

DOI: 10.5846/stxb201511052249

蒋锦刚, 邵小云, 万海波, 齐家国, 荆长伟, 程天佑. 基于语谱图特征信息分割提取的声景观中鸟类生物多样性分析. 生态学报, 2016, 36(23): 7713-7723.

Jiang J G, Shao X Y, Wan H B, Qi J G, Jing C W, Cheng T Y. Bird diversity research using audio record files and the spectrogram segmentation method. Acta Ecologica Sinica, 2016, 36(23): 7713-7723.

基于语谱图特征信息分割提取的声景观中鸟类生物多样性分析

蒋锦刚¹, 邵小云¹, 万海波^{1,*}, 齐家国^{1,2}, 荆长伟¹, 程天佑¹

¹ 浙江大学, 海洋学院, 杭州 310058

² 密歇根州立大学, 全球变化与对地观测研究中心, 东兰辛市 48823

摘要: 声学手段是监测和研究生态系统生物活动规律、评价生态系统健康状况的一种新方法, 声景观生态学也是景观生态学的一个新兴研究领域。声景指数是描述复杂的音频数据生态学特征的有效方法, 但是, 单一的声景指数并不能有效的指示物种的真实丰度。在充分挖掘音频文件时频结构特征的基础上, 将遥感领域常用的面向对象图像分割技术引入语谱图分割, 并提出了适合于自然界鸟类生物多样性提取的知识规则和斑块统计分析方法。研究实验在杭州植物园的不同区域布点采集音频数据, 研究结果表明: 鸟类的多样性与地物景观类型和人类活动影响程度密切相关, 鸟类叫声的中心频率集中分布在 2.5—4.5kHz 之间, 最低频率分布在 0.67—2.1kHz 之间, 最高频率分布在 7.6—8.9kHz 之间, 人类活动较多的区域, 鸟类活动较少且叫声更短促。此外, 该方法提取的斑块面积周长比参数, 可以定量的反映鸟类叫声的婉转程度。

关键词: 声景观生态学; 音频文件; 傅立叶变换; 语谱图; 鸟类多样性

Bird diversity research using audio record files and the spectrogram segmentation method

JIANG Jingang¹, SHAO Xiaoyun¹, WAN Haibo^{1,*}, QI Jianguo^{1,2}, JING Changwei¹, CHENG Tianyou¹

¹ Ocean College, Zhejiang University, Hangzhou 310058, China

² Center for Global Change and Earth Observations, East Lansing, MI 48823, USA

Abstract: Soundscape ecology is an emerging area in landscape ecology. Soundscape characteristics can provide information that complements landscape characteristics. The study of soundscapes is based on an understanding of how sound from various biological, geophysical, and anthropogenic sources can be used to understand coupled natural-human dynamics across different spatial and temporal scales. The audio characteristics emitted from an ecosystem provide unique insights into the spatial and temporal patterns of ecosystem responses to human disturbances. The biological composition and the audio characteristics of different parts of the ecological environment, such as wetland, forest, and city parks. Furthermore, the soundscape changes throughout the day and throughout the seasons. Acoustic signal diversity, which refers to the patterns of frequency and temporal use of the acoustic spectrum, reflects biocomplexity and human activities. Previous research divided the soundscape into three primary components within the acoustic frequency spectrum, defined as biophony, anthrophony, and geophony. Biophony is the part of soundscape that consists of the sounds produced by all the organisms in a particular landscape. The biophony composition can be used to indicate community biodiversity. Several sound (or soundscape)

基金项目: 国家自然科学基金青年基金项目(41401404); 国家“千人计划”人才资助项目(188020-193810101/061); 浙江大学基本科研业务费资助项目(2014QNA4029)

收稿日期: 2015-11-05; 修订日期: 2016-04-25

* 通讯作者 Corresponding author. E-mail: hwan@zju.edu.cn

diversity indicators have been developed. In this study, we introduced an object-oriented classification approach for analyzing sound diversity using bird communities as a case study. The bird sound characteristics were used to extract a patch index that included the number of patches, the frequency range and standard deviation, and the area-to-perimeter ratio. This approach was able to distinguish between the soundscape scenarios recorded at different locations in Hangzhou Botanical Garden, China. The results showed that bird diversity was closely related to the particular features of different landscapes. The central frequency for birds ranged between 2.5 and 4.5 kHz; the lowest frequency was between 0.67 and 2.1 kHz, and the highest frequency between 7.6 and 8.9 kHz. The stronger the human interference, the shorter the bird songs and calls. In addition, bird sound characteristics quantitatively reflected the complexity of the bird songs and calls using the patch area-to-perimeter ratio parameter.

Key Words: soundscape ecology; audio record files; Fourier transform; spectrogram; bird diversity

声景是指自然景观中所有声音的综合,声景观生态学是景观生态学的一个新兴研究领域,可以为景观生态的综合特征提供重要的参考信息,同时也是生态健康的重要指示因子^[1]。声景观生态学研究的的主要内容是研究不同时空尺度下,自然声响、人类声响、生态环境声响的耦合与表现特征^[2],并从声音的组成、时空分布、时间变化特征来揭示生态环境的多维格局和变化过程。20世纪60年代晚期,加拿大的 R. Murray Schafer 创立了一个以教育和科研为目的的“世界声景计划(World Soundscape Project, WSP)”团队,首次提出声景观生态学的概念,倡导人类与声环境的和谐,推广环境声音景观和噪音污染方面的知识^[3]。传统的声景观研究与声景观生态学研究存在较大的差异,传统的声景观研究是从生活噪声、环境心理学和景观规划的角度为噪声控制、城市规划和建筑设计服务^[4-5]。声景观生态学研究主要从“人类-自然”生态系统声音频率的结构特征、空间分布来进行分析,并认为音频构成来自于生物活动、地球物理过程和人类活动的综合。在相关研究中,声音频率的范围被分为3个主要组成部分,生物频谱即:来自于自然界的发声生物如鸟类、昆虫、青蛙等;地球频谱即:来自于地球表面过程如河流流水、海洋潮汐、沙丘移动、风云雷电等;人类频谱即:来自于人类活动如车辆行驶、道路建设、人们行走等^[1-2,6]。显然,不同的区域(如湿地、森林,或城市公园),生态环境、生物组成和其音频特征也是不同的。与此同时,声景观生态特征还存在日变化和季节变化的特征。目前,声景观生态学的新理论和方法,已被广泛应用于生态环境的各种领域,如城市规划、环境监测与管理、生态系统健康评估、生物多样性及生物的活动行为规律等^[7-15]。

