较高的准确性,能够准确识别。塑料类的分类结果较为理想,只有不到3%的样本被错分为玻璃板类或金属类。而玻璃板类和金属类的分类结果相对较低,主要因为两类样本在时域上及频域上的相似性所致。上述结果与人耳在主观听音时得出的结论相符,听音时发现,塑料类与其他两类冲击声相异性较大,容易辨识,而玻璃板和金属类的冲击声之间相似性较高,容易产生分类误判。

### 3 结束语

本文研究了塑料、玻璃及金属类材质物体敲击时产生的冲击声音色特征,主要提取了12维MFCC系数以及谱质心、谱通量、谱下降值、振幅因数作为声波的特征,并建立了特征库。在此基础上,利用BP神经网络进行了船舶生活垃圾算法的验证,得到了较高的分类准确率,通过分类实验确定了在冲击声分类任务中占据重要地位的多种音色感知特征,为利用声波识别技术在船舶生活垃圾等相关领域智能分类方法的进一步研究打下良好基础。

### 参考文献:

- [1]刘翠伟,李玉星,王武昌,等.输气管道声波法泄漏检测技术的理论与实验研究[J].测试技术学报,2013(3):373-378.
- [2]赵毅,尹雪飞,陈克安,等.一种新的基于倒谱的共振峰频率检测算法[J].应用声学,2010(6):416-419.
- [3]林玉志.基于声学特征的乐器识别的研究[D].广州:华南理工大学,2012.
- [4]靳舜.基于声音和震动信号特征融合的车辆识别研究[D].西安:长安大学,2014.
- [5]侯雪梅.基于改进LP倒谱特征和神经网络的语音识别算法研究[D].太原:太原理工大学,2006.
- [6]王燕.基于倒谱分析的单水听器目标运动参数估计[J].物理学报,2014(3):63-70.

## Research on Classification Methods for Marine Life Rubbish Based on Cepstrum and BP Network

LIU Qiang<sup>1</sup>, CHEN Fu-lan<sup>2</sup>

- (1.Dept. of Ship & Ocean Engineering, Nantong Vocational & Technical Shipping College, Nantong 226010, China;  
2.Dept. of Mechatronics, Nantong Vocational & Technical Shipping College, Nantong 226010, China)

**Abstract:** This article introduces the feature extraction methods of object percussion sound waves and establishes the intelligent recognition model and algorithm of object material based on Cepstrum and neural network. In addition, based on the classification test of such ordinary material as glass, plastic and metal foil, it verifies the effectiveness of the classification methods, which is expected to offer a new angle for the intelligent classification of marine life rubbish.

**Key words:** Cepstrum; Material classification; Neural network; Intelligent recognition