# Acoustic features for multi-level classification of Australian frogs: Family, Genus and Species

Author

*Address*

**Abstract**

Over the past few decades, frogs have been experiencing dramatical population decline around the world. There are many reasons for this decline, including habitat loss, invasive species, pollution and climate change. To protect and increase the frog populations, it is important to study frogs. In this study, acoustic features are investigated for multi-level classification of Australian frogs: family, genus and species. Three families, ten genera and eighty five species collected from Queensland, Australia, are analysed in this experiment. For each frog species, six instances are first selected from which eleven acoustic features are extracted. Then a decision tree (DT) classifier is used to visually and explicitly determine which acoustic features are relatively high important for classifying family, which for genus and which for species. Finally, a weighted support vector machines (SVMs) classifier is used for the family, genus and species classification with three most important acoustic features. Our experimental results indicate that different level classification needs different acoustic feature sets. With selected acoustic features, average classification accuracy can be up to 85.68%, 75.58% and 64.07% for family, genus and species respectively.

*Keywords:* Frog call classification, Acoustic feature, Feature selection, Decision tree, Support vector machines

## 1. Introduction

In the past decades, frog populations have rapidly declined around the world due to habitat loss, invasive species, pollution and climate change. Therefore, it it becoming increasingly important to monitor and assess the frog [1]. Acoustic survey is often the only possible method for studying frogs, because it's often a lot easier to hear frogs than to see them. With acoustic survey methods, large volumes of acoustic data are collected. Subsequently developing semi-automatic or automatic method for investigating frogs with acoustic data is in high demand.

With frog calls, several papers have been proposed for frog species recognition. Lee et al. introduced a recognition method based on spectrogram analysis to detect each syllable and calculate the Mel-frequency cepstrum coefficients (MFCCs). All averaged MFCCs of each frame were defined as features. The linear discriminant analysis was used for classifying 30 kinds of frog calls and 19 kinds of cricket calls [2]. Chen et al. proposed a method based on syllable duration and multi-stage average spectrum for frog call recognition. Syllable duration is used for the pre-classification, then a multi-stage average spectrum is proposed for the frog species recognition with the accumulation distance evaluation [3]. Bedoya et al. used Mel-frequency cepstral coefficients (MFCCs) as the acoustic feature for the recognition of anuran species with fuzzy clustering [4].

For the frog call classification, Chen et al. combined spectral centroid, signal bandwidth and threshold crossing rate to do frog classification. The k-NN and support vector machines classifiers were then introduced for frog call classification [5]. Han et al. introduced a k nearest neighbour (k-NN) classifier to classify frog calls, spectral centroid and two entropy features were extracted from syllables as the input to the classifier [6]. Xie et al. extracted syllable features (syllable duration, dominant frequency, oscillation rate, frequency modulation and energy modulation) based on the advertisement call. Then a k-NN classifier was used for frog call classification [7].

All the prior work achieves a high accuracy rate in recognition and classification of frog species. However, a frog's classification can be determined to three levels, including family level, genus level and species level. Few work have been done to analysis the frogs in genus level. For the genus level classification, Glaw et al. proposed a method for the genus level classification of family Mantellidae based on published phylogenetic information and on a new analysis of molecular data [8]. Gingras et al. introduced a three-parameter model for classifying anurans into four genera based on advertisement calls [9]. To our knowledge, no study has yet been published that utilize acoustic features for family level classification of frogs.

Since the vocalizations of frogs are mostly genetically determined and do not show evidence of vocal learning as birds [10], it is possible to utilize frog calls for the classification of frogs in multi-level: family, genus and species. Furthermore, advertisement calls of closely related phylogenetic species are more similar than those of species that are distant. Therefore, extracting acoustic features from advertisement calls can be possible for investigating the multi-level classification of frogs.

The goal of this study is to provide the qualitative and comparative analysis of acoustic features for the multi-level classification of frogs. For this, three families, eleven genera and eighty five species in Queensland, Australia are studied. For each species, six instances are selected, from which ten acoustic features are extracted: spectral centroid, spectral flatness, spectral roll-off, zero crossing rate, Shannon entropy, spread, skewness, kurtosis, root mean square value and averaged energy. To investigate the relationship between acoustic features and the correlation between frog families, genera and species, a DT classifier is used to intelligently select three most important features for final classification. Finally, a weighted SVMs classifier is conducted for the multi-level classification of frogs with selected features.

