

引用格式: 陈德昊, 林建恒, 衣雪娟, 等. 三种鱼类发声信号频带能量特征提取与分类研究[J]. 声学技术, 2021, 40(2): 254-259. [CHEN Dehao, LIN Jianheng, YI Xuejuan, et al. Research on band energy extraction and classification of three kinds of fishes sound signals[J]. Technical Acoustics, 2021, 40(2): 254-259.] DOI: 10.16300/j.cnki.1000-3630.2021.02.017

三种鱼类发声信号频带能量特征提取与分类研究

陈德昊^{1,2}, 林建恒², 衣雪娟², 江鹏飞²

(1. 中国科学院大学, 北京 100049; 2. 中国科学院声学研究所北海研究站, 山东青岛 266114)

摘要: 海洋渔业资源的开发和利用对经济鱼类的分类识别提出了迫切的技术需求。根据鱼类的不同发声特征, 文章采用有监督机器学习方法实现了三个不同鱼种发声信号的分类。基于小波包分解技术提取了黄花鱼、大米鱼和黄姑鱼三种鱼类发声信号的频带能量特征, 并利用不同的分类器进行分类。结果表明: 三种鱼的发声信号频率主要集中在 300~800 Hz 之间, 基于小波包分解的频带能量特征可实现 3 种鱼的有效分类。其中, 线性判别分类器和随机子空间判别分类器的分类效果较好。该方法可为海洋渔业资源的开发和利用提供服务。

关键词: 鱼类发声信号; 频带能量; 小波包分解; 特征提取; 分类

中图分类号: TN911.7

文献标志码: A

文章编号: 1000-3630(2021)-02-0254-06

Research on band energy extraction and classification of three kinds of fishes sound signals

CHEN Dehao^{1,2}, LIN Jianheng², YI Xuejuan², JIANG Pengfei²

(1. University of Chinese Academy of Sciences, Beijing 100049, China;

2. Qingdao Branch, Institute of Acoustics, Chinese Academy of Sciences, Qingdao 266114, Shandong, China)

Abstract: The exploitation and utilization of marine fishery resources put forward urgent technical requirements for classification and recognition of economic fishes. According to the different acoustic features of vocal fishes, the classification of vocal signals of three kinds of fishes is implemented by supervised machine learning in this paper. Based on wavelet packet decomposition technique, the band energies of the vocal signals of yellow croaker, rice fish and yellow drum are extracted and classified by different classifiers. The results show that the vocal signals of the three kinds of fishes are mainly concentrated in the frequency bands of 300–800 Hz, and the band energies of vocal signals based on wavelet packet decomposition can achieve effective classification for the three kinds of fishes. In different classifiers, linear discriminant classifier and random subspace discriminant classifier have better classification results. This method can provide services for the exploitation and utilization of marine fishery resources.

Key words: fish sound signal; band energy; wavelet packet decomposition; feature extraction; classification

0 引言

随着人类对海洋渔业资源需求的日益增加, 利用声学技术开采海洋渔业资源越来越受到人们的重视。目前各国已经广泛采用各种主动声呐技术探测、开发和利用海洋渔业资源。同时, 被动声学技术也成为海洋渔业资源评估和海洋生态系统监测的重要手段和工具。

对于鱼类发声信号, 可在时域、频域或梅尔频率倒谱域等变换域下进行分析。文献[1-7]采用了带

通滤波技术、时域分析、频域分析、时频分析和混合神经网络算法等信号处理方式对鱼类发声信号进行分析。然而, 传统的时域或频域分析无法同时兼顾频域和时域信息, 故无法全面地分析信号。陈功等^[8]于 2013 年利用仿声技术发现, 可将鱼类发声信号类比说话人语音信号进行分析。他们将梅尔频率倒谱系数(Mel Frequency Cepstrum Coefficient, MFCC)作为鱼类被动声信号的特征信息, 该特征具有维数少、区分度大、鲁棒性强、计算量小的优点。时频分析虽可以同时兼顾时域和频域, 但选取的方法不同, 结果会存在差异。常见的时频分析方法有短时傅里叶变换(Short-Time Fourier Transform, STFT)、维格纳分布(Wigner Distribution Function, WDF)、希尔伯特-黄变换(Hilbert-Huang Transform, HHT)以及小波变换(Wavelet Transform, WT)等。短时傅里叶变换计算简单, 但无法同时兼顾时域分辨

