BRIDGE: BIOACOUSTICS RESEARCH, INTERACTION DESIGN, AND GAMIFICATION IN ECOLOGY

Isak Herman
University of Cambridge
Computer Laboratory
15 J J Thompson Ave
Cambridge CB3 0FD

Alan F. Blackwell
University of Cambridge
Crucible Network
Computer Laboratory
15 J J Thompson Ave
Cambridge CB3 0FD

Chris Sanbrook
UNEP – WCMC
University of Cambridge
Department of Geography
Downing Place
Cambridge CB2 3EN

ABSTRACT

The burgeoning roles of bioacoustics and citizen science within biodiversity research gives rise to a number of questions. What information can we obtain from a collection of soundscape recordings? Is bioacoustic data sufficient to form the basis of a conservation biology metric? Do data collected by citizen scientists allow comparable analysis to that done by experts? And with sufficient soundscape data, can we emulate the analysis of experts using computational methods to a level relevant to policy-makers? We believe that the information in bioacoustic data can ultimately suffice to augment policy decisions.

We intend to develop a crowd-sourced, participatory sensing, biodiversity monitoring system combining acoustic with collaboratively filtered environmental data to provide unprecedented levels of information regarding the correlation between biodiversity shifts, acoustic environmental characteristics, and ecological change. This research falls at the intersection of biological acoustics, eco-informatics and human-computer interaction (HCI) and encompasses the design and implementation of an end user data-collection interface with inference based data analysis providing the basis for predictive ecological models.

1. INTRODUCTION

We will develop a mobile, gamified, participatory sensing, client application for time-stamped, GPS localized, optionally tagged, acoustic data collection as the input to a machine learning infrastructure for detection and classification of bioacoustic and environmental acoustic data. Tagging will allow citizen scientists to identify the source of the audio data at the species level or otherwise. Development of the application will require finding an optimal balance between acoustic processing on the mobile clients and the server which will maintain sufficiently light data on the client machines and sufficiently low data transmission requirements so that amateur naturalists and citizens who appreciate the outdoors will remain willing participants. The initial focus of our research is on the development of a mobile application that balances the technical demands of bandwidth and computation with features that motivate broad environmental acoustic data collection and user involvement and retention.

Once the data are collected and delivered to the server and the audio is fully processed, we will explore inference techniques for feature extraction from the acoustic and collaboratively filtered data, clustering the acoustic data and then cross-referencing the clusters to collaboratively filtered tags. We are interested in applying machine learning techniques from the music information retrieval (MIR) domain and automatic speech recognition (ASR) to the tasks of acoustic identification and classification that will mitigate the need for classification by expert ecologists and biologists and provide heretofore unavailable data to policy makers and local stakeholders alike. We will explore means of developing intrinsic as well as collective motivation on the part of the public for performing species identification, noise-pollution assessment and other tasks within the client application that can be incorporated into the server-side classification procedure. We intend to research the application of machine learning techniques to the identification of species from audio data, the clustering of species to communities, and the correlation of climate change data to biodiversity shifts over time based on observations of species and communities.

In subsequent sections of this paper we will discuss our identification of acoustic space in the context of ecological space as relevant to conservation biology. We will examine the role of citizen science, gamification, and participatory design in distributed data collection. We will broadly consider the organizational structures necessary for mining big acoustic data sets as well as the existent signal processing methodology that may be applicable to bioacoustic research as applied to biodiversity assessment. We will further examine prior work in machine learning approaches to acoustic data including the incorporation of crowd-sourced human input. Finally, we will conclude with a discussion of the potential for incorporating our results into the current state of conservation biology and the potential for acoustic representation of ecological health as the basis for biodiversity metrics that are sufficient to play a role in defining policy.

2. BIOACOUSTIC DIVERSITY AND CONSERVATION BIOLOGY IN ECOLOCICAL SPACES

Conservation biology encompasses the practice of conservation science as well as the application of said science to management and policy. Practitioners of conservation biology consider maintenance and restoration of biological diversity to be of paramount importance for the health of the earth as describe by the Society of Conservation Biology on their website [28]. Through our research, we intend to augment conservation biologists' policy positions by providing data that is constructed to survive and function in multiple settings and for which the provenance is explicit and transparent from our robust distributed, user driven sensor network modeled in part on the system described for stationary sensor nodes in [25].