The rest of this paper is organized as follows: Section 2 explained the used materials and the related techniques including signal pre-processing, feature extraction, feature selection and classification. Section 3 reports the experimental results. In Section 4, the discussion of result is given. The conclusion and future work are offered in Section 5.

## 2. Materials and methods

In this study, the frog call classification system including five sections is shown in Fig.1.

### 2.1. Materials

In this study, frog calls are obtained from two sources: David Stewart's CD (*http : //amphibiaweb.org/maps/index.html*) and one public website (*http : //www.naturesound.com.au/cd$_f$rogsSE.htm*). All the recordings are in stereo, re-sampled at 44.1 KHz and saved in WAV format.

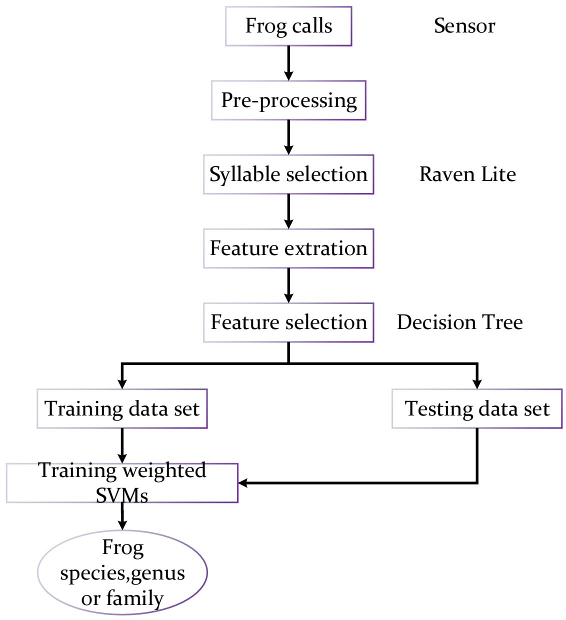



Figure 1: Flowchart of frog call classification system

In total, 503 syllables of good quality frog calls from 3 families, 10 genera and 85 species, are selected for experiment. Each species includes six instance except five species which are Cophixalus bombiens (5), Cophixalus concinnus (4), Cophixalus crepitans (4), Cophixalus exiguus (5) and Cophixalus hosmeri (5).

### 2.2. Pre-processing

In this study, the audio data is segmented with a software tool named "Raven Lite". Before feature extraction, a first-order high-pass filter with finite impulse response (FIR) is applied to the original frog calls for reducing the low-frequency components as follows:

$$y(n) = s(n) - \alpha s(n) \quad (1)$$

where $s(n)$ is the original frog call, $y(n)$ is the output after pre-emphasis filtering. Here $\alpha$ means the cutoff frequency of the high-pass filter and was set at 0.97.

After pre-filtering, a hamming window is used to minimize the maximum sidelobe in the frequency domain which can be defined as

$$w(n) = 0.54 - 0.46cos(\frac{2n\pi}{L-1}), 0 \leq n \leq L-1 \quad (2)$$

where L is the length of the frame and is set at 128 in this study. After windowing, the signal can be represented as

$$x(n) = w(n)s(n) \tag{3}$$

### 2.3. Feature extraction

For the feature extraction, ten features are extracted from each frog syllable in this work. They are spectral centroid, spectral flatness, spectral roll-off, zero-crossing rate, Shannon entropy, spread, skewness, kurtosis, root mean square value and averaged energy.

#### 2.3.1. Spectral centroid

spectral centroid ($S_c$) is the centre point of spectrum distribution. In terms of human audio perception, it is often associated with the brightness of the sound. With the magnitudes as the weight, it is calculated as the weighted mean of the frequencies.

$$S_c = \frac{\sum_{k=0}^{N-1} f_k X_k}{\sum_{k=0}^{N-1} X_k} \tag{4}$$

where $x_k$ is the DFT of the signal syllable of the k-th sample, N is the half size of DFT.