生物频谱作为声景观理论模型中的3个重要组成部分之一,利用声音频谱分析手段进行生物多样性的监测,在国内外已有一定的研究。Klaus Riede^[16]早在20世纪90年代年就开始利用声学手段来监测亚马逊雨林的生物多样性特征,为生物多样性的声学调查开创了先河,国内较早开展相关研究的姜仕仁教授^[17-18]团队在本世纪初开始积极探索利用声学手段研究昆虫等生物群落结构和多样性特征。同时,各种声景指数是描述复杂的音频数据生态学特征的有效方法^[19]。不同的学者从不同的研究角度提出的声景特征指数包括:声音复杂度指数(ACI)^[20]、归一化声景观指数(NDSI)、声音多样性指数(ADI)、声音均匀性指数(AEI)等^[21]。Towsey等^[22]人在研究鸟类的多样性和丰度时,选择了14种声景指数,研究发现,单一的声景指数并不能有效的指示物种的真实丰度,几种指数的有效组合才能全面的表现生物多样性和丰度。

不同声景指数的提出,针对的通常是声景结构或功能的某一个方面,具有一定的局限性。语谱图是一种描述声音时间-频率-能量谱密度变化的图像,是声音分析的重要方法,在以往基于频谱分析的研究中,分析手段主要为统计分析方法,分析的目标侧重在频谱特征的复杂性和差异性上,而在图像的时间-频率结构特征分析上的方法相对较少。面向对象分类技术是遥感图像分析领域的一种新的分类方法,是图像纹理和几何结构信息提取的有效方法,同时,该方法可以融合先验知识要素参与信息提取。为了充分挖掘音频文件的时频结构特征,本文在音频记录文件语谱图绘制的基础上,将遥感领域常用的面向对象图像分割技术引入语谱图分割与图像几何结构信息提取,并提出了适合于自然界鸟类生物多样性提取的知识规则和斑块统计分析聚类方

法。研究实验选择 2015 年春季在杭州植物园的不同景观区域布点采集音频数据,通过与“载频-周期图”法对比分析表明:该方法对于利用音频记录文件研究和分析自然界生物特别是鸟类多样性和活动规律是非常有效的,可以为生态系统多样性的监测、研究和分析提供了一种新的方法。

1 时频结构特征提取流程与关键算法

1.1 时频结构特征提取流程

生物频谱、地球频谱、人类频谱是声景分析和划分的最为经典模型,其划分的物理机理是频率特征分布的差异,其中,生物声响的主要频率分布范围在较高频率区间^[1-2]。图 1 是一个典型的含有生物声响的音频记录文件波形图和语谱图,可以看出声音频率和能量的分布主要集中在高频区间,符合上述频率特征描述。

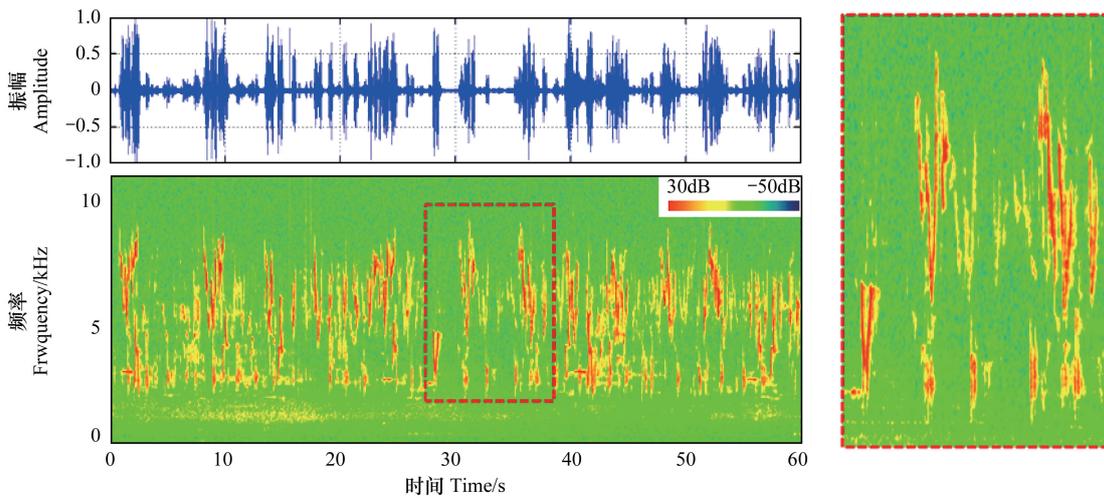


图 1 声音信号的波形图和语谱图特征

Fig.1 The waveform and spectrogram of an acoustic signal

语谱图的精细分析是挖掘声景信息和生物多样性特征最为关键的手段,本文提出的一套基于时频结构特征提取的生物多样性分析方法流程如图 2 所示。

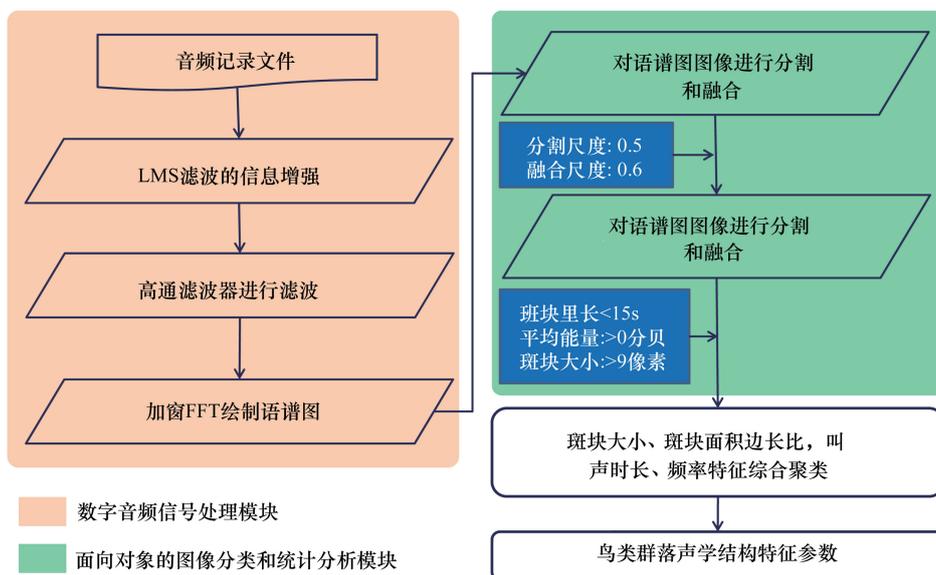


图 2 声景时频结构特征提取流程图

Fig.2 The flow chart of soundscape time-frequency structure's extraction

从图中可以看出该流程主要分为 2 个关键处理模块,即:数字音频信号处理模块和面向对象的图像分类和统计分析模块。数字信号处理模块中主要涉及的算法有 LMS 音频信息滤波增强方法、高通滤波方法和加窗傅里叶变换(FFT)的语谱图绘制方法;面向对象的图像分类和统计分析模块主要借助遥感图像处理中的图像分割、融合技术提取蕴含生物信息的特征斑块,并结合区域统计手段得出生物信息的数据统计特征,具体统计指标包括斑块大小、斑块面积周长比、叫声时长、频率特征等。根据上述生物学信息,采用层次聚类算法确定典型鸟类种群类型数目与特征。

