



Using a Novel Visualization Tool for Rapid Survey of Long-Duration Acoustic Recordings for Ecological Studies of Frog Chorusing

Sheryn Brodie^{1*}, Michael Towsey², Slade Allen-Ankins¹, Paul Roe² and Lin Schwarzkopf¹

¹ College of Science and Engineering, James Cook University, Townsville, QLD, Australia, ² Faculty of Science, Queensland University of Technology, Brisbane, QLD, Australia

OPEN ACCESS

Edited by:

Gianni Pavan,
University of Pavia, Italy

Reviewed by:

Karl-Heinz Frommolt,
Museum of Natural History Berlin
(MfN), Germany
Marco Gamba,
University of Turin, Italy

*Correspondence:

Sheryn Brodie
sheryn.brodie@my.jcu.edu.au

Specialty section:

This article was submitted to
Population, Community,
and Ecosystem Dynamics,
a section of the journal
Frontiers in Ecology and Evolution

Received: 19 August 2021

Accepted: 30 December 2021

Published: 20 January 2022

Citation:

Brodie S, Towsey M,
Allen-Ankins S, Roe P and
Schwarzkopf L (2022) Using a Novel
Visualization Tool for Rapid Survey
of Long-Duration Acoustic
Recordings for Ecological Studies
of Frog Chorusing.
Front. Ecol. Evol. 9:761147.
doi: 10.3389/fevo.2021.761147

Continuous recording of environmental sounds could allow long-term monitoring of vocal wildlife, and scaling of ecological studies to large temporal and spatial scales. However, such opportunities are currently limited by constraints in the analysis of large acoustic data sets. Computational methods and automation of call detection require specialist expertise and are time consuming to develop, therefore most biological researchers continue to use manual listening and inspection of spectrograms to analyze their sound recordings. False-color spectrograms were recently developed as a tool to allow visualization of long-duration sound recordings, intending to aid ecologists in navigating their audio data and detecting species of interest. This paper explores the efficacy of using this visualization method to identify multiple frog species in a large set of continuous sound recordings and gather data on the chorusing activity of the frog community. We found that, after a phase of training of the observer, frog choruses could be visually identified to species with high accuracy. We present a method to analyze such data, including a simple R routine to interactively select short segments on the false-color spectrogram for rapid manual checking of visually identified sounds. We propose these methods could fruitfully be applied to large acoustic data sets to analyze calling patterns in other chorusing species.

Keywords: acoustic monitoring, Ecoacoustics, frog chorusing, acoustic data analysis, acoustic data visualisation, chorus detection

INTRODUCTION

Passive acoustic monitoring is now a standard technique in the ecologist's toolkit for monitoring and studying the acoustic signals of animals in their natural habitats (Gibb et al., 2019; Sugai et al., 2019). Autonomous sound recorders provide significant opportunities to monitor wildlife over long time frames, and at greater scale than can be done physically in the field. Recording over extended periods and locations provides insight into species' activity patterns, phenology and distributions (Nelson et al., 2017; Wrege et al., 2017; Brodie et al., 2020b), and allows for the study and monitoring

of whole acoustic communities (Wimmer et al., 2013; Taylor et al., 2017). Continuous and large-scale acoustic monitoring has become feasible as technological advances have provided smaller, cheaper recording units with improved power and storage capacities. However, the large streams of acoustic data that can be collected must be mined for ecologically meaningful data, and so the problem of scaling of observations has been translated into a problem of scaling data analysis (Gibb et al., 2019). Ecologists using acoustic approaches require effective and efficient sound analysis tools that enable them to take advantage of the scaling opportunities in large acoustic data sets.

Often, computer-automated approaches to detecting and identifying the calls of target species, such as pattern recognition and machine learning, are put forward as the solution to analyzing big acoustic data (Aide et al., 2013; Stowell et al., 2016; Gan et al., 2019). However, developing automated detection pipelines requires a high level of signal processing, computational and programming expertise, as well as considerable time and effort in labeling call examples to train classifiers, and then test and refine their performance (e.g., Brodie et al., 2020a). Unsupervised machine learning methods circumvent the need for labeled data but still require a large amount of data, and considerable time and expertise, to compute learning features and interpret the results (Stowell and Plumbley, 2014). Long-duration field recordings often contain intractable amounts of noise and variability in the quality of calls, and achieving accurate species identification is challenging in large-scale studies or studies of multiple species (Priyadarshani et al., 2018). The focus of automated acoustic analysis has been on detection of individual calls, but this granularity of data is often not what is required in studies of population chorusing activity, and call detections are instead aggregated into calls per unit of time. The time and effort in developing automated species detection methods to create results of limited practical use means this approach is not feasible in many studies and monitoring programs. Thus, manual sound analysis continues to be used in the majority of ecological studies using acoustic methods (Sugai et al., 2019), while automated call detection methods continue to be developed and improved (e.g., Ovaskainen et al., 2018; Marsland et al., 2019; Brooker et al., 2020; Kahl et al., 2021; Miller et al., 2021).

The manual approach to analyzing environmental sound recordings for studies of vocal animals, is to inspect each sound file using specialized software with both spectrogram and playback functions (e.g., Audacity¹; Raven, Cornell Lab of Ornithology). An observer familiar with the calls of target species will typically scan the spectrogram visually for candidate sounds, and may use playback to confirm the species when uncertainty exists. In this way, an expert observer does not need to playback and listen to the entire recording to analyze it and identify the species present. This can be more efficient than designing automated call recognizers for short-term studies where there are few target species. However, manual analysis of sound recordings becomes impractical for large-scale studies (long-term or many species). As a consequence, many acoustic surveys are still designed with restricted sampling regimes

that permit manual analysis. By programming recording units to record for a limited time at regular intervals throughout the study period, the temporal and spatial scale of surveys can still be kept large while keeping manual analysis feasible. Restricted sampling regimes have disadvantages over continuous recordings, however, such as a reduced likelihood of detecting rare species or species that vocalize infrequently (Wimmer et al., 2013), as well as a narrower temporal sampling resolution, which may miss ecological patterns of interest. Therefore, techniques that allow analysis of long, continuous audio recordings that do not rely on statistical techniques that are beyond the expertise of many users (e.g., machine learning), and that do not restrict the amount of time sampled, are required.