Traditional biodiversity monitoring is performed by collating the field work of trained practitioners. This has the benefit that the data can safely be assumed to be of a high standard. However, it suffers in that the scope of such monitoring is limited by the availability of such expert practitioners, and curtailed by funding limitations. This has led to the prevalent practice of extrapolating biodiversity metrics from the health of populations of indicator species as a means of overcoming the traditional method's scalability limitations. We are working under the assumption that a diverse ecology not only exhibits an acoustic signature with a strong aesthetic component, but relevant concomitant ecological information too.

As participants in our environment, we are constantly exposed to acoustic features which can be considered signifiers of biodiversity. Bioacoustic feature diversity may in and of itself be perceived as an environmental asset by some some populations and has been shown in [21] to correlate with other metrics of biodiversity. In addition, research to be published in the Journal of Environmental Psychology presents anecdotal evidence that certain bird songs contribute to features of emotional wellbeing [17]. Widespread ecological changes, and increasing extremes in climate have led to biodiversity shifts being observed and documented around the world. With the ubiquity of mobile computing devices now capable of providing high quality audio data, finding motivations to invoke the assistance of citizen scientists in the collection of audio data for furthering our understanding of the ongoing effects of climate change on biodiversity has never been more relevant. Given the complexity of ecological systems when viewed at a large spacio-temporal scale, a data-driven approach to analysis is necessary. In order for our proposed system to successfully generate causal and predictive models for ecological science and ecosystem management it must combine: a versatile client that motivates public participation and ensures the capacity for big data generation; consistent audio processing; accurate bioacoustic detection, classification and clustering; and coherent, relevant correlation to external biodiversity assessment data.

3. CITIZEN-SCIENCE, PARTICIPATORY DESIGN AND GAMIFICATION FOR DISTRIBUTED DATA COLLECTION

Much conservation science takes place in the field, in engagement with local populations who live among threatened species and ecosystems. New forms of engagement are often achieved via the use of digital technologies. When technology is being designed to encourage broader engagement with scientific research, there are two levels at which this engagement might occur. One is the recruitment of members of the public to assist with data analysis work. This is increasingly well understood, under the name "citizen science". The second is engagement of the public in design of the technology itself - an enterprise known as "participatory design".

Where conventional science recognizes the concerns that are shared between experiment and theory, the relationship between citizen science as a variety of extended theory on one hand, and the participatory design of technology platforms for citizen science as a kind of extended experiment on the other, is not so clearly understood. Both can be perceived as abstractive - capturing local or indigenous knowledge in a manner that might later be exploited by commercial and colonial interests. Local populations are often very much aware of the fact that genetic material can be patented and exploited. Where techniques such as gamification are used to incentivise data collection and analysis work, players may well seek opportunities for financial conversion of quantitative scores. Finally, even participatory design fieldwork and consultation may be perceived as a professional contribution to co-design of a saleable product (for example, if mobile applications are distributed via Google or Apple AppStores).

The challenge in our research will be to identify modes of engagement that escape these threats, leading to ways of using appropriate technologies in local context that respect the concerns of those contexts, and are perceived as meaningful contributions to the lives of the people who use them. An acoustic biodiversity assessment system will benefit from crowd-sourcing the sensor network, enabling the collection of a large and growing data set for comparison with other pre-existing data but must remain sensitive to the needs and motivations of the users.