1.2 自适应 LMS 滤波和高通滤波算法

音频记录文件的噪声干扰普遍存在且不可避免,噪声的存在给音频信号中“人类—自然”系统中的各种声音信息的理解带来了困难。音频信号的滤波增强技术就是在保证减少特征声音失真度的同时,尽可能多地提取出有效的声音信号,抑制背景噪声。自适应滤波法是目前广泛应用的一种音频信息滤波增强方法^[23],而自适应滤波算法中的最小均方误差算法(LMS)以其运算简单、稳健且易于实现等优点成为了自适应滤波技术的首选算法^[24]。假设 $v(k)$ 为原始含噪声的声音信号, $x(k)$ 为 $v(k)$ 延时后的参考信号,算法的计算迭代过程如下:

$$y(k) = X^T(k) \times W(k) \quad (1)$$

$$e(k) = v(k) - y(k) \quad (2)$$

$$W(k+1) = W(k) + 2\mu \times e(k) \times X(k) \quad (3)$$

式中, $X(k) = [x(k), x(k-1), \dots, x(k-n+1)]^T$ 为 n 阶滤波器在 k 时刻的参考输入, $y(k)$ 为滤波器的估计输出, $W(k) = [w(k), w(k-1), \dots, w(k-n+1)]^T$ 对应滤波器权重系数矢量, μ 为步长因子,对 LMS 算法性能的优劣起着决定性作用,可通过调整步长因子来改善 LMS 算法性能。要使得该算法收敛,应满足如下条件:

$$0 < \mu < \frac{1}{\lambda_{\max}} \quad (4)$$

式中, λ_{\max} 为输入信号的自相关矩阵 R 的最大特征值,对于 μ 的取值,本文参考了文献^[24]提出的方法。

由于采集音频文件的信息构成主要为生物活动、地球物理过程和人类活动的综合,文件中必然包含多种声源信息,因此,剔除不相关信息的干扰,也是数字音频信号处理模块必须要考虑的一块内容。根据前文叙述,鸟类等生物频谱多分布在高频区,而地球物理过程和人类活动的频谱多分布在低频区。这里在数字音频信号处理模块中加入了高通滤波器模块,将地球物理过程和人类活动的信息成分滤除。高通滤波器是容许高频信号通过、但减弱(或减少)频率低于截止频率信号通过的滤波器。对于数字音频信号,高通滤波在时域中的数理解释如下:

$$y(n) = x(n) - f(n) \quad (5)$$

$$f(n) = \frac{1}{M}x(n) + \left(1 - \frac{1}{M}\right)f(n-1) \quad (6)$$

式中, M 为根据截止频率而预先设定的时间常量, $x(n)$ 为第 n 帧的输入, $f(n)$ 为第 n 帧的低通输出, $y(n)$ 为校正后的输出。

1.3 语谱图绘制算法

“人类-自然”系统中的各种声音具有不同的频率特征,声景生态定量分析的关键取决于声波频谱结构的精细提取。语谱图是一种描述声音时间-频率-能量谱密度变化的图像,是声音分析的重要方法,由 *W. Koenig* 于 1946 年提出^[25]。绘制语谱图的关键算法包括傅立叶变换、短时傅立叶变换和快速傅立叶变换,傅立叶变换是分析平稳信号稳态特性的强有力手段,短时傅立叶变换就是在短时平稳的假设下,用稳态分析方法处理非平稳信号的一种方法。语音信号是一种典型的非平稳信号,但是在极短的时间内(10—30ms)内声音信号是符合平稳性假设的。短时傅立叶变换是一种基于加窗的傅立叶变换,通过窗口函数 $w(n)$ 将分析信号划分

成很多小的间隔(也叫帧)^[26],对于信号 $x(n)$ 的短时傅立叶变换可以被定义为以下公式^[27-28]。

$$X_n(e^{j\omega k}) = \sum_{m=-\infty}^{+\infty} w(n-m)x(m)e^{-j\omega km} \quad (7)$$

式中, $w(n)$ 为窗序列, m 是帧同步的时间序列号, 信号 $x(n)$ 的离散时域傅立叶变换为:

$$X(n, e^{j\omega}) = \sum_{m=0}^{N-1} x_n(m)e^{-j\omega m} \quad (8)$$

式中, N 为帧长(一帧内的采样点数), 进一步进行变换可得:

$$\begin{cases} X(n, k) = \sum_{m=0}^{N-1} x_n(m)e^{-j\frac{2\pi km}{N}} \\ 0 \leq k \leq N-1 \end{cases} \quad (9)$$

式中, $|X(n, k)|$ 就是 $x(n)$ 的短时幅度谱估计, 而时间 m 处的频谱能量密度函数 $P(n, k)$ 为:

$$P(n, k) = |X(n, k)|^2 = (X(n, k)) \times (\text{conj}(X(n, k))) \quad (10)$$

将 $P(n, k)$ 的值表示为灰度级所构成的二维图像就是语谱图, 语谱图的 dB 表示方法是通过 $10 \log_{10}(P(n, k))$ 变换得到的, 而语谱图声压级分贝值则是通过如下计算方法得到。

$$L_p(k) = 20 \times \log(P(k)/P_0) \quad (11)$$

式中, $P(k)$ 为 k 中心频段短时傅立叶后的谱密度, P_0 为参考声压, 空气中为 $20 \mu\text{pa}$ 。由于本文实验仪器采集的原始数据信息为声压信号, 为了便于理解, 语谱图的能量密度信息采用声压级分贝值结果。

1.4 面向对象的语谱图分割与特征斑块提取方法

面向对象分类技术是遥感图像分析领域的一种新的分类方法, 该方法面向的分类目标不是单个的图像像素, 而是具有相似特性的斑块单元, 而这一特点正好符合生物的声响特征在语谱图上的表现。面向对象图像分类技术的关键算法包括图像分割和融合算法、特征选择分析、分类规则和分类算法等,