Recent developments in computational approaches to the analysis of environmental sound recordings have led to software tools being made available that generate visual representations of sound recordings at scales of 24-h or more. Towsey et al. (2014) developed a method of representing a long sound file in a single spectrogram that can be viewed whole on a standard computer monitor screen. This was achieved by using acoustic indices, which are numerical summaries of the sound signal calculated at coarse time scales, and which can be considered a form of data compression (Sueur et al., 2008; Pieretti et al., 2011). The compressed spectrograms were generated using three different acoustic indices calculated at 1-min resolution and mapping the values to three color channels (red, green, and blue) to form a “false-color” spectrogram. The sound content of the recording is reflected in the visual patterns which highlight dominant sound events. While these false-color spectrograms were devised to visualize general patterns in the soundscape, exploration of the patterns revealed that the calls of some species could be identified in the images (Indraswari et al., 2018; Towsey et al., 2018b). An example of a false-color spectrogram for a recording used in this study is presented here (Figure 1).

This manuscript presents a method of using long-duration false-color spectrograms to navigate and sample a large set of environmental recordings to detect species in a chorusing frog community. The impetus for applying this method was to collect data on the chorusing phenology and nightly chorusing activity of frog species at multiple breeding sites. We present simple R (R Core Team, 2021) routines for generating false-color spectrograms and for interactive selection of time segments to automate the process of finding and opening the segments of interest in the audio for manual analysis. We also test the accuracy of an observer, after some learning experience, to visually identify the frog species present at the study sites from patterns on the false-color spectrograms. Our aim is to outline and describe a protocol that will be useful to ecologists looking for an easily implemented method of navigating acoustic recordings and identifying the calls of target species.

MATERIALS AND METHODS

Acoustic Recordings

Long-duration sound recordings were made at frog breeding sites near Townsville, north Queensland, Australia (19.357° S,

¹<https://www.audacityteam.org>

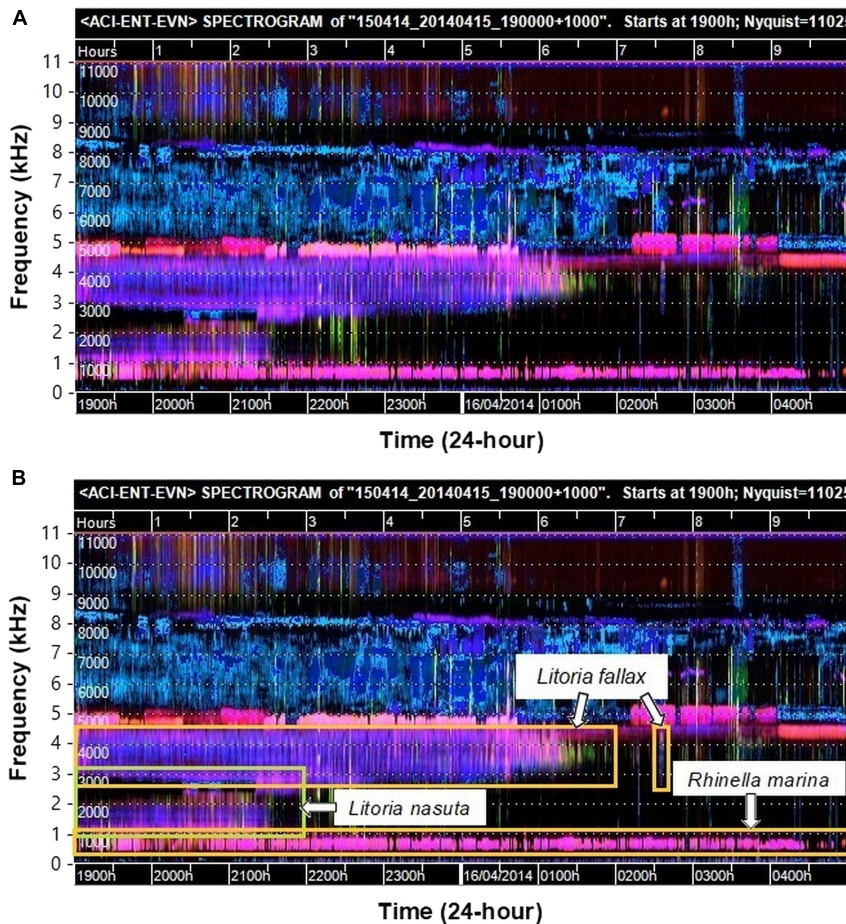


FIGURE 1 | (A) Example false-color spectrogram of a recording used in this study with chorusing of three frog species dominant. The frog choruses show as pink and purple tracks below 4.5 kHz. Sporadic birds calls occur (green/yellow), and insect choruses are prominent as pink and blue tracks above 4 kHz. The horizontal dotted lines delineate 1000 Hz frequency intervals (labeled in kHz on the axis outside the figure for clarity). Colors are derived from three acoustic indices (Acoustic complexity – red; Entropy – green; and Event count – blue) which are defined in the text. **(B)** The same false-color spectrogram image with the chorusing frog species identified inside the labeled boxes.

