Bioacoustic Unit Recognizer Protocol

May 2016

Table of Contents

1 Bioacoustic Unit

The Bioacoustic Unit (BU) is a partnership between the Alberta Biodiversity Monitoring Institute (ABMI) and the University of Alberta. The BU is a full-service organization that provides everything a client requires for wildlife acoustic studies. We provide advice, supply standardized protocols, provide equipment, conduct fieldwork, process audio recordings, and report on the results. Our team is actively conducting leading-edge research to improve methods and to understand acoustic wildlife communities better.

The recording equipment typically used for BU studies are manufactured either by Wildlife Acoustics (Song Meters SM2+ or SM3, www.wildlifeacoustics.com) or River Forks (www.riverforks.com).

Clients regularly collaborate with us to assist with their wildlife monitoring needs. Our involvement varies from client to client and spans the full range of services from simply providing information to conducting a full research project on their behalf.

Suggested Citation: McLeod, L. 2016. Bioacoustic Unit Recognizer Protocol Bioacoustic Unit, University of Alberta and Alberta Biodiversity Monitoring Institute. Edmonton, Alberta.

2 Protocol Background

Sound recognition software allows user to automate species identification and greatly reduces the time required to process large acoustic data sets. The software can also provide information on song phenology, calling rates, and other singing behaviours. Accurate and efficient processing methods maximizes the utility of large-scale and long-term acoustic datasets. This is especially true for rare or uncommon species because the amount of effort required to manually identify those species can be substantial.

The Bioacoustic Unit (BU) at the University of Alberta has developed several species- and vocalization-specific classifiers or 'recognizers' using the Wildlife Acoustics software Song Scope. These recognizers are freely available to the public and are included with this downloaded package.

3 Software

Song Scope™ is a digital signal processing application designed to scan audio recordings collected during field projects. Included in the software is a supervised learning algorithm that allows a user to create recognizers to automatically detect vocalizations made by specific bird species' and other wildlife. This shifts the burden of sifting through hours of audio recordings from the technician to the computer. The user provides Song Scope with annotated training

data which exemplify the vocalizations of interest and the recognizer filters through the audio recordings to find the closest matches to those vocalizations. Where high quality training data is available the recognizers can be very accurate in their ability to discriminate signal from the noise.

Refer to the Song Scope Users Manual[1] for more additional information.

3.1 Contents

The BU has developed more than a dozen recognizers to date (Table 1). These recognizers primarily target 'species at risk' birds and amphibians from Alberta[2]. Our recognizers have either been developed using field recordings made using ARU data from our study areas or high quality recordings sourced from online repositories (e.g., Xeno-Canto[2] or Dendroica[3]).

These recognizers are available to use 'as-is' and can simply be loaded into a Song Scope application and immediately applied to batch process your audio dataset. The settings used for each of these tools will be published here and these make for a useful starting point for anyone interesting in constructing their own recognizer for that species. All recognizers can be supplemented with additional training data to generalize them to new populations with regional variability in their vocalizations.

Table 1: List of available recognizers provided by the Bioacoustic Unit for download.

Recognizer	Target Species
BADO_Recognizer	Barred Owl
BBWA_Recognizer	Bay-breasted Warbler
BOOW_Recognizer	Boreal Owl
CATO_Recognizer	Canadian Toad
GHOW_Recognizer	Great Horned Owl
GRSG_Recognizer	Greater-sage Grouse
OSFL_Recognizer	Olive-sided Flycatcher
RUBL_Recognizer	Rusty Blackbird
WETO_Recognizer	Western Toad
YERA_Recognizer	Yellow Rail
CONI_Recognizer	Common Nighthawk

[1] http://www.wildlifeacoustics.com/images/documentation/Song-Scope-Users-Manual.pdf
[2] http://xeno-dns.org/
[3] http://www.natureinstruct.org/dendroica/

4 Recognizer Use

Note: The applicability of a recognizer to new data sets will vary depending on the specific objectives they originally were designed to fulfill. Our recognizers are provided free of charge but with no guarantee that they are suitable for your needs. Please contact the BU for recognizer-specific information prior to use.

The output from Song Scope includes a text file that lists each of the 'flagged' potential matches to the target vocalization. When the output is loaded in Song Scope the sonogram associated with each flag is displayed. These flags require a user to manually verify their accuracy which can be time consuming when the number of flagged hits is large. Also included in the output file are:

- the file name of the audio recording,
- the time stamp within the recording of each flag, and
- the duration of the flagged call or song.

Most simply, a recognizer can be applied to make a rapid assessment of the presence of a target species at a site. The output from a recognizer can also be used in evaluating variability in a species probably of detection over different hours or days at a location. Having a clear sense of your objective can allow for time saving efficiencies. Complete verification of flags may not be necessary. For example, validation until the first true detection is sufficient to assess presence or absence at a site while more effort must be allocated to achieve other goals.

4.1 Tradeoffs

Sensitivity: A primary dichotomy when using a recognizer to achieve multiple objectives is in the specificity of the recognizer. This can be managed using the 'Quality' and 'Score' settings within Song Scope even after the recognizer is created. Depending on your purpose you should consider whether you need the recognizer to be more sensitive or more specific:

1. Highly sensitive recognizers

   These recognizers are designed to be more sensitive to reduce the number of missed detections. This could be a priority for a variety of reasons including: rarity of the target species, vocalizations with low detectability (infrequent or often masked), and a desire to detect every vocalization.
2. Highly specific recognizers

   For more common species, or vocalizations which are more readily detectable, the recognizer may have been designed to be more specific and limit the number of false positives. This is often desirable because it reduces the amount of post-processing required when validated flags.

BU recognizers are generally named with a suggested quality and score setting indicated in parenthesis.

Training Data: Another critical decision is in the sourcing of the training data for the recognizer. Our recognizers are developed using (lower quality) ARU field recordings and/or (higher quality) recordings sourced from online repositories (e.g., Xeno-Canto or Dendroica). This reflects another purpose-based tradeoff:

1. Higher Quality Recordings

   Recordings made using specialized recording equipment (e.g., parabolic microphones) can have very high signal to noise ratios relative to ARU field recordings. This data can be very useful in developing a recognizer and are especially useful where existing examples from the dataset are in short supply.

2. Lower Quality Recordings

   ARU field recordings tend to better reflect the signals found in additional data sourced in the same way. Example vocalizations from our ARU datasets are justified by the need to include local and regional variations in the training data. Often these are impossible to find elsewhere!

Some BU recognizers were developed as a first-pass using external sources and subsequently incorporated examples from our own dataset as they were identified.

4.2 Limitations

The process by which a Song Scope recognizer classifies a vocalization is two-stepped. At each stage there are some functional limitations to acknowledge:

Signal Detection: In cases where the species vocalization on the recording is faint, masked by background noise, or the recording quality is poor, it is common to have difficulty with signal detection. This can result in high rates of false negative detections as these vocalizations cannot be classified. Therefore, poor signal detection might preclude you from using recognizer data for certain objectives such as determining singing rate.

Classification: Problems at the classification stage generally result from either overtraining or using training data with insufficient information. If the training data does not adequately reflect the variability of the vocalization the classifier may not generalize well when confronted with novel variation. Alternatively, if the training data is not of sufficient quality important characteristics of the vocalization may not be incorporated into the recognizer model.