Analysing Big Audio Data for Environmental Monitoring

Jinglan Zhang

School of Computer Science

Why Environmental Monitoring?

Australia is far away from other continents.

- One of the most biologically diverse regions on Earth.
- Unique fauna and flora.
- However, species are undergoing exceptional loss of habitat due to climate change and human activities.

QUT

Queensland University of Technology

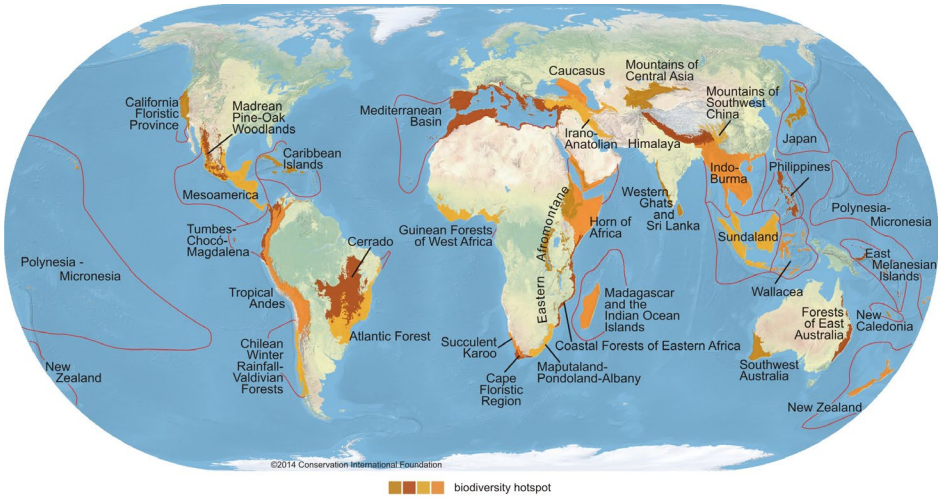
Megadiverse Countries

- Species are not distributed evenly
- 17 countries are 'megadiverse’
- has less than 10% of the global surface, but support more than 70% of the biological diversity on earth.
- represent more than two-thirds of all life forms and the majority of tropical rainforests, coral reefs and other priority systems.
- Australia has 600,000 and 700,000 species, many of which are endemic
    - 84% of our plant species, 83% of mammals, 89% of reptiles, 90% of fish and insects, 93% of amphibians, and 45% of birds are endemic.

•Australia
•The Congo
•Madagascar
•South Africa
•China
•India
•Indonesia
•Malaysia
•Papua New Guinea
•Philippines
•Brazil
•Colombia
•Ecuador
•Mexico
•Peru
•United States
•Venezuela

Biodiversity Hotspots

- Terrestrial regions with exceptional concentrations of endemic species that are undergoing exceptional loss of habitat: **irreplaceable** and **threatened**
- Southwest Australia and Eastern coastal Australia are two of the world's Biodiversity hotspots
- The extinction crisis is vast
  - Human activities (hunting, land management, non-native species), Climate change, Disease
  - Since European settlement, more than 50 species of Australian animals and over 60 species of Australian plants have become extinct.
- The conservation funds are limited
- Targeted protection in nature's most important places
  - Biodiversity underpins all life on Earth.
  - Biodiversity Hotspots also hold some of the highest human population
  - human-biodiversity impacts lies not in human density but rather in human activity
  - For our own sake, we need to protect bio-diversity



Conservation International (conservation.org) defines 35 biodiversity hotspots — extraordinary places that harbor vast numbers of plant and animal species found nowhere else. All are heavily threatened by habitat loss and degradation, making their conservation crucial to protecting nature for the benefit of all life on Earth.

Paradise Parrot

Mount Glorious day frog

Extinct from Queensland

Why sensing?

- Regular field visit and call count
  - Time consuming
  - High cost
  - Influenced by weather or observation bias

- Passive acoustic sensing
  - Time saving for people
  - Low cost: price keeps falling
  - Large spatiotemporal scales
  - Less invasive

Servick, K. (2014). Eavesdropping on ecosystems.

Why acoustic sensing?

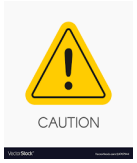
- Many animals actively produce sound for communication
- Some animals are very small or criptic

- Affordable recording devices
- Sound travels
  - Wide range and not directional
- Archive of nature sound at large space over long time
- Tractable analysis

Servick, K. (2014). Eavesdropping on ecosystems.

What Noises Do Koalas Make? https://www.youtube.com/watch?v=fq74aS4irmc

Environmental Audio

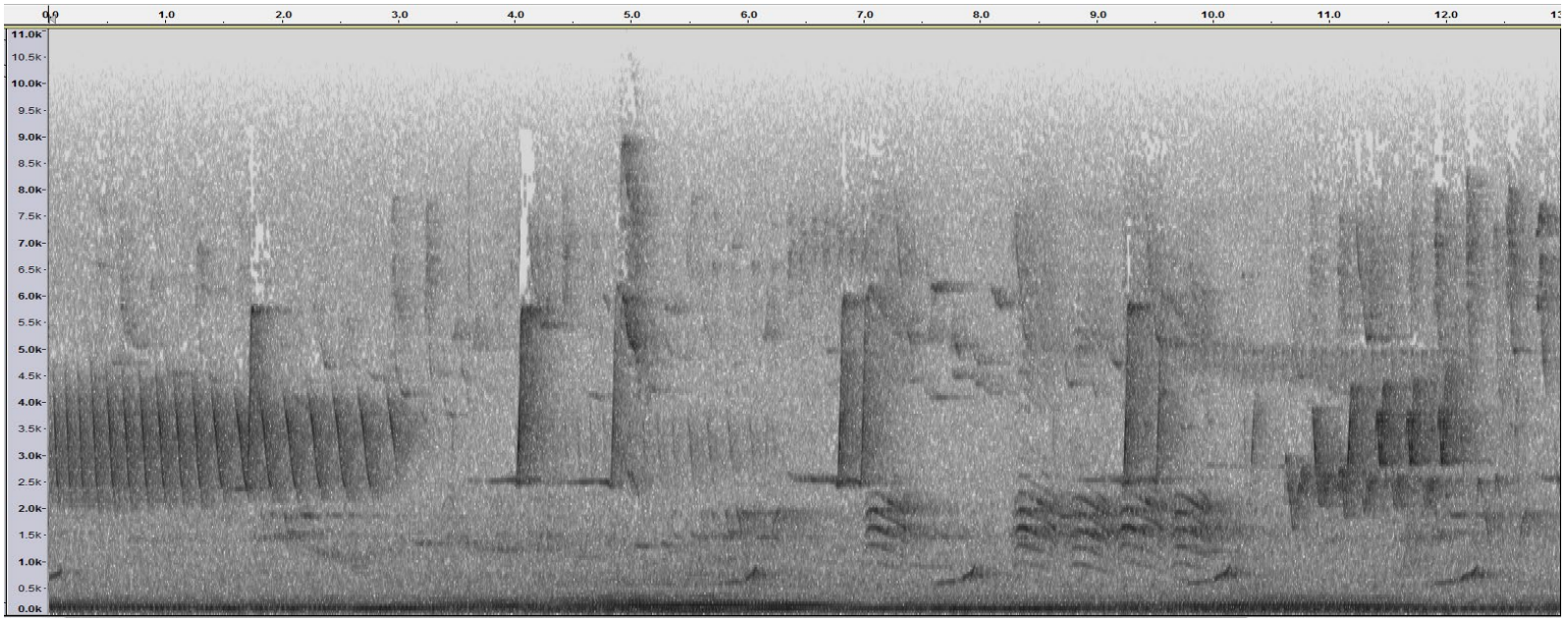


Problem: Recordings can be 1000s of hours long, making them infeasible to listen through manually.

13sec long audio recording

Time(s)

Frequency(Hz)



What is EcoAcoustics?

- Ecoacoustics is an interdisciplinary science that investigates natural and anthropogenic sounds and their relationship with the environment over a wide range of study scales, both spatial and temporal, including populations, communities, and landscapes.

eScience: The Fourth Paradigm
- Data-Intensive Scientific Discovery

- New mantra:
  - Data is the new IntelInside
  - Gather whatever data you can whenever and wherever possible.
- Assumption: Data has/will have value
- New Phenomenon: Data collected at enormous speeds
  - Remote sensors on a satellite:
    - NASA archives over petabytes of earth science data / year
  - Sensors in the sky, on the ground, under the water
    - A2O
- Expectations: Data and software help scientists
  - manipulate and explore massive datasets
  - e-Science: Big data, complex computation, visualization

We aim to understand and address the most pressing threats to the diversity of wildlife through data and software engineering.

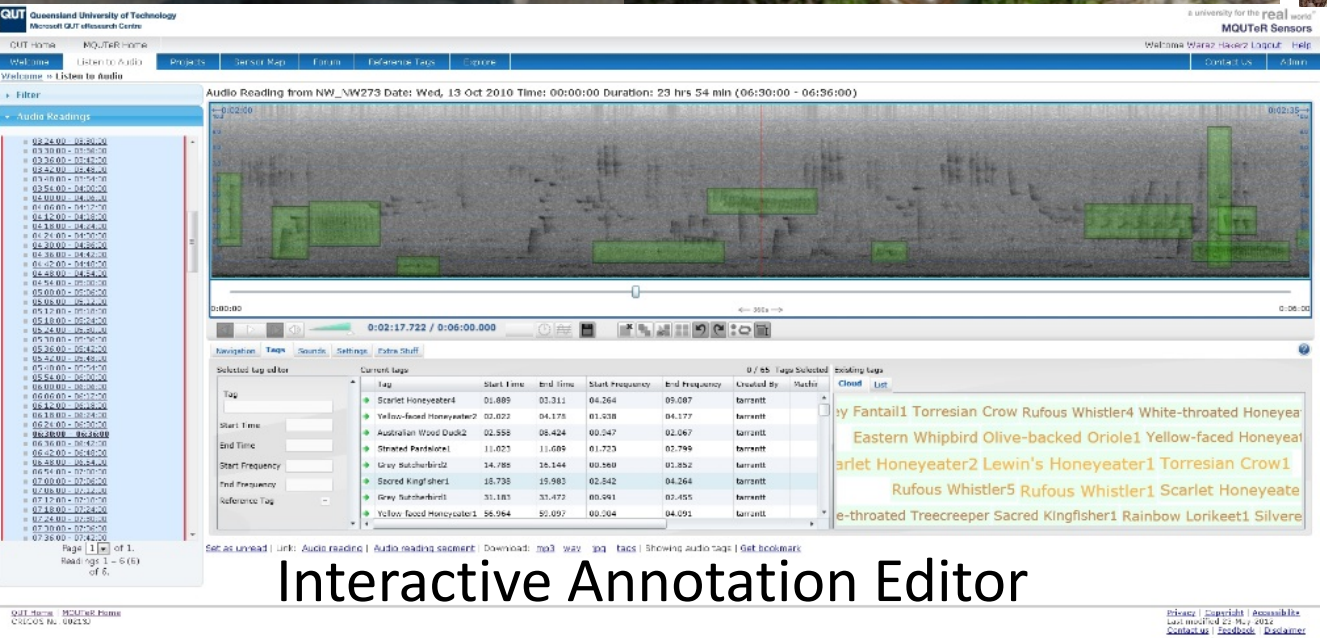
Acoustic Biodiversity Sensing

https://www.ecosounds.org/









Interactive Annotation Editor

We aim to develop new tools to enable new forms of ecological research via acoustic sensing.

