Ecological Informatics 21 (2014) 110–119

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

The use of acoustic indices to determine avian species richness in audio-recordings of the environment

Michael Towsey *, Jason Wimmer, Ian Williamson, Paul Roe

Science and Engineering Faculty, Queensland University of Technology, QLD, Australia

ARTICLE INFO

Article history:
Received 21 June 2013
Received in revised form 30 October 2013
Accepted 14 November 2013
Available online 22 November 2013

Keywords:
Acoustic sensing
Biodiversity monitoring
Species richness
Acoustic sampling
Automated bird call analysis
Ecological indices

ABSTRACT

Interpreting acoustic recordings of the natural environment is an increasingly important technique for ecologists wishing to monitor terrestrial ecosystems. Technological advances make it possible to accumulate many more recordings than can be listened to or interpreted, thereby necessitating automated assistance to identify elements in the soundscape.

In this paper we examine the problem of estimating avian species richness by sampling from very long acoustic recordings. We work with data recorded under natural conditions and with all the attendant problems of undefined and unconstrained acoustic content (such as wind, rain, traffic, etc.) which can mask content of interest (in our case, bird calls).

We describe 14 acoustic indices calculated at one minute resolution for the duration of a 24 hour recording. An acoustic index is a statistic that summarizes some aspect of the structure and distribution of acoustic energy and information in a recording. Some of the indices we calculate are standard (e.g. signal-to-noise ratio), some have been reported useful for the detection of bioacoustic activity (e.g. temporal and spectral entropies) and some are directed to avian sources (spectral persistence of whistles). We rank the one minute segments of a 24 hour recording in descending order according to an "acoustic richness" score which is derived from a single index or a weighted combination of two or more. We describe combinations of indices which lead to more efficient estimates of species richness than random sampling from the same recording, where efficiency is defined as total species identified for given listening effort. Using random sampling, we achieve a 53% increase in species recognized over traditional field surveys and an increase of 87% using combinations of indices to direct the sampling. We also demonstrate how combinations of the same indices can be used to detect long duration acoustic events (such as heavy rain and cicada chorus) and to construct long duration (24 h) spectrograms.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

The analysis of acoustic recordings is an increasingly important technique for ecologists wishing to monitor the terrestrial and aquatic environments. Rapid advances in electronic hardware and computing power now make it possible to leave unattended acoustic sensors in exposed locations for several weeks of continuous recording. It is clearly impossible for ecologists to listen to even a small fraction of this audio data. Some degree of automated assistance is essential.

Recorded audio data can contribute to a number of ecological investigations, most obviously the identification of vocal animals. Bird species in particular are regularly surveyed because of their importance as indicator species of environmental health (Gregory and Strien, 2010). There is now a considerable body of published work on the detection of bird vocalizations (Acevedo et al., 2009; Agranat, 2009; Anderson et al., 1996; Brandes, 2008; Chen and Maher, 2006; Digby et al., 2013; Juang and Chen, 2007; McIlraith and Card, 1997; Somervuo et al., 2006). However vocal frog and insect species are also of interest (Brandes et al., 2006) and, in the Australian context, the koala (*Phascolarctos cinereus*, Ellis et al., 2010, 2011) and the cane toad (*Bufo marinus*, Hu et al., 2010) have received particular attention.

In contrast to the bioacoustic interest in individual species, there is a growing interest in *soundscape ecology*, that is, the study of the temporal and spatial distribution of sound through a landscape, reflecting important ecosystem processes and human activities (Kasten et al., 2012; Pijanowski et al., 2011a, 2011b). From this perspective, the soundscape is a finite resource in which organisms (including humans) compete for spectral space (Krause, 2008).

Although this work does not depend on the theoretical perspective of soundscape ecology, it does address the ecological problem of estimating species richness using acoustic recordings. In theory it might be possible to automate this task by preparing individual recognizers for the expected vocal species (which could number 100 or more) but

---

* Corresponding author at: Science and Engineering Faculty, QUT, GPO Box 2434, Brisbane, QLD 4000, Australia. Tel.: +61 7 3138 9340; fax: +61 7 3138 4438.

E-mail addresses: m.towsey@qut.edu.au (M. Towsey), j.wimmer@qut.edu.au (J. Wimmer), i.williamson@qut.edu.au (I. Williamson), p.roe@qut.edu.au (P. Roe).

1574-9541/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ecoinf.2013.11.007

the preparation of call recognizers is not an easy task. Lack of suitable training data can be a significant constraint and, even if a recognizer is successfully trained for one species in one locality, the natural geographic variation of calls may render it less effective in a new locality. Our research group has previously addressed the problem of recognizing vocal species by steering a middle path between "one-recognizer-fits-all-species" and "one-recognizer-for-each-species". The former strategy can sacrifice accuracy for generality but the latter is cumbersome and difficult to maintain. We have built a number of recognizers for generic features shared by many bird calls (Towsey et al., 2012).

This paper investigates the problem of determining species richness by approaching it as a problem of computer assisted sampling from long duration audio recordings. We illustrate our approach by focusing on bird species. The traditional method to determine avian richness at a specific location is the *point count* — one or more appropriately skilled persons count all species heard and/or seen within a specified area over a fixed period of time. Clearly this is a time consuming task where sampling effort is constrained by cost. A typical protocol is to visit a site for 20 min each at morning, noon and dusk over several days (Wimmer et al., 2013) but many other protocols are in use (Bibby et al., 1992).