收稿日期: 2020-01-09; 修回日期: 2020-03-03

基金项目: 国家重点研发计划资助课题(2019YFD0901301)、国家自然科学基金资助项目(11804361)。

作者简介: 陈德昊(1995—), 男, 山东济宁人, 硕士研究生, 研究方向为信号与信息处理。

通信作者: 林建恒, E-mail: linjh@mail.ioa.ac.cn

率和频域分辨率,不太适合处理具有瞬态非平稳特性的鱼类发声信号。维格纳分布会出现“交叉项效应”,且复杂度较高。希尔伯特-黄边换在检测微弱突变信号具有一定优势,但其结果不具有解析形式^[8-9]。对于小波变换,在选取合适的小波基条件下,它的优势才能更好地发挥出来,且其结果具有解析形式。小波包变换是小波变换的推广,张静远等^[10]验证了基于小波包变换的特征提取对主动探测水下目标声信号的有效性。Harakawa 等^[11]提出了一种在水下环境中自动检测鱼类声音的方法,这种方法是结合K近邻(K-Nearest Neighbor, KNN)、支持向量机(Support Vector Machine, SVM)和新型逻辑回归(New Logistic Regression, NLR)的混合方法,他们表示这种混合分类方法性能优于其中的单个分类方法,他们通过对部分石首鱼类展开实验,验证了这种混合分类方法的有效性。文献[12-14]中将深度学习中的深度卷积神经网络应用于鱼类图像的分类识别。文献[15-18]中采用了线性频率倒谱系数(Linear Frequency Cepstrum Coefficient, LPCC)、梅尔频率倒谱系数(MFCC)、香农熵(Shannon Entropy, SE)、音节长度(Syllable Length, SL)、信号波形、信号时频图等特征,并结合了支持向量机、随机森林(Random Forest, RF)、人工神经网络(Artificial Neural Network, ANN)和深度卷积神经网络(Deep Convolutional Neural Networks, DCNN)等分类器,得出了不同分类方案下鱼类发声信号的分类结果。Lin 等^[19]对比了不同情况下有监督和无监督方案在鱼类被动声监控中的应用。焦奎奎^[20]将小波包变换特征应用到了淡水鱼鱼体回波信号的主动声识别研究,得出了小波包变换特征在支持向量机分类器下的性能总体优于反向传播(Back Propagation, BP)神经网络的结论。综上可知,对鱼类发声信号的分类识别已经开始由传统的智能分类方法逐步过渡到用人工神经网络分类甚至是深度学习方法。尽管如此,一些经典的“特征提取+分类器”方法仍然在鱼声识别中发挥着重要作用。本文分别将黄花鱼(yellow croaker)、大米鱼(rice fish)和黄姑鱼(yellow drum)三种鱼类发声信号经过小波包变换得到的频带能量特征作为分类特征,尝试分析了该特征在不同分类器下的有效性。

三种鱼均属于石首鱼类,均是我国重要的发声经济鱼类, Ramcharitar 等^[21]介绍了部分石首鱼类的生物声学机制。对黄花鱼、大米鱼和黄姑鱼的发声信号进行特征提取与分类研究,可为改善这三种海洋鱼类资源的开发和利用环境提供研究思路,同时也为其他鱼类识别方法提供技术参考。

本文首先概述了小波变换和小波包变换的基本理论,然后说明了数据采集和预处理方法,分析了用小波包分解得到的实验环境背景噪声时频图和三种鱼发声信号时频图,尝试利用小波包分解提取三种鱼类发声信号的频带能量特征并对其进行分类判别。最后依据总体准确率和预测得分评判准则对比分析了基于小波包分解的频带能量特征在不同分类器下的分类性能。