QUT ecoAcoustics

- data science, computer science, and CHI

- ecoInformatics: IT meets ecology through sensors and ubicomp
- Acoustic sensing: scale biodiversity observations
  - Birds, frogs, koalas, quolls …
  - Marine life (future)
- Hardware and software tools for sensing the environment
  - Ecological apps, participatory sensing, citizen science
- Novel interfaces: mobile, big screens

Research Problem

ecoAcoustic Analysis for Environmental Monitoring

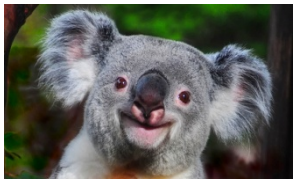
Ecology problems

- Species presence/absence analysis (e.g. Endangered/Intrusive species detection)
- BioDiversity analysis
- Abundance/population analysis
- Habitat health assessment
- Correlation analysis (spatial, temporal, species, environment …)
- Behavioural studies

IT problems

- **Big data collection, storage, management**
- **Big data analysis**
  - Event detection
  - Find event like this
  - Data mining: clustering, classification, association rules, outlier detection
  - Visualization
- **Human data interaction**
  - **Interface, exploration**
- **Techniques to support citizen science**
  - Engaging community: web apps, mobile apps for participatory sensing/analysis
  - Empowering community: sound and annotation library

Data Flow in Acoustic Sensing



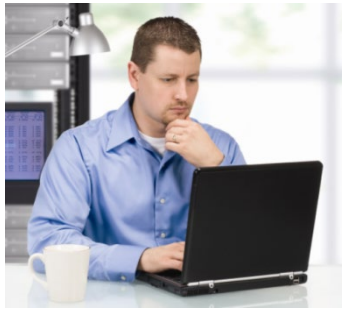
Nature

Sensing – sensors, recorders, human

Nature's iTunes

Ecologists

semi- or auto- analysis
- identifying species
- Measuring biodiversity
- Tracking changes

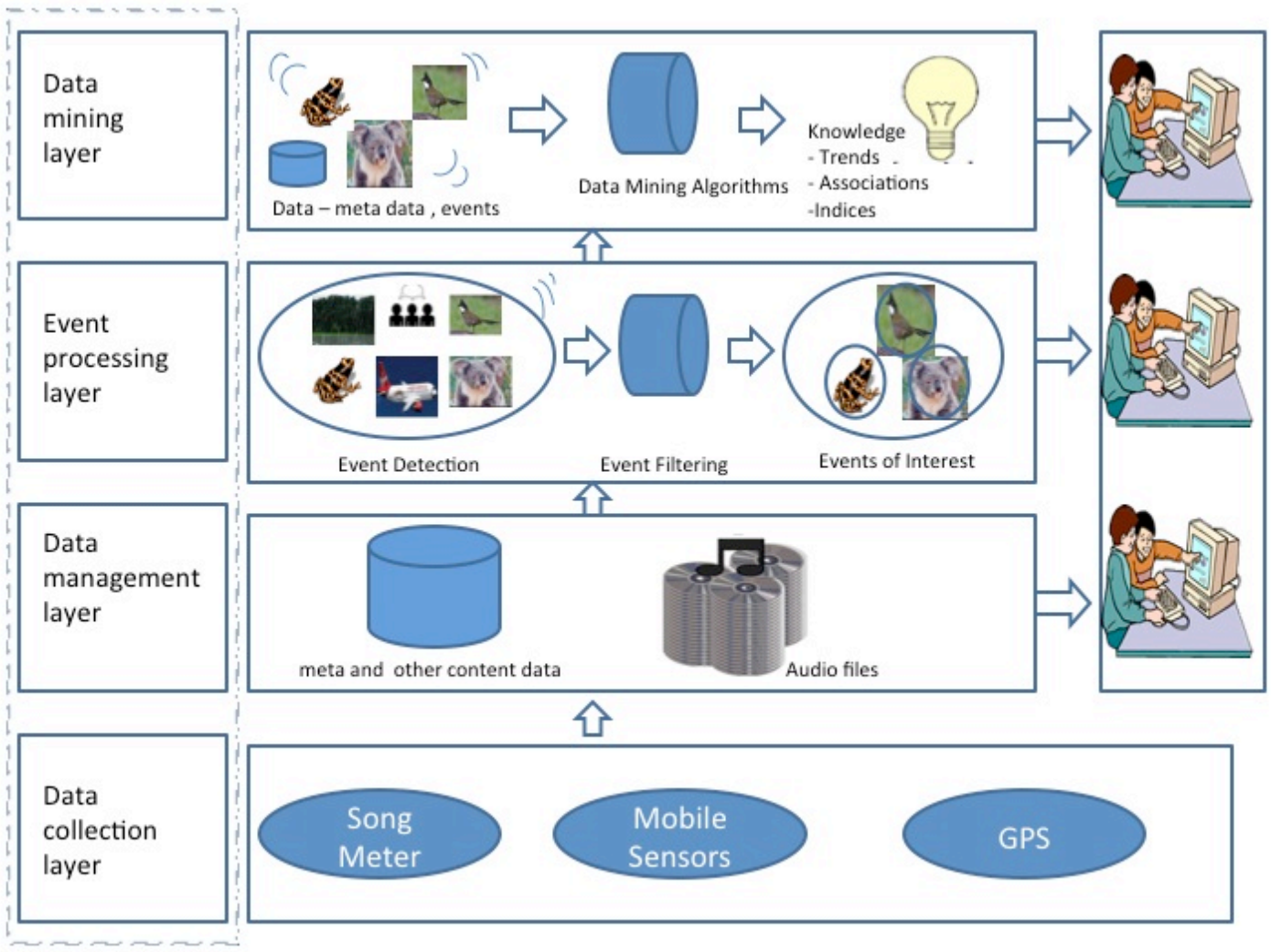


IT Scientists
Develop tools

Ecologists

Manual analysis

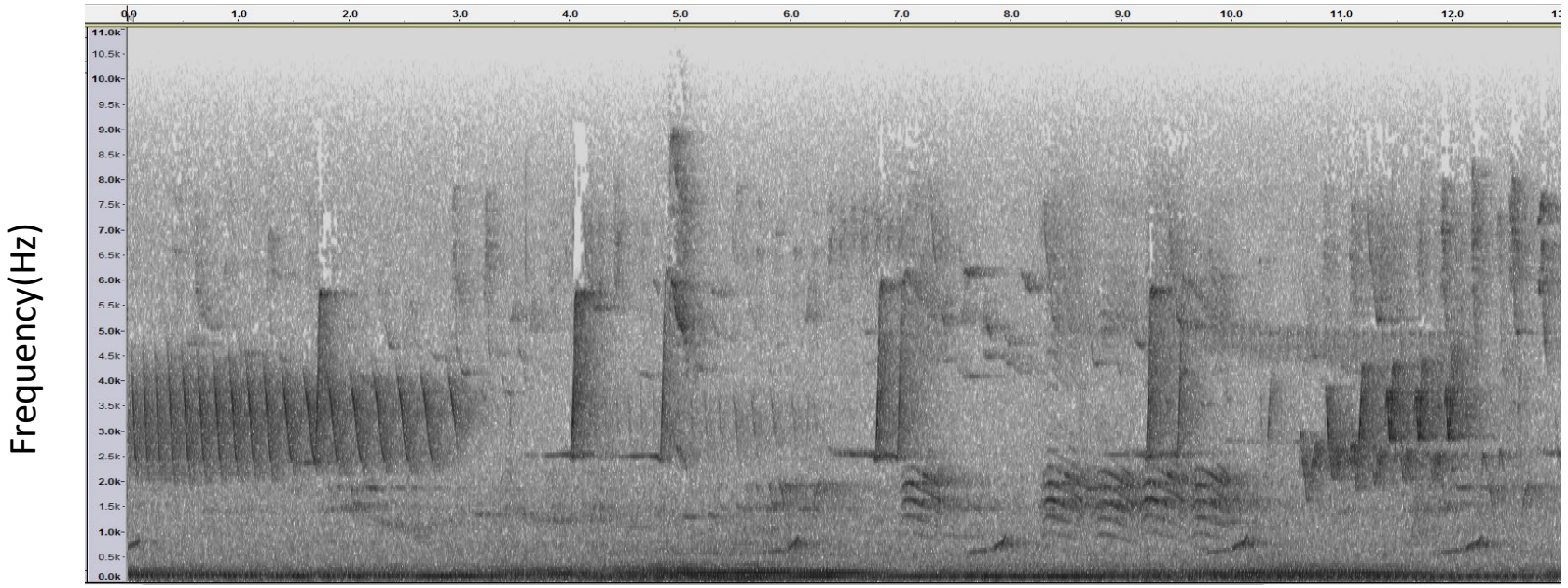
Data Management Framework



Multi-layer data management for environmental monitoring via Acoustics

Environmental Audio Data

- Lab audio data vs. environmental audio data
  - location: quiet room vs. wild area
  - sound sources : unique vs. multiple (e.g. multiple bird species ,rain, wind, thunder, human speech, traffic, machinery, frogs, cicada)
  - An example



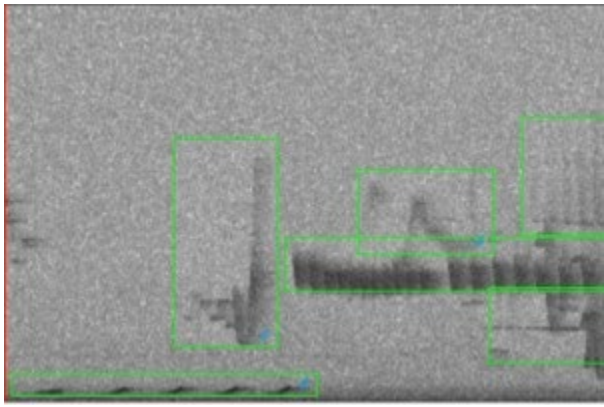
A spectrogram of a 13-second long field recording.

Challenges of Audio Analysis in Environmental monitoring

- Sound are recorded from field directly over large area and long term Big Data: Large volume, high increment velocity
  - Sensors work 24/7
  - 24TB of audio ≈ 50 years to listen to all of it
  - Less than 0.4% of audio analysed by experts
- Noise: weather noise, human-made noise, recording device noise,
- Overlapping sounds from different sources: fauna calls, environment sound, human made sound e.g. car
- Variation: time, temperature, individuals, species, region, distance from microphone, environment, equipment
- Dynamic, unstructured and unpredictable
- Opaque: Audio data is much more difficult to handle than textual or scalar values

EcoAcoustic Analysis Approaches

Analysis = input audio → output annotations



- Manual (human computation)
  - Simple
  - Humans are exceptional classifiers / pattern recognisers
  - Slow, difficult, doesn't scale, requires huge workforce
- Automated Analysis with algorithms
  - Algorithmic, efficient, ideal, scales well, some success
  - Hard and time consuming to develop and test
  - Needs lots of training data
- Semi-automated (human-in-the-loop)
  - Combine complementary aspects of humans and machines
  - Machines and algorithms can
    - process large amounts of data
    - learn off human input
    - reduce work load for a human
  - Humans
    - Amazing visual recognition ability
    - Advanced classification skills
    - Social logic – discuss with others
    - Creative reasoning – can apply knowledge, skill and reasoning without needing to encode it into an algorithm

Manual EcoAcoustic Analysis

- Human computation
  - Listen through
  - Manually scan spectrogram (faster than listening)