146.454° E). Recording units (HR-5, Jammin Pro, United States) were set to record continuously at 10 sites each night throughout a 19-month period from October 2012 to April 2014. Recorders were housed in water-proof metal boxes with external microphones in plastic tubing, and recordings were made in MP3 file format (128 kbps bit rate; 32000 Hz sampling rate). The study area is in a tropical savanna ecoregion and frogs in this habitat are nocturnal, so recordings were only made during the night. Most recordings were between 10- and 13-h duration, typically commencing between 1800 and 1930 h and ending after sunrise. The number of nights recorded at each site during the study period ranged from 375 to 473 nights (audio was not obtained for all nights because of recording equipment failure). At the end of recording period we had collected 3,965 nightly recordings totaling approximately 46930 h of audio.

Generation of False-Color Spectrograms

False-color spectrograms were produced using the QUT Ecoacoustics Audio Analysis Software v.17.06.000.34

(Towsey et al., 2018a) following the methods detailed in Towsey et al. (2015) and Towsey (2017). Audio recordings were divided into 1-min segments, re-sampled at a rate of 22,050 samples per second and processed into standard spectrogram form using a fast-Fourier transform with Hamming window and non-overlapping frames of 512 samples per frame (~23.2 ms per frame). Acoustic indices were calculated for each minute segment in each of 256 frequency bins from 0 to 11025 Hz (bandwidth ~43.1 Hz). False-color spectrograms can be produced using a combination of any three of the calculated acoustic indices, which are mapped to red, green and blue colors. We used the default combination of indices output by the software – the acoustic complexity index (ACI), spectral entropy (ENT) and acoustic events (EVN) (Table 1). This combination of acoustic indices best displays biotic sounds of interest in the false-color spectrograms, because they are minimally correlated and highlight different features (Towsey et al., 2018b).

The analysis to calculate acoustic indices and produce false-color spectrograms for our large set of recordings was done by the

TABLE 1 | Definitions of acoustic indices used in composing false-color spectrograms, calculated for each minute in each frequency bin (Towsey et al., 2014; Towsey, 2017).

Acoustic index	Description
Acoustic complexity index (ACI)	The average relative change in sound amplitude from one frame to the next.
Entropy (ENT)	A measure of temporal concentration of acoustic energy.
Event count (EVN)	The number of acoustic events exceeding 3 dB.

QUT Ecoacoustics Research Group's data processing lab using multiple computers which were dedicated to research analyses. However, the QUT Ecoacoustics Audio Analysis program is available as open source software and can be run on a personal computer. The program is downloaded as an executable file and run from the command line which provides flexibility for scripting and batch processing on different platforms (Truskinger et al., 2014). R code to run the open source version on multiple sound files in a single process is provided on GitHub (Brodie, 2021). When tested on a desktop PC (16 GB RAM, Intel(R) Core(TM) i7-7700 CPU @ 3.60 GHz) and running analyses in parallel, a 12-h recording took on average 7 min to analyze. For large data sets where this rate of output is inadequate, dedicated high-performance computing facilities or professional support may be required.

Navigation and Inspection of False-Color Spectrograms

The QUT Ecoacoustics Audio Analysis software output a set of files for each separate audio file including the raw acoustic index values in CSV files, and the false-color spectrograms as PNG image files. All the PNG image files of the ACI-ENT-EVN indices combination were placed into a single directory for each site for ease of navigation through each set of images. Each pixel on the false-color spectrogram images represented 1 min on the time scale and approximately 43 Hz frequency range. A time scale is included on the image displaying the time since the start of the recording or, if a valid date and time is included the audio filenames, the time of the recording. Therefore, the position of a pixel on a false-color spectrogram informs the time position

within the audio recording (x-axis), and the approximate frequency range (y-axis). We used the XnView image viewer application (v 2.43²) to view the PNG image files, as this application displays the position coordinates of the mouse pointer on the image. This allowed identification of the precise point, in number of minutes, from the start of the recording.

The patterns in the false-color spectrograms reflect the dominant sound sources in each time segment and frequency bin. Learning to relate visual patterns to sound events was done by identifying potential sounds of interest on the false-color spectrogram images and then manually inspecting the corresponding minute in the audio file to identify potential sound sources. We used Audacity audio software (see text footnote 1) for playback of the raw audio and viewing in standard spectrogram format.

R Routines for Efficient Analysis

Although sound analysis software packages such as Audacity can open and display long sound files, opening and navigating long recordings is inefficient when short segments from many separate recordings need to be analyzed. We made the analysis more efficient using R routines in the RStudio environment (RStudio, 2021) to slice short segments of the recordings at specified time points using the "Audiocutter" function in the QUT Ecoacoustics Audio Analysis software (Towsey et al., 2018a). The R routines included user-defined functions to select and cut audio segments using two alternative methods of selection:

- (i) direct user input - the user entered the start minute (the *x*-value identified on the false-color spectrogram image) and desired length of audio segment as variables into the R script;
- (ii) interactive selection - the user invoked a function from the R 'imager' package (Barthelme, 2021) which opened a graphic window displaying the false-color spectrogram and prompted the user to select the desired minute(s) (*x*-value) interactively on the image (the 'grabPoint' function to select a single *x*-value, or the "grabRect" function to select a range of *x*-values).

²<https://www.xnview.com/en>

TABLE 2 | Results of accuracy test of visual identification of frog species in 321 test minutes using false-color spectrograms.

Species	Minutes present	Correct (TP)	Incorrect (FP)	Missed (FN)	Not present (TN)	Precision TP/(TP + FP)	Recall (sens) TP/(TP + FN)	Specificity TN/(TN + FP)
<i>Rhinella marina</i>	132	125	1	7	189	99.2%	94.7%	99.5%
<i>Litoria fallax</i>	149	140	3	9	172	97.9%	94.0%	98.3%
<i>Litoria nasuta</i>	100	88	3	12	221	96.7%	88.0%	98.7%
<i>Limnodynastes convexiusculus</i>	30	28	0	2	291	100.0%	93.3%	100.0%
<i>Limnodynastes terraereginae</i>	21	17	2	4	300	89.5%	81.0%	99.3%
<i>Limnodynastes peronii</i>	15	13	0	2	306	100.0%	86.7%	100.0%
<i>Litoria rubella</i>	7	4	0	3	314	100.0%	57.1%	100.0%

Minutes present, number of randomly selected minutes in which the species was confirmed to be calling; TP, True positive; FP, False positive; FN, False negative; TN, True negative.

