
© Copyright 2008 IEEE

*Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.*
Towards an Acoustic Environmental Observatory

Richard Mason, Paul Roe, Michael Towsey, Jinglan Zhang, Jennifer Gibson, Stuart Gage
Queensland University of Technology, Michigan State University
{r.mason, p.roe, m.towsey, jinglan.zhang, jennifer.gibson}@qut.edu.au, gages@msu.edu

Abstract

The need for large scale environmental monitoring to manage environmental change is well established. Ecologists have long used acoustics as a means of monitoring the environment in their field work, and so the value of an acoustic environmental observatory is evident. However, the volume of data generated by such an observatory would quickly overwhelm even the most fervent scientist using traditional methods. In this paper we present our steps towards realising a complete acoustic environmental observatory – i.e. a cohesive set of hardware sensors, management utilities, and analytical tools required for large scale environmental monitoring. Concrete examples of these elements, which are in active use by ecological scientists, are also presented.

1. Introduction

Climate change and human activity are subjecting the environment to unprecedented rates of change, so much so that traditional manual approaches to monitoring ecological dynamics are no longer adequate. The scale of task requires an automated approach. The environmental observatory metaphor is appropriate in this context because it implies remote sensing of multiple environmental targets over a broad spectrum of modalities and over an extended period of time.

In order to provide insights into ecosystem dynamics, an environmental observatory must monitor the four fundamental ecosystem processes: cycling of water, cycling of nutrients, energy transfers, and species relationships including biodiversity [1]. Of course, these four processes are strongly interconnected [2], but distinguishing them is useful if only to highlight the diversity of environmental targets that must be monitored and therefore the diversity of sensors that must be deployed. The first three of these processes can be monitored by low level sensors recording observations such as soil moisture, dissolved nutrients, solar intensity, etc. Indeed, it was for the monitoring of such data that wireless sensor networks were first proposed [3].

The automated monitoring of species relationships and biodiversity is a far more difficult task, requiring more sophisticated sensors and analytical methods; effectively the scientist’s eyes and ears must be emulated. Audio, image and video recordings are used by ecologists, but while these methods do lend themselves to automation, they pose higher order difficulties in bandwidth, data management, and analysis. This paper is concerned with just one part of the bigger challenge of an environmental observatory – the monitoring and analysis of the acoustic environment (which we refer to as a soundscape).

Acoustic data is a particularly rich source of information. Apart from the obvious use for monitoring specific species of fauna from their vocalisations, sound can be used for measuring general environmental features such as precipitation intensity, soil wetting, and atmospheric turbulence [4]. Furthermore, it is possible to localise sound and, in some cases, to identify specific individuals.

Environmental monitoring has attracted considerable interest from the sensor network community. A variety of sensors and sensor networks have been developed which can gather environmental measurements; however, gathering the data is only the first step in assessing environmental health. A complete environmental observatory requires tools for processing and analysing data and, unlike the legacy tools scientists have been using, these tools must be capable of supporting the large amount of data which sensor networks make available.

In this paper we describe a terrestrial acoustic environmental observatory and our steps towards realising this vision. The goal of the system is to enable new scales of environmental monitoring through acoustic sensor networks and it has been designed in partnership with ecologists, ensuring it is kept relevant and useful to the field. We identify three parts to the
observatory: sound capture and management, acoustic analysis and annotation, and environmental analysis. All three parts of the observatory are discussed in detail, including the tools we have developed to assist scientists during each of these stages.

The remainder of this paper is organised according to these parts, working from scientist to sensor. These sections are followed by a related work section and finally by a section on conclusions and further work.

2. Environmental analysis

2.1 Acoustics for science

Acoustics, particularly when coupled with other data, can yield rich insights into the environment and environmental change. Whilst this paper concerns the technical hurdles towards realising an acoustic environmental observatory, the system described has been developed in close consultation with ecological scientists and is already in use by those scientists. By working with the target audience, we have ensured the system is realistic, and able to address real world problems. Some typical questions which an acoustic observatory can answer are shown below.