Mobile, participatory sensing is a burgeoning field with input from the HCI community helping to identify a number of the challenges and considerations in developing technology to facilitate community participation and foster political change. Researchers with Intel Research presented both insights to help "frame meaningful and effective [policy] interventions" as well as design implications for HCI research in their field study exploring the role of novel technological interventions [1]. They describe a process of using consumer electronics to capture, process and disseminate sensor data and the longer-term problems of applying the output of such data-collection efforts to modeling and policy. In addition to identifying the technological infrastructure for their project, they also explored the development of visualization tools and community features to support collaborative interpretation of the data. They concluded that while simple sensors may remain less accurate and precise, greater quantities of data would overcome the precision loss and more samples would allow the more sophisticated statistical techniques necessary to produce useful results. Furthermore they noted that "[s]ystems may be most effective in the environmental action context if they provide a unified interface for exploring data and taking actions."

This will necessarily inform our client design as we endeavor to create a system which motives participant response and retention. Additional, albeit botanical, research of interest regarding the development of a mobile, environmental, participatory sensing application comes from the Center for Embedded Networked Sensing at UCLA [10]. They contend that their research provides "substantial and unique datasets that have the potential for greatly increasing the volume of observational data for research" and are investigating gamification methods for motivating participants. Their research indicates that allowing volunteers to start contributing with simple tasks before allowing them to perform more complex contributions is important for ensuring useful participation and that seeding observations at the outset of the project with preexisting data to overcome initial scarcity encourages participation. They concluded that incorporating game activities as well as social aspects into their client application served to positively influence participation, improving recruitment and retention whilst increasing the volume of data collected.

Initially, we will iteratively design the mobile client, testing features to determine how best to obtain and retain satisfied citizen scientists while balancing the processor and bandwidth loads on data being sent to the servers. We propose to pursue the gamification of data collection on mobile platforms as a means to this end and as a way to maintain meaningful interaction with our citizen scientists.

4. DIGITAL SIGNAL PROCESSING FOR BIOACOUSTIC DATA

Research into bioacoustic analysis of individual species and communities is an active research discipline with numerous approaches being taken to species identification and broader assessments of ecological health. Research has been presented at a variety of conferences, from acoustics to engineering and artificial intelligence. One facet of our research question with system design implications is in selecting the structures created to allow the categorization of data points in a way that allows rapid accurate comparison and retrieval given the assumption that the user base may identify the same organism in different ways. While the raw acoustic data will need to be retained to manage robustness, our data-structures must also maintain a copy of the output of the short time Fourier transfer of the audio waveform, as well as timbral information derived from Mel Frequency Cepstral Coefficients (MFCCs), along with any further parameterized representations of the sounds we determine to be of use, as well as text tags optionally provided by the user.

Prior and ongoing research into biodiversity monitoring using acoustic sensors has been performed in Australia [22],Tanzania [21], and Greece [27], but in all circumstances tends to be focused on limited networks of stationary sensors rather than mobile, interactive sensors allowing public verbal input to augment the audio data. Some research approaches to species identification and biodiversity assessment incorporate analysis from machine learning techniques, however the applications have generally been limited to small sets of related species [3][19][23]. Prior work in Tanzania chose to forgo species or morphospecies identification and consider the acoustic signatures of communities as a mechanism for determining biodiversity within communities [21]. They selected the total acoustic entropy, which was shown to logarithmically correlate to the number of species within a community, and the temporal and spectral dissimilarities, shown to link linearly to the number of unshared species between communities, as acoustic indices into the biodiversity of localized communities. While limited in spatial scope, this approach demonstrated that in a healthy community, the species tend to share acoustic space equitably, with spectral profiles showing a high dispersion of amplitude peaks along the frequency axis. This result is of benefit to us as one problem for species identification will be source separation filtering.