1 小波变换与小波包变换概述

小波变换作为一种经典的时频变换方法,实现了泛函分析、傅里叶分析、调和分析以及数值分析理论之间的紧密结合,成为继傅里叶变换后,另一个重要的信号处理方式。小波变换中的多分辨率分析具有良好的时频局部化能力,能够通过时域或频域逐步精细步长,实现对信号细节的聚焦,从而能更加全面细致地描述信号特征,尤其是非平稳信号。鱼类发声信号作为一种典型的瞬态非平稳信号,应用小波变换有时要比傅里叶变换和短时傅里叶变换等一些传统方法而言具有更显著的优势。

小波变换分为连续小波变换和离散小波变换。考虑数字化,实际应用通常采用离散小波变换。离散小波正变换和重构逆变换可以表述为式(1)和式(2),信号 S 的三层小波分解树示意图如图1所示。

$$T_{w,f}(m,n)=\oint_{\mathbb{R}} f(t)\cdot\overline{\Psi_{m,n}(t)}dt, \quad m=a_0^j, n=k\tau_0 \quad (1)$$

$$f(t)=\sum_j \langle f, \Psi_{j,k}(t) \rangle \tilde{\Psi}_{j,k}(t) \approx \sum_{j,k} T_{w,f}(j,k) \Psi_{j,k}(t), \quad c \in \mathbb{R}, j \in \mathbb{N}, k \in \mathbb{Z} \quad (2)$$

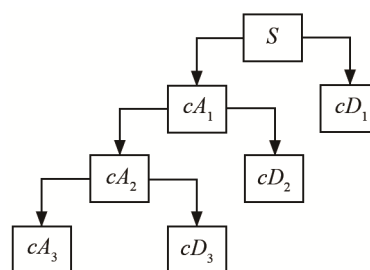


图1 小波分解树示意图

Fig.1 Schematic diagram of wavelet decomposition tree

小波包分解,又称为小波包变换或最优子带树结构,是对小波变换的进一步优化。小波包分解由于保留了小波分解舍弃的高频细节分量信息,故相比小波分解具有更强的时频分析能力。图2给出了对信号 S 进行三层小波包分解的示意图。小波包树每一层的节点个数等于 2^n , n 为分解层数。总的来说,小波包分解要比小波分解更加精细化。

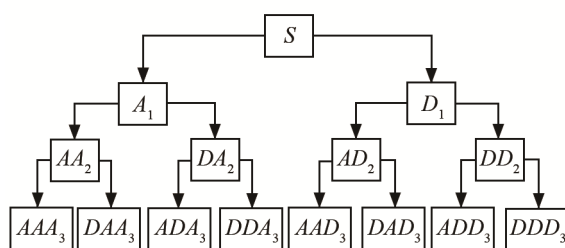


图2 小波包分解树示意图

Fig.2 Schematic diagram of wavelet packet decomposition tree

2 数据采集与分析方法

2.1 数据采集与预处理

2.1.1 数据采集

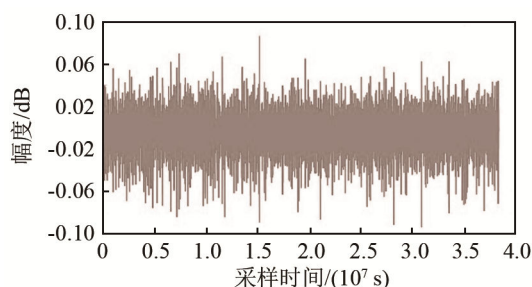
本次实验采集时间于秋季傍晚,采集地点位于中国东海的某养殖场网箱,具体细节如下:

黄花鱼发声信号的采集时间为 18:21:36~22:41:36。采集信号分为大黄花鱼和小黄花鱼发声信号。小黄花鱼鱼龄为 4~5 个月,体重不详。大黄花鱼体重和鱼龄均不详。

大米鱼发声信号的采集时间为 18:10:32~23:40:32。被测的大米鱼重 1 000 g~1 500 g,鱼龄 20 个月。黄姑鱼发声信号的采集时间是 15:34:06~23:34:06。被测的黄姑鱼重 500 g 左右,鱼龄 24 个月。