图像分割和融合算法是将整个图像区域根据同质性和异质性标准分割成若干互不交叠的非空子域的过程, 同一区域内部在灰度、纹理、颜色、形状等上面具有相同或相似的特性。图像分割可以借助数学集合的概念进行理解^[29]。假设图像 I 中具有相同属性的条件表示为 H , 图像分割是将图像 I 分割成 n 个区域 $R_i, i = 1, 2, \dots, n$, 要求满足:

$$\begin{cases} \bigcup_{i=1}^n R_i = I, R_i \cap R_j = \Phi, \forall i, j, i \neq j \\ \forall i, j, = 1, 2, \dots, n, H(R_i) = \text{True} \\ \forall i, j, i \neq j, H(R_i \cup R_j) = \text{False} \end{cases} \quad (12)$$

上式所表述的含义是分割后的图像要求全覆盖且互不重叠, 相同区域的像元属性相似, 不同区域斑块的像素属性相异。图像分割的算法目前有成百上千种, 新的算法也在不断被提出, 从图像分割的数学原理角度, 图像分割可以分为阈值分割、边缘检测、区域灰度、形状、纹理分割等类别。由于图像分割算法多样, 这里将不详细介绍具体分割融合算法, 本文的分割融合过程的实现是在 ENVI 5.0 的 Feature Extraction 模块下实现的。

图像分割是进行语谱图特征斑块提取的第一步, 对特征斑块提取的关键是知识规则的确定, 对于声音文件语谱图信息中蕴含的生物叫声、特别是鸟类叫声的信息, 本文在结合文献^[30-31]和实际数据综合分析的基础上提出的知识规则如下:

鸟类的叫声是一种非常重要的鸟类行为学特征, 不同的鸣叫声具有不同的行为意义^[30], 文献^[31]中对鸟类求偶鸣唱句子时长记录结果表明, 其时长可达 7—8 s, 考虑到有些善鸣唱的鸟类的叫声可能持续十几秒, 规定特征斑块的时长 $< 15\text{s}$;

鸟类等生物的叫声, 能够被传感器监测并记录, 是具有一定的能量信息, 根据人耳正常听到的轻微声音分贝值为 30—40 分贝, 规定斑块的平均能量 > 30 分贝;

由于采集仪器会受到外界电磁环境的干扰,语谱图图像中会存在电磁等噪声信息,因此对斑块大小进行条件设置,规定斑块大小>9个像素,这里的像素信息的物理含义是时间频率的分辨率,与傅立叶分析是参数有关,本文分析时像素时长为 15ms,像素频率带宽为 43Hz。

提取出来的斑块数据斑块大小,斑块形状复杂程度等等属性,如何将斑块的几何属性与鸟类生物的叫声特征联系起来是定量分析鸟类生物多样性需要解决的问题之一,本文提取分析的斑块属性包括:斑块大小、中心频率、频率分布特征(分布范围)、面积周长比、叫声时长等,结合图 3 来进一步说明各斑块的几何属性对应的生物特征信息。

从图 3 中可以看出,提取斑块的几何属性对应的鸟类叫声生物学特性,包括声音频率分布、叫声长短的特点,此外提取斑块的数目是鸟类生物叫声和多样特性的最直接反映,面积周长比是斑块复杂程度的参数,参数值越小,对应鸟类叫声婉转丰富。就图 3 而言,斑块 b 的面积周长比要小于斑块 a,从图中可以看出 b 斑块的复杂程度要大于 a 斑块,鸟类的叫声也更婉转。

不同鸟类鸣叫声因种类而有所差异,在前文的鸟类典型叫声信息提取的基础上,如何将不同鸟类叫声信息进行凝炼,得到不同鸟类的种群信息,本文在上述处理的基础上,提出了利用层次聚类方法进行鸟类叫声种的聚类。聚类分析的实质是建立一种分类方法,它能够将一批样本数据按照他们在性质上的亲密程度在没有先验知识的情况下自动进行分类,以便对不同类的样本进行分析。本文层次聚类主要运用系统聚类方法,其基本思想是:首先,将要分类的 n 个变量各自看作一类,然后分别计算各个个体之间的对应程度或密切程度。可以用两种方式来测量:1)采用描述个体对(变量对)之间的接近程度的指标,例如距离,距离越小的个体(变量)越具有相似性。2)采用表示相似程度的指标,例如相关系数,相关系数越大的个体(变量)越具有相似性。

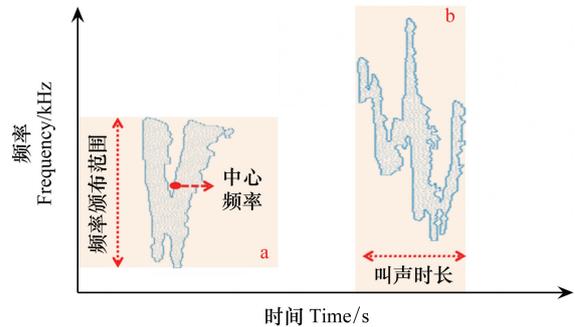


图 3 斑块几何属性与生物特征信息对应示意图

Fig.3 Corresponding schematic of patch geometric properties and biometric information

2 试验及结果对比分析

2.1 试验研究区、仪器及材料选择处理

杭州植物园位于杭州市西湖区桃源岭,区域占地 115.6 hm^2 ,根据不同的功能可分为:观赏植物区(专类园),植物分类区,经济植物区,森林公园。观赏植物区由木兰山茶园、杜鹃槭树园、桂花紫薇园、桃花园、灵峰梅园、百草园、山水园、竹类植物区等 8 个专类园组成,园内绿树成荫,良好的生态环境使得园内鸟类多样,为本文方法的研究提供了良好的实验数据采集场地。图 4 是本文选择的植物园内 6 个典型样点,其中 A、B 点选择为植物园内沿道路较近的区域,C、D、E 点为植物分类区的林地区域,F 点为经济植物区的林地区域。

实验记录声音文件的仪器选择美国 Wildlife Acoustics 公司生产的 Song Meter SM2+ 声音记录仪器。该公司是专门从事野外生物声音采集、记录和分析的仪器公司,选择的 SM2+ 款仪器是专门采集记录鸟类和陆