The user input was then passed to the “Audiocutter” function which cut the selected minute(s) from the audio file and opened the selected segment in the Audacity program. The R code files have been made available on GitHub (Brodie, 2021).

Validation of False-Color Spectrograms as a Species Identification Tool

To validate that the false-color spectrograms were a reliable tool for visual identification of the frog species in this data set, a random selection of minutes was analyzed by an observer (SB) before being validated by inspecting the raw audio. Fifty false-color spectrograms (i.e., for recordings of different nights) from three sites were chosen which had not been previously analyzed. A random selection of up to 20 one-minute segments was made from each spectrogram using a random number generator. The presence of frog species was predicted for each randomly selected minute in each recording solely from visual inspection of the false-color spectrogram and prior to any inspection of the audio file. The visually based predictions were then validated by inspecting the corresponding audio segment using the Audacity program. A total of 321 separate minutes were randomly selected for validation, and the identification precision, recall and specificity metrics were calculated for each species identified. The frog species present at the study sites aggregate at water bodies to breed and males call in choruses. We did not distinguish between times when only one individual was calling and more than one individual was calling, since the ultimate aim is to use calling or chorusing as an indicator of breeding activity. It should be noted that this test was performed after the observer (SB) had gained some familiarity with the species’ patterns in the false-color spectrograms of the data set, and had an expert level of ability to identify the calls of the frog species present in the raw audio.

RESULTS

Using the false-color spectrograms as a visual guide to the sound content of long environmental recordings, we were able to efficiently collect data on the presence and timing of chorus activity of multiple species of frogs in a large set of acoustic recordings. This method greatly reduced the manual listening effort required when compared to scanning entire recordings and increased the detectability of species over a method using a restricted sampling regime of regular time intervals. The time taken to survey the nightly recordings using the R routine to select, cut and open short segments of audio ranged from a few seconds (on nights with no frog chorusing) to 90 min (a full 13-h continuous recording with 11 species of frogs identified and extensive chorus activity). The average time taken to survey each night was 14 min.

In the test of species identification accuracy, 9 false-positive identifications were made, and 39 false-negative identifications (species missed) out of a total of 454 occurrences of any frog species (Table 2). As a result, precision (the percentage of correct identifications) was very high for all species present. Recall (the percentage of actual species occurrences detected) was high for

the most common species, but low for *Litoria rubella* which was present in only 7 of the minutes selected for validation.

Inspection of the possible reasons for the identification errors revealed that other noises in the same frequency band caused the false-positive detections (Table 2). *Litoria fallax* and *Litoria nasuta* were falsely detected occasionally because they were confused with visual patterns made by splashing water. *L. fallax* was also falsely detected in one instance when insect noise was present. *L. fallax* has a call in the frequency range 2–6 kHz, which overlaps with some insect sounds. *L. nasuta* was falsely detected in one instance when *L. rubella* was calling, and once when *L. fallax* was calling. *L. nasuta* has a short, broadband call in the frequency range 1–4 kHz which entirely overlaps the calls of *L. rubella* and partly that of *L. fallax*. *Rhinella marina*, which has a long, low-frequency call made up of a trill of rapid pulses, was misidentified only once when rapid dripping of water onto the recorder housing created a similar pattern on the false-color spectrogram. In two instances, *R. marina* was mistakenly identified as *L. terraereginae*. The calls of these two species overlap in the frequency range of approximately 500–900 Hz. False-negative identifications (species missed) occurred either because the missed species was obscured by other dominant noise (vehicles, wind, other frogs or insects) or because the calls were very faint and distant, very short bouts or one individual calling at a very slow rate.

DISCUSSION

Visualization of long sound recordings is an innovative approach for providing insight into the acoustic structure of environmental soundscapes, and to aid detection of wildlife vocalizations. We found that false-color spectrograms generated using acoustic indices were a reliable and accurate method of identifying the chorus activity of individual species in a large community of chorusing frogs. A routine using the R programming environment was developed that automates searching and opening segments of sound files after interactive selection on the false-color spectrogram image. This method provided an easily implemented and practical tool for biological researchers to explore and navigate sound recordings for species of interest, and provides opportunities for increasing the scale of acoustic analysis with open-source software tools. False-color spectrograms allowed easy identification of which recordings contained large amounts of vocal activity and those that did not. For example, recordings with no frog chorusing had false-color spectrograms with very little color pattern in the frequency range below 4 kHz (e.g., Figure 2). This allowed us to quickly eliminate nights with no frog chorus activity without the need to manually check the audio file, and focus on those recordings with high vocal activity (e.g., Figure 3).