- How is global warming affecting the environment?
- How is land use change affecting the environment?
- What value does a particular environmental bank have?
- How is agriculture affecting the environment?
- How is a pest species spreading?

Shorter term an observatory can measure human disturbances, and the health of an environmental site through comparisons with others. Monitoring can also act as an early warning system to determine if a sudden change has occurred (e.g. frogs suddenly disappearing from a river due to pollution).

Acoustics can be used to measure and recognise several phenomena including:

- Different species and species density
- Identification of species e.g. indicator species
- Species behavior
- Localisation of individuals
- Identification of individuals

In addition to fauna, acoustics can also be used to measure human disturbance e.g. traffic, weather (e.g. wind, rain) and other specialised phenomena e.g. soil moisture through microphones placed in the soil. We focus on species detection; typical fauna include: birds, frogs, insects, and some mammals.

Traditionally acoustics is used by scientists to study a particular species, especially birds and frogs, through the placement of acoustic data loggers (e.g. tape recorders) or through manual call-response monitoring. This research differs in being directed towards general environmental health assessment using a variety of sensors at scales requiring automation.

An important characteristic of general environmental health monitoring is that scientists do not necessarily know a priori what they are looking for and they may want to refine their data collection based on gathered information. Thus data collection and analysis must be flexible to accommodate scientists’ needs.

2.2 A web based observatory

A natural way for scientists to interact with acoustic, and acoustic derived, data is through the web. The web enables ubiquitous access to data from almost any connected device. It also allows data and insight to be shared for research collaboration and publication.

In general the combination of acoustic events, such as the detection of a particular species, with other data will lead to an explosion of possible correlations. Comparisons may involve:

- Other similar locations (similar environment, vegetation, similar acoustic events, etc.)
- Other close locations
- Historical data
- Data from similar times e.g. climate, weather, time
- Other data e.g. remote sensing data

A variety of techniques can be used to analyse such data including: data mining, OLAP cubes, and visualisation tools. Some observations may be relatively simple to interpret others may be considerably more complex. In general we expect scientists to have a repertoire of standard assessment tools and custom ones based on discoveries or hypotheses. Observations may also be linked to models for example to calibrate models concerning the spread of pests or the return of fauna to a re-vegetated site. Finally data must be made accessible to other external applications and collaborators through web services and a querying interface.

Thus a web based interface to an acoustic environment must support:
Analysis of acoustic events
Annotation of events
Sharing of data
Correlation of data with other data
Spatial display of data e.g. virtual earth
Advanced visualisation tools
Data freedom (data upload and download into other tools)
Sensor management and control

Once sound has been captured and annotated, the environmental investigation can begin. This investigation will likely be an iterative process with the scientist investigating the data, then perhaps returning to import further data of interest, perform new annotation tasks, or even reschedule the actual sensors to better direct the data capture.

The partnership between the environmental scientists and the information technology tools continues during the investigation. Appropriate tools are required to allow the scientists to adequately visualise and interpret the data. In some cases these tools already exist and the scientists will be familiar with their use. To support this all the data stored in the workbench must be made available for interchange with these tools.

There are three stages in the environmental assessment pipeline (see Figure 1):

- Data collection,
- Data augmentation,
- Analysis and visualisation

This entire process should be controlled by the domain scientist; however, each stage requires different kinds of software tools. It is these tools which we have developed in consultation with environmental scientists.

In addition we have developed a web based data management and provision system. This allows the scientists to access the raw data. Although this raw data is difficult to use due to its overwhelming quantity, we have found the ability to browse the raw data is essential for allaying scientist’s fears in automating their data collection. As scientists become more familiar with automated sensors and the analysis tools provided, we expect this browsing feature to decrease in importance. This system also acts as an interface for providing data to other third party applications.