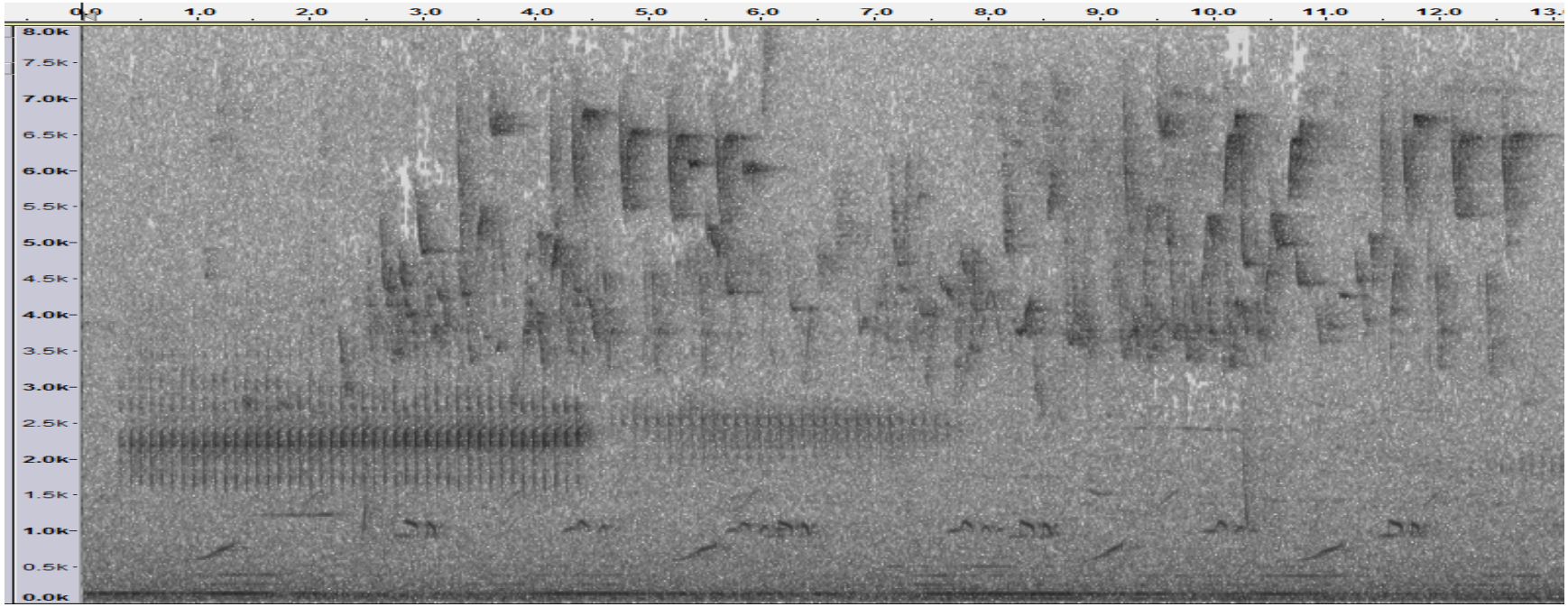
  Simple, context aware, but slow
- Citizen science approach
- Problem: Data quality
- Solution 1: Reputation modelling
- Solution 2: Collaborative Analysis
- Solution 3: Statistics

Yang, Hao-Fan, Zhang, Jinglan, & Roe, Paul (2013) Reputation modelling in Citizen Science for environmental acoustic data analysis. *Social Network Analysis and Mining*, *3*(3), pp. 419-435.

Truskinger, Anthony, Yang, Hao-Fan, Wimmer, Jason, Zhang, Jinglan, Williamson, Ian, & Roe, Paul (2011) Large scale participatory acoustic sensor data analysis: Tools and reputation models to enhance effectiveness. In Laure, E & Henningson, D (Eds.) *Proceedings of the 2011 Seventh IEEE International Conference on eScience.* IEEE Computer Society Conference Publishing Services, United States, pp. 150-157.

Collaborative exploration and sensemaking of Audio

- Human scan through the spectrogram
- Human discuss with each other and cross check the result



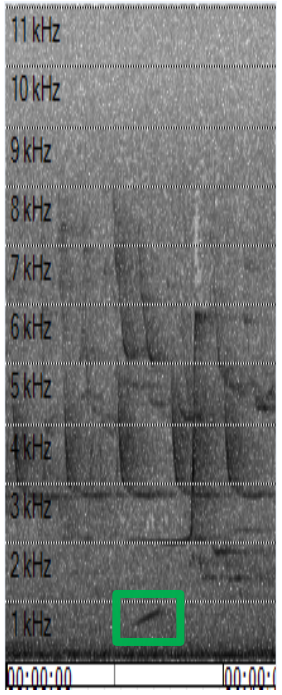
Dema, Tshering, Brereton, Margot, Cappadonna, Jessica Lea, Roe, Paul, Truskinger, Anthony, & Zhang, Jinglan (2017) Collaborative exploration and sensemaking of big environmental sound data. *Computer Supported Cooperative Work*, *26*(4 - 6), pp. 693-731.

EcoAcoustic Analysis Approaches (Cont.)

- Semi-Automated (human-in-the-loop)

Human creative thinking and reasoning + machine fact gathering and computation but requires some human intervention

  - Online species identification library
  - Suggestion tool: Information filtering using machine intelligence
  - Spectrogram scanning
  - Tag validation, correction, and linking
  - **Visualization**: easy to use and understand
  - Online collaboration tool (future work)



Brown Cuckoo-dove

Audio Visualization

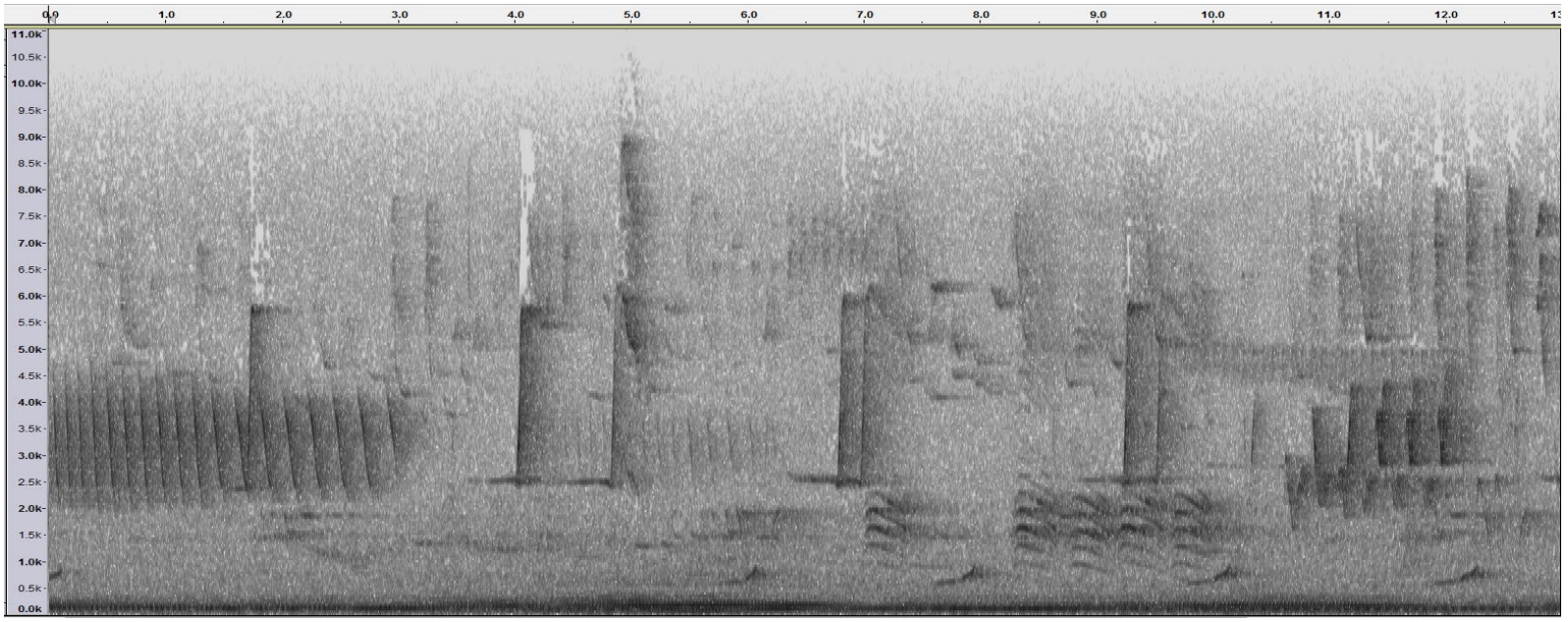
It help people:

- name different kinds of sounds for clearer communication;
- increase “ear-recognition” skill by engaging visual memory in building a library of known sounds;
- evaluate recordings.

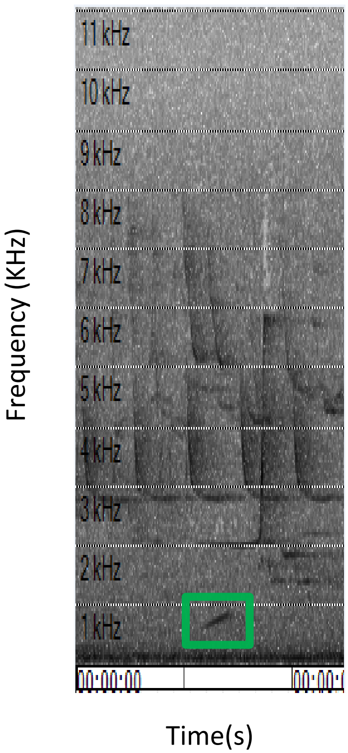
13sec long audio recording

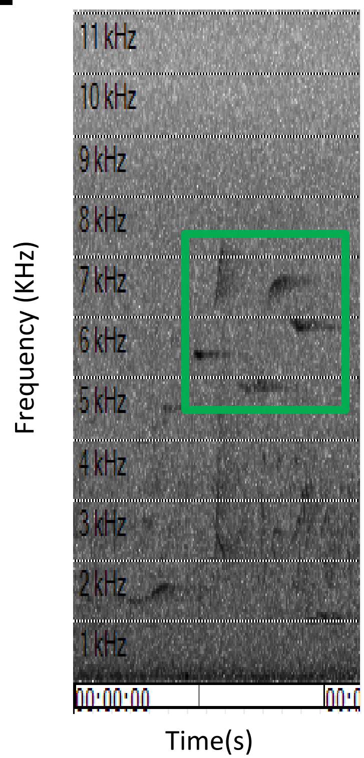
Time(s)

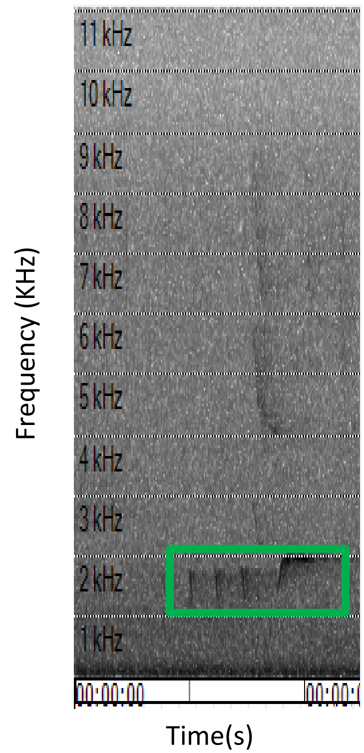
Frequency(Hz)