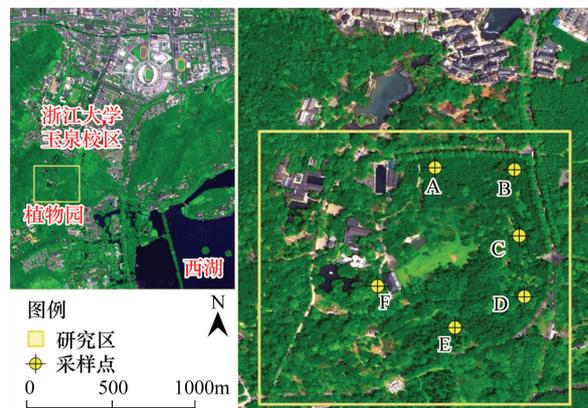


图 4 采样点点位分布图(影像来自 IKONOS,空间分辨率 1m)

Fig.4 Location map of sampling sites (Image from IKONOS satellite data, and the spatial resolution is 1m)

A、B 点选择为植物园内沿道路较近的区域,C、D、E 点为植物分类区的林地区域,F 点为经济植物区的林地区域

地生物叫声的仪器,可记录 2 路声音文件,采样频率最高可达 96kHz,数据分辨率 16bit。考虑到鸟类晨间鸣叫是鸟类的重要行为之一,实验采集选择时间为 2015 年 4 月 1—14 日 5:00—7:00,每个地点在每天整点时刻开始自动同步采集记录声音数据,每次采集文件时长为 5 min,实验共收集文件个数为 252 个,采集声音总时长为 1260 min。

对于选择的记录音频文件材料,按照前文叙述的处理流程和关键算法进行处理计算。数字信号处理模块中的 LMS 滤波增强算法、高通滤波算法和基于傅里叶变换的语谱图绘制是在 MATLAB 平台下编程实现,其中,高通滤波的截止频率的设定为 300Hz^[32],绘制的语谱图数据保存为 TIFF 数据格式。面向对象的特征斑块提取是在 ENVI5.0 的 Feature Extraction 模块下实现的,特征斑块的信息统计是在 ArcGIS10.1 平台下进行的。典型斑块图斑信息的层次聚类也是在 MATLAB 平台下实现的,其中聚类密切程度的指标选择相关系数。

2.2 与“载频-周期图”法对比分析

“载频-周期图”法是由 Klaus Riede 最早提出来方法,用于研究鸟类的生物多样性特征,为了对比说明本文提出的方法在鸟类生物多样性提取和分析上的有效性,这里将本文的方法和“载频-周期图”法进行了对比试验分析。载频-周期图”法的处理过程需要通过人为干预,首先需要绘制语谱图,然后结合语谱图的结构特征反复聆听记录的声音文件信息,找出能够描述特定鸟类叫声的载频和周期信息,通过分析载频-周期的数据信息以及绘制载频-周期图来研究生物多样性的规律。

为了直观的展现两种方法的有效性,试验选取了一段含有丰富鸟类叫声的 1min 音频文件片段,对该声音文件进行绘制语谱图,及反复聆听得该声音文件中含有 9 中鸟类的叫声(图 5),利用“载频-周期图”法得到的鸟类群落结构的信息如图 6(a)所示,图 6(b)为典型的布谷鸟叫声的频率-周期特征,其频率分布集中在 692Hz 附近,叫声时长在 1.05s 附近。

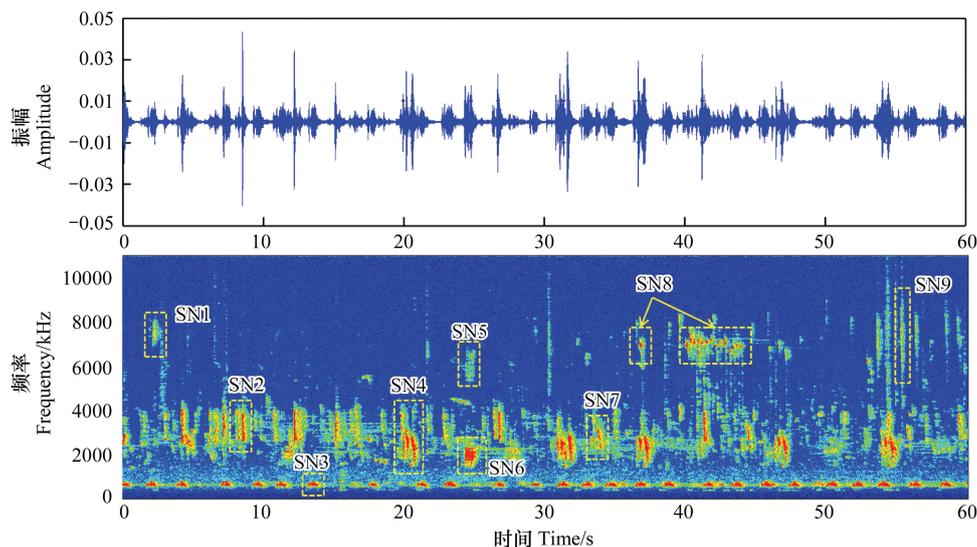


图 5 对比试验声音片段数据波形图、语谱图及典型鸟类叫声时频区域框选

Fig.5 The waveform and spectrogram of an acoustic signal and typical bird calls

图 7 是本文方法得到的鸟类多样性的结果,从图中的结果可以看出通过语谱图的斑块信息提取,特征信息的层次聚类,共得到特征鸟类的信息 10 种,各种鸟类的斑块分布图如图 8 所示。

表 1 是两种不同方法提取的鸟类群落信息的频率及鸣叫时长结果,从表中可以看出两种方法在鸟类的种类特征上存在一个种类差异,即:“载频-周期图”法得到的鸟类群落数为 9 中,本文方法得到的鸟类群落数为 10 种,两者的相对误差在 12.5%。在中心频率提取的最大误差体现在 SN4,相对误差 3.6%,鸣叫时长最大误差体现在 SN9,相对误差为 7.1%。

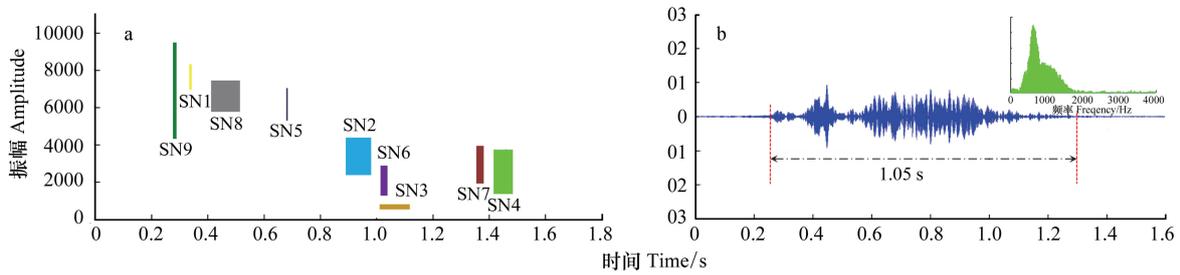


图 6 基于载频-周期图的鸟类群落结构绘制结果

Fig.6 The figure of bird community structure based on frequency-periodogram method