2.3 Data interfaces

We provide a number of methods for exchanging data. As already described, all data is initially submitted to the system through web services. We also provide the data, both raw and annotations, for extraction via web services; however, where possible we try to make the data available in formats which can be directly imported into existing tools. The acoustic data can be exported using a REST style interface in a variety of formats including compressed MP3 files and the widely supported PCM encoded WAV format. Acoustic annotations are made available in text based Comma-Separated Values (CSV) format which is commonly supported by spreadsheets and statistical applications.

Despite the availability of many tools, there is still considerable opportunity for innovation in this area. The web service interface provides the perfect source for investigating and displaying the data in web mashups. Mashups are designed for easily combining and processing data from multiple sources and could provide a useful model for scientist to use when combining and analysing data.

Figure 1 – QUT Environmental Observatory

Figure 2 – Mashup of sensor data on map
A GeoRSS feed is also available which allows annotations to be fed with their location and time information into visualisation tools such as Virtual Earth [5] (see Figure 2). Reports from ecologists indicate this visualisation is particularly useful during manual field work as it enables them to quickly recall and interrogate existing data while planning where to concentrate their studies.

3. Acoustic analysis and annotation

Before scientists can extract useful observations from acoustic data, the data must be preprocessed and augmented with higher level information. For example, if acoustic data is being used to monitor a specific bird species, the preprocessing stage would identify calls made by that species and record information about those calls, such as their time, frequency, duration, or volume. It is the output of this preprocessing which the scientists will use to make their observations.

Previously this type of processing has been performed by the scientists in the field. Rather than using microphones and recording raw audio, the scientists would simply record the data of interest. In this process they utilise their own expertise to identify and classify environmental sound. Using this approach with a sensor network is impossible there is simply too much audio data to be processed manually. Instead, software tools are required which enable large scale acoustic analysis through automated and semi-automated recognition and classification of environmental sound.

There are two different kinds of environmental sound: deliberate communication (e.g. bird calls) and unintentional noise (e.g. rain, footsteps, tree fall, traffic). Our focus is on fauna communication.

Two approaches may be taken to acoustic analysis: the first is to attempt recognition of particular sounds (e.g. an owl screech). We refer to this as acoustic pattern matching. The second is to extract information from the audio stream without prior knowledge of the precise form of that information. These are described in the next two sections. The following section describes the hybrid approach we have taken.

3.1 Acoustic pattern matching - classification

Standard machine learning casts bird call recognition as a classification task, that is, training a classifier to recognise a fixed number of call categories. This presupposes prior knowledge of the expected birds and access to suitable recordings of their calls in sufficient number for training.

There are several reasons why this approach is not appropriate for many real-world applications: (a) there are many cases where obtaining tagged training instances is too expensive or simply not possible because the bird under investigation is rare or elusive, (b) in general, classifiers should be trained on data whose statistical characteristics accurately reflect those of the intended operational environment. Training a classifier to distinguish the calls of several species of bird may be a challenging machine learning task, but of little practical benefit if some of those birds are not found in the operational environment or if many other sounds and calls intrude. Consequently, we need a simpler, more effective method to identify specific bird calls even as acoustic samples are uploaded.

Template matching has been used successfully for pattern recognition in cases where the patterns are well-defined, that is, they have distinctive discriminating characteristics. This is true for many animal and bird calls. The advantage is that templates can be constructed with just a few, even one, training instance. Template matching also circumvents the problem of choosing appropriate negative training instances.

We have implemented a template matching algorithm based on a set of parameterised models of call syllables. The templates are independently compared to the sample data and any matches are recorded. The statistical significance of a match score is determined by reference to an appropriate noise model. This algorithm has been used in a pilot program to identify the call of the Lewins Rail from data captured by acoustic sensors. The output from the algorithm is used to greatly reduce the amount of data which must be listened to manually when monitoring the Lewins Rail.

3.2 Soundscape analysis

Soundscape analysis endeavours to examine acoustic samples without requiring any preexisting templates. Our system derives a number of indices which indicate the amount of acoustic activity in different frequency bands.

