DOI: http://dx.doi.org/10.7551/978-0-262-31709-2-ch120

Acoustic sensor arrays for understanding bird communication. Identifying Cassin's Vireos using SVMs and HMMs

Julio G. Arriaga [1], George Kossan[2], Martin L.Cody[2], Edgar E. Vallejo[1] and Charles E. Taylor[2]

[1]ITESM Campus Estado de México, Atizapán de Zaragoza, 52926, México
[2]University of California, Los Angeles, Los Angeles, CA, 90095, USA
taylor@biology.ucla.edu

Abstract

In this paper, we present a series of experiments on the automated classification of Cassin's Vireo individuals from song phrases using support vector machines and from sequences of song phrases using hidden Markov models. Experimental results show that accurate classification of bird individuals can be achieved using these two different levels of description of bird songs.

Introduction

Understanding the structure and function of bird songs is a long-sought goal in ecology research. Recent advances in sensor arrays, machine learning and computational linguistics finally make the achievement of this goal feasible. Understanding bird songs may also prove helpful in guiding the construction of artifacts that possess high-level communication abilities.

Over the last few years we have collected very large amounts of bird song recordings from acoustic sensor arrays in a variety of natural settings. This data have been processed by localizing source with beamforming, then filtering out noise, identifying events of interest, and then classifying them according to species and individual, and combining that with behavioral observations in a large database.

Our previous work on acoustic classification of birds has been successful at recognizing several species of antbirds and antbird individuals in a Mexican rainforest, Vallejo and Taylor (2009), Trifa et al. (2008). These birds possess quite simple, but distinctive, songs which are thought to be innate. In contrast, songbirds have a vocal organ highly developed that normally produce the relatively long and complicated vocalizations which are usually learnt, Catchpole and Slater (2008).

Particularly, the work presented here aims at exploring to what extent the methods we have used in the past are able to address and conduct the classification of songbird individuals. Further, it would be very useful for our research goals to understand the classification capabilities and limitations of sensor arrays when dealing with different levels of description of bird songs –song phrases, sequences of song phrases, etc. Toward this end, here we explore on classification of Cassin's Vireo (CaVi) individuals from song phrases using support vector machines (SVMs) and from sequences of song phrases using hidden Markov models (HMMs).

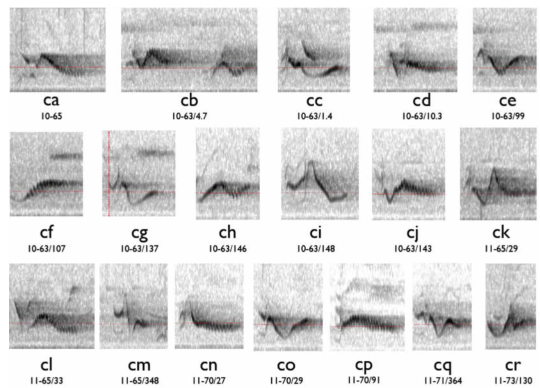

Figure 1: A subset of the CaVi phrases

Identifying CaVi individuals from song phrases using SVMs

The species of birds in our analysis have been the Cassin's Vireo (*Vireo cassinii*) a North American songbird, ranging from southern British Columbia in Canada through the western coastal states of the United States. The song consists of sequences of short, rough whistled phrases of several notes. The songs used in this work were recorded from April 2010 to July 2012, by Martin L. Cody. Examples of the CaVi songs are posted on http://taylor0.biology.ucla.edu/al/bioacoustics/.

A collection of 65 different phrases was identified by visual inspection of the sonograms. The sonograms of some of the CaVi phrases are in Figure 1, above. An example of extracted song grammar for a sample of the dataset is described by the Markov chain of figure 2. Samples of 12-53 phrases from each of the 12 individuals were included in the dataset. The sonogram of each phrase was measured for 124 traits using the Marsyas software package, Tzanetakis and Cook

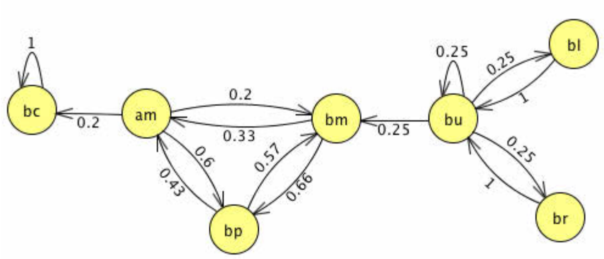

Figure 2: Markov chain of CaVi song. The states correspond to song phrases and the arrows indicate the transition probabilities among phrases

(1999) so that each song was represented by a vector. From these vectors, principal components were extracted and represented by a vector of 26 principal components.