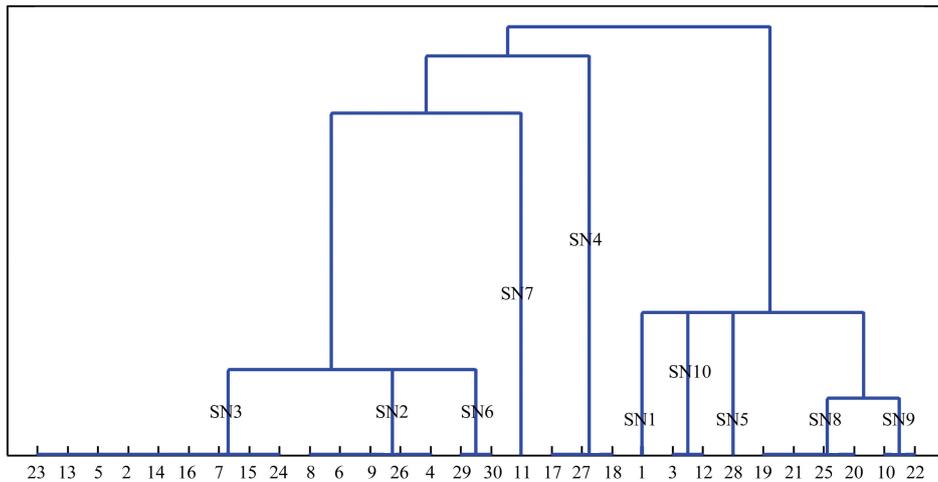


图 7 鸟类群落结构层次聚类结果

Fig.7 The figure of bird community structure using hierarchical clustering method

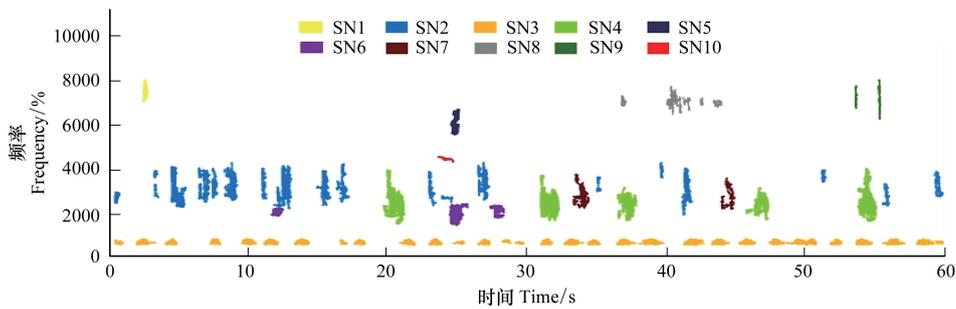


图 8 基于语谱图分割的鸟类群落结构绘制结果

Fig.8 The figure of bird community structure using spectrogram segmentation method

2.3 试验区鸟类多样性初步分析

图 9 选取的是 6 个采样点典型声音片段 3 维语谱图图像,从图中可以看出 6 个采样点中,A、D、E、F 采样点中的声音强度和斑块形状特征最为丰富,反映的生物多样性信息也最为突出,鸟类叫声此起彼伏,B、C 两点中的生物叫声相对较少,在低频段能量信息较突出,含有人类活动的信息成分。

为了进一步对比分析各采样点的生物多样性信息的综合特征,表 2 给出了 6 个典型采样点的鸟叫声斑块信息提取和统计结果。从表中可以看出,6 个实验样点中,提取出来的鸟类种群数量最多的是 E 点,最少的是 C 点,反映了各采样点鸟类多样性丰度的顺序为 E>F>D>A>B>C,其中 E 采样点为植物分类区的林地区域,该

处主要为林地地物类型,且离道路和人类活动区域较远,是鸟类典型的栖息地。各采样点的中心频率分布在 2.5—4.5kHz 之间,最低频率分布在 0.67—2.1kHz 之间,最高频率分布在 7.6—8.9kHz 之间。斑块的面周长比在 3.33—4.46 之间,面积周长比越小,反映的是提取斑块的形状越复杂,即生物叫声的规律越丰富,实验结果表明 E 点的鸟类叫声更婉转动听。统计斑块的叫声时长在 0.43—0.57s 之间,不同的时长反映的是不同鸟类叫声的长短,6 个采样点的统计结果表明,F 点鸟类生物叫声时间较长,C 点生物叫声时间较短,本实验的 F 点为经济植物区的林地区域,而 C 点离道路较近,生物叫声的长短与所处的景观类型和人类干扰程度存在一定关系。

表 1 两种不同方法提取的鸟类群落信息频率及叫声时长结果

Table 1 Comparison of two methods for extracting bird diversity information

序号 Sample number	“载频-周期图”法 Frequency-periodogram method				本文方法 Spectrogram segmentation method			
	中心频率/Hz Center frequency	最低频 率/Hz Minimum frequency	最高频 率/Hz Maximum frequency	鸣叫时 长/s Bird calls duration	中心频 率/Hz Center frequency	最低频 率/Hz Minimum frequency	最高频 率/Hz Maximum frequency	鸣叫时 长/s Bird calls duration
SN1	7635	6983	8287	0.34	7569	7012	8127	0.32
SN2	3422 ± 212	2335 ± 55	4472 ± 142	0.94 ± 0.03	3347 ± 908	2269 ± 103	4228 ± 411	0.97 ± 0.11
SN3	692 ± 12	556 ± 14	827 ± 27	1.05 ± 0.05	679 ± 11	561 ± 11	833 ± 34	1.06 ± 0.07
SN4	2608 ± 29	1444 ± 51	3782 ± 36	1.43 ± 0.02	2513 ± 64	1546 ± 33	3699 ± 142	1.41 ± 0.04
SN5	6092	5281	6903	0.68	6199	5487	6911	0.66
SN6	2152 ± 77	1382 ± 47	2922 ± 41	1.02 ± 0.01	2044 ± 73	1364 ± 86	2724 ± 33	1.03 ± 0.02
SN7	2971 ± 53	1989 ± 33	3953 ± 28	1.36 ± 0.01	2996 ± 41	2024 ± 13	3968 ± 57	1.37 ± 0.02
SN8	6640 ± 173	5821 ± 63	7460 ± 233	0.44 ± 0.03	6895 ± 224	6123 ± 74	7667 ± 286	0.45 ± 0.07
SN9	6979 ± 199	4522 ± 413	9436 ± 364	0.28 ± 0.005	7184 ± 111	5433 ± 214	8935 ± 161	0.26 ± 0.005
SN10	—	—	—	—	4352	4265	4439	0.31