The simplest approach is to measure the ratios of acoustic energy in different one kHz bands. This approach is used by the Acoustic Habitat Quality Index (AHQI) developed by Gage et al. [6]. The AHQI provides a measure of the health of a habitat by comparing the amount of sound energy in the frequencies typically occupied by fauna with the energy in the frequencies typically occupied by manmade noises. This gives an estimation of the prevalence of either natural sounds or manmade noise pollution. The AHQI values are displayed with the individual readings.
and can also be exported in a CSV file for importing into spreadsheets and statistical packages.

We also provide a more fine-grained approach which identifies individual acoustic events, such as call syllables, in the soundscape. Closely adjacent syllables are grouped as one event. We then use a clustering algorithm to group similar acoustic events. The algorithm should produce clusters consisting of syllables uttered by the same species of animal. The identification and clustering of distinct acoustic events serves two purposes. In addition to the amount of biological activity, the number of clusters indicates the diversity of species involved. In addition, the characteristics of a cluster gives clues as to the kind of species present and which call templates should be scanned if exact species identification is required.

Our system samples the soundscape over a 24 hour cycle. By combining the results for several days’ analysis, it is possible to observe how the available acoustic bandwidth in any habitat is utilised. For example, in an environment near Brisbane Airport we observe a sharp increase in both the quantity and diversity of calls around 5am. Quantity and diversity of calls gradually declines through the day. However as can be seen in Figure 3, the greatest diversity of calls is to be found in the 6-8 kHz frequency band occupied by the calls of insects and some birds.

3.3 A Hybrid annotation approach

Whilst automatic acoustic analysis is desirable, such processes are inherently sub-optimal and error prone; however, full manual analysis of sound is infeasible even with the aid of students and the community for listening to sound. Consider that even if sensors do not continuously stream sound there is likely to be several hours sound per day per sensor, so purely manual approaches will not scale.

Additionally, scientists currently employ considerable specialist knowledge on field trips when identifying fauna. It is important that this knowledge is utilised even when making use of sensor networks, particularly in the development and training of automated systems.

We therefore provide a hybrid system with both manual and automated analysis. Our system allows scientists to manually listen to, recognise, and mark events when automated classifiers are unavailable. These manual annotations can be used by the scientists for their research, but, perhaps more importantly, can also be used to train and develop automated classifiers. These automated classifiers will then be able to scale the scientists’ specialist knowledge across the masses of data gathered by the sensor network.

As well as extracting features from the raw data, external data may be required for complete understanding of the environment. For example, many animals have characteristics which are dependent on weather or lunar patterns. For the scientist to recognise and understand these characteristics external data describing such patterns should be made available in a comprehensive interface alongside the acoustic features. Therefore augmenting the raw data becomes an iterative process where the scientist searches for acoustic features (either manually or by selecting an automated tool) then draws upon external data sources to help discover or explain patterns in the data. This may prompt the scientist to search for further acoustic features or gather other external data.

As part of our environmental observatory we provide a rich interface for playing back and annotating audio data. The interface makes use of Microsoft Silverlight to allow scientists to visualise the acoustic signal (see Figure 4). When the scientists recognise an event of interest they can select the event and insert an annotation describing it. Annotations are stored in the database to allow further querying and analysis. We have chosen a tagging (folksonomic) approach to data annotation. This enables users to build lightweight tag ontologies.
We have integrated our analysis tools into the web-based environmental observatory workbench. This workbench combines the tools for automated analysis, manual analysis, calculation of environmental indices like the AHQI, and importing external data such as weather into a single unified interface. This workbench enables scientists to investigate acoustic environmental data in a variety of ways.

4. Sound capture and management

Sound may be captured in several different ways including manual loggers (e.g. recording MP3 players or professional recorders), wireless sensor networks, and participatory sensing (community uploaded sound recordings). There are challenges associated with the quality and reputation of such data. In a real system there are also trade-offs of quality against breadth – recording MP3 players are cheap, professional recording equipment expensive. An observatory should support a variety of capture methods.