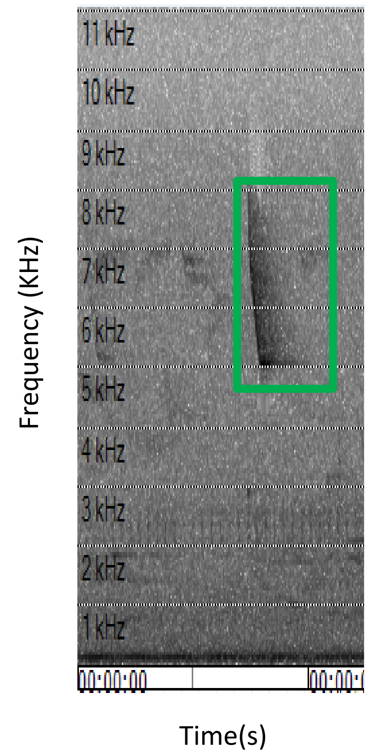


Sample Birdcalls



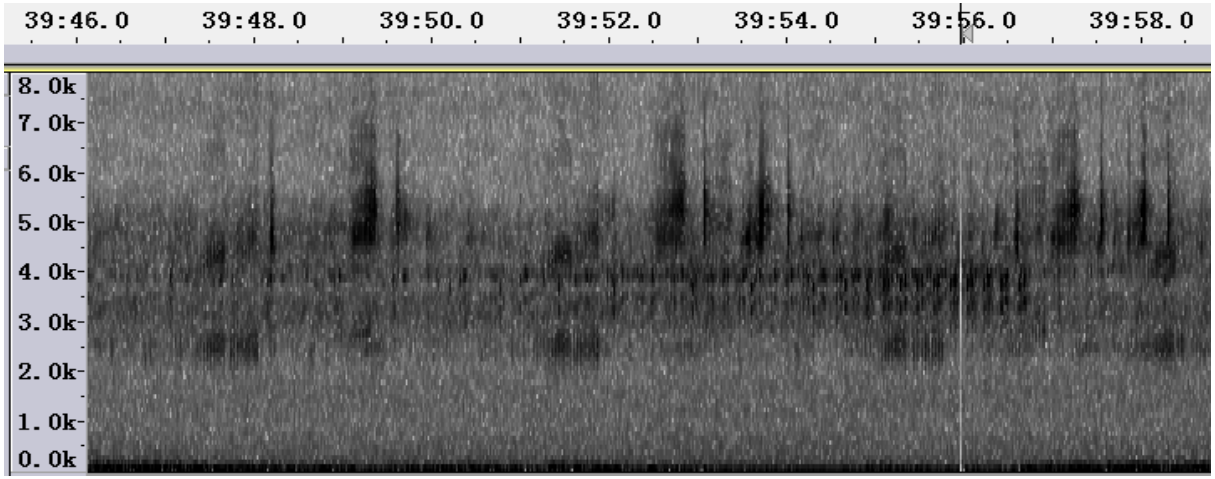


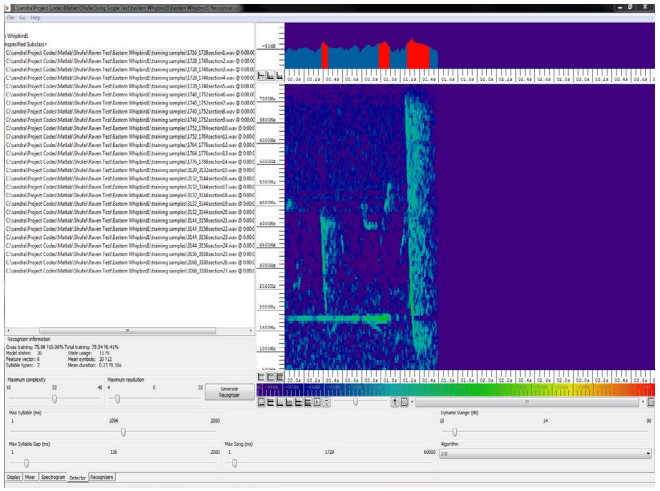




1. Brown Cuckoo-dove
2. Grey Fantail
3. Grey Shrike-thrush
4. Scarlet Honeyeater

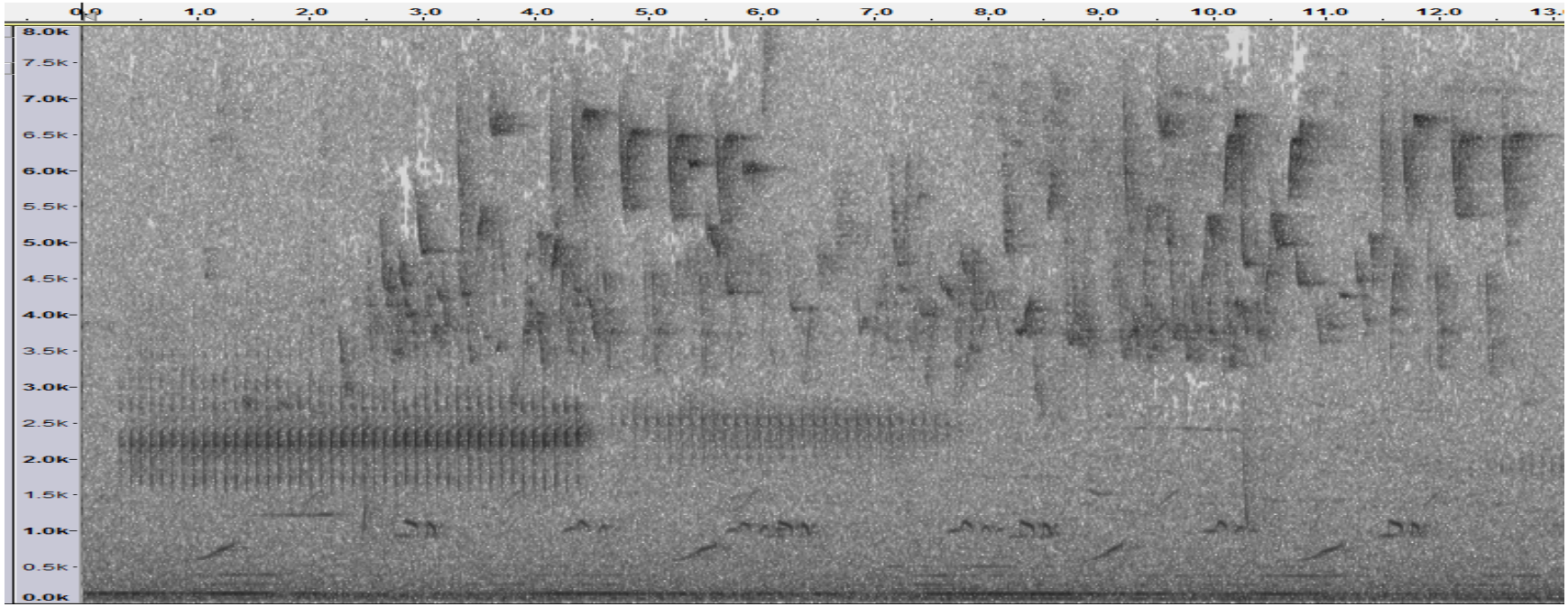
Audio Visualization
- Spectrogram





Human Machine Teaming To Make Sense Of Audio

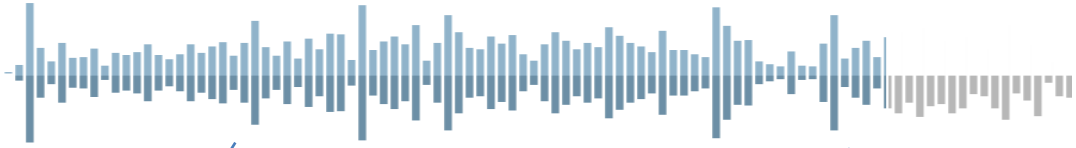
- Human scan through the spectrogram
- Machine scan through the spectrogram
- Human cross check the result



Dema, Tshering, Zhang, Liang, Towsey, Michael, Truskinger, Anthony, Zhang, Jinglan, Brereton, Margot, et al. (2017) An investigation into acoustic analysis methods for endangered species monitoring: A case of monitoring the critically endangered white-bellied heron in Bhutan. In Bubendorfer, K (Ed.) *Proceedings of the 2017 IEEE 13th International Conference on e-Science (e-Science).* Institute of Electrical and Electronics Engineers (IEEE), United States of America, pp. 177-186.

Acoustic index

Weeks, Months or even years of recordings



Acoustic indices

Describe the acoustic environment

Ratio of natural sound and anthrogogenic sound

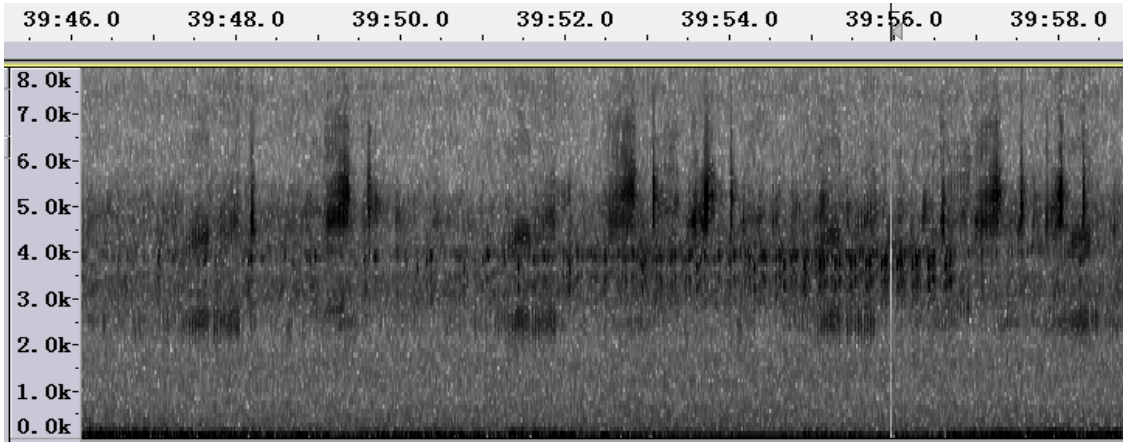
To characterize biodiversity

Acoustic index

- Acoustic indices are designed to reflect acoustic properties of audio recordings of the natural environment.
- **A summary index** is a scalar representing the entire segment of a recording.
- **A spectral index** is a vector (256) representing a summary spectrum for a segment, each element representing the index value for a frequency bin.

What they measure

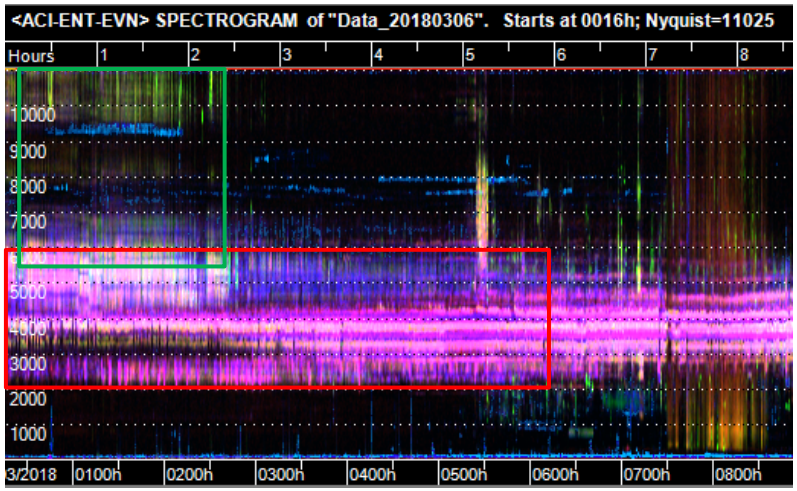
- **PMNsp,** Power Minus Noise of each frequency bin equals to the maximum decibel value minus decibel background noise value.
- **ACIsp**, Acoustic Complexity Index quantifies the relative change in acoustic intensity.
- **EVNsp**, is the the number of acoustic events per minute in each frequency bin.
- **CVRsp**, is the the proportion of cells in each frequency bin of the spectrogram where the spectral power exceeds 3 dB



Towsey, M., Wimmer, J., Williamson, I., & Roe, P. (2014). The use of acoustic indices to determine avian species richness in audio-recordings of the environment. Ecological Informatics, 21, 110-119.