2.1.2 预处理

对三种鱼的发声信号每间隔 10 min 采集一次,原始采样率为 64 kHz。三种鱼 10 min 的时域信号样本如图 3 所示。由于鱼类发声信号主要分布在低



(c) 黄姑鱼

图3 三种鱼的时域发声信号

Fig.3 Time domain diagrams of the vocal signals of the three kinds of fishes

频,所以为了在低频范围内观察和分析信号,将每个采集到的原始 10 min 时域信号做降采样处理,降采样率设为 4 000 Hz。

2.2 小波包分解与特征提取

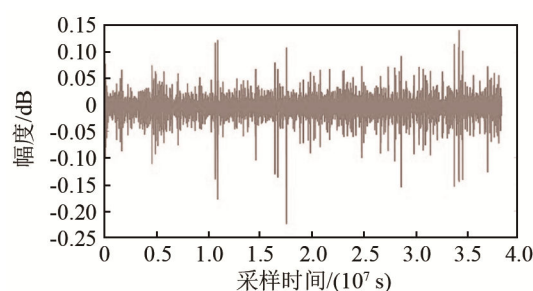
2.2.1 小波包分解

将预处理后的三种鱼发声时域信号进行分帧,帧长设为 1 s,对每帧信号用“db6”小波基进行 7 层小波包分解。在小波包分解树的第 7 层将得到的 128 个节点按频率段递增的顺序依次排序,形成 128×600 维的时频分布矩阵,128 表示 7 层小波包树最底层节点的频带数、600 表示时长。矩阵中的每一列表示每帧信号用 7 层小波包分解后,第 7 层节点对应的 128 个频带在排序后各自的能量分布。在此基础上得到不同样本 10 min 的时频图,结果如图 4 所示。其中,图 4(a)为鱼声稀疏条件下的实验环境背景噪声时频图,图中的鱼声信息成分较少。图 4(b)给出了黄花鱼密集发声时的时频图,可见,黄花鱼的发声频率主要集中在 300~800 Hz,且随时间变化保持一致性。这和任新敏等^[1]在利用快速傅里叶变换分析大黄花鱼发声信号时所得出的“大黄花鱼发声信号幅度峰值在 800 Hz”的结论相近。同时,本文的分析结果表明,大米鱼和黄姑鱼发声信号的主要频率范围 and 黄花鱼基本一致,结果如图 4(c)和 4(d)所示。

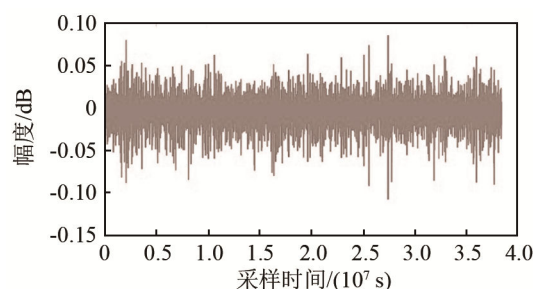
2.2.2 特征提取方法

本文将基于小波包分解的频带能量作为三种鱼发声信号的分类特征。具体的提取方法为:在小波包分解时频图(或时频分布矩阵)的基础上,取每一行的均值,也就是每个频带在 10 min 内的能量均值,从而得到 128 维的特征向量。将这 128 维的频带能量均值特征作为基于小波包分解的频带能量特征。