#### 2.3.2. Spectral flatness

Spectral flatness ($S_f$) provides a way to quantify the tonality of a sound. A high spectral flatness indicates a similar amount of power of the spectrum in all spectral bands. Spectral flatness is measures by the ratio between the geometric mean and the arithmetic mean of the power spectrum and defined as:

$$S_f = \frac{\sqrt{\frac{1}{N}\sum_{k=0}^{N-1} InX(k)}}{\frac{1}{N}\sum_{k=0}^{N-1} X(k)} \tag{5}$$

#### 2.3.3. Spectral roll-off

Spectral roll-off ($S_r$) is a measure of spectral shape. It is defined as the frequency $H$ below which $\theta$ of the magnitude distribution is concentrated.

$$\sum_{k=1}^{H} X(k) = \theta \sum_{n=1}^{N-1} X(k) \tag{6}$$

Here $\theta$ is 0.85.

#### 2.3.4. Zero-crossing rate

Zero-crossing rate ($Z_c$) means the rate of signal change along a signal. When adjacent signals have different signs, a zero-crossing occurs. It can be defined as

$$Z_c = \frac{1}{2}\sum_{k=0}^{N-1}[sgn(X(k)) - sgn(X(k+1))] \tag{7}$$

#### 2.3.5. Shannon entropy

Shannon entropy ($E$) is the expected information content of a sequence of signal. It describes the average of all the information contents $C$ weighted by their probabilities $p_i$.

$$E = -\sum_{i=1}^{L} p_i C(p) \tag{8}$$

where $L$ is the length of a frog syllable.

#### 2.3.6. Spread, skewness and kurtosis

Spread is used to measure the flatness or the spikiness of a signal. Skewness means the asymmetry of the probability distribution of a real-valued random variable about its mean. Kurtosis is defined as the measure of the "peakedness" of a distribution. Here spread, skewness and kurtosis is calculated based on the Hilbert envelope of the signal.

#### 2.3.7. Root mean square value

Root mean square value ($RMS$) is the square root of the arithmetic mean of the squares of the values and defined as

$$RMS = \sqrt{\frac{1}{L}\sum_{i=0}^{L-1} x(i)^2} \tag{9}$$

#### 2.3.8. Averaged energy

Averaged energy ($E_{avg}$) is defined as the sum of intensity in each frame times weights.

$$E_{avg} = \frac{1}{f}\sum_{k=0}^{f-1}\sum X(k) \tag{10}$$

where f is the number of frames of a frog syllable.