Work done in Australia produced an autonomous biological sound monitoring system [22]. The project is ongoing and they are working to devise a machine learning software system to automatically categorize biological sounds into largely homogeneous categories with no a priori information about the sound sources to build an unlabeled library for post hoc identification and labeling. Although this implementation is a work in progress, the goal of correctly classifying the various calls of a single species correctly could benefit from the identification of meta-features from ASR vocal models such as MFCCs[18]. Their requirement that categories, once clustered, still need to be manually labeled could be superseded by my approach, where probabilistic reasoning will be applied to matching user tags for the sound samples to my clusters. An ongoing project in Greece amalgamates data from a predefined set of stationary sensors but offers a number of viable techniques for our approach as well (Report A.4)[27]. They incorporate thresholding into their sensors as a means of sound activity detection, reducing transmitted file sizes without loss of sound events in the raw data (Report A.5)[27]. They identify a number of algorithms, including subspace algorithms, spectral subtractive algorithms and statistical model based algorithms for audio de-noising which are needed to make recordings device neutral; the latter will be of especial import to us as we deploy our data collection software to fun on a variety of mobile devices. They propose furthermore that statistical machine learning algorithms be used for sound source separation and enumeration, sound segmentation, and sound recognition (Report A.7)[27]. While these previously discussed projects share our goals in the breadth of their data, they are yet to bridge successfully the gap between broad biodiversity assessment and species specific information, and likewise they are spatially limited by virtue of using stationary sensors rather than leveraging the potential of mobile devices.

A number of researchers are simultaneously exploring automated bioacoustic species identification and tracking on narrowly defined sets of target species. Work being done at the Intelligent Systems Research Group at the University of York describes the development of a novel bioacoustic signal recognition system and its application to the recognition of Orthoptera species in Britain using time domain signal coding (TDSC) coupled with a multilayer perceptron artificial neural network (ANN) classifier with backpropagation training [3]. Prototyping in high signal to noise ratio (SNR) environments, their system performed nearly perfectly, while the results of field trials with variable SNR, as naturally occurs in a noisy environment, were consistent given sufficient thresholding [3]. For testing and training purposes, the recordings were manually examined for species' syllables and songs which were extracted to separate files, however, the pattern analysis methods described in [8] might serve to reduce the need for such expert human examination. TDSC processing produced a stream of codewords based on the zero-crossing rate of the recorded waveform, which were coded into an A-matrix, showing the frequency occurrence of pairs of codewords. This matrix was used as input to the ANN which was trained to produce the species as output. It was shown that recognizing sound on a short time scale, using randomized 2s clips, without any prior knowledge of the signals was possible.

In another project from UCLA, researchers developed an acoustic habitat monitoring sensor network that can recognize and locate specific animal calls in real time [23][24]. Working from the premise that each species generates calls with a specific spectrogram, which are the inputs to the system, the authors perform identification by determining the maximum cross-correlation coefficient between the sensor input spectrogram and the training spectrogram for a small set of predetermined target species. They describe their sensor network model and discuss the question of processing distribution between the server and clients. In their model, signal intensity monitoring is performed on the sensor, while the server performs the computationally expensive task of target classification [23]. The UCLA researchers further develop an approach to efficiently balancing the computational and data transmission requirements of their system, demonstrating that preprocessing at the sensor for acoustic event recognition significantly limits the data transmission requirements of their system. Due to hardware limitations, they moved from cross-correlation of spectrograms to zero-crossing rates mapped to S-code from sequence of segment lengths (SSL), thus losing some information [24]. With the processing power of current mobile devices however, this should no longer be necessary, although the need to identify a set of minimum sufficient statistics and a suitable data encoding method remains.

Research from the Ecosystem Informatics Lab at Oregon State University discusses the role of data-driven analysis in ecological science in the context of vision rather than acoustics [6]. The machine learning techniques described for the counting and classification of arthropod specimens are transferable to model fitting and optimization with an acoustic feature set as well. A k-means clustering approach is defined to generate a dictionary for each class, in this case a target species. Due to high error rates with a standard dictionary approach, they pursued two further augmentations to the process, the first using stacked random forests, and the second using AdaBoost and weak classifiers; the latter approach provided the most significant drop in error rates. While this success was with image feature vectors, results from the MIR domain have shown that similar models have performed well at audio classification tasks.