表 2 6 个典型采样点的生物斑块信息提取和统计结果

Table 2 Patch properties for example soundscape clips at different sites

点位 Sites	种群个数 Population size	中心频率/kHz Center frequency	最低频率/kHz Minimum frequency	最高频率/kHz Maximum frequency	平均面积周长比 Mean area- perimeter ratio	平均时长/s Mean bird calls duration
A	27	3.6	1.3	7.6	3.82	0.48
B	23	2.9	1.2	8.7	3.64	0.56
C	22	2.5	2.1	7.6	4.47	0.43
D	31	4.5	0.69	8.7	4.46	0.44
E	38	3.9	0.67	8.9	3.33	0.52
F	35	4.1	0.69	8.8	3.51	0.57

3 讨论与结论

随着数字录音、通讯和计算机分析技术的不断提高,声学手段成为了研究人类-自然生态系统多样性,生物行为活动规律的一种特殊手段。如何充分挖掘声音记录文件的内在信息,为分析研究生态系统的生物多样性特征和生物活动规律服务,是声音文件分析研究的重要内容。本文提出的基于音频文件时频结构特征提取的声景观中鸟类生物多样性分析方法,通过将面向对象图像分割技术引入语谱图分割,并提出了适合于自然界生物多样性特别是鸟类叫声特征提取的知识规则和斑块统计分析方法,与经典的“载频-周期图”法进行对比试验结果表明,该方法在提取的种群数目及种群参数信息上与经典方法保持了较好的一致性,新方法结合了传统的数字信号的处理手段和先进的数字图像处理技术,提出的信息参数更加丰富多样。同时本文实验采

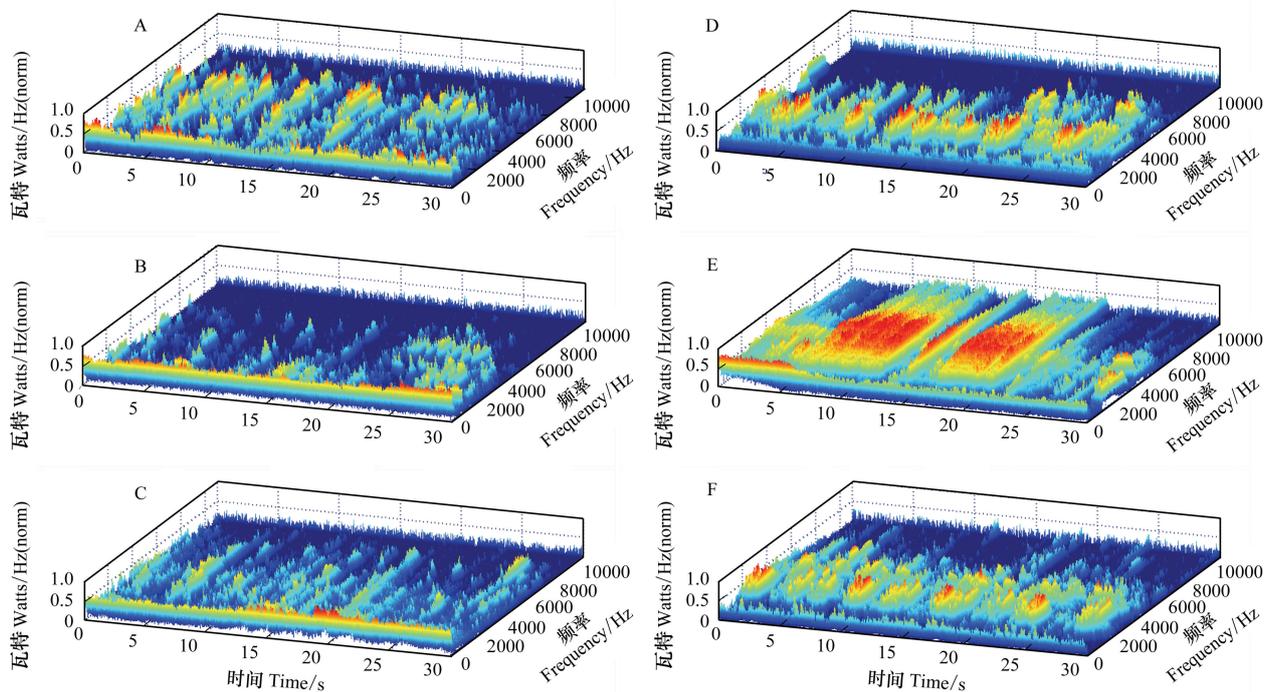


图9 不同点位典型声音片段的3维语谱图

Fig.9 The 3D spectrograms of example soundscape clips at different sites

样数据分析结果佐证了前人的研究成果,即:本文分析数据结果进一步揭示杭州植物园春季鸟类叫声的中心频率集中分布在 2.5—4.5kHz 之间,最低频率分布在 0.67—2.1kHz 之间,最高频率分布在 7.6—8.9kHz 之间。根据监测与分析结果,其鸟类种群的数目在 38 种左右。

提取的鸟类种群数目是鸟类多样性或丰度的直接证据,鸟类的多样性与地物类型密切相关。实验中提取的斑块数目表明鸟类最多且最活跃的区域是 E 点,该处地物类型为林地,且离人类活动的区域较远。斑块面积周长比这个参数,反映的是生物叫声的复杂程度,面积周长比越小,其提取斑块的形状越复杂,生物叫声的规律越婉转丰富。通过实验采集样点数据的分析结果可以看出,E 点的鸟类叫声更婉转动听。斑块时长反映的是不同鸟类叫声的长短,采样点的统计结果表明,生物叫声的长短与所处的景观类型和人类干扰程度存在一定关系,人类活动干扰大的区域,鸟类活动较少,叫声时长也更短,其可能原因是鸟类路过该区域或临时休憩时发出的叫声。