Table 1: Frog list of family, genus and species

| No. | Species | Genus | Family | No. | Species | Genus | Family |
|---|---|---|---|---|---|---|---|
| 1 | Austrochaperina fryi | Austrochaperina | Microhylidae | 43 | Litoria inermis | Litoria | Hylidae |
| 2 | Austrochaperina robusta | Austrochaperina | Microhylidae | 44 | Litoria infrafrenata | Litoria | Hylidae |
| 3 | Cophixalus bombiens | Cophixalus | Microhylidae | 45 | Litoria jervisiensis | Litoria | Hylidae |
| 4 | Cophixalus concinnus | Cophixalus | Microhylidae | 46 | Litoria latopalmata | Litoria | Hylidae |
| 5 | Cophixalus crepitans | Cophixalus | Microhylidae | 47 | Litoria lesueuri | Litoria | Hylidae |
| 6 | Cophixalus exiguus | Cophixalus | Microhylidae | 48 | Litoria littlejohni | Litoria | Hylidae |
| 7 | Cophixalus hosmeri | Cophixalus | Microhylidae | 49 | Litoria longirostris | Litoria | Hylidae |
| 8 | Cophixalus infacetus | Cophixalus | Microhylidae | 50 | Litoria microbelos | Litoria | Hylidae |
| 9 | Cophixalus monticola | Cophixalus | Microhylidae | 51 | Litoria nannotis | Litoria | Hylidae |
| 10 | Cophixalus neglectus | Cophixalus | Microhylidae | 52 | Litoria nasuta | Litoria | Hylidae |
| 11 | Cophixalus ornatus | Cophixalus | Microhylidae | 53 | Litoria nigrofrenata | Litoria | Hylidae |
| 12 | Crinia deserticola | Crinia | Microhylidae | 54 | Litoria nyakalensis | Litoria | Hylidae |
| 13 | Crinia parinsignifera | Crinia | Microhylidae | 55 | Litoria olongburensis | Litoria | Hylidae |
| 14 | Crinia remota | Crinia | Microhylidae | 56 | Litoria pallida | Litoria | Hylidae |
| 15 | Crinia signifera | Crinia | Microhylidae | 57 | Litoria pearsoniana | Litoria | Hylidae |
| 16 | Crinia tinnula | Crinia | Microhylidae | 58 | Litoria peronii | Litoria | Hylidae |
| 17 | Cyclorana alboguttata | Cyclorana | Hylidae | 59 | Litoria phyllochroa | Litoria | Hylidae |
| 18 | Cyclorana brevipes | Cyclorana | Hylidae | 60 | Litoria rheocola | Litoria | Hylidae |
| 19 | Cyclorana cryptotis | Cyclorana | Hylidae | 61 | Litoria rothii | Litoria | Hylidae |
| 20 | Cyclorana manya | Cyclorana | Hylidae | 62 | Litoria rubella | Litoria | Hylidae |
| 21 | Cyclorana novaehollandiae | Cyclorana | Hylidae | 63 | Litoria subglandulosa | Litoria | Hylidae |
| 22 | Cyclorana verrucosa | Cyclorana | Hylidae | 64 | Litoria tyleri | Litoria | Hylidae |
| 23 | Limnodynastes convexiusculus | Limnodynastes | Myobatrachidae | 65 | Litoria verreauxii | Litoria | Hylidae |
| 24 | Limnodynastes dumerilii dumerilii | Limnodynastes | Myobatrachidae | 66 | Litoria xanthomera | Litoria | Hylidae |
| 25 | Limnodynastes dumerilii grayi | Limnodynastes | Myobatrachidae | 67 | Mixophyes schevillii | Mixohyes | Myobatrachidae |
| 26 | Limnodynastes fletcheri | Limnodynastes | Myobatrachidae | 68 | Mixophyes fasciolatus | Mixohyes | Myobatrachidae |
| 27 | Limnodynastes ornatus | Limnodynastes | Myobatrachidae | 69 | Mixophyes fleayi | Mixohyes | Myobatrachidae |
| 28 | Limnodynastes peronii | Limnodynastes | Myobatrachidae | 70 | Mixophyes iteratus | Mixohyes | Myobatrachidae |
| 29 | Limnodynastes tasmaniensis | Limnodynastes | Myobatrachidae | 71 | Philoria kundagungan | Philoria | Myobatrachidae |
| 30 | Limnodynastes terraereginae | Limnodynastes | Myobatrachidae | 72 | Philoria loveridgei | Philoria | Myobatrachidae |
| 31 | Litoria aurea | Litoria | Hylidae | 73 | Philoria sphagnicolus | Philoria | Myobatrachidae |
| 32 | Litoria bicolor | Litoria | Hylidae | 74 | Pseudophryne australis | Pseudophryne | Myobatrachidae |
| 33 | Litoria brevipalmata | Litoria | Hylidae | 75 | Pseudophryne bibronii | Pseudophryne | Myobatrachidae |
| 34 | Litoria caerulea | Litoria | Hylidae | 76 | Pseudophryne coriacea | Pseudophryne | Myobatrachidae |
| 35 | Litoria chloris | Litoria | Hylidae | 77 | Pseudophryne covacevichae | Pseudophryne | Myobatrachidae |
| 36 | Litoria dentata | Litoria | Hylidae | 78 | Pseudophryne major | Pseudophryne | Myobatrachidae |
| 37 | Litoria eucnemis | Litoria | Hylidae | 79 | Pseudophryne raveni | Pseudophryne | Myobatrachidae |
| 38 | Litoria ewingii | Litoria | Hylidae | 80 | Taudactylus liemi | Taudactylus | Myobatrachidae |
| 39 | Litoria fallax | Litoria | Hylidae | 81 | Taudactylus rheophilus | Taudactylus | Myobatrachidae |
| 40 | Litoria freycineti | Litoria | Hylidae | 82 | Uperoleia altissima | Uperoleia | Myobatrachidae |
| 41 | Litoria genimaculata | Litoria | Hylidae | 83 | Uperoleia fusca | Uperoleia | Myobatrachidae |
| 42 | Litoria gracilenta | Litoria | Hylidae | 84 | Uperoleia lithomoda | Uperoleia | Myobatrachidae |
| | | | | 85 | Uperoleia littlejohni | Uperoleia | Myobatrachidae |

