Adaptive sensor arrays for acoustic monitoring of bird behavior and diversity: Preliminary results on source identification using Support Vector Machines

Edgar E. Vallejo
Computer Science Dept.
ITESM, Campus Estado de México
Atizapán de Zaragoza, 52926, México
vallejo@itesm.mx

Charles E. Taylor
Dept. of Ecology and Evolutionary Biology
University of California, Los Angeles
Los Angeles, CA, 90095, USA
taylor@biology.ucla.edu

Abstract

We summarize the work in our laboratories toward developing adaptive sensor arrays for monitoring bird vocalizations. We have focused on four species of antbirds in a tropical rainforest of Mexico. Preliminary results on individual identification using Support Vector Machines are presented. Also, we describe our initial attempts at higher order processing of information about the identification and localization of each source.

1 Introduction

Adaptive sensor arrays provide excellent platforms for testing hypothesis on different aspects of adaptive behavior such as collective and social behavior, communication and language, emergent structures and behaviors, among others. This technology holds the potential to produce a major paradigm shift in the way we interact with the physical environment. Furthermore, understanding the capabilities and limitations of sensor arrays will be important for guiding the construction of artifacts that possess problem-solving abilities.

In this study we are concerned with developing acoustic sensor arrays so that they will be useful for observing and analyzing bird diversity and behavior. We would like each sensor to see and "understand" part of the situation – depending on its own location – then to fuse their experiences with other such sensors to form a single, coherent understanding by the ensemble [1]. The ideal is that the array will act something like a living membrane, sensitive to what is going on within it, around it and passing through it.

Toward that goal we have developed and tested sensor arrays that can identify their own location and sense bird vocalizations in real-world settings. We have developed filters to identify species (in some instances individual birds) and software tools to localize those individuals in natural environments. More recently, we are beginning to explore how we can identify the meaning of these vocalizations in the social context of the vocalizing animals. Separate aspects of this work has been described elsewhere, indicated below. In this paper we will briefly touch on those topics, but focus on the issue of individual recognition by Support Vector Machines [2].

2 Methods and tools

2.1 Biological context

The principal field site for our work has been the rainforest environment at the Estación Chajul, Reserva de la Biósfera Montes Azules, in Chiapas México (approximately $16°6'44''$ N and $90°56'27''$ W). The species of birds in our analysis have been the Barred Antshrike (BAS) (*Thamnophilus doliatus*), Dusky Antbird (DAB) (*Cercomacra tyrannina*), Great Antshrike (GAS) (*Taraba major*), and the Mexican Antthrush (MAT) (*Formicarius analis*). A sonogram of two MATs is in Figure 1, below, and those for the other species are illustrated in [3]. Examples of the songs from these species are posted on http://taylor0.biology.ucla.edu/al/bioacoustics/.

2.2 Sensor arrays

We have developed and tested an acoustic platform with small microphone sub-arrays that can be deployed 10-30m apart. They are automatically calibrated, to determine their location and orientation, then activated to perform streaming event recognition, and acquire data when triggered by animal vocalizations. Details on the development and implementation of the Acoustic ENSBox platform are described in [4].

2.3 Event recognition

Event recognition a critical first step to processing that follows, triggering source recognition and localization. We find that streaming adaptive statitical classifiers are a good approach in most cases.

We have implemented a marmot alarm call detector which runs in real time on the Acoustic ENSBox platform. Since the nodes are connected to each other via wireless ethernet, we make one additional improvement. If one node detects an event, that node tells all the others to trigger. In this way, the thresholds can be set quite high since only the node nearest to the event needs to detect it. Specific performance results are reported in [5].

2.4 Acoustic bird species recognition

We have developed filters to identify species, and individual birds in natural environments. We have taken several approaches. We have obtained promising results by extracting a sonogram of the vocalization, then look at particular features of those sonograms that might be particular to the species.

We have found it most helpful is to adapt methods from human voice recognition to create a Markov Transition Matrix appropriate to the vocalizations of each individual or species. We are also looking at other methods that appear promising, especially data mining and Self-Organizing Maps.

Trifa [3] describes in detail our experience with using HMMs to discriminate among different species of antbirds. In general, discrimination is at least 90% successful. We are currently directing efforts at identifying individuals, with quite positive preliminary results.

Similarly, we have explored with the use of data mining for the classification of bird species. The main goal has been to understand the importance of particular features of the acoustic signal that are distinctive for the accurate discrimination of bird species. A secondary goal has been to reduce the dimensionality of the acoustic signal in order to minimize the computational resources required for its manipulation and analysis [6].

Escobar [7] employed Self-Organizing Maps (SOMs) for the acoustic classification of bird species. The overall goal has been to examine the scope in which unsupervised learning is capable of conferring meaningful categorization abilities and increasing autonomy to sensor arrays.