提取得到的 3 种鱼各自所有样本的频带能量特征分布如图 5 所示。以图 5(a)为例,它表示了 27



(a) 黄花鱼



(b) 大米鱼

个黄花鱼时域发声信号样本基于小波包分解的频带能量特征分布情况。在这 3 幅图中，三种鱼所有

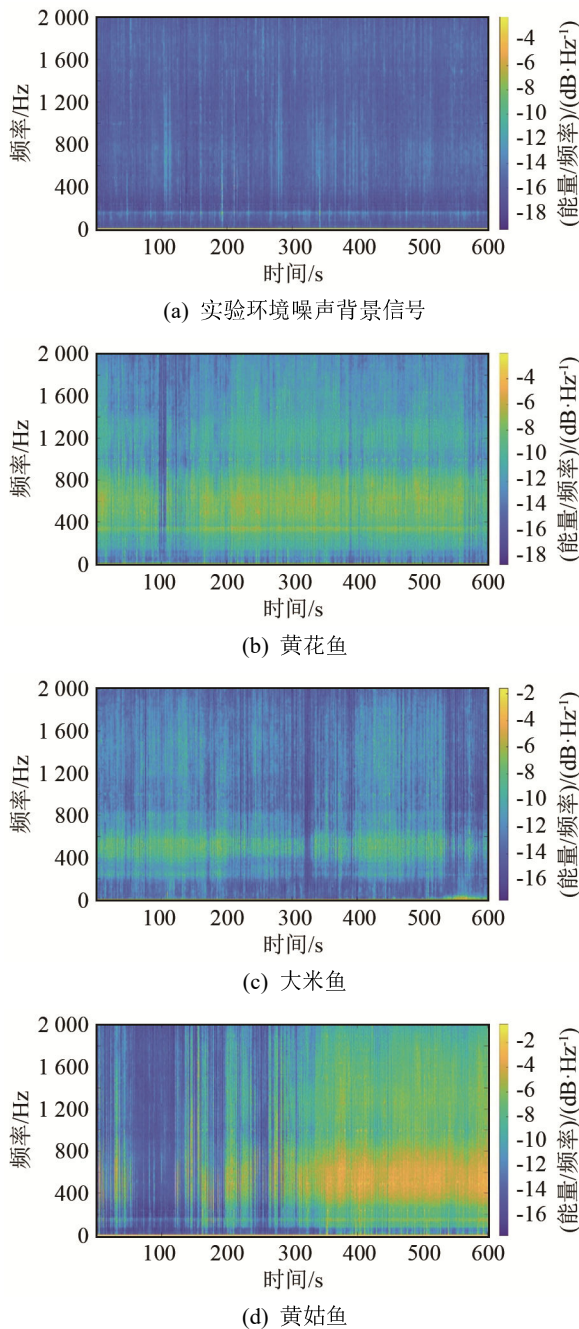


图 4 实测背景噪声和三种鱼类发声信号的时频分布图
Fig.4 Time-frequency maps of the measured background noise and vocal signals of the three kinds of fishes

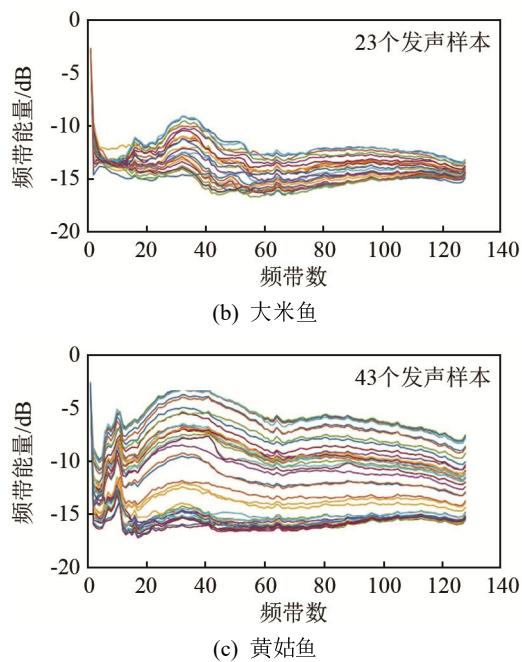
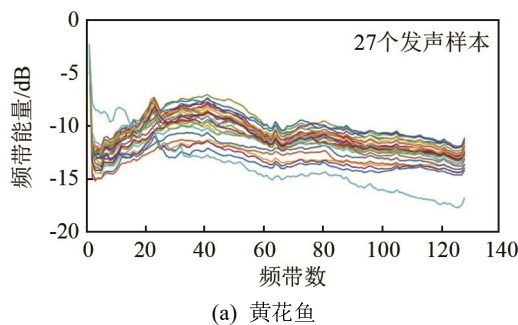