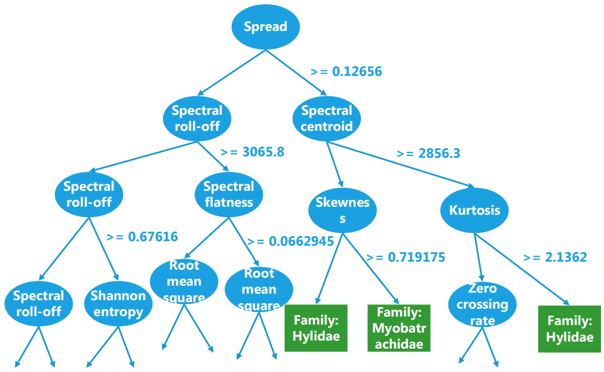
### 2.4. *Feature suggested by Decision Tree*

After feature extraction, a decision tree (DT) classifier is used to evaluate the feature important for classifying family, genus and species. The input to the DT classifier is ten features, The output is the decision tree, which is shown in Fig.2.

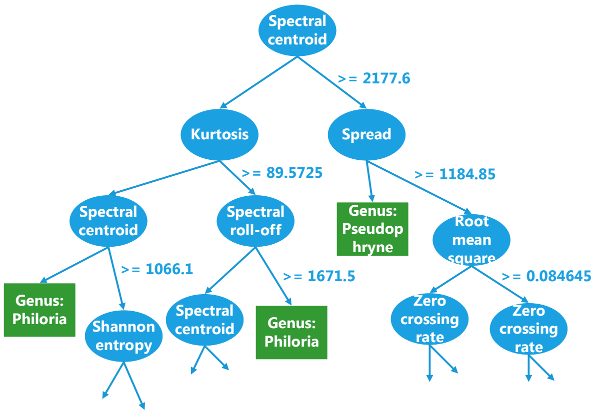
Referring to Fig.2, it is found that the three dominant features and their corresponding significance are different. For classifying frog families, spread, spectral roll-off and spectral centroid are three dominant features. For genus, spectral centroid, kurtosis and spread are three dominant features. Kurtosis, spectral centroid and zero-crossing rate are three dominant features.

### 2.5. *Classification*

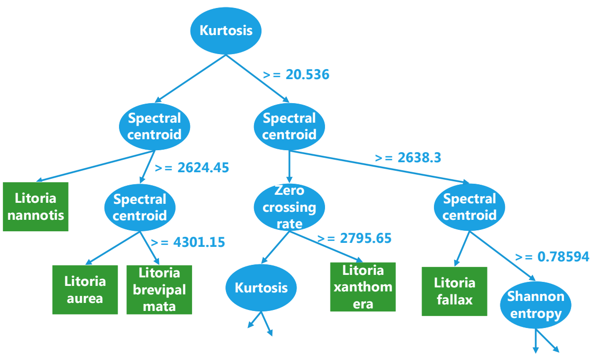
In the next step referring to the Fig.1, a weighted support vector machines classifier is employed to identify frog vocalizations. Due to the high accuracy and superior generalization properties, SVMs have been widely used for the classification of animal sounds [5] [11]. In this study, the 50% of data is used as the training data, the rest for testing. Selected features in Section 2.4 is used to construct the pairs $(v_l^n, L_l^n)$, $l = 1, ..., C_l$, where $C_l$ is the number of frog instance in the training data, $v_l^n$ is the feature vector obtained from the l-th frog species in the training data, and $L_l^n$ is the frog label. As for the classification, the decision function for the classification




(a) Family classification tree




(b) Genus classification tree of Family Myobatrachidae




(c) Species classification tree in Litoria genus

Figure 2: A part of Decision Tree for frog call classification

problem is defined as

$$f(v) = sgn(\sum_{sv} \alpha_l^n L_l^n K(v, v_l^n) + b_l^n) \tag{11}$$

where $K(.,.)$ is the kernel function whose kernel is Gaussian, $\alpha_l^n$ is the Lagrange multiplier, and $b_l^n$ is the constant value.