Raw acoustic data must be processed to produce acoustic events for environmental analysis. This processing may include compression, noise removal, event classification, and soundscape analysis. Since sound may be re-analysed it is preferable to keep all data. This yields data management challenges due to the voluminous nature of acoustic data. An approximate figure for moderate quality sound with moderate compression is a megabyte per minute; thus a handful of sensors recording data for several hours per day can easily produce a gigabyte of data per day.

All acoustic data needs to be tagged with the time and place of the recording, the recording instrument, and in the case of participatory sensing we also need to know the person who has undertaken the recording. This metadata needs to be stored in the database along with the sound data.

Data collected using data loggers and participatory collection can be manually uploaded onto the observatory server through a web interface. Wireless sensors can use web services in a similar way; however, wireless sensors are typically constrained in terms of bandwidth and power. Thus in the case of sensors it can be effective to push processing, analysis, and querying into sensors: generally processing uses less power than communication and bandwidth limits the degree of centralised data collection which is possible.

Data collection must also be managed and controlled. This applies for automatic (autonomous sensor) and manual data collection. Sensors must be managed and controlled and they must adaptively utilise power and bandwidth available to them. Manual collection entails scientists communicating a sensing plan to those deploying data loggers or requests for participatory sensing.

4.1 QUT System architecture

We have developed custom sensors based on Windows Mobile Smartphones for capturing acoustic data [7]. The sensors sample the acoustic landscape according to a schedule defined by a central server. The audio data is collected using external microphones and transmitted over WiFi or 3G mobile phone connections to a central data store. Solar panels are used to power the devices allowing them to be deployed indefinitely.

The sensor software is flexible allowing sensors to also work in disconnected mode. This can be used by scientists for shorter deployments, up to a week. In this case the sensors are entirely contained in a small polycarbonate box with an external microphone. At the end of the week the boxes are collected and brought back to base where their data is uploaded.

In addition, data can be uploaded to our database from other sources such as manual data capture or alternative acoustic sensors such as the acoustic sensors developed by our collaborators at Michigan State University. Web services are used for all data transfers, allowing for simple interoperation with other systems. We expect that scientists and community members will also use their existing recording MP3 players or their own mobile phones to sample the environment and contribute to the database.

The audio data collected is associated with the time and location of the recording. All acoustic data is stored enabling long-term trend detection and analysis. This is particularly important when trying to monitor longer effects such as changes to the environment due to climate change. In addition to data capture we have implemented a sophisticated management system for sensors, enabling them to be remotely controlled and monitored. This is particularly important for real world deployments. Schedules may be adjusted on the central server and sent to sensors and logging data collected and viewed on the server for sensor health monitoring.

5. Related work

Work related to this project can be divided into five main areas: environmental observatories, acoustic and video environmental monitoring, acoustic sensors,
sound analysis including sound tagging, and bird recognition.

In recent years, several environmental observing systems have been designed and implemented [8]; however, none of these observatories target acoustic observation. These include the Oceans Observatory Initiative (OOI), the National Ecological Observatory Network (NEON), the Collaborative Large-scale Engineering Analysis Network for Environmental Research (CLEANER), the Hydrologic Observatory Initiative (HOI), the Southern California Earthquake Center (SCEC) (http://www.scec.org/), the Real-time Observatories, Applications, and Data management Network (ROADNet) (http://roadnet.ucsd.edu/), Science Environment for Ecological Knowledge (SEEK) (http://seek.ecoinformatics.org/), and Laboratory for the Ocean Observatory Knowledge INtegration Grid (LOOKING) (http://lookingtosea.ucsd.edu/). The monitoring targets vary including sea [9], sky, the earth (e.g. SCEC), the galaxies, plants, and animals [10].