Classification of individuals by SVMs was conducted using the Weka package, Witten et al. (2011). A radial basis function (RBF) kernel was used for the experiments. 10-fold cross-validation was conducted to find appropriate kernel parameters. Training was performed using the obtained kernel parameters on the training set. Testing was conducted using data samples not included in the training set. The classification results obtained in our experiments are in Figure 3.

Identifying Cassin's Vireo individuals from sequences of song phrases using HMMs

Samples of 44 songs represented by sequences of phrases from each of the 12 individuals were included in the data set. Classification of individuals by HMMs was conducted using the Accord package http://accord.googlecode.com/. The Baum-Welch training algorithm was used for the experiments. A collection of training and validation experiments was conducted using this dataset. The classification results obtained in our experiments for each of the CaVi 12 individuals are presented in Figure 3. Results for HMMs are shown in gray, while those for SVMs in black. In nearly all cases the precision of classification was high - 90% or better. The HMMs appear to achieve slightly higher precision for practically all phrase types except for two, where there results were quite poor, only 40-50%, while the SVM results remained high throughout.

Conclusions

Here we show that is feasible to discriminate songbirds individuals with similar accuracy from different levels of descriptions of birds songs –on the one hand, using their spectral and temporal acoustic features, and on the other hand, using the composition of their sequences of phrases.

The results presented are currently being analyzed by computational-linguistic tools to identify the syntax of the songs, and combined with information about the context in

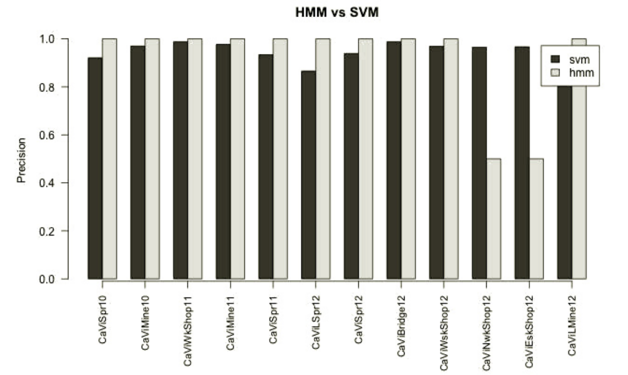

Figure 3: Classification of CaVi individuals. HMMs are shown in gray, those for SVMs in black

which they occurred, then analyzed by new software methods to identify the meaning of those songs. These methods will draw inferences from those meanings and explore consequences for individual and community ecology.

We believe this work will contribute to the recognition of very sophisticated signaling strategies and syntactic structures in non-human species. In addition, the work reported here can contribute to expand the range of engineering with voice recognition and classification, which so far has been restricted almost exclusively to humans. They hold promise to elucidate the fundamental properties of bird language. These results could then be useful to make progress on enabling high-level communication in artificial agents.

Acknowledgements

This work was supported by the US National Science Foundation under Award Number 1125423 and by Consejo Nacional de Ciencia y Tecnología under Award Number I010/214/2012.

References

Catchpole, C. K. and Slater, P. J. B. (2008). *Bird Song. Biological Themes and Variations. Second Edition.* Cambridge University Press.

Trifa, V. M., Kirschel, A. N. G., Taylor, C. E., and Vallejo, E. E. (2008). Automated species recognition of antbirds in a mexican rainforest using hidden markov models. *The Journal of Acoustical Society of America*, 123:2424–2431.

Tzanetakis, G. and Cook, P. (1999). Marsyas: a framework for audio analysis. *Organized Sound*, 14:169–175.

Vallejo, E. E. and Taylor, C. E. (2009). Adaptive sensor arrays for acoustic monitoring of bird behavior and diversity: preliminary results on source identification using support vector machines. *Artificial Life and Robotics*, 14:485–488.

Witten, I. H., Frank, E., and Hall, M. A. (2011). *Data Mining. Third Edition.* Morgan Kaufmann.