

APPROACHES TO MULTIPLE CONCURRENT SPECIES BIRD SONG RECOGNITION

Jonathan Springer, Zhiyao Duan, Bryan Pardo

Electrical Engineering and Computer Science, Northwestern University, Evanston, IL, USA

ABSTRACT

The ability to automatically distinguish one bird species from another based on acoustic recordings of their song has attracted researchers for some time. Existing approaches have concentrated on the case where only a single species is present in a recording. This limits real-world utility of such systems since multiple species of birds frequently sing concurrently and are captured in the same recording. We explore two approaches to identifying bird species present in recordings with multiple concurrent species: first, perform source separation through binary masking and then do species identification on the separated recordings; second, apply a multi-label classification system on the original audio mixture. We compare approaches on mixtures of two to four concurrent bird species.

Index Terms— source separation, multi-label classification,

1. INTRODUCTION

The recognition of bird species in a given environment has been of interest to ornithologists and others for some time [7]. Birds have significant direct contact with humans, and because of this, monitoring their environments ensures that we as humans fully understand and evaluate the changes in our environment. For a formal review of the problem statement see [7].

In recordings made in the field, multiple species of birds are often captured into the same recording. Existing research makes the assumption of one bird per recording [1,2,3,6,7] making these approaches inappropriate for many real-world situations. In this work, we explore two approaches to identifying bird species in recordings containing multiple concurrent species. The first approach is to perform source separation through binary masking, guided by a multi-pitch tracking algorithm and then perform species identification on the separated recordings. The second approach is to apply multi-label classification on the original audio mixture.

2. DATA SET

We used a dataset provided by Cornell [4] consisting of 10 different species of birds: Bar Winged Wood Wren, Elegant Crescentchest, Mouse Colored Tyrannulet, Plain

Tailed Wren, Rufous Vented Tapaculo, Speckle Breasted Wren, Stripe Backed Antbird, Three Striped Warbler, White Bellbird, and White Throated Toucan. Each file contained a single bird species and all were recorded with very little background noise. To create multi-species mixtures, we combined recordings, linearly adjusted the amplitude of the file(s) so that all bird recordings had equal amplitude in the final mixture. This resulted in a total of 797 mixtures, consisting of 111 two-species mixtures, 311 three-species mixtures, and 375 four-species mixtures.

Recordings were then either encoded as feature sets for roughly 30ms time frames (in the case of direct multi-label classification) or separated into individual audio files by a source separation algorithm and then encoded as audio features. The features we used were based off those used in [7]: twelve Mel-Frequency Cepstral Coefficients (MFCCs), zero crossing rates, and spectral centroid, flux, and rolloff power. Features were calculated using the Marsyas framework [8].

3. SOURCE SEPARATION

Since all bird species in this study produced highly harmonic sounds, source separation was performed using harmonic masking, guided by the multi-pitch tracker described in [5]. The individual pitch tracks for each bird were calculated using the ground truth pitches calculated from single species files. Then, a harmonic mask was created for each tracked source, isolating energy at integer multiples of the fundamental frequency. If multiple sources occur in the same time-frequency bin, an even portion of the sources is distributed to the separate harmonic masks. An overlap and add method [21] was used to recombine these spectrograms into our desired time amplitude waveforms.

4. CLASSIFICATION

We took two high level approaches to recognizing multiple bird species in a single recording. First: apply source separation to the audio and feed the resulting individual sources to a single-label trained classifier. Second: feed a mixture of multiple recordings to a multi-label classifier [10] trained on single species recordings. In the second case, the classifier makes the decision as to which and how many of the species were in each individual file. In this work, all classifiers were implemented using weka [11], for single

label classification, or meka [9], for multi-label classification.

We used two classifiers – Support Vector Machine (SVM) with radial basis function kernel & a Multilayer Perceptron (MLP) - trained on single species data and tested on the source separated data after a mixture of 2, 3, and 4 species. The same two classifiers (SVM and MLP) are trained and tested on non-source separated data using Classifier Chains [10]. Essentially, classifier chains are an extension of a binary relevance method in the sense that classifier chains are able to model inter-label correlations while maintaining acceptable computational complexity; a more formal review can be found [10].