Audio Visualization
- False Color Spectrogram

- Generated by mapping three spectral indices to the primary colors, red, green and blue (RGB) respectively.



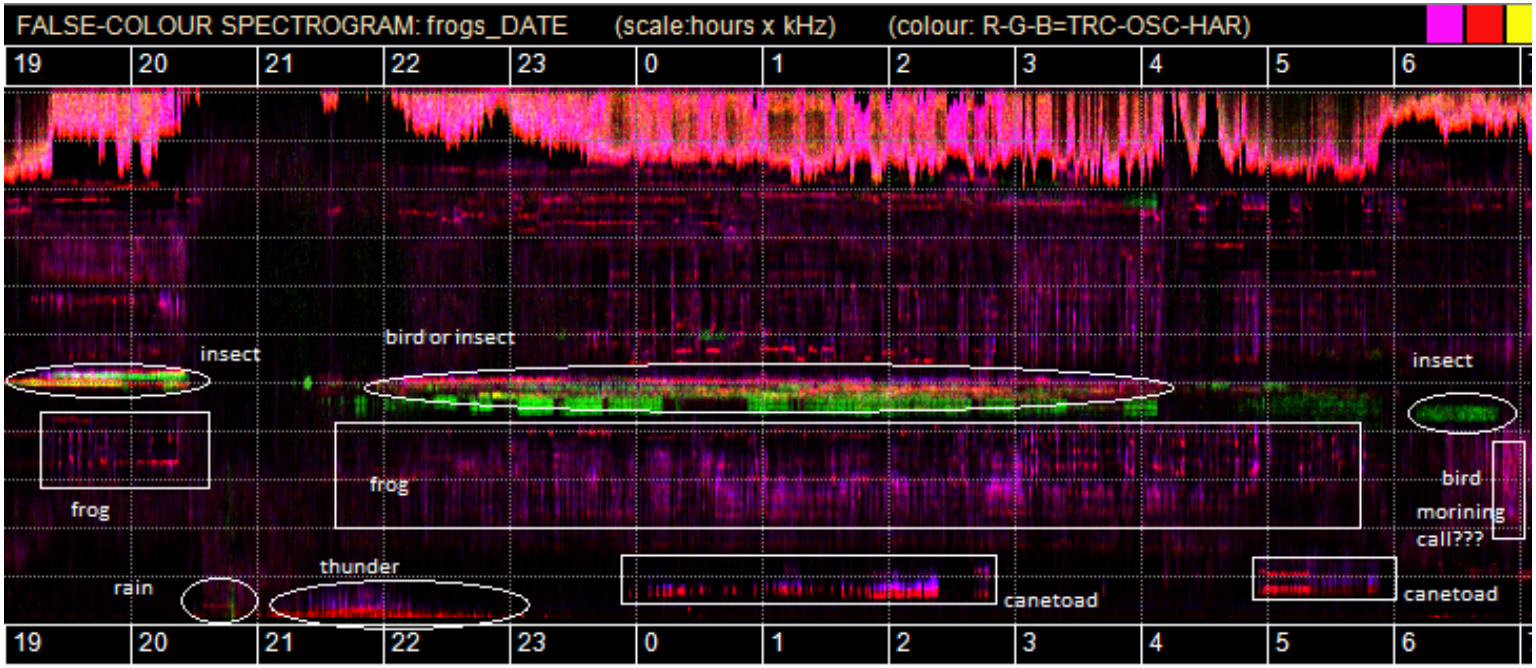
A false color spectrogram of 8.5 hours recording
(ACI-R, ENT-G, EVN-B)

From 00:00 am to 2:30 am, there is green vertical lines at the top (9500-11025 Hz), synchronising with the pink lines (mixed with some green and yellow lines) below. That indicates the intense chorusing of two frog species.

After 02:30 am, L. fallax reduced their call frequency and L. olongburensis dominant the acoustic space.

Gan, Hongxiao, Zhang, Jinglan, Towsey, Michael, Truskinger, Anthony, Stark, Debra, Van Rensburg, Berndt, et al. (2019) Recognition of frog chorusing with acoustic indices and machine learning. In Gupta, A (Ed.) *Proceedings - IEEE 15th International Conference on eScience, eScience 2019.* IEEE, United States of America, pp. 106-115.

False-color Spectrogram

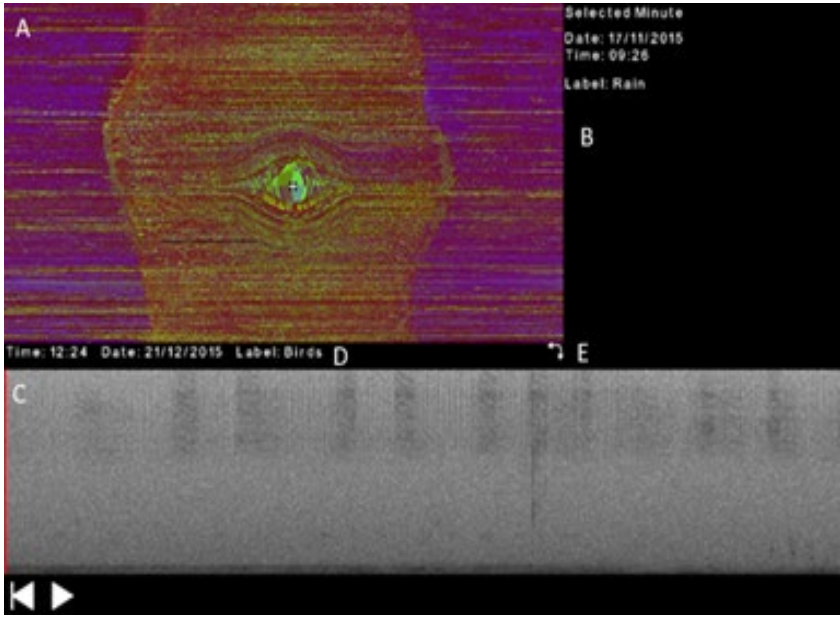


Towsey, M., Zhang, L., Cottman-Fields, M., Wimmer, J., Zhang, J., & Roe, P. (2014). Visualization of long-duration acoustic recordings of the environment. Procedia Computer Science, 29, 703-712.

Audio Visualization
- Multi-scale Multi-view Exploration

-Use different scales to compliment each other to overcome the issue time-window size

- -Utilises a large time window diel plot for the purpose of navigation and to view long-duration trends
- -small time window spectrogram is utilised for viewing smaller time-scale events
- -issues that need to be overcome
  - -not being able to view events that require a medium sized time window
  - -confusing to navigate

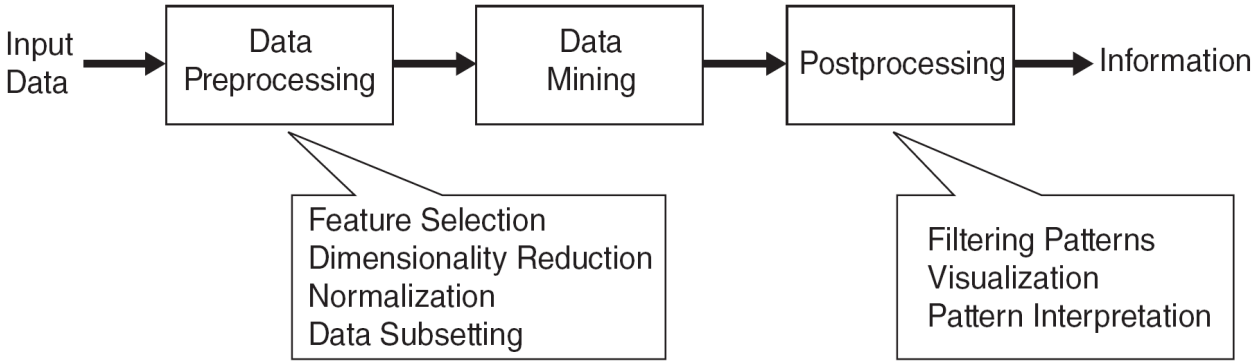


Rowe, Benjamin, Zhang, Jinglan, Towsey, Michael, Roe, Paul, & Brereton, Margot (2018) Ecosound-explorer: a method for large scale interactive visual navigation of environmental acoustic data. In Choi, J H J, McKay, D, Kelly, R, Waycott, J, Lugmayr, A, Morrison, A, et al. (Eds.) *Proceedings of the 30th Australian Conference on Computer-Human Interaction (OzCHI 2018).* Association for Computing Machinery, United States of America, pp. 539-543.

Automated EcoAcoustic Analysis

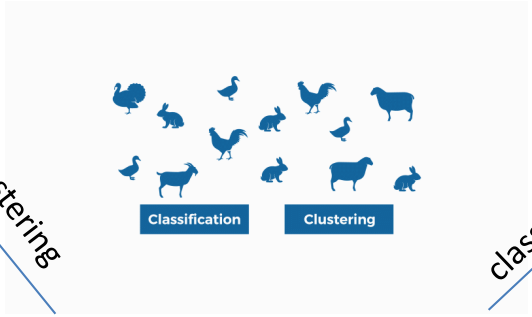
- Machine learning and pattern recognition techniques
  - Feature Extraction: MFCC , ZC, energy, entropies, ……
  - Classification/clustering: DT, RF, SVM, …
  - Recognition: statistic, syntactic, template-based

  efficient, scalable but low accuracy



Zhang, Liang, Towsey, Michael, Xie, Jie, Zhang, Jinglan, & Roe, Paul (2016) Using multi-label classification for acoustic pattern detection and assisting bird species surveys. *Applied Acoustics*, *110*, pp. 91-98.

Automated Audio analysis



clustering

classification

Data: collection of objects defined by attributes

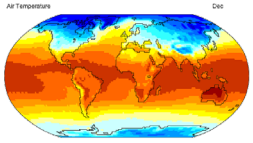
Similarity Search

audio_id	start_date	start_time	project_name	site_name	event_start_date	event_start_time	event_start	event_end	event_duration	low_frequency	high_frequency	common_name_tags
352967	25/11/2015	13:31:39	1029:Cooloola	Gympie NP	25/11/2015	18:34:34	18175.7601	18187.602	11.8422	516.7969	1205.8594	236:Wonga Pigeon
352967	25/11/2015	13:31:39	1029:Cooloola	Gympie NP	25/11/2015	18:31:42	18003.0047	18010.435	7.4304	473.7305	1248.9258	236:Wonga Pigeon
461823	20/01/2017	2:45:00	1029:Cooloola	Gympie NP	20/01/2017	5:01:00	8160.209	8168.87	8.661	86.1328	1464.2578	4:Australian Brush-turkey
352956	23/11/2015	6:45:52	1029:Cooloola	Gympie NP	23/11/2015	6:52:14	382.6982	395.6782	12.98	86.1328	516.7969	137:Painted Button-quail
352952	28/11/2015	6:45:54	1029:Cooloola	Gympie NP	28/11/2015	10:03:30	11856.1161	11866.751	10.6347	86.1328	516.7969	137:Painted Button-quail
352952	28/11/2015	6:45:54	1029:Cooloola	Gympie NP	28/11/2015	10:16:38	12644.6378	12647.68	3.0418	679.5898	4694.2383	111:Leaden Flycatcher
333239	27/09/2015	0:00:03	1029:Cooloola	Gympie NP	27/09/2015	5:42:12	20529.7524	20534.722	4.9691	5531.4453	5340.2344	229:White-throated Honeyeater
333239	27/09/2015	0:00:03	1029:Cooloola	Gympie NP	27/09/2015	5:39:51	20388.7617	20396.749	7.9877	5727.832	8010.3516	167:Rufous Fantail
333239	27/09/2015	0:00:03	1029:Cooloola	Gympie NP	27/09/2015	5:38:28	20305.7974	20309.536	3.7384	5124.9023	8785.5469	167:Rufous Fantail
333239	27/09/2015	0:00:03	1029:Cooloola	Gympie NP	27/09/2015	5:36:56	20213.5218	20215.612	2.0898	4134.375	8785.5469	127:Mistletoebird
333239	27/09/2015	0:00:03	1029:Cooloola	Gympie NP	27/09/2015	5:36:54	20211.4088	20213.777	2.3684	602.9297	3057.7148	99:Grey Shrike-thrush
331813	8/09/2015	6:45:52	1029:Cooloola	Gympie NP	8/09/2015	12:43:03	21431.5635	21433.375	1.8112	3832.9102	9130.0781	97:Grey Fantail
354764	22/10/2015	6:45:53	1029:Cooloola	Gympie NP	22/10/2015	13:16:00	23407.1858	23436.791	29.6054	7235.1563	10637.4023	
354776	19/10/2015	6:45:51	1029:Cooloola	Gympie NP	19/10/2015	12:39:59	21248.0573	21255.279	7.2214	861.3281	6330.7617	111:Leaden Flycatcher
353001	3/12/2015	0:00:04	1029:Cooloola	Gympie NP	3/12/2015	5:51:00	21056	21060.18	4.1796	4014.6484	8182.6172	127:Mistletoebird
353001	3/12/2015	0:00:04	1029:Cooloola	Gympie NP	3/12/2015	5:50:34	21030.6672	21033.291	2.6239	689.0625	1205.8594	236:Wonga Pigeon
352952	28/11/2015	6:45:54	1029:Cooloola	Gympie NP	28/11/2015	10:18:48	12774.6688	12776.689	2.0202	689.0625	3402.2461	99:Grey Shrike-thrush