## 3. Experiment results and discussion

In this experiment, the following classification accuracy is used to examine the performance:

$$Classification\ Acc(\%) = \frac{N_c}{N_s} \tag{12}$$

where $N_c$ is the number of correctly classified instances and $N_s$ is the total number of test species, genus or family.

For frog family classification, spread, spectral roll-off and spectral centroid are put into classifier. The confusion matrix of family classification is shown in Table.2. Frogs used in this experiment are distributed in three families: Microhylidae, Hylidae and Myobatrachidae.

Table 2: Confusion matrix of frog family classification

| **Family** | **Hylidae** | **Microhylidae** | **Myobatrachidae** |
|---|---|---|---|
| **Hylidae** | 109 | 4 | 10 |
| **Microhylidae** | 12 | 25 | 5 |
| **Myobatrachidae** | 5 | 0 | 81 |

For frog genus classification (seven genera in Family Myobatrachidae), spectral centroid, kurtosis and spread are used. The confusion matrix of genus classification is shown in Table.3.

Table 3: Confusion matrix of frog genus classification

| **Genus** | **Crinia** | **Limnodynastes** | **Mixohyes** | **Philoria** | **Pseudophryne** | **Taudactylus** | **Uperoleia** |
|---|---|---|---|---|---|---|---|
| **Crinia** | 15 | 0 | 0 | 0 | 2 | 0 | 1 |
| **Limnodynastes** | 0 | 17 | 4 | 2 | 1 | 0 | 1 |
| **Mixohyes** | 0 | 6 | 7 | 1 | 0 | 0 | 0 |
| **Philoria** | 0 | 1 | 1 | 6 | 0 | 0 | 0 |
| **Pseudophryne** | 0 | 0 | 0 | 0 | 15 | 1 | 0 |
| **Taudactylus** | 0 | 0 | 0 | 0 | 0 | 4 | 0 |
| **Uperoleia** | 0 | 0 | 0 | 0 | 0 | 1 | 10 |

Thirty six frog species of Litoria genus is used for frog species experiment. Kurtosis, spectral centroid and zero-crossing rate are used as selected features. The confusion matrix of species classification is shown in Fig.3.

In this study, ten acoustic features are evaluated for multi-classification of Australia frogs: Family, Genus and Species. For family classification, the classification accuracy is 86.5%, 86.2% and 84.35% for family Hylidae, Microhylidae and Myobatrachidae respectively. Spread calculates statistics of the signal distribution. Different spread value of the frog call shows different signal amplitude variation. Since family Hylidae is termed as tree frogs, family Microhylidae is often known as narrow-mouthed frogs and family Myobatrachidae is known as Australian ground frogs, different habitat area and different physiological structure make a different for the frog call amplitude. For genus classification, the average classification classification accuracy

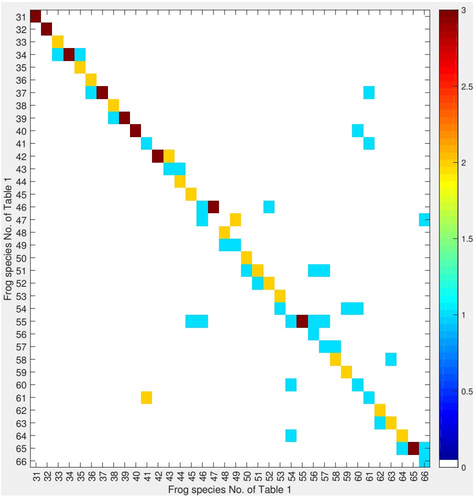
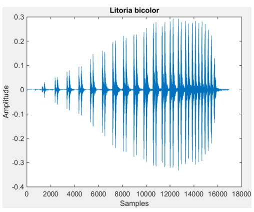
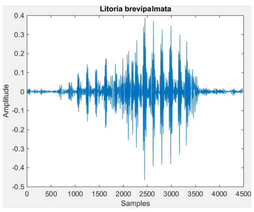