My research comprises the development of a mobile application based sensor network feeding into an algorithmic, inference based, data and text mining back end. This system will benefit from collaborative filtering, but needs some direction from standard audio analysis; this as a multi-modal problem. Some of the features are standard audio analysis features, whilst others are linguistic descriptors obtained from user tagging and the probabilistic meta-features that can be heuristically identified from them. There are a number of components required for generating useful output, each with their own problems. The input audio signals will be recorded on a variety of devices, each with their own acoustic signatures and the signals will also be influenced by environmental noise. This leads to the need for a coherent approach to automate filter tuning to standardize the bioacoustic data prior to processing. Initially a noise floor will have to be filtered out, after which we will investigate using Spectral Mean Subtraction (SMS), Cepstral Mean Subtraction (CMS) and other ASR techniques for pre-processing. There remains a decision regarding what portion of the audio processing will be performed on the mobile client devices prior to transmission and whether the application might make this optional with some motivating factors introduced to encourage public participation at this stage. While previous bioacoustic analysis has tended to use lossy features to limit the data transmission requirements, we are exploring alternatives.

There is discussion in the MIR field regarding which of the standard audio features are best for audio classification tasks; generally a combination of several of the following are used: Fourier transform coefficients and the related MFCCs; a collection of spectral features and the zero-crossing rate; wavelet analysis; and autoregression coefficients. However temporal information is diminished or lost in the extraction of some of these features. We will explore possible parallels to research in timbral analysis and vocal models to aid in sound separation when recordings reflect songs simultaneously on overlapping frequencies, perhaps by combining clustered tones of similar timbre into probabilistic models for single species songs. We propose to design a system to pursue a model for onset detection and sound separation based on the research done in [8] and [24] while learning sets of meta-features, much as [11] describes it being done with our collaboratively filtered tag data. It is of note that prior results in HCI with regard to the mapping of MFCCs from musical tones to verbal descriptors has yielded inconclusive results [9], however, by removing this later stage analysis from the domain of citizen science, it is hoped that the feature vectors obtained will remain relevant as inputs to the machine learning architecture as demonstrated in [14].

5. MACHINE LEARNING FOR ACOUSTIC FEATURE EXTRACTION WITH CROWD-SOURCED INPUT

In the MIR domain, the tasks of clustering and classification tend to manifest as techniques for identification of artists and genres. As with bioacoustic signals, clusters are formed around similar spectral patterns, timbral features, and temporal patterns. Recently, hybrid algorithms and probabilistic models which combine a content-based approach with collaborative filtering have started to show promise for music classification and it is in this direction that we see parallels with our goals for analysis of tagged bioacoustic data. A summary of current trends in multi-class genre classification, clustering for musical style identification and semi-supervised learning for music similarity was published by [12].

A group from LabROSA at Columbia University have developed a number of support vector machine implementations for music information retrieval. The LabROSA group describe a system for flexible audio similarity queries using relevance feedback, specifically Support Vector Machine (SVM) active learning, with selected MFCCs from the ASR domain forming their feature vectors [14]. Concurrently they devised a SVM classifier to identify melodic notes which, after smoothing with a hidden Markov model to capture temporal constraints, was successful at recovering a series of pitch labels from a complex audio signal [7]. This is of particular value for the application of content-based audio retrieval to the identification of birds or other species' calls where the species' produce pitched vocalizations. More recently, they presented a comparison of several SVM multiple instance learners (MIL), algorithms which train classifiers from partially supervised data (i.e. labeled collections of items rather than labeled items) concluding that the decreased cost of collecting training data within the MIL framework could benefit a broad selection of MIR tasks [13]. We can use this technique to cluster different bioacoustic songs by individuals tagged as belonging to a given species helping to identify meta-characteristics present in the vocal models of the species [18].

Further work into unifying content-based and collaboratively filtered data has been presented by [26] who use a Bayesian network, the three-way aspect model first proposed in [16], to introduce latent variables that describe characteristics that cannot be observed. As the initial work with this approach was in document search, the authors first devise a mapping for a set of weights of polyphonic timbres to represent their feature vector in which all dimensions are semantically equivalent. I will experiment with incorporating this research into a model for timbral features present in bioacoustic data.