Feature selection

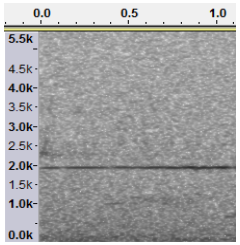
Association Rule

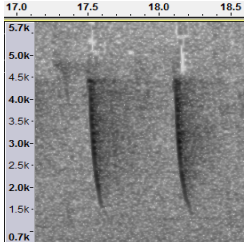
Anomaly Detection

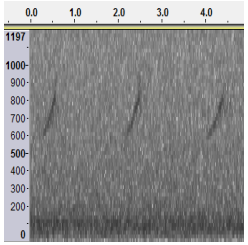


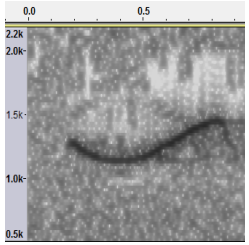
Surface Temperature of Earth

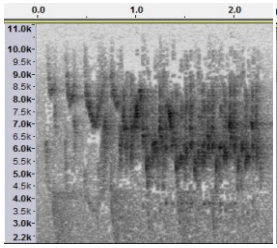
Acoustic Component Detection

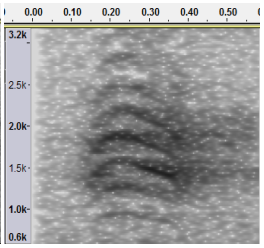

Whistle

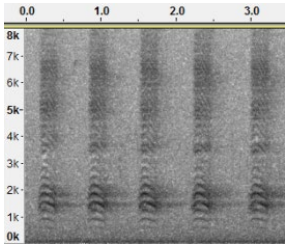

Click


Slur

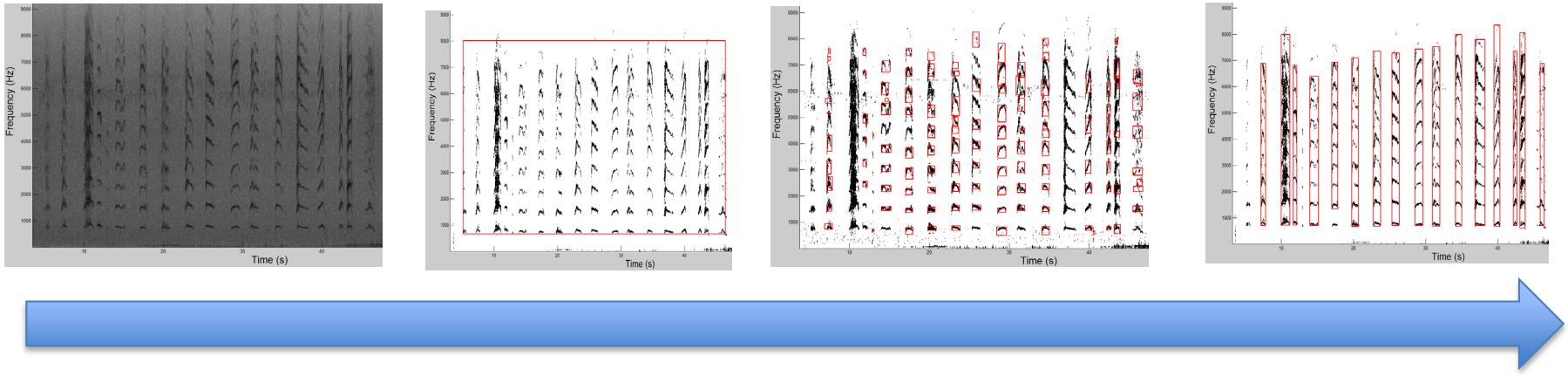

Warble


Block


Harmonic

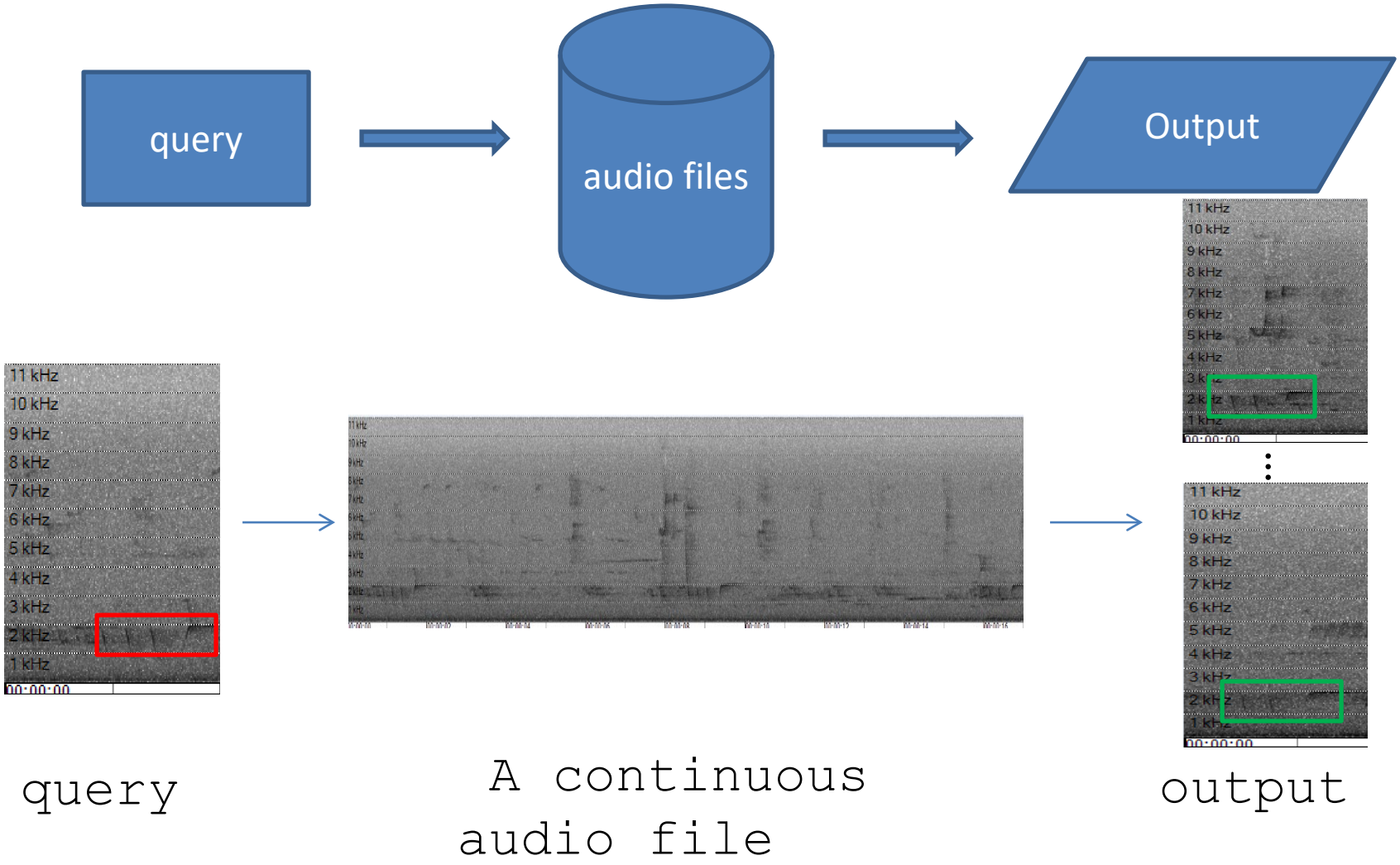

Oscillation

Acoustic Component Definition



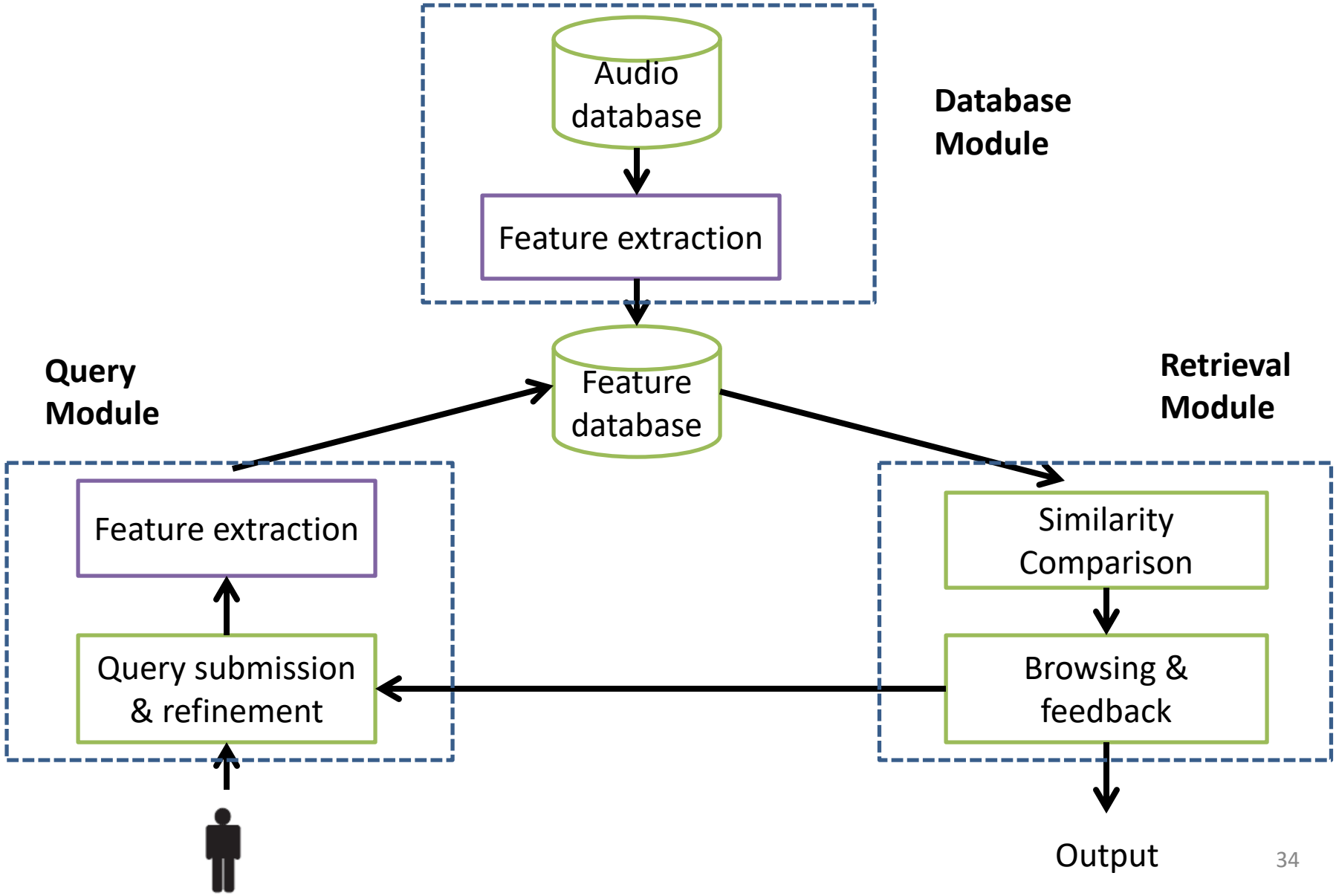
Duan, S., Towsey, M., Zhang, J., Truskinger, A., Wimmer, J., & Roe, P. (2011, 6-9 Dec. 2011). *Acoustic component detection for automatic species recognition in environmental monitoring.* Paper presented at the Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2011 Seventh International Conference on.