The use of visualization as a tool to analyze long recordings in ecological studies was developed independently by several researchers (Wiggins and Hildebrand, 2007; Towsey et al., 2014) but, despite its demonstrated usefulness, it has not been applied extensively in practice. Wiggins and Hildebrand (2007) first devised a method of visualizing sound recordings by averaging

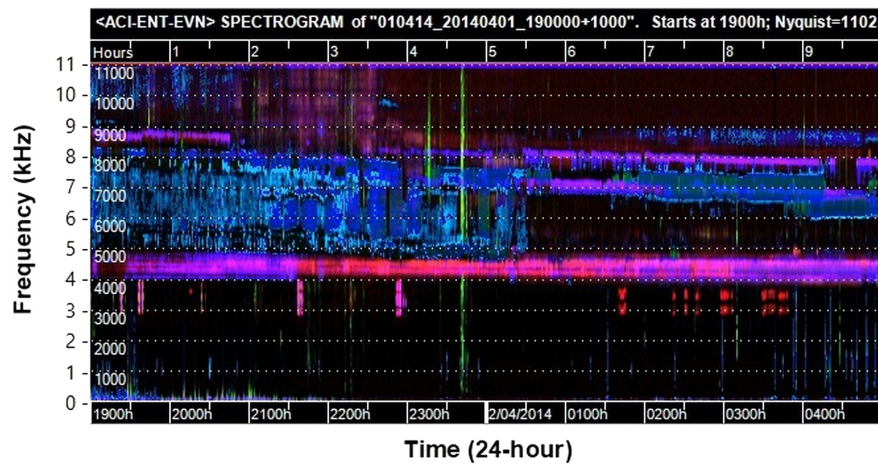


FIGURE 2 | False-color spectrogram of a recording used in this study showing a night with no frog vocal activity. The dominant sounds are insect choruses above 4 kHz. The occasional pink and red tracks at 3–4 kHz are also insects. Sporadic sounds below 3 kHz which occur include wind, passing vehicles and occasional bird calls. The obvious green broadband mark at approx. 23:40 hrs is made by water birds splashing and flapping wings close to the microphone. The horizontal dotted lines delineate 1,000 Hz frequency intervals (labeled in kHz on the axis outside the figure for clarity).

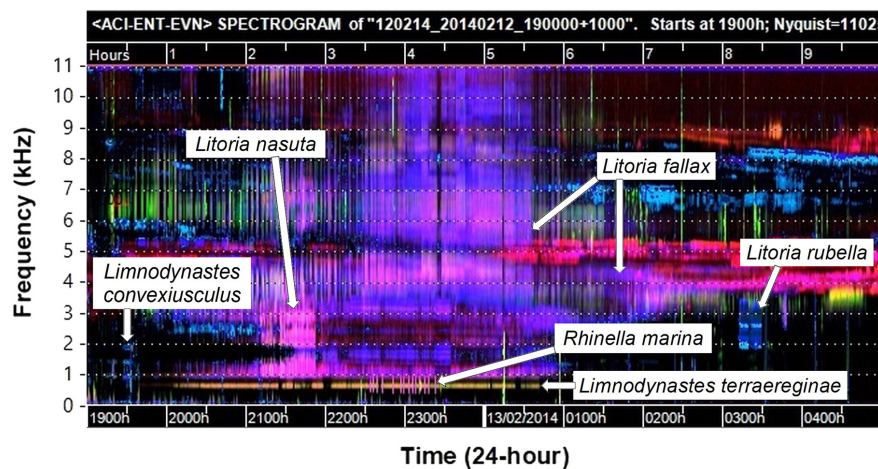


FIGURE 3 | False-color spectrogram of a recording used in this study which features the choruses of six frog species calling simultaneously. The horizontal dotted lines delineate 1,000 Hz frequency intervals (labeled in kHz on the axis outside the figure for clarity). Patterns are sometimes obscured by other dominant species but can be distinguished at other times.

spectral power values over chosen time frames to generate compressed spectrograms. Their method was implemented in the MATLAB programming environment using the Triton software package (Wiggins, 2007) which also facilitates navigation to specific segments of the raw audio for manual analysis. Published examples have applied visualization using the Triton package in marine environments, for which it was designed, to detect whale calls (Soldevilla et al., 2014) and describe marine soundscapes (Rice et al., 2017), but it has also been used in freshwater environments to detect chorusing of an underwater-calling frog (Nelson et al., 2017), and in terrestrial environments to detect chimpanzee vocalizations (Kalan et al., 2016).

The false-color spectrograms developed by Towsey et al. (2014), Towsey et al. (2015), and demonstrated here, progressed

the concept of soundscape visualization, by using acoustic indices that highlight biological sounds. The method of visualization using three color channels based on different metrics enables display of more complex patterns, and highlights a greater variety of sound sources than using the single spectrogram power values. False-color spectrograms have been used in ecological studies to describe and compare soundscapes, by using the visual images to detect the dominant sounds in the environment (Dema et al., 2018; Campos et al., 2021). Several studies have shown the calls of individual species can be detected visually using false-color spectrograms. Towsey et al. (2018b) and Znidersic et al. (2020) were able to visually detect the presence of cryptic marsh birds. Brodie et al. (2020b) used the method to confirm the nightly presence of invasive toad calling activity.

The general advantage of visualization of environmental recordings is that it allows rapid detection of candidate sounds of interest without relying on complex computational methods, and reduces the effort required to manually scan sound files. While previous published studies have used false-color spectrograms for detecting species presence and characterizing soundscapes, here we have demonstrated the method can scale to studies of communities of chorusing frogs over extended time periods and multiple locations. The constant choruses of several frog species left unique traces on the false-color spectrograms which, in many instances, could be confidently identified without the need to analyze the raw audio, decreasing further the manual analysis required. Manual inspection of the audio, either by listening or viewing the standard spectrogram, was still required for many of the recordings where the noise source of presence of a frog species was unclear. Therefore, there is a limit to the scalability of using this method for very large data sets.

The high precision and specificity of frog species identification achieved in the test cases (Table 2) reflects the low rate of false-positive detections. That is, patterns on the false-color spectrograms were only very occasionally incorrectly identified as another species. The majority of identification errors were missed species' presence, in cases where the frog calls were distant and low-quality in the recording, or there was a low rate of calling. Low-quality, background calls will always be difficult to detect regardless of the method used. The accuracy results presented here are better for the 5 species that are shared with a previous study investigating the use of automated classification using acoustic indices and machine learning (Brodie et al., 2020a). This suggests that even with considerable time and effort to label training data and train classification models, automated methods may still not perform as desired, and manual methods such that presented here may be more suitable.