Research has also been done studying the application of ASR methods to detection and classification of the echolocation signals of five species of bat [19][20]. Using machine learning methods, they reduced the error rates by an order of magnitude compared to prior research. While standard mobile devices are unable to record echolocation frequencies, the techniques described are applicable to signals in the audible range. Using a Gaussian mixture model (GMM) for detection, their error rates were exponentially lower than when a baseline energy detector model was used [19]. Furthermore, classification techniques using GMMs and hidden Markov models (HMM) also showed significant reduction in error compared to previous methods [19]. The features and classifiers in the machine learning models are more robust to noise, an inevitable component of bioacoustic data. Prior to analysis, normalization of the broadband energy estimate was achieved by subtracting the noise floor. In addition, prior to feature extraction, SMS, similar to CMS which is common in ASR, was performed to remove the recording environment transfer function of each data acquisition system [18][20]. These techniques will be invaluable, given the variety of recording devices and environments for our sensors.

We will investigate incorporating collaboratively filtered tags with content based features for the clustering algorithm and compare the output with an approach that first generates unlabeled clusters from content-based features of audio data before applying the information in the collaboratively filtered tags to the unlabeled clusters. Labeled clusters will be essential for species classification, and community identification as well as provide the possibility to return hypothetical data groups to the clients for assessment and analysis. As the most successful models for audio analysis tend to juxtapose a number of learning algorithms, we intend to design a series of experiments to SVMs, with various kernels, with probabilistic Gaussian Process models. Initially, we hope to integrate a set of learning models into a complete system for audio feature extraction and begin testing content-based and hybrid models, for clustering. With labeled clusters, we can then begin to analyze temporal shifts in location of species and communities. Once we have obtained spatio-temporal data regarding species' movements and changes in communities, we will investigate methods for correlating these with geo/spacio-temporal trends in light of other data. Ultimately we hope to develop a set of causal and predictive models for ecological science and ecosystem management.

6. CITIZEN-SCIENCE DERIVED BIOACOUSTIC DIVERSITY METRICS FOR CONSERVATION BIOLOGY AND POLICY

The field of conservation biology is beginning to incorporate the results of citizen science initiatives into the corpus of knowledge available for guiding policy. However, acceptance of such practice is delayed in part by doubts about the reliability of data collected by non-experts. Our perspective on the role of citizen science in conservation biology is that the potential for big data is sufficient to warrant exploration, and that inference mechanisms in machine learning will, given sufficient ground truth, be capable of providing data with a high degree of reliability.

A review of the role of citizen science in conservation efforts is provided in [15] wherein the the following is observed: "[e]merging technologies influence the scientific research research process by streamlining data collection, improving data management, [and] automating quality control." The authors further note, however, that while "[n]ew technologies and skills (e.g. mobile applications, sensor networks, and gaming) will appeal to a diverse set of citizen-science participants, [it] could potentially marginalize those unwilling or unable to adopt them." Thus, in the context of our research, interaction design methodology forms the underlying approach we take to mitigating this marginalization and directs our development of multiple means for users to interact with our data collection platform.

A broad assessment of "environmental monitoring schemes to asses whether participation in data collection and analysis influences the speed and scale of decision making and action" is presented in [4] with more detailed local case studies supporting these claims in [5]. Their results support our premise that involving citizen scientists in data collection, with the opportunity for them to dynamically interact with the data they are collecting, is essential for rapid implementation of policy shifts at a local level. They found that "[a]t the village scale, monitoring schemes that involve local people … are much more effective at influencing decisions" and "[i]nvolving local stakeholders in monitoring enhances management responses at local spatial levels, and increases the speed of decision-making to tackle environmental challenges." Thus our proposed system will work simultaneously to offer the knowledge necessary to induce policy shifts at local levels by involving local stakeholders and providing them with data feedback through our gamified interface, as well as generate the broader data set necessary to provide a basis for policy at higher levels.