Bird Call Similarity Search



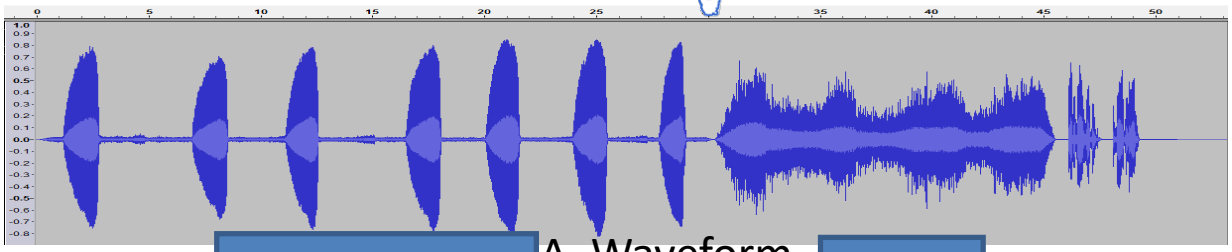
Dong, Xueyan, Towsey, Michael, Truskinger, Anthony, Cottman-Fields, Mark, Zhang, Jinglan, & Roe, Paul (2015) Similarity-based birdcall retrieval from environmental audio. *Ecological Informatics*, *29*(Part 1), pp. 66-76.

Content-based audio retrieval



Frog call analysis for environmental monitoring

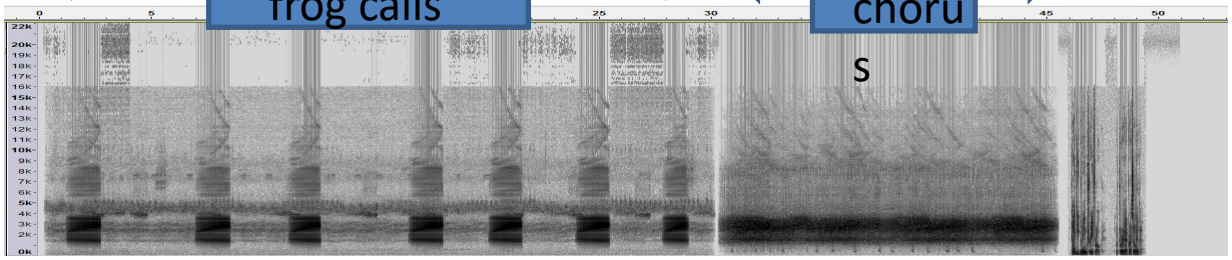
Litoria gracilenta(Dainty green tree frog)



A. Waveform

Individual frog calls

Frog chorus



B. Spectrogram

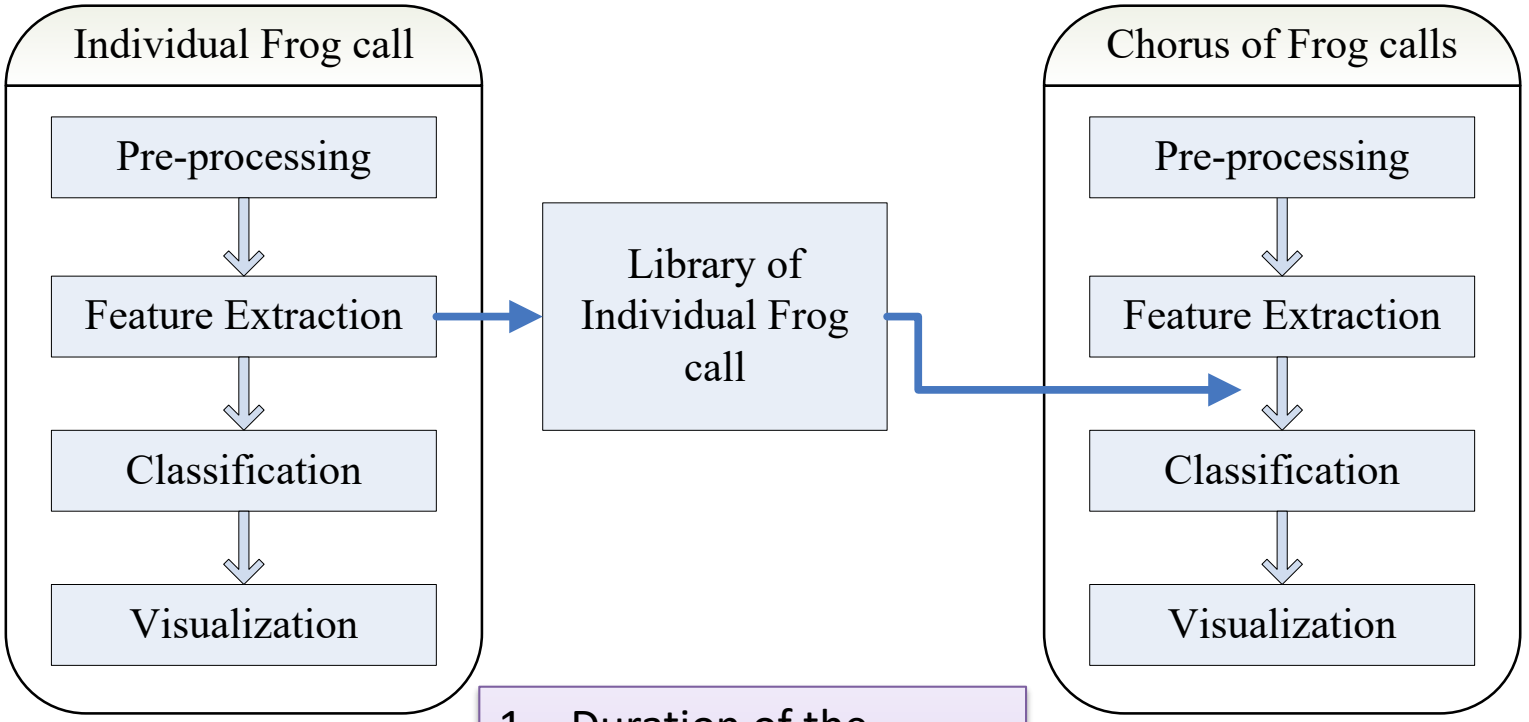
Xie, Jie, Towsey, Michael, Zhang, Jinglan, & Roe, Paul (2016) Adaptive frequency scaled wavelet packet decomposition for frog call classification. *Ecological Informatics*, *32*, pp. 134-144.

Xie, Jie, Towsey, Michael, Zhang, Jinglan, & Roe, Paul (2016) Acoustic classification of Australian frogs based on enhanced features and machine learning algorithms. *Applied Acoustics*, *113*, pp. 193-201.

Xie, Jie, Towsey, Michael, Truskinger, Anthony, Eichinski, Phil, Zhang, Jinglan, & Roe, Paul (2015) Acoustic classification of Australian anurans using syllable features. In Tan, H P & Palaniswami, M S (Eds.) *Proceedings of the 2015 IEEE Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP 2015).* IEEE, United States of America, pp. 1-6.

Xie, Jie, Towsey, Michael, Zhang, Jinglan, Dong, Xueyan, & Roe, Paul (2015) Application of image processing techniques for frog call classification. In Labeau, F & Thiran, J P (Eds.) *Proceedings of the 2015 International Conference on Image Processing (ICIP).* IEEE, United States of America, pp. 4190-4194.