Our aim in this study was not to compare the accuracy of species identification using false-color spectrograms with automated detection methods, as these are different approaches to data reduction and analysis of acoustic data. The use of false-color spectrograms to survey acoustic recordings for target species can reduce the amount of manual analysis required, but still requires significant manual effort and time to learn to identify patterns of interest. In addition, computing time to calculate acoustic indices and generate the images is considerable for large data sets. Automated species detection for environmental sound recordings is a rapidly advancing field, however, may not be feasible or practical for all acoustic studies. The most successful automated detection algorithms are for species with well-described calls which are distinct from the calls of other species (e.g., Walters et al., 2014 for bats; reviewed in Kowarski and Moors-Murphy, 2020 for fin and blue whales) or for which large sets of training data are available (e.g., Kahl et al., 2021; Miller et al., 2021). Nonetheless, the challenge of automating analysis of acoustic data is far from solved for many research questions. Automated animal call detection is now the domain of computer scientists and computational experts, and there is considerable time and expertise required in developing accurate detection algorithms. Further, recent reviews have revealed that the majority of studies utilizing automated call detection methods

incorporate manual human intervention in post-processing stages, such as manual validation and cleaning of call detection results (Sugai et al., 2019; Kowarski and Moors-Murphy, 2020). There are inevitable trade-offs in time, cost and effort when researchers decide whether to utilize automated or manual methods in their acoustic data analysis.

Several factors combine to render frog choruses visually distinct and readily identifiable on the false-color spectrograms. Frog choruses tend to be persistent through time, often continuing for several hours, and are the dominant sound at breeding sites during breeding periods. Frog calls are repetitive and consistent in structure within species, but vary in both structure and frequency range among species. This method of using visualization to analyze long-duration audio is, therefore, highly suited to monitoring frog communities where species form persistent, loud choruses at breeding sites. This approach would also be applicable to other chorusing species, such as soniferous insects. Sounds that occur over short periods may also be visible on the false-color spectrograms but are less obvious than patterns that extend through a large portion of the recording. Some nocturnal birds that call continuously for at least a few minutes, such as owls and cuckoos, can also be identified (Phillips et al., 2018; personal observation). Short bursts of sounds may also be highlighted on the false-color spectrograms if they are louder than other sounds in the same minute segment, so this technique of detecting sounds is not limited to species with long-duration calls. However, it became clear from our experience analyzing this data set that the representation of sounds in the false-color spectrograms is dependent on other sounds present in the same minute segment and frequency band. The loudest sounds in each segment are highlighted so that the choruses of several frog species were sometimes obscured, or masked, in periods of high chorus activity dominated by other frog species. On the other hand, soft short calls may be identified in other periods when there are no competing noises in the same frequency range (Znidarsic et al., 2020; personal observation). We found that the masking by dominant frog species could be somewhat overcome by using long, continuous recordings rather than shorter, intermittent recordings. Having a complete, continuous recording for each study night meant we could detect most of the chorusing frog species at some point in the false-color spectrogram when masking was reduced. Whether false-color spectrograms are a suitable tool for the detection of a species depends on the likelihood of capturing calling individuals within range of the microphone and the level of competing noise in the target frequency range.

A further advantage to the approach described is that all software used was open source and does not require a specialized platform. The QUT Ecoacoustics Audio Analysis program³ automatically performs all processing of raw audio, calculation of acoustic indices and generation of the false-color spectrograms. Some knowledge of running programs from a command-line environment is required, but user input requirements are limited to defining the input and output files, with some configuration options. The interactive selection functions were implemented in

³<https://github.com/QuTEcoacoustics/audio-analysis/releases>

R using R Studio (Brodie, 2021), and are simple to run for users with basic knowledge of the R programming environment. R is now widely used in ecological research (Lai et al., 2019) and easily accessible for most researchers.

The false-color spectrograms can be a useful tool to analyze long recordings, even without the R routine program, simply by manually opening the corresponding sound file and navigating to the time-point of interest indicated on the false-color spectrograms. The interactive R routine was created to increase time efficiency, as shorter sound files are quicker to open than longer files, and when opened can be immediately inspected without having to navigate through a long recording to the relevant time point. In addition to increased efficiency, the R routine reduces the risk of human error. When dealing with large sets of sound files there is a risk of choosing the wrong file if many files have similar names with the same date, or of navigating to the wrong time point in long recordings.

Although automated methods of identifying species in acoustic data is an advancing field of research, many researchers continue to use manual analysis methods in acoustic monitoring studies. Our aim in this paper was to demonstrate a work-flow including the practical application of false-color spectrograms (Towsey et al., 2014) as a navigation aid to streamline the manual analysis of acoustic data. The process described here takes this innovative method of visualizing sound and incorporates it into an efficient routine for detecting the chorusing of multiple species of frogs in large acoustic data sets. The accuracy achieved in identifying multiple species of frogs from field recordings taken at different times and locations confirms this can be a reliable method of species detection and identification. Used as a means to quickly scan the content of recordings for target sounds, the amount of manual analysis is greatly reduced. There is potential for its use in increasing the coverage of ecological monitoring programs, particularly where automated methods of analysis are not practical or feasible. In describing and outlining our process of utilizing false-color spectrograms to analyze long-duration

recordings, we seek to make this method accessible and practical for use by other researchers using acoustic monitoring methods.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

SB and LS conceived the ideas and designed methodology. MT conceived and designed the original methodology of audio visualization with false-color spectrograms with input from PR. SB developed and wrote the R routines with input from SA-A. SB collected and analyzed the data and led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

FUNDING

Funding for this work was provided by the Australian Research Council (Linkage grant number LP150100675 “Call Out and Listen In: A New Way to Detect and Control Invasive Species”) and an Australian Government Research Training Program Scholarship.