图 5 三种鱼发声信号的频带能量特征分布图
Fig.5 Band energy distribution of the vocal signals of the three kinds of fishes

发声信号样本的频带能量特征分布差异较为明显。具体来讲，大米鱼发声信号的频带能量特征分布较为集中，黄花鱼的次之，黄姑鱼的最为分散。此外，由图 5(a)、5(b)、5(c)也可发现，三种鱼的发声信号在 128 个频带下的能量变化趋势也有较大差异。因此，在有合适分类器的条件下，这些差异将有助于得到较好的分类效果。

2.3 分类决策方法与分类结果

将这三种鱼所有的特征样本分为训练集和测试集。其中，训练集采用 15 折交叉验证，得到对应的交叉验证集。三种鱼的特征样本划分方式如表 1 所示。

表 1 大黄鱼、大米鱼、黄姑鱼的特征样本划分				
Table 1 Feature sample division of the three kinds of fishes				
特征样本类型	黄花鱼	黄姑鱼	大米鱼	总数
训练集	16	16	15	47
测试集	11	27	8	46

将提取好的特征样本输入到 8 种分类器中进行分类。在分类过程中，为了进一步验证分类结果的可靠性，本文将训练集和测试集进行了一次交换。交换前后采用相同的方法处理。两次分类验证得到如表 2 和表 3 所示的总体准确率。

在表 2 分类器的参数选择上，线性判别(Linear discriminant)采用全协方差结构(Full covariance structure)。随机子空间维度设为 64，学习器的数量设为 30。支持向量机采用线性核(Linear kernel)和

二次多项式核(Quadratic polynomial kernel), 多分类拆分策略采用一对一(one-versus-one)间接方案。

表 2 输出结果比较好的 5 种分类器及对应的总体准确率
Table 2 Five classifiers with better output results and corresponding overall accuracy

分类器	总体准确率/%	
	交叉验证集	测试集
线性判别	100 (100)	100 (100)
随机子空间判别	100 (100)	100 (100)
线性 SVM	100 (100)	100 (83)
二次多项式 SVM	95.7 (100)	100 (89)
随机子空间 KNN	95.7 (89.1)	20 (85)

表 3 输出结果比较差的 3 种分类器及对应的总体准确率
Table 3 Three classifiers with worse output results and corresponding overall accuracy

分类器	总体准确率/%	
	交叉验证集	测试集
提升决策树	31.9 (58.7)	24 (34)
K 近邻	31.9 (58.7)	24 (34)
高斯核 SVM	66.0 (58.7)	43 (34)

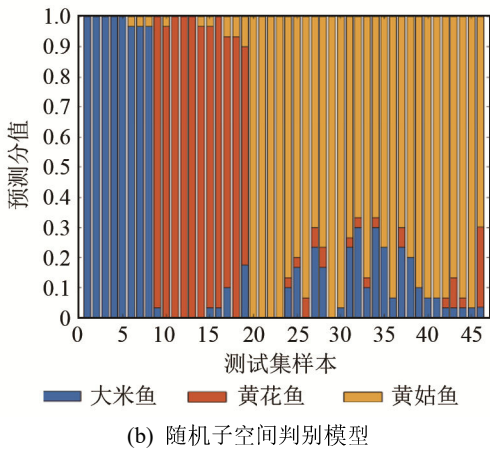
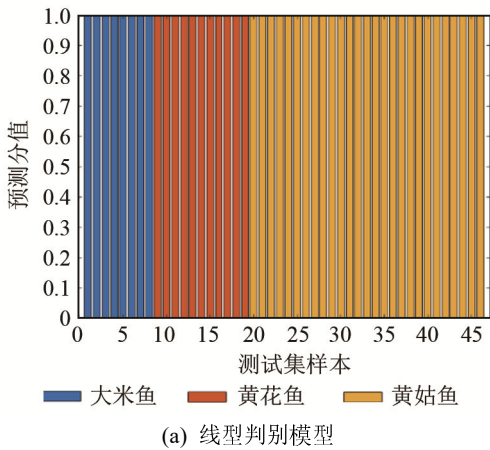