Block diagram of the frog call recognition system

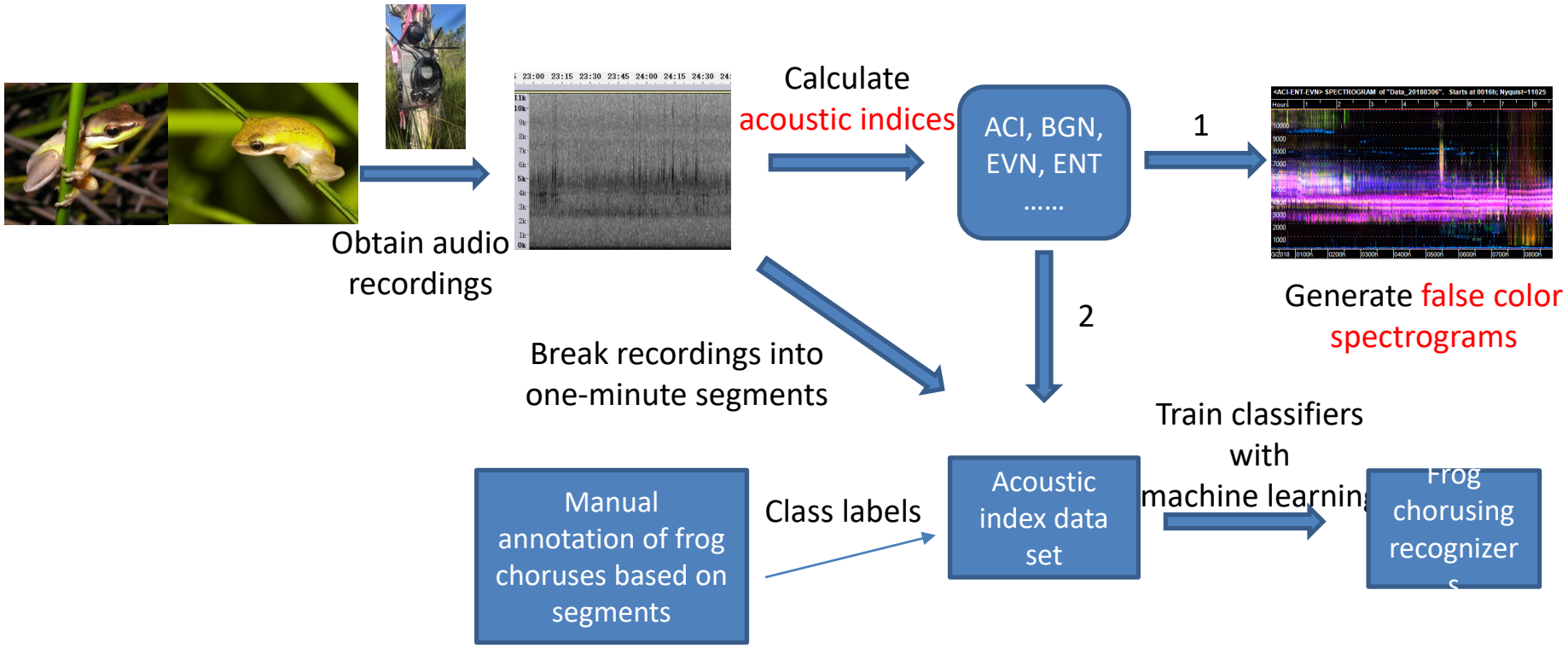


1. STFT
2. Noise reduce

1. Duration of the syllable
2. Oscillation structure
3. Harmonic structure
4. Entropy
5. ……

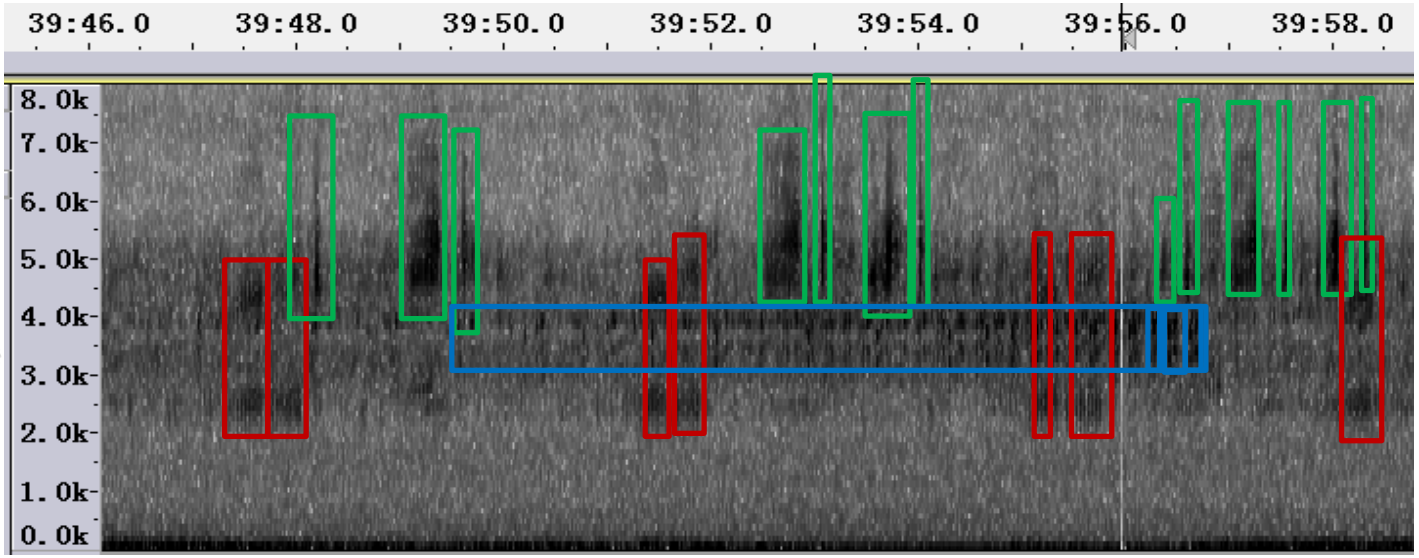
1. KNN
2. HMM
3. SVM
4. ……

Frog Chorusing Recognizers with Indices



Gan, Hongxiao, Zhang, Jinglan, Towsey, Michael, Truskinger, Anthony, Stark, Debra, Van Rensburg, Berndt, et al. (2019) Recognition of frog chorusing with acoustic indices and machine learning. In Gupta, A (Ed.) *Proceedings - IEEE 15th International Conference on eScience, eScience 2019.* IEEE, United States of America, pp. 106-115.

Frog calls of target species



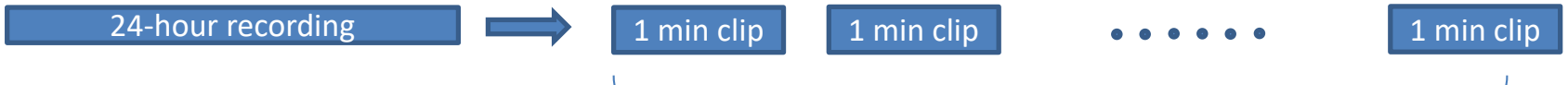
- *Olong* calls with two dominant frequency bands in red boxes.
- *Fallax* calls in green boxes.
- Both of their calls have oscillations, and *L. fallax* calls have high a oscillation rate.

Wallum Sedgefrog, *Litoria olongburensis*

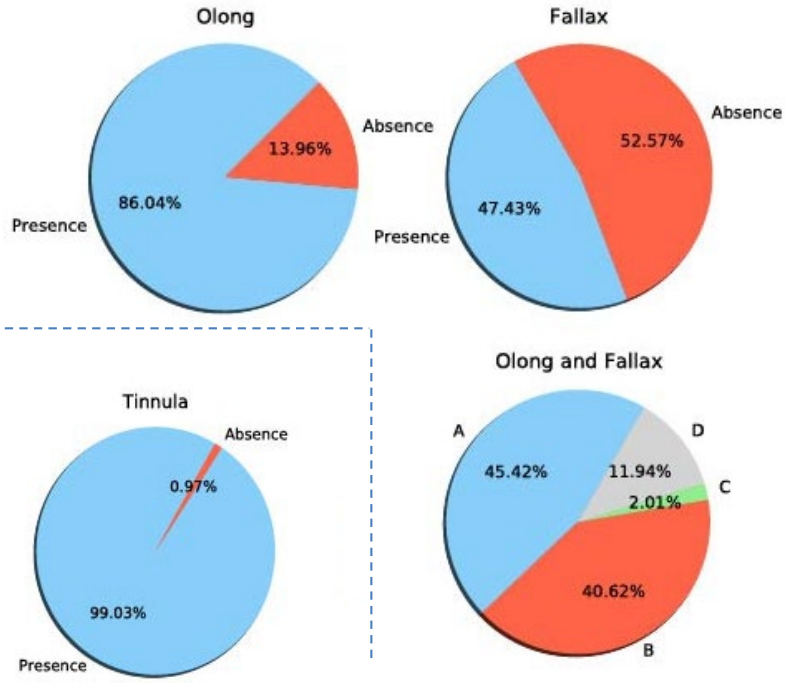
Eastern Sedgefrog, *Litoria fallax*

Wallum Froglet, *Crinia tinnula*

Index-based frog classification



1440 clips



A	Both present
B	Only Olong
C	Only Fallax
D	Neither present

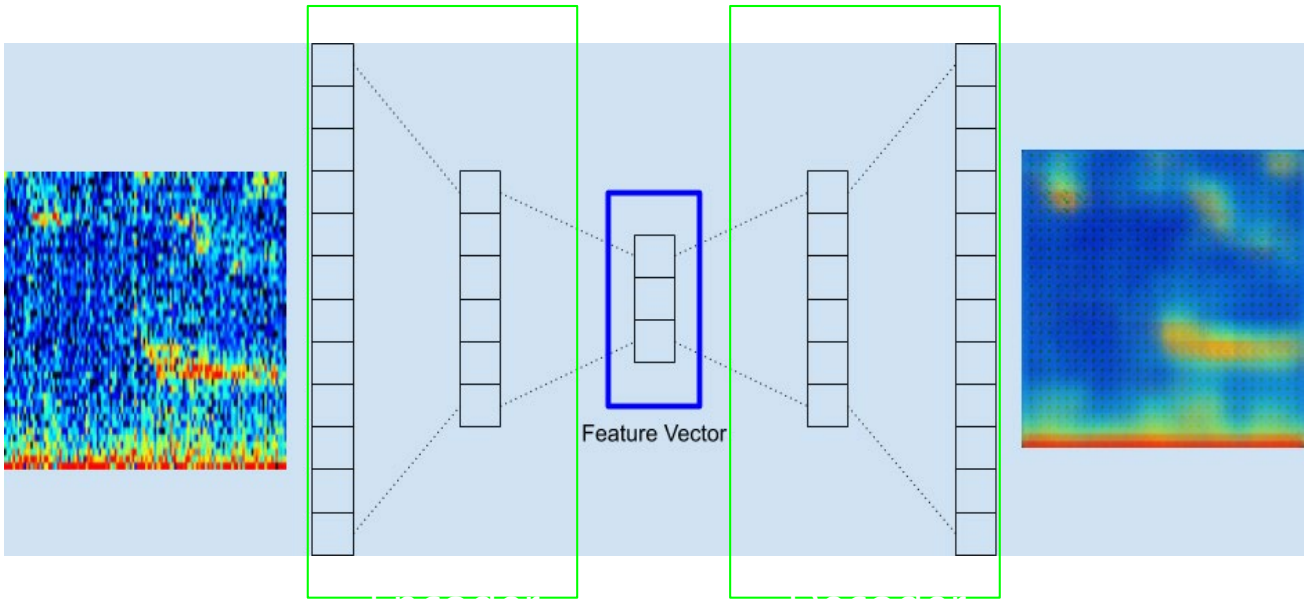
Audio clip class label

Summary indices	Spectral indices
BGN	BGN_{sp} 0
SNR	BGN_{sp} 1
ACT	… …
… …	BGN_{sp} 255

256 values for each spectral index

14 features 3072 features

Deep learning for Audio Analysis



J Xie, R Zeng, C Xu, J Zhang, P Roe, Multi-label classification of frog species via deep learning, 2017 IEEE 13th International Conference on e-Science (e-Science), 187-193

Conclusion

- Manual analysis

  - accurate and comprehensive
  - high cost: highly trained experts, time consuming
  - limited spatiotemporal scale

- Automated analysis
  - powerful
  - not very accurate yet: difficult to deal with noise and variations
  - Specific recognisers perform better than generic ones
- Semi-automated annotation
  - needs some human intervention
  - allows quality data analysis now
  - allows harvest of citizen intelligence

Conclusion (Cont.)

- We can leverage on both human and machine intelligence
  - Providing searching and suggestion tools based on machine intelligence
  - Providing visualization tools for easy interaction
  - Providing data linking and sharing tools for collaboration
  - Exporting clean and rigorous data for human to further interpret
- Still plenty of research questions to be answered

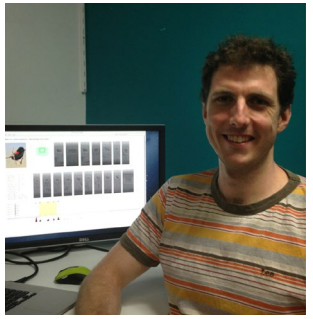
Future work

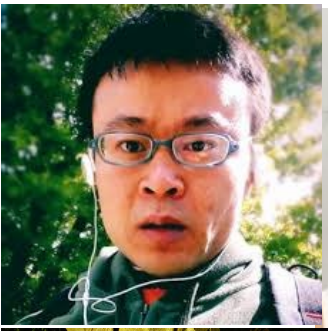
- Data mining
- Human data interaction
- Engaging citizen scientists
- Web Apps
- Mobile apps

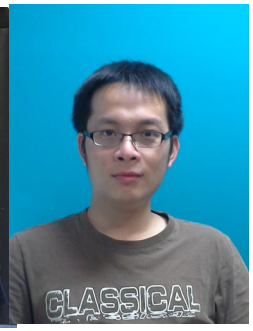
Thanks to my colleagues and students











[See us online: https://www.ecosounds.org/](https://www.ecosounds.org/)

Thanks for listening

Questions?

Dr. Jinglan Zhang
Email: jinglan.zhang@qut.edu.au
Ph:61 7 3138 9353
Web: http://staff.qut.edu.au/staff/zhang2/
Queensland University of Technology