5. EXPERIMENTAL RESULTS

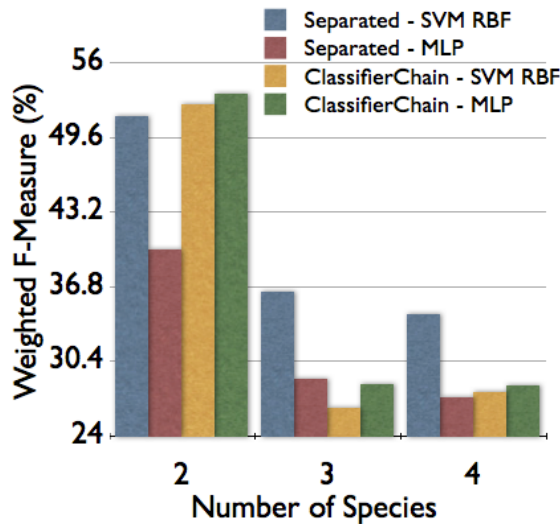


Figure 1. Weighted F-Measure Bird species ID

Figure 1 shows average weighted F-measures for classification of the testing set of 100 mixtures of each number of species in a file. We also experimented with a majority rule based classifier. For each (how many millisecond) time frame within a file, assign a decision (or multiple decisions for the multi-label classifiers) and sum the number of decisions for each label in a file. If there were three species in the mixture, the three labels that occur most are considered the returned answer.

Figure 2 shows these results.

6. CONCLUSIONS

Based on these results, it appears that frame-wise majority rule voting using classifier chains is a better approach to bird species identification in multi-source mixtures than application of harmonic source separation prior to identification. We believe that, moving forward, the main problem to solve is not development of new classification

approaches, but to develop better acoustic feature representation and source separation algorithms.

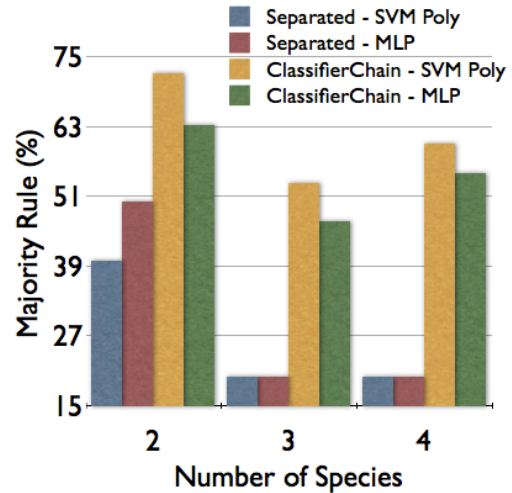


Figure 2. Majority Rule Bird Species ID

7. REFERENCES

- [1] Chih-Hsun Chou; Pang-Hsin Liu; Bingjing Cai; , "On the Studies of Syllable Segmentation and Improving MFCCs for Automatic Birdsong Recognition," *Asia-Pacific Services Computing Conference, 2008. APSCC '08. IEEE* , vol., no., pp. 745-750, 9-12 Dec. 2008
- [2] Chih-Hsun Chou; Pang-Hsin Liu; , "Bird Species Recognition by Wavelet Transformation of a Section of Birdsong," *Ubiquitous, Autonomic and Trusted Computing, 2009. UIC-ATC '09. Symposia and Workshops on* , vol., no., pp.189-193, 7-9 July 2009
- [3] Chu, W.; Blumstein, D.T.; , "Noise robust bird song detection using syllable pattern-based hidden Markov models," *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on* , vol., no., pp.345-348, 22-27 May 2011
- [4] Cornell Lab of Ornithology – Macaulay Library, <http://birds.cornell.edu/MacaulayLibrary/>
- [5] Duan, Z.; Pardo, B.; Changshui Zhang; , "Multiple Fundamental Frequency Estimation by Modeling Spectral Peaks and Non-Peak Regions," *Audio, Speech, and Language Processing, IEEE Transactions on* , vol.18, no.8, pp.2121-2133, Nov. 2010
- [6] Fagerlund, S. "Bird species recognition using support vector machines", *EURASIP Journal on Applied Signal Processing*, v. 2007 n.1, p.64-64, 1 January 2007
- [7] Lopes, M.T.; Gioppo, L.L.; Higushi, T.T.; Kaestner, C.A.A.; Silla, C.N.; Koerich, A.L.; , "Automatic Bird Species Identification for Large Number of Species," *Multimedia (ISM), 2011 IEEE International Symposium on* , vol., no., pp.117-122, 5-7 Dec. 2011
- [8] Marsyas Software Package, <<http://marsyas.info/>>, accessed in July, 2012.
- [9] Meka Software Package, < <http://meka.sourceforge.net/>>, accessed in November 2012.
- [10] Read, J.; Pfahringer, B.; Holmes, G.; Frank, E. Classifier Chains for Multi-label Classification. *Machine Learning Journal*. Springer. Vol. 85(3), pp 333-359. (May 2011).
- [11] Weka Software Package, <<http://www.cs.waikato.ac.nz/ml/weka/>> , accessed in September 2012.