ACKNOWLEDGMENTS

We thank Ross Alford, Richard Duffy and Kiyomi Yasumiba for obtaining the original set of field recordings used in this study. We are also grateful to Jodie Betts for assistance with audio analysis and feedback on the R routine.

REFERENCES

- Aide, T. M., Corrada-Bravo, C., Campos-Cerqueira, M., Milan, C., Vega, G., and Alvarez, R. (2013). Real-time bioacoustics monitoring and automated species identification. *PeerJ* 1:e103. doi: 10.7717/peerj.103
- Barthelme, S. (2021). *imager: Image Processing Library Based on 'CImg'. R package version 0.42.8*. Available online at: <https://CRAN.R-project.org/package=imager> (accessed May 23, 2021).
- Brodie, S. (2021). *sherynbrodie/fcs-audio-analysis-utility (Version 21.08.0) [Computer software]*. Available online at: <https://doi.org/10.5281/zenodo.5220459> (accessed August 19, 2021).
- Brodie, S., Yasumiba, K., Towsey, M., Roe, P., and Schwarzkopf, L. (2020b). Acoustic monitoring reveals year-round calling by invasive toads in tropical Australia. *Bioacoustics* 30, 125–141. doi: 10.1080/09524622.2019.1705183
- Brodie, S., Allen-Ankins, S., Towsey, M., Roe, P., and Schwarzkopf, L. (2020a). Automated species identification of frog choruses in environmental recordings using acoustic indices. *Ecol. Indic.* 119:106852. doi: 10.1016/j.ecolind.2020.106852
- Brooker, S. A., Stephens, P. A., Whittingham, M. J., and Willis, S. G. (2020). Automated detection and classification of birdsong: an ensemble approach. *Ecol. Indic.* 117:106609. doi: 10.1016/j.ecolind.2020.106609
- Campos, I. B., Fewster, R., Truskinger, A., Towsey, M., Roe, P., Vasques Filho, D., et al. (2021). Assessing the potential of acoustic indices for protected area monitoring in the Serra do Cipó National Park, Brazil. *Ecol. Indic.* 120:106953. doi: 10.1016/j.ecolind.2020.106953
- Dema, T., Towsey, M., Sherub, S., Sonam, J., Kinley, K., Truskinger, A., et al. (2018). Acoustic detection and acoustic habitat characterisation of the critically endangered white-bellied heron (*Ardea insignis*) in Bhutan. *Freshw. Biol.* 65, 153–164. doi: 10.1111/fwb.13217
- Gan, H., Zhang, J., Towsey, M., Truskinger, A., Stark, D., Van Rensburg, B., et al. (2019). “Recognition of frog chorusing with acoustic indices and machine learning,” in *Proceedings of the 2019 15th International Conference on eScience (eScience)*, (Piscataway, NJ: IEEE), 106–115. doi: 10.1109/eScience.2019.00019
- Gibb, R., Browning, E., Glover-Kapfer, P., Jones, K. E., and Börger, L. (2019). Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods Ecol. Evol.* 10, 169–185. doi: 10.1111/2041-210X.13101
- Indraswari, K., Bower, D. S., Tucker, D., Schwarzkopf, L., Towsey, M., and Roe, P. (2018). Assessing the value of acoustic indices to distinguish species and quantify activity: a case study using frogs. *Freshw. Biol.* 65, 142–152. doi: 10.1111/fwb.13222