There are primarily two types of environmental observatories [9]: data grids and service grids. A data grid is a grid computing system that deals with data — the controlled sharing and management of large amounts of distributed data. A service grid helps realise the business potential of Web services. Large scale environmental observatory data grids include SCEC, SEEK, and ROADNet etc. Other grid projects (e.g. LOOKING and NEON) are both data and service grids. Our research aims to integrate both data and services, addressing a key challenge faced by similar projects [11].

Acoustic and video monitoring has long been practiced. There are many webcams around the world pointing to interesting objects ranging from the Eiffel tower to African waterholes [12]. Typically, these comprise a few large permanent installations which stream data for human viewing. They do not persist data nor do they provide tools for data analysis. Our system is designed to support larger numbers of relatively cheap sensors with associated data storage and analysis tools. Security surveillance systems also provide audio and video monitoring; however, they are designed for detecting gross human presence and movement and hence typically they utilise low resolution data.

A number of systems exist for monitoring animals’ behaviour, for example, Gage’s system for monitoring birds’ calls [13], strapping cameras to crows [14], monitoring seabird nesting environment on Great Duck Island off the coast of Maine [15] and Deer Net [16]; however, none of these are acoustic observatories.

The Owl project at MIT [17] used mobile phones to capture bird calls via two different approaches. The first approach allowed a phone to be called and sound to be captured through a VOIP connection. The second approach allows sound capture through a custom microphone array connected to smartphones via Bluetooth. We have found the sound quality of ordinary phone calls (which are filtered and compressed by the telephony network) to be too poor to use. Other work using mobile phones as sensors involves using phones as a basis for personal medical sensing or for social networking [18][19]. The latter can include environmental applications where people report on the environment using mobile phones.

Ornithologists usually listen to acoustic signals and identify birds using their expert knowledge, where most ecological sampling is conducted at small spatial scales or is infrequent or one-time sampling [20]. As this manual process is very tedious and time consuming, it is impractical to carry out large scale surveys or studies. Recently, automated pattern recognition techniques have been developed for bird species recognition [21][22][23]. Sensor networks bring ornithologists and pattern recognition researchers together to make some applications possible, for example, assessing bird striking risks in aviation industry, and the study of spatial and temporal variation in biological processes [20].

6. Conclusions and future work

An acoustic observatory is proposed as a valuable tool for environmental monitoring. To realise this vision several challenges must be addressed including data capture, data analysis, and environmental analysis. We have presented our steps towards realising this vision based on a network of sensors constructed from smartphones, a hybrid template matching signal processing system for sound analysis, and a sophisticated web based tool for undertaking environmental analysis and sensor management.

The system has been successfully used to monitor the environment at several locations including: Brisbane Airport for the rare Lewins Rail bird, and QUT’s Samford Ecological Reserve Facility. We are currently preparing for the deployment of more sensors to enable koala monitoring and the general monitoring of several sites of environmental value.

There are several areas where our current system needs further research and development to complete the vision of a true environmental observatory. Firstly data fusion – currently we can only support simple data correlation based on time and spatial dimensions.
Correlation with weather and lunar information should not be difficult; however, correlation with other data including qualitative observations made by scientists in the field will be much harder; as will be the incorporation of participatory observations. A second area for investigation is around tagging and semantics. With greater semantic information it will be possible to undertake more complex environmental queries. The use of Ecological Metadata Language (EML) may facilitate this. Finally our initial system has a deliberately centralised architecture with a central server and data being uploaded through 3G or WiFi networks. As the number of sensors is increased the centralised server will become a bottleneck and a different architecture will be required where data is kept and processed closer to the sensors. The system may be tried out from here: www.mquter.qut.edu.au/sensor/demo/.

7. Acknowledgements

This work was undertaken by the Microsoft QUT eResearch Centre (www.mquter.qut.edu.au). We would like to thank Microsoft Research, the Queensland Government, Brisbane Airport Corporation and Telstra for their support.

8. References