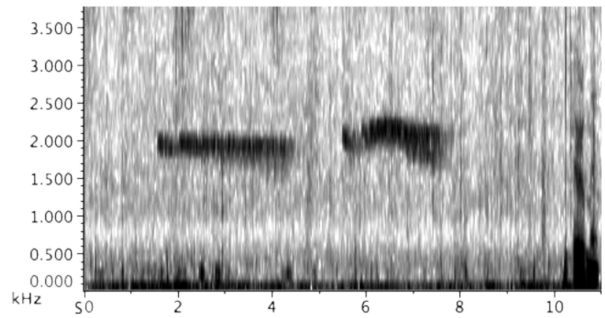

Figure 1: Spectrograms for two MATs from one territory

2.4.1 Acoustic bird individual recognition

It is sometimes possible to distinguish individual singers. For example, Figure 1 shows sonograms for the songs of we inferred to be two MATs singing on the same territory. Songs were recorded from each of 10 birds during December 2006, by Martin Cody. The identitication of each singer was inferred from timing and location. Samples of 20 - 50 songs from each of the 6 territories they occupied were included. The sonogram of each song was measured for 20 traits, including length and maximum or minimum frequency at various parts of the song, so that each song was represented by a vector. The standardized variance-covariance matrix for all songs was calculated and principal components extracted. Each song is plotted in the first two principal axes of Figure 2. The convex hull of songs for what we identified as each individual shows the clustering. It is apparent that some individuals are clearly distinguished while others are much less so, at least when plotted in these two dimensions. We are currently exploring ways to automate this procedure and increase the power of discrimination, with the goal of identification being done in real time on each node in the array.

Particularly, we have recently explored with the use of Support Vector Machines (SVMs) for the classification of individual MATs. Using feature selection we reduced the dimensionality of the original vector to 7

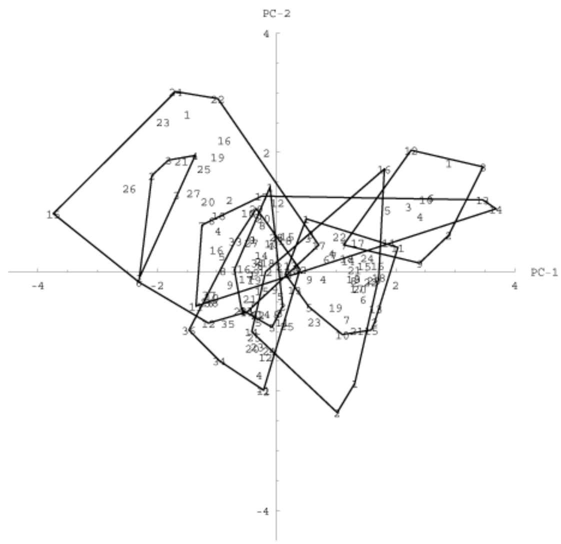

Figure 2: Convex hulls of principal component scores from 10 putatively individual MATs

individual	samples	training	testing
PMPa	28	21	7
PMPb	22	16	6
LGEa	12	9	3
PBEa	25	19	6
AVEa	38	30	8
LCNb	17	13	4
SNWa	20	15	5

Table 1: Individual MATs data set

features. Additional data selection yielded the data set used in our experiments with SVMs (Table 1).

We conducted simple scaling to the data (z-scores). A radial basis function (RBF) kernel was used for the experiments. N-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. This procedure is fully described in [8].

The classification results obtained in our preliminary experiments are presented in Table 2.

2.5 Localizing sound sources

When an array of sound sensors are employed, localizing the source of a sound should be possible in any of several ways – including comparison of sound energy, comparison of time of arrival of the sounds, and analysis of phase relations of the sound waves. In the rainforest, comparing sound energy is difficult because of reflection and interference from the vegetation. Comparing time of arrival is made difficult when sensors are widely spaced because of drifting time synchronization among widely spaced processors that need to process the sounds. Consequently, we have focused our efforts on comparing phase relations among the several microphones on the sub-arrays described in section 2.2 above. Within the sub-array there are expected to be differences in the phases that arrive at the several sensors, but time synchronization is achieved by using the same or closely coupled processors. While not permitting localization as such, this method does permit estimation of direction of arrival (DOA) to any one sub-array. Triangulation of estimated DOA from several sub-arrays can then be used to identify the location, itself, of the source.Our colleague Kung Yao and his students have developed algorithms for estimating DOA in these circumstances. Their method, termed "Approximate Maximum Likelihood" (AML), is described in [9] and [5].

procedure	accuracy	classified	misclassified
training	94.30	116	7
testing	84.62	33	6

Table 2: Classification results

2.6 Emergent understanding

Our long term goal is to provide sensor arrays with the adaptation capabilities required to identify the meaning of bird vocalizations in the social context of the vocalizing animals. This requires event recognition, symbol grounding and adaptive communication in order for the array to arrive at a collective understanding [10]. Previous studies have established plausible scenarios for the emergence of these capabilities in sensor arrays [11].