图 6 线性判别模型和随机子空间判别模型分类的预测得分
Fig.6 Prediction scores of linear discriminant classifier and random subspace discriminant classifier

表 2、3 中, 交换后的总体准确率用括号里的粗体数字表示。交换后的线性判别模型和随机子空

间判别模型在 46 个测试集样本中的预测得分分布如图 6 所示。图 6 是堆叠式的条形图, 每个条形段的高度代表判别模型输出的每个样本类别的预测得分, 每个样本所有类别的预测得分之和均为 1, 预测得分最高的类别即为该模型最终预测的类别。其中, 在横坐标对应的样本分布上, 前 8 个样本的真实类别是 大米鱼, 第 9~19 个是 黄花鱼, 后 27 个是 黄姑鱼。

通过对比 8 种分类器在测试集的总体准确率和预测得分分布可知, 基于小波包分解的频带能量特征在表 2 中的线性判别和随机子空间判别分类器表现出较好的分类性能。但通过进一步对比两者的预测得分分布, 得知随机子空间判别的预测得分稳定性相比线性判别较差, 故可靠性不如线性判别。此外, 表 3 中的三种分类器表现较差。将训练集和测试集交换后, 仍然可得出相同的分类结果。

3 总 结

本文基于小波包分解提取了黄花鱼、大米鱼和 黄姑鱼三种鱼发声信号的频带能量特征, 然后对该特征进行分类验证。研究结果表明: 三种鱼发声信号的主要频率范围在 300~800 Hz 之间; 在已有的样本数量和参数条件下, 将基于小波包分解的频带能量特征作为分类特征, 线性判别和随机子空间判别分类器在交叉验证集和测试集中均达到了 100% 的总体准确率, 而由于随机子空间判别相比线性判别预测得分不稳定, 故可靠性相比线性判别较差。即便如此, 基于小波包分解的频带能量特征在结合线性判别分类器和随机子空间判别分类器的情况下, 仍均可实现三种鱼发声信号的有效分类。

本次研究是在信噪比相对较高的网箱环境下对指定的三种石首鱼发声信号进行特征提取与分类研究, 没有考虑海洋环境噪声干扰, 也没有考虑不同鱼龄、不同鱼体行为和不同时间、不同海域的分类差异。这些内容将成为下一步研究的重点。

参 考 文 献

[1] 任新敏, 高大治, 姚玉玲, 等. 大黄鱼的发声及信号特性研究[J]. 大连水产学院学报, 2007, 22(2): 123-128.
REN Xinmin, GAO Dazhi, YAO Yuling, et al. Occurrence and characteristic of sound in large yellow croaker (*Pseudosciaena crocea*)[J]. Journal of Dalian Fisheries University, 2007, 22(2): 123-128.

[2] 王巍巍. 鱼类声信号的分析及特征提取研究[D]. 哈尔滨: 哈尔滨工程大学, 2009.
WANG Weiwei. Research of analysis and characteristics extraction on fish sound[D]. Harbin: Harbin Engineering University,