- Kahl, S., Wood, C. M., Eibl, M., and Klinck, H. (2021). BirdNET: a deep learning solution for avian diversity monitoring. *Ecol. Inform.* 61:101236. doi: 10.1016/j.ecoinf.2021.101236
- Kalan, A. K., Piel, A. K., Mundry, R., Wittig, R. M., Boesch, C., and Kuhl, H. S. (2016). Passive acoustic monitoring reveals group ranging and territory use: a case study of wild chimpanzees (*Pan troglodytes*). *Front. Zool.* 13:34. doi: 10.1186/s12983-016-0167-8
- Kowarski, K. A., and Moors-Murphy, H. (2020). A review of big data analysis methods for baleen whale passive acoustic monitoring. *Mar. Mamm. Sci.* 37, 652–673. doi: 10.1111/mms.12758
- Lai, J., Lortie, C. J., Muenchen, R. A., Yang, J., and Ma, K. (2019). Evaluating the popularity of R in ecology. *Ecosphere* 10:e02567. doi: 10.1002/ecs2.2567
- Marsland, S., Priyadarshani, N., Juodakis, J., Castro, I., and Poisot, T. (2019). AviaNZ: a future-proofed program for annotation and recognition of animal sounds in long-time field recordings. *Methods Ecol. Evol.* 10, 1189–1195. doi: 10.1111/2041-210X.13213
- Miller, B. S., The IWC-SORP/SOOS Acoustic Trends Working Group, Balcazar, N., Nieuwkerk, S., Leroy, E. C., Aulich, M., et al. (2021). An open access dataset for developing automated detectors of Antarctic baleen whale sounds and performance evaluation of two commonly used detectors. *Sci. Rep.* 11:806. doi: 10.1038/s41598-020-78995-8
- Nelson, D. V., Garcia, T. S., and Klinck, H. (2017). Seasonal and diel vocal behavior of the northern red-legged frog, *Rana aurora*. *Northwest. Nat.* 98, 33–38. doi: 10.1898/NWN16-06.1
- Ovaskainen, O., Moliterno de Camargo, U., and Somervuo, P. (2018). Animal sound identifier (ASL): software for automated identification of vocal animals. *Ecol. Lett.* 21, 1244–1254. doi: 10.1111/ele.13092
- Phillips, Y. F., Towsey, M., and Roe, P. (2018). Revealing the ecological content of long-duration audio-recordings of the environment through clustering and visualisation. *PLoS One* 13:e0193345. doi: 10.1371/journal.pone.0193345
- Pieretti, N., Farina, A., and Morri, D. (2011). A new methodology to infer the singing activity of an avian community: the acoustic complexity index (ACI). *Ecol. Indic.* 11, 868–873. doi: 10.1016/j.ecolind.2010.11.005
- Priyadarshani, N., Marsland, S., and Castro, I. (2018). Automated birdsong recognition in complex acoustic environments: a review. *J. Avian Biol.* 49:jav-1447. doi: 10.1111/jav.01447
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rice, A., Soldevilla, M., and Quinlan, J. (2017). Nocturnal patterns in fish chorusing off the coasts of Georgia and eastern Florida. *Bull. Mar. Sci.* 93, 455–474. doi: 10.5343/bms.2016.1043
- RStudio (2021). *RStudio (version 1.4.1106)*. [Software]. [Accessed 5 March 2021].
- Soldevilla, M. S., Rice, A. N., Clark, C. W., and Garrison, L. P. (2014). Passive acoustic monitoring on the north Atlantic right whale calving grounds. *Endanger. Species Res.* 25, 115–140. doi: 10.3354/esr00603
- Stowell, D., and Plumbley, M. D. (2014). Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning. *PeerJ* 2:e488. doi: 10.7717/peerj.488
- Stowell, D., Wood, M., Stylianou, Y., and Glotin, H. (2016). “Bird detection in audio: a survey and a challenge,” in *Proceedings of the 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, (Piscataway, NJ: IEEE). doi: 10.1109/MLSP.2016.7738875
- Sueur, J., Pavoine, S., Hamerlynck, O., and Duvail, S. (2008). Rapid acoustic survey for biodiversity appraisal. *PLoS One* 3:e4065. doi: 10.1371/journal.pone.0004065
- Sugai, L. S. M., Silva, T. S. F., Ribeiro, J. W., and Llusia, D. (2019). Terrestrial passive acoustic monitoring: review and perspectives. *Bioscience* 69, 15–25. doi: 10.1093/biosci/biy147
- Taylor, A., McCallum, H. I., Watson, G., and Grigg, G. C. (2017). Impact of cane toads on a community of Australian native frogs, determined by 10 years of automated identification and logging of calling behaviour. *J. Appl. Ecol.* 54, 2000–2010. doi: 10.1111/1365-2664.12859
- Towsey, M. (2017). *The Calculation of Acoustic Indices Derived from Long-Duration Recordings of the Natural Environment*. [Online]. Available online at: <https://eprints.qut.edu.au/110634> (accessed February 11, 2020).
- Towsey, M. W., Truskinger, A. M., and Roe, P. (2015). “The navigation and visualisation of environmental audio using zooming spectrograms,” in *Proceedings of the ICDM 2015: International Conference on Data Mining*, (Atlantic City, NJ: IEEE). doi: 10.1109/ICDMW.2015.118
- Towsey, M., Znidersic, E., Broken-Brow, J., Indraswari, K., Watson, D. M., Phillips, Y., et al. (2018b). Long-duration, false-colour spectrograms for detecting species in large audio data-sets. *J. Ecoacoustics* 2:6. doi: 10.22261/jea.Iuswui
- Towsey, M., Truskinger, A., Cottman-Fields, M., and Roe, P. (2018a). *Ecoacoustics Audio Analysis Software v18.03.0.41*. (Version v18.03.0.41). Zenodo. Available online at: <http://doi.org/10.5281/zenodo.1188744> (accessed May 27, 2021).
- Towsey, M., Zhang, L., Cottman-Fields, M., Wimmer, J., Zhang, J., and Roe, P. (2014). Visualization of long-duration acoustic recordings of the environment. *Procedia Comput. Sci.* 29, 703–712. doi: 10.1016/j.procs.2014.05.063
- Truskinger, A., Cottman-Fields, M., Eichinski, P., Towsey, M., and Roe, P. (2014). “Practical analysis of big acoustic sensor data for environmental monitoring,” in *Proceedings of the 2014 IEEE Fourth International Conference on Big Data and Cloud Computing*, (Piscataway, NJ: IEEE), 91–98. doi: 10.1109/BDCLOUD.2014.29
- Walters, C. L., Freeman, R., Collen, A., Dietz, C., Fenton, M. B., Jones, G., et al. (2012). A continental-scale tool for acoustic identification of european bats. *J. Appl. Ecol.* 49, 1064–1074. doi: 10.1111/j.1365-2664.2012.02182.x
- Wiggins, S. (2007). *Triton Software Package*. San Diego, CA: Scripps Whale Acoustic Lab.
- Wiggins, S. M., and Hildebrand, J. A. (2007). “High-frequency acoustic recording package (HARP) for broad-band, long-term marine mammal monitoring,” in *Proceedings of the Symposium on Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies, 17–20 April 2007 2007 Tokyo, Japan*, (Piscataway, NJ: IEEE), 551–557. doi: 10.1109/UT.2007.370760
- Wimmer, J., Towsey, M., Roe, P., and Williamson, I. (2013). Sampling environmental acoustic recordings to determine bird species richness. *Ecol. Appl.* 23, 1419–1428. doi: 10.1890/12-2088.1
- Wrege, P. H., Rowland, E. D., Keen, S., Shiu, Y., and Matthiopoulos, J. (2017). Acoustic monitoring for conservation in tropical forests: examples from forest elephants. *Methods Ecol. Evol.* 8, 1292–1301. doi: 10.1111/2041-210x.12730
- Znidersic, E., Towsey, M., Roy, W. K., Darling, S. E., Truskinger, A., Roe, P., et al. (2020). Using visualization and machine learning methods to monitor low detectability species - the least bittern as a case study. *Ecol. Inf.* 55: 101014.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Brodie, Towsey, Allen-Ankins, Roe and Schwarzkopf. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.