The current state of the art in devising biodiversity metrics, as set forth by the Biodiversity Indicators Partnership, involves the combination of expert knowledge on various indicator species from which the prevalence of other species, and the general state of biodiversity is assessed. Unfortunately, the limited means of gaining expert scientific knowledge means that there are significant gaps in knowledge and a general inverse correlation between the prevalence of data, and the prevalence of species. In [2], the authors attempt to form metrics by "determin[ing] the trend, and timing and direction of significant inflections in the trend for individual indicators and calculat[ing] aggregated indices relating to the state of biodiversity." However, they note that the current methods of data collection mean that "there are considerable gaps and heterogeneity in geographic, taxonomic, and temporal coverage of existing indicators, with fewer data for developing countries [and] nonvertebrates." Leveraging the potential of citizen scientists through gamified collection of acoustic data will serve to reduce these gaps.

While the practice of incorporating human input from citizen scientists in qualified results is still in the process of becoming broadly accepted, and there remain issues with validation, the current state of machine learning methodology is such that increases in available data will serve to diminish the errors with regard to expert observation while making available exponentially more analyzed data for the design of biodiversity metrics. Information obtained from our project will be valuable to environmental policymakers and stakeholders as it has the potential to provide data on an unprecedented scale which can form the basis for new and novel predictive models to shape our understanding of the surrounding world. The classification methods developed in the course of this research will contribute to the meta-analysis of how field recordings are processed as well as provide insight into the effects of anthropogenic impacts on the environment.

7. REFERENCES

[1] Paul M. Aoki, J. Honicky, Alan Mainwaring, Chris Myers, Eric Paulos, Sushmita Subramanian, and Allison Woodruff, *A Vehicle for Research: Using Street Sweepers to Explore the Landscape of Environmental Community Action*, Proc. ACM SIGCHI Conf. On Human Factors in Computing Systems, Apr. 2009, 375-384.

[2] Stuart H.M. Butchard *et al, Global Biodiversity:Indicators of Recent Declines*, Science 328, 1164, 2010

[3] David Chesmore, *Automated Bioacoustic Identification of Species*, Annals of the Brazilian Academy of Sciences, 2004, 76(2):435-440

[4] Finn Danielsen, Neil D. Burgess, Per M. Jensen, and Karin Pirhofer-Walzl, *Environmental Monitoring: The Scale and Speed of Implementation Varies According to the Degree of Peoples* Involvement, Journal of Applied Ecology, doi:10.1111/j.1356-2664.2010.01874.x

[5] Finn Danielsen, *et al*, *Taking Stock of Nature in Species-Rich but Economically Poor Areas: An Emerging Discipline of Locally Based Monitoring*, from *Taking Stock of Nature: Participatory Biodiversity Assessment for Policy, Planning and Practice*, ed. Anna Lawrence, Cambridge University Press, 2010.

[6] Thomas G. Deiterich, *Machine Learning in Ecosystem Informatics and Sustainability*, Proc. Of the 21st Int'l Joint Conf. On Artificial Intelligence, 2009

[7] Daniel P. W. Ellis and Graham E. Poliner, *Classification-based Melody Transcription*, Machine Learning, 2006, 65:439 – 456.

[8] Charles Fox, Iead Rezek and Stephen Roberts, *Drum'N'Bayes: On-Line Variational Inference for Beat Tracking and Rhythm Recognition,* Proceedings of the Int'l Computer Music Conference (ICMC), 2007.

[9] Claudia Fritz, Alan F. Blackwell, Ian Cross, Jim Woodhouse, and Brian C. J. Moore, *Exploring Violin Sound Quality: Invistigating English Timbre Descriptors and Correlating Resynthesized Acoustical Modifications with Perceptual Properties*, Acoustical Society of America, 2012. DOI: 10.1121/1.3651790

[10] K. Han, E.A. Graham, D. Vassallo, and D. Estrin, Enhancing Motivation in Mobile Participatory Sensing Project through Gaming, in Proceeding of the Workshop in Social Connection in the Urban Spaces (SocialUrb 2011), October, Boston, USA, 2011.