- 2009.
- [3] 刘贞文, 许肖梅, 覃柳怀. 大黄鱼发声信号特性研究[C]//2010年全国声学学术会议, 2010.
 - [4] WOOD M, CASARETTO L, HORGAN G, et al. Discriminating between fish sounds—a wavelet approach[J]. *Bioacoustics*, 2002, **12**(2-3): 337-339.
 - [5] STOLKIN R, RADHAKRISHNAN S, SUTIN A, et al. Passive acoustic detection of modulated underwater sounds from bio-logical and anthropogenic sources[C]//OCEANS 2007. Vancouver, BC, Canada. IEEE, 2007: 1-8.
 - [6] ROUNTREE R A, GILMORE R G, GOUDEY C A, et al. Listening to fish[J]. *Fisheries*, 2006, **31**(9): 433-446.
 - [7] 陈功, 王平波, 鲍玉军, 等. 希尔伯特-黄变换在微弱被动瞬态鱼声信号中的检测[J]. *海洋科学*, 2016, **40**(10): 91-96.
CHEN Gong, WANG Pingbo, BAO Yujun, et al. Research on the detection of weak transient passive fish acoustic signals based on Hilbert-Huang Transform[J]. *Marine Sciences*, 2016, **40**(10): 91-96.
 - [8] 陈功, 朱锡芳, 许清泉. 仿声技术在海洋鱼类被动声信号特征提取中的应用[J]. *海洋技术*, 2013, **32**(3): 50-54.
CHEN Gong, ZHU Xifang, XU Qingquan. Application of imitated voice technology on feature extraction from marine passive fish acoustic target[J]. *Ocean Technology*, 2013, **32**(3): 50-54.
 - [9] 杨露, 沈怀荣. 希尔伯特-黄变换与小波变换在故障特征提取中的对比研究[J]. *兵工学报*, 2009, **30**(5): 628-632.
YANG Lu, SHEN Huairong. Research and comparison on the application of Hilbert-Huang transform and wavelet transform to fault feature extraction[J]. *Acta Armamentarii*, 2009, **30**(5): 628-632.
 - [10] 张静远, 张冰, 蒋兴舟. 基于小波变换的特征提取方法分析[J]. *信号处理*, 2000, **16**(2): 156-162, 155.
ZHANG Jingyuan, ZHANG Bing, JIANG Xingzhou. Analyses of feature extraction methods based on wavelet transform[J]. *Signal Processing*, 2000, **16**(2): 156-162, 155.
 - [11] HAKAWA R, OGAWA T, HASEYAMA M, et al. Automatic detection of fish sounds based on multi-stage classification including logistic regression via adaptive feature weighting[J]. *J. Acoust. Soc. Am.*, 2018, **144**(5): 2709.
 - [12] MONTALBO F J P, HERNANDEZ A A. Classification of fish species with augmented data using deep convolutional neural network[C]//2019 IEEE 9th International Conference on System Engineering and Technology (ICSET). Shah Alam, Malaysia. IEEE, 2019: 396-401.
 - [13] OLSVIK E, TRINH C M D, KNAUSGÅRD K M, et al. Biometric fish classification of temperate species using convolutional neural network with squeeze-and-excitation[C]//Advances and Trends in Artificial Intelligence from Theory to Practice, 2019.
 - [14] LU Y C, TUNG C, KUO Y F. Identifying the species of harvested tuna and billfish using deep convolutional neural networks[J]. *ICES Journal of Marine Science*, 2020, **77**(4): 1318-1329.
 - [15] NODA J, TRAVIESO C, SÁNCHEZ-RODRÍGUEZ D. Automatic taxonomic classification of fish based on their acoustic signals[J]. *Applied Sciences*, 2016, **6**(12): 443.
 - [16] SHANG Y, LI J L. Study on echo features and classification methods of fish species[C]//2018 10th International Conference on Wireless Communications and Signal Processing (WCSP). Hangzhou, China. IEEE, 2018: 1-6.
 - [17] MALFANTE M, DALLA MURA M, MARS J I, et al. Automatic fish sounds classification[J]. *J. Acoust. Soc. Am.*, 2016, **139**(4): 2115-2116.
 - [18] KOTTEGE N, KROON F, JURDAK R, et al. Classification of underwater broadband bio-acoustics using spectro-temporal features[C]//WUWNet '12: Proceedings of the Seventh ACM International Conference on Underwater Networks and Systems. 2012: 1-8.
 - [19] LIN T H, TSAO Y, AKAMATSU T. Comparison of passive acoustic soniferous fish monitoring with supervised and unsupervised approaches[J]. *J. Acoust. Soc. Am.*, 2018, **143**(4): 278-284.
 - [20] 焦奎奎. 基于小波包变换的鱼类识别研究[D]. 大连: 大连理工大学, 2018.
JIAO Kuikui. Research on fish identification based on wavelet packet transform[D]. Dalian: Dalian University of Technology, 2018.
 - [21] RAMCHARITAR J, GANNON D P, POPPER A N. Bioacoustics of fishes of the family Sciaenidae (croakers and drums)[J]. *Transactions of the American Fisheries Society*, 2006, **135**(5): 1409-1431.