本文提出的基于音频文件时频结构特征图像分割提取与鸟类多样性分析方法,主要包括数字音频信号处理、面向对象的图像分类和统计分析两个主要处理模块,其中数字音频信号处理模块采用的是经典的音频数字处理技术,处理过程中人为干预较少,人为因素对后期的结果影响较小。面向对象的图像分类和统计分析模块的处理中人为参与的过程较多,如分割融合尺度的设定,特征斑块提取的知识规则等很多经验性的因素都有可能对结果造成影响。然而,提取的生物叫声时间-频率特征斑块信息,具有更多的统计分析指标,包括种群数目、中心频率、斑块叫声的时间长度、斑块的复杂程度等,可以更加全面的分析生物叫声的综合特征,为利用声学手段研究自然生态的规律和特征提供了一种好的方法。

参考文献 (References):

- [1] Qi J G, Gage S H, Joo W, Napoletano B, Biswas S. Soundscape characteristics of an environment: a new ecological indicator of ecosystem health// Ji W, ed. Wetland and Water Resource Modeling and Assessment. New York, USA: CRC Press, 2008: 201-211.
- [2] Pijanowski B C, Farina A, Gage S H, Dumyahn S L, Krause B L. What is soundscape ecology? An introduction and overview of an emerging new science. *Landscape Ecology*, 2011, 26(9): 1213-1232.

- [3] Schafer R M. The New Soundscape. Toronto; BMI Canada Limited, 1969.
- [4] 葛坚, 赵秀敏, 石坚韧. 城市景观中的声景观解析与设计. 浙江大学学报: 工学版, 2004, 38(8): 994-999.
- [5] 宋剑玮, 马蕙, 冯寅. 声景观综述. 噪声与振动控制, 2012, (5): 16-20.
- [6] Gage S H, Napoletano B M, Cooper M C. Assessment of ecosystem biodiversity by acoustic diversity indices. Journal of the Acoustical Society of America, 2001, 109(5): 2430-2430.
- [7] Jennings P, Cain R. A framework for improving urban soundscapes. Applied Acoustics, 2013, 74(2): 293-299.
- [8] Raimbault M, Dubois D. Urban soundscapes: experiences and knowledge. Cities, 2005, 22(5): 339-350.
- [9] Liu J, Kang J, Luo T, Behm H, Coppack T. Spatiotemporal variability of soundscapes in a multiple functional urban area. Landscape and Urban Planning, 2013, 115: 1-9.
- [10] Rodriguez A, Gasc A, Pavoine S, Grandcolas P, Gaucher P, Sueur J. Temporal and spatial variability of animal sound within a neotropical forest. Ecological Informatics, 2014, 21: 133-143.
- [11] Farina A, Pieretti N, Piccioli L. The soundscape methodology for long-term bird monitoring; a Mediterranean Europe case-study. Ecological Informatics, 2011, 6(6): 354-363.
- [12] Farina A, Pieretti N. The soundscape ecology: a new frontier of landscape research and its application to islands and coastal systems. Journal of Marine and Island Cultures, 2012, 1(1): 21-26.
- [13] 李淑玲, 包军, 王文峰, 白晓杰, 崔卫国. 丹顶鹤性活动的声行为研究. 生态学报, 2004, 24(3): 503-509.
- [14] 陈敏, 冯江, 李振新, 周江, 赵辉华, 张树义, 盛连喜. 普氏蹄蝠 (*Hipposideros pratti*) 回声定位声波、形态及捕食策略. 应用生态学报, 2002, 13(12): 1629-1632.
- [15] 扈军, 葛坚, 李东浩. 基于 GIS 的声景观地图制作与分析—以杭州柳浪闻莺公园为例. 浙江大学学报: 工学版, 2015, 49(7): 1295-1304.
- [16] Riede K. Monitoring biodiversity: analysis of Amazonian rainforest sounds. Ambio, 1993, 22(8): 546-548.
- [17] 姜仕仁. 杭州植物园秋季夜间昆虫群落的声音监测. 浙江科技学院学报, 2003, 15(2): 72-76.
- [18] 姜仕仁, 陈水华. 同一生境中强脚树莺鸣声的个体差异及多样性. 动物学研究, 2006, 27(5): 473-480.
- [19] Kasten E P, Gage S H, Fox J, Joo W. The remote environmental assessment laboratory's acoustic library: an archive for studying soundscape ecology. Ecological Informatics, 2012, 12: 50-67.
- [20] Pieretti N, Farina A, Morri D. A new methodology to infer the singing activity of an avian community: the Acoustic Complexity Index (ACI). Ecological Indicators, 2011, 11(3): 868-873.
- [21] Villanueva-Rivera L J, Pijanowski B C, Doucette J, Pekin B. A primer of acoustic analysis for landscape ecologists. Landscape Ecology, 2011, 26(9): 1233-1246.
- [22] Towsey M, Wimmer J, Williamson I, Roe P. The use of acoustic indices to determine avian species richness in audio-recordings of the environment. Ecological Informatics, 2014, 21: 110-119.
- [23] 陈素芝, 李英. 一种基于变步长 LMS 算法的语音增强方法. 声学技术, 2005, 24(1): 42-45.
- [24] 王瑜琳, 田学隆, 高雪利. 自适应滤波语音增强算法改进及其 DSP 实现. 计算机工程与应用, 2015, 51(1): 208-212.
- [25] Koenig W, Dunn H K, Lacy L Y. The sound spectrograph. The Journal of the Acoustical Society of America, 1946, 18(1): 19-49.
- [26] 李振春, 刁瑞, 韩文功, 刘力辉. 线性时频分析方法综述. 勘探地球物理进展, 2010, 33(3): 239-246.
- [27] Schafer R, Rabiner L. Design and simulation of a speech analysis-synthesis system based on short-time Fourier analysis. IEEE Transactions on Audio and Electroacoustics, 1973, 21(3): 165-174.
- [28] Allen J B, Rabiner L R. A unified approach to short-time Fourier analysis and synthesis. Proceedings of the IEEE, 1977, 65(11): 1558-1564.
- [29] 张德丰. 数字图像处理 (MATLAB 版). 北京: 人民邮电出版社, 2009: 250-260.
- [30] 伟利国, 张小超, 吴文彪. 鸟鸣叫声特征提取方法的研究. 辽宁师范大学学报: 自然科学版, 2010, 33(1): 108-111.
- [31] 张宪德, 徐婷婷, 赵婷婷, 万冬梅, 殷江霞. 配偶选择中“听众效应”及鸣唱对雌性灰文鸟行为的影响. 生态学报, 2016, 36(16), doi: 10.5846/stxb201501290238.
- [32] 姜仕仁, 丁平, 诸葛阳, 邹艳春. 白头鹎繁殖期鸣声行为的研究. 动物学报, 1996, 42(3): 253-259.