[11] Su-In Lee, Vassil Chatalbashev, David Vickrey and Daphne Koller, *Learning a Meta Level Prior for Feature Relevance from Multiple Related Tasks,* Proceedings of the International Conference on Machine Learning (ICML), 2007, pp. 489-496

[12] Tao Li, Mitsunori Ogihara, Bo Shao and Dingding Wang, *Machine Learning Approaches for Music Information Retrieval,* In M. J. Er and Y. Zhou, eds., *Machine Learning*, In-Tech Education and Publishing, 2009, pp. 259-278

[13] Michael I. Mandel and Danial P. W. Ellis, *Multiple-Instance Learning for Music Information Retrieval,* Session 5a, Int'l Symposium on Music Information Retrieval (ISMIR), 2008.

[14] Michael I. Mandel, Graham E. Poliner and Daniel P. W. Ellis, *Support Vector Machine Active Learning for Music Retrieval,* Multimedia Systems, Vol 12, No. 1, 2006, pp. 3-13

[15] Greg Newman, Andrea Wiggins, Alycia Crall, Eric Graham, Sarah Newman, and Kevin Crowston, *The Future of Citizen Science: Emerging Technologies and Shifting Paradigms*, Front Ecol Environ 2012; 10(6): pp 298-304

[16] Alexandrin Popescul, Lyle H. Ungar, David M. Pennock, and Steve Lawrence, *Probabilistic Models for Unified Collaborate and Content-based Recommendation in Sparse-Data Environments*, Proc. Of the 17th Conf. In Uncertainty in Artificial Intelligence, 2001, pp. 437- 444.

[17] Eleanor Ratcliff, Birgitta Gatersleben, and Paul T. Snowden, *Bird Sounds and their Contributions to Perceived Attention Restoration and Stress Recovery*, Journal of Environmental Psychology, Vol. 36, December 2013, pp 221-228.

[18] Giampiero Salvi, *Mining Speech Sounds*, Doctoral Thesis, KTH School of Comp. Sci. and Comm., 2006

[19] Mark D. Skowronski and John G. Harris, *Acoustic Detection and Classification of microchiroptera Using Machine Learning: Lessons Learned From Automatic Speech Recognition*, J. Acoust. Soc. Am., 119 (3), March 2006, pp 1817-1833

[20] Mark D. Skowronski and John G. Harris, *Automatic Speech Processing Methods for Bioacoustic Signal Analysis: A Case Study of Cross Disciplinary Acoustic Research*, IEEE Int'l Conf. On Acoustics, Speech and Signal Processing, vol 5, May 2006, pp 793-796.

[21] Jerome Sueur, Sandrine Pavoine, Olivier Hamerlynck, and Stephanie Duvail, *Rapid Acoustic Survey for Biodiversity Appraisal*, PloS ONE3(12):e4065. doi:10.1371.journal.pone.0004065

[22] Andrew Taylor and Gordon Grigg, *Development of an Autonomous (Acoustic) Biodiversity Monitoring System*, Final Report to CERF, 31st January, 2011, downloaded on 6/10/12 from http://www.environment.gov.au/about/programs/cerf/pubs/fellow-acoustic.pdf

[23] Hanbiao Wang, Jeremy Elson, Lewis Girod, Deborah Estrin, and Kung Yao, *Target Classification and Localization in Habitat Monitoring*, IEEE Proceedings of the International Conference on Speech and Signal Processing, April, 2003, pp.597-600

[24] Hanbiao Wang, Deborah Estrin, and Lewis Girod, *Preprocessing in a Tiered Sensor Network for Habitat Monitoring*, EURASIP Journal on Applied Signal Processing, 2003:4, 392-401

[25] Fan Ye, Gary Zhong, Songwu Lu and Lixia Zhang, *GRAdient Broadcast: A Robust Data Delivery Protocol for Large Scale Sensor Networks,* Wireless Networks 11, 285-298, 2005

[26] Kazuyoshi Yoshii, Masataka Goto, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno, *Hybrid Collaborative and Content-based Music Recommendation Using Probabilistic Model with Latent User Preferences,* Proceedings of ISMIR, 2006.

[27] AmiBio Action Reports, AMIBIO Project, downloaded on 8/10/12 from: **http://www.amibio-project.eu/en/resources/documents/report**

[28] Society of Conservation Biology website http://www.conbio.org/about-scb/who-we-are accessed 17/10/13