Figure 3: Confusion matrix of frog species classification

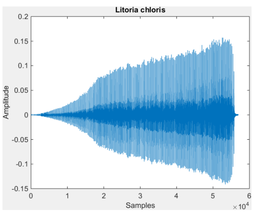
is 75.58%. Spectral centroid which is highly correlated to the dominant frequency (used to represent the advertisement call in our previous work [7]) is the most important feature for genus classification. It is because that advertisement calls of closely related species in phylogenetic are predicted to be more similar than those of distant species. For species classification, kurtosis is the most important which shows that different frog species tend to have different shapes of frog calls (Fig.4). The average classification accuracy is It is worth to mention that spectral centroid is in the most three important features for all level classification, which shows the important of advertisement call in analysing frogs due to the high correlation between spectral centroid and the advertisement call.



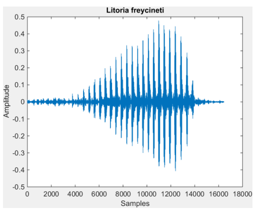

(a) Litoria bicolor




(b) Litoria brevipalmata




(c) Litoria chloris




(d) Litoria freycineti

Figure 4: Waveform of four frog species in Litoria genus

## 4. Conclusion

In this study, we evaluate ten acoustic features for multi-level frog classification:family, genus and species. For all ten features, spectral centroid is the most three important features for all level frog classification, which demonstrates the importance of advertisement call in classifying frog calls due to their high correlation. The highest classification accuracy is achieved by the family level and the lowest is species level. It shows that the call difference between frogs in higher phylogenetic level is larger than lower level.

## Reference


[1] J. Wimmer, M. Towsey, P. Roe, I. Williamson, Sampling environmental acoustic recordings to determine bird species richness, Ecological Applications 23 (6) (2013) 1419–1428.

[2] C.-H. Lee, C.-H. Chou, C.-C. Han, R.-Z. Huang, Automatic recognition of animal vocalizations using averaged mfcc and linear discriminant analysis, Pattern Recognition Letters 27 (2) (2006) 93–101.

[3] W.-P. Chen, S.-S. Chen, C.-C. Lin, Y.-Z. Chen, W.-C. Lin, Automatic recognition of frog calls using a multi-stage average spectrum, Computers & Mathematics with Applications 64 (5) (2012) 1270–1281.

[4] C. Bedoya, C. Isaza, J. M. Daza, J. D. López, Automatic recognition of anuran species based on syllable identification, Ecological Informatics 24 (2014) 200–209.

[5] C.-J. Huang, Y.-J. Yang, D.-X. Yang, Y.-J. Chen, Frog classification using machine learning techniques, Expert Systems with Applications 36 (2) (2009) 3737–3743.

[6] N. C. Han, S. V. Muniandy, J. Dayou, Acoustic classification of australian anurans based on hybrid spectral-entropy approach, Applied Acoustics 72 (9) (2011) 639–645.

[7] J. Xie, M. Towsey, A. Truskinger, P. Eichinski, J. Zhang, P. Roe, Acoustic classification of australian anurans using syllable features, in: 2015 IEEE Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (IEEE ISSNIP 2015), Singapore, Singapore, 2015.

[8] F. Glaw, M. Vences, Phylogeny and genus-level classification of mantellid frogs (amphibia, anura), Organisms Diversity & Evolution 6 (3) (2006) 236 – 253.

[9] B. Gingras, W. T. Fitch, A three-parameter model for classifying anurans into four genera based on advertisement calls, The Journal of the Acoustical Society of America 133 (1) (2013) 547–559.

[10] W. E. Duellman, L. Trueb, Biology of amphibians.

[11] M. A. Acevedo, C. J. Corrada-Bravo, H. Corrada-Bravo, L. J. Villanueva-Rivera, T. M. Aide, Automated classification of bird and amphibian calls using machine learning: A comparison of methods, Ecological Informatics 4 (4) (2009) 206–214.