Automated and semi-automated methods offer the advantage that recording devices can be deployed in the field for days or weeks obviating the need for regular field visits by a trained ecologist. However, the use of acoustic recordings to determine avian species richness is a relatively new technology and there are few well-established protocols or even comparisons of automated methods with traditional (Acevedo and Villanueva-Rivera, 2006). Our research group is investigating protocols for the use of environmental recordings (Digby et al., 2013; Wimmer et al., 2010, 2013). Wimmer et al. (2013) have compared a number of acoustic sampling protocols and demonstrated that they can be significantly more efficient than traditional point counts, where efficiency is defined as the number of species identified for equivalent listening effort. They also found that an effective sampling strategy is to select one minute audio samples at random from the 3 h after civil-dawn which encompasses the morning chorus when most birds are most likely to sing.

In this paper we investigate the use of a variety of acoustic indices to direct sampling from recordings of the environment. An acoustic index is a statistic that summarizes some aspect of the distribution of acoustic energy and information in a recording. We present one minute sound segments to a person skilled in bird identification, in an order ranked by indices that describe the acoustic content of the segments. Success is achieved if an estimate of avian species richness is obtained more efficiently (number of species identified for a given listening effort) than using *either* traditional on-site point-counts *or* random sampling from the recordings.

There is a growing body of work on the ecological uses of acoustic indices. It is convenient to divide the indices into three categories: waveform indices, spectral indices and second order indices. Waveform indices include traditional measures such as signal amplitude and signal-to-noise ratio. More recently, temporal entropy ($H[t]$) was introduced to characterize the temporal dispersal of acoustic energy within a recording (Sueur et al., 2008).

Spectral indices include spectral entropy ($H[s]$), a measure of acoustic energy dispersal through the spectrum (Sueur et al., 2008), and spectral peak count (NP), a measure of the average number of peaks in the spectra of the frames through a recording (Gasc et al., 2013). NP was shown to reflect acoustic activity as determined by ear. Pieretti et al. (2011) have introduced the *acoustic complexity index* (ACI), which is a measure of the average absolute fractional change in signal amplitude from one frame to the next through a recording.

The above indices show varying degrees of correlation with bio-acoustic activity. To obtain better correlations, a number of second order indices have been proposed. Sueur et al. (2008) demonstrated that $H[t] * H[s]$ is weakly correlated with "acoustic heterogeneity", and that an *acoustic dissimilarity* index, D_f, between two spectra $S1$ and $S2$, where:

$$D_\mathrm{f} = \Sigma_\mathrm{f} |S1(\mathrm{f}) - S2(\mathrm{f})|/2,$$

correlates with *differences* in "acoustic heterogeneity" between recordings.

A convenient property of $H[t]$, $H[s]$ and ACI is that their values are naturally normalized in [0, 1] and can therefore be used to compare recordings of quite different content and amplitude. It is possible to combine non-normalized indices, such as amplitude, by first converting them to a ranked index. For example, Depraetere et al. (2012) calculate the index *Acoustic Richness* (AR) given by:

$$\mathrm{AR} = ((\mathrm{rank}(H[\mathrm{t}]) \times \mathrm{rank}(M)))/\mathrm{n}^2, \quad \text{with } 0 \le \mathrm{AR} \le 1,$$

which combines $H[t]$ and M (median of the recording's amplitude envelope) by combining their ranks rather than their values. AR correlates with avian species richness.

Working on the assumption that acoustic activity in the 1–2 kHz and 2–11 kHz bands is likely to be technophony (sound due to machine sources) and biophony (sound due to animal sources) respectively, Joo et al. (2011) have proposed an *acoustic health quality index* (AHQI), more recently called the *normalized difference soundscape index* (NDSI):

$$\mathrm{NDSI} = (\mathrm{biophony} - \mathrm{technophony}) \;/\; (\mathrm{biophony} + \mathrm{technophony}),$$

where biophony and technophony are the summed power spectral densities (PSD) in the appropriate bands (McLaren, 2012).

In this work we investigate the hypothesis that *combinations* of indices will be more useful than single indices to characterize the acoustic content of one minute recordings. Our hypothesis is that a single index cannot capture all that is acoustically relevant in a recording. For example, $H[t]$ is not sensitive to frequency content and none of $H[t]$, $H[s]$, NP and ACI is sensitive to signal amplitude since their calculation 'normalizes' amplitude information. We apply combinations of acoustic indices to two tasks: 1. the efficient estimation of avian species richness and; 2. the detection of common acoustic "regimes" in Australian sub-tropical environmental recordings, namely rain and cicada choruses. A particular feature of our work is that we directly analyze real field-data recorded under normal environmental conditions and with all the attendant problems of unconstrained and undefined acoustic content. In particular, we do not remove audio segments containing wind and rain "noise" prior to analysis.

In this context, the issue of what constitutes "noise" in recordings of the environment requires some clarification. In a non-technical sense, "noise" is a sound where it is not wanted (adopting the classical definition of a *weed*). Because our focus is bird vocalizations, geophony (sounds due to wind, rain, leaf rustle, etc.), anthrophony (sounds due to human sources, traffic etc.) and biophony (sounds due to other animal vocalizations) can be considered noise. However in this study, we use the term "noise" in a technical sense to mean that acoustic energy which remains constant through the duration of a one-minute audio segment regardless of its source. Thus it is possible that the same acoustic source may contribute to both "noise" and "signal". For example, if we assume that crickets are evenly distributed in the landscape around a sensor, there will be a background "murmur" of crickets but the chirps of those crickets closest to the microphone will register as specific acoustic events within the background. Likewise, wind gusts will stand out as specific acoustic events within the constant noise generated by a background of moving air.

Download English Version:

https://daneshyari.com/en/article/4374901

Download Persian Version:

https://daneshyari.com/article/4374901

Daneshyari.com