Identification of Songbird Species in Field Recordings

Hsiao-Yu Tung
htung@andrew.cmu.edu
htung

De-An Huang
deanh@andrew.cmu.edu
deanh

Xiao-Feng Xie
xfxie@cs.cmu.edu
xfxie

Yurui Zhou
yuruiz@andrew.cmu.edu
yuruiz

Joseph Russino
jrussino@rec.ri.cmu.edu
jrussino

Given the documented declines in many migrant songbird populations, there is a pressing need to gain a more complete understanding of all aspects of migration patterns. During migratory periods, many species of songbirds use flight calls, which are species-specific and are distinct from other vocalizations. Therefore, flight calls information can be used to determine the relative abundance of species and is important to understand long-term population trends. Identification of bird species from continuous audio recordings has been a hot topic in recent conference competitions.[1].

Data Set Large amounts of audio data (about 20 terabytes) for bird calls have been collected. The data contain flight calls from approximately 20-30 species of songbirds, examples of other background noises, and long audio files that contain flight calls (some of them might be too faint to identify them to species) and other background sounds. The software developed can be tested using long audio files where all the flight calls of songbirds have been detected, and with manual identification by Amy Tegeler, an Avian Ecologist from the Carnegie Museum of Natural History.

Project Ideas & Software In this project, we will focus on two critical aspects of this problem. First, we will work on segmenting flight call vocalizations within long audio files that contain other types of sounds. Second, given labeled segments of bird flight calls, we will explore possible algorithms to extract useful features and classify the corresponding songbird species automatically.

We plan to start from some state-of-the-art algorithms [1, 2, 3, 4], adapt components to the data we have, and finally develop a scalable, integrated software tool. The audio data are first preprocessed into spectrograms using Fourier transformation or existing sound analysis software (e.g., Raven). The spectrograms are further cleaned by applying background noise reduction and image processing techniques. Afterward, connected pixels (acoustic patterns) in the spectrograms are labeled into rectangle segments. We then extract and select features from different sources, e.g., file statistics, segment statistics and probabilistics, and mel-frequency cepstral coefficients, etc. Finally, the classification can then be done using multiple algorithms, e.g., decision trees, support vector machines, and multi-instance multi-label learning algorithms, as well as some ensemble methods to further improve the overall performance, based on existing libraries (e.g., scikit-learn).

Work Division Yurui Zhou will work on preprocessing and segmentation;Joseph Russino will tackle feature extraction and selection; Hsiao-Yu Tung and De-An Huang will focus on classification algorithms; and Xiao-Feng Xie will work on ensemble learning methods. Each group member will do assistance work on at least one task of others to gain experience and improve on the overall system.

Midterm Milestone Develop an initial workable software version and perform basic tests based on indoor voice data, after selecting suitable algorithm components that might work scalable.

Project Outline The software developed for this project will be used by the Carnegie Museum of Natural History, and possibly shared with other land managers, researchers, and educators to enhance the use of flight calls as a method to study the populations of migratory songbirds.

[1] ICML 2013: The Bird Challenge; NIPS 2013: Multi-label Bird Species Classification; MLSP 2013: Bird Classification Challenge

References

[1] F. Briggs, X. Z. Fern, R. Raich, and Q. Lou. Instance annotation for multi-instance multi-label learning. *ACM Transactions on Knowledge Discovery from Data*, 7(3):14, 2013.

[2] M. Lasseck. Bird song classification in field recordings: Winning solution for NIPS4B 2013 competition. In *Workshop on Neural Information Processing Scaled for Bioacoustics*, pages 176–181, 2013.

[3] L. Massaron. Ensemble logistic regression and gradient boosting classifiers for multilabel bird song classification in noise. In *Workshop on Neural Information Processing Scaled for Bioacoustics*, pages 190–194, 2013.

[4] E. Stattner, W. Segretier, M. Collard, P. Hunel, and N. Vidot. Song-based classification techniques for endangered bird conservation. In *Workshop on Machine Learning for Bioacoustics*, pages 67–73, 2013.