Symbol grounding, identifying and binding semantically meaningful events to symbols, then communicating that information among parts of the arrays is of great importance. We are currently examining methods based on information theory [12].

Once events have been recognized then we can use self-organizing maps to categorize the songs . A problem has been that new events might be attached to one symbol in one part of the array, but to another symbol in other parts of the array. We have determined, to some extent, the conditions under which the different "meanings" will converge or remain separate [13].

Finally, we are developing the linguistic structure that is necessary to describe these songs and events in an expressive, learnable manner, based on the ideas developed by Stabler [14].

3 Conclusions

Overall, adaptive sensor arrays seem promising platforms for habitat monitoring applications. In the near future, our efforts will be directed towards enabling sensor arrays with increasing adaptability and cognitive abilities. To accomplish this we will build largely on the results reported here.

Acknowledgements

This work was supported by the US National Science Foundation under Award Number 0410438 and by Consejo Nacional de Ciencia y Tecnología under Award Number REF:J110.361/2007. The authors would like to thank Martin L. Cody, Kung Yao, Edward P. Stabler, Lewis Girod and Travis C. Collier for their invaluable contribution to this paper.

References

[1] Taylor, C. E., From cognition in animals to cognition in superorganisms, In M. Bekoff, C. Allen and G. Gurghardt, (eds.), *The Cognitive Animal. Empirical and Theoretical Perspectives on Animal Cognition* The MIT Press, 2002.

[2] Cristianini, N. and Shawe-Taylor, J., *An Introduction to Support Vector Machines.* Cambridge University Press, 2000.

[3] Trifa, V. M., Kirschel, A. N., Taylor, C. E., Vallejo, E. E: Automated species recognition of antbirds in a Mexican rainforest using hidden Markov models, *The Journal of Accoustical Society of America* (In press)

[4] Girod, L., Lukac, M., Trifa, V., Estrin, D., The Design and Implementation of a self-calibrating distributed acoustic sensing platform, In *ACM SenSys, 2006.*

[5] Ali, A. M. S., Asgari, T. C. Collier, M. Allen, L. Girod, R. E. Hudson, K. Yao, C. E. Taylor, Blumstein, D. T., An empirical study of collaborative acoustic source localization *Wireless Sensor Networks*, 2008.

[6] Vilches, E., Escobar, I. A., Vallejo, E. E., Taylor, C. E., Data mining applied to acoustic bird species recognition, In *18th International Conference on Pattern Recognition, ICPR 2006* Volume 3, pp. 400-403., IEEE, 2006.

[7] Escobar, I. A., Vilches, E., Vallejo, E. E., Cody, M. L., Taylor, C. E., Self-organizing acoustic categories in sensor arrays, In F. Almeida e Costa, M. L. Rocha, E. Costa et al (eds.) *Advances in Artificial Life, 9th European Conference, ECAL 2007.* LNAI 4648, pp. 1161-1160, Springer-Verlag, 2007.

[8] Chang, C.-C. and Lin, C.-J., LIBSVM a library for support vector machines, 2001. `http://www.csie.ntu.edu.tw/~cjlin/libsvm`

[9] Chen, C. E., Lorenzelli, F., Hudson, R. E. and Yao, K. Acoustic source DOA estimation using the cross-entropy method, In ICASSP, 2007.

[10] Lee Y., Riggle, J., Collier, T.C. et al., Adaptive communication among collaborative agents: Preliminary results with symbol grounding, In M. Sugisaka, H. Tanaka (eds), *Proceedings of the Eighth International Symposium on Artificial Life and Robotics* (AROB8th), 2003.

[11] Collier, T. C., Taylor C.E.: Self-Organization in Sensor Networks. *Journal of Parallel and Distributed Computing* **64**:7 (2004) pp.866–873.

[12] Kobele, G. M., J. Riggle, R. Brooks, D. Friedlander, C. Taylor, E. Stabler., Induction of Prototypes in a Robotic Setting Using Local Search MDL, in M. Sugisaka and H. Tanaka, (eds.), *Proceedings of the Ninth International Symposium on Artificial Life and Robotics*, (AROB9th), 2004.

[13] Vallejo, E. E. and Taylor, C. E., A simple model for the evolution of a lexicon, In M. Sugisaka and H. Tanaka, (eds.), *Proceedings of the International Symposium on Artificial Life and Robotics* (AROB9th), 2004.

[14] Stabler, E. P., Collier, T. C., Kobele, G. M., et al, The learning and emergence of mildly context sensitive languages, In W. Banzhaf, T. Christaller, P. Dittrich, et al (Eds.) *Advances in Artificial Life, 7th European Conference, ECAL 2003.* Springer-Verlag, 2003.