Journal of Tropical Resources and Sustainable Science

journal homepage: jtrss.org

Effect of Natural Background Noise and Man-Made Noise on Automated Frog Calls Identification System

Haryati Jaafar*, Dzati Athiar Ramli

Intelligent Biometric Group, School of Electrical and Electronic, Universiti Sains Malaysia, 14300 Nibong Tebal, Pulau Pinang, Malaysia

Available online 25 July 2015

Keywords:
Frog identification system, natural background noise, man-made noise, syllables segmentation, feature extraction and classification

⌧*Corresponding author:
Haryati Jaafar,
Intelligent Biometric Group,
School of Electrical and Electronic,
Universiti Sains Malaysia, 14300
Nibong Tebal, Pulau Pinang,
Malaysia
Email: haryati.jaafar@yahoo.com

Abstract

Frog identification based on their calls becomes important for biological research and environmental monitoring. However, identifying particular frog calls becomes challenging particularly when the frog calls are interrupted with noises either in natural background noise or man-made noise. Hence, an automatic identification frog call system that robust in noisy environment has been proposed in this paper. Experimental studies of 675 audio obtained from 15 species of frogs in the Malaysian forest and recorded in an outdoor environment are used in this study. These audio data are then corrupted by 10dB and 5dB noise. A syllable segmentation technique i.e. short time energy (STE) and Short Time Average Zero Crossing Rate (STAZCR) and feature extraction, Mel-Frequency *Cepstrum Coefficients (MFCC)* are employed to segment the desired syllables and extract the segmented signal. Subsequently, the Local Mean k-Nearest Neighbor with Fuzzy Distance Weighting (LMkNN-FDW) are employed as a classifier in order to evaluate the performance of the identification system. The experimental results show both of natural background noise and man-made noise outperform by 95.2% and 88.27% in clean SNR, respectively.

1. Introduction

In recent years, a study on the frog species recognition has become crucial as frogs also play an important role in the ecological system. These amphibian species have survived for the past 250 million years in countless ice ages, asteroid crashes and other environmental disturbances. Yet, one-third of their species are on the verge of extinction nowadays [1]. They are also considered as one of main links in many ecosystem food webs. Often unseen, they are the most abundant, diverse group of vertebrate organisms in forested for high trophic levels and considered the top predators of invertebrates [2]. Apart from that, these amphibians are also useful to be a bio-indicator of environment stress. The health of frog population indicates the health of the whole ecosystem due to their bi-phasic life [3]. Frog can be found in an aquatic environment. Some species are widespread, but others are very localised. Although a few frog species are flourishing in human environments and has adapt to the noise, many species have suffered dramatic population declines. Hence, an intelligent frog identification system is needed so as to preserve the world from frog species elimination.

Generally, frog use acoustic communication for a wide range of essential functions, not only for territorial defence and mating ritual, but also for navigation, nurturing, detection of predators, and foraging [4]. Their sound can receive over varying distance that allows an obstructive detection of their existence [5]. Therefore, identifying frog species based on their calls is more effective for environmental monitoring. Nonetheless, identifying particular frog calls becomes challenging, in the case where background noise often interferes the process. In a

ISSN Number: 2289-3946

natural source of background noise include wind, rain, and waterfalls that have moderate and low frequencies (i.e. under 4 kHz), frogs dwell in noisy environment by produces the higher level of their call rates and call durations [6] makes the identification of frog species is more convenience. In contrast, the identification of frog species in high frequencies noise that usually creates by human activities i.e. industry, construction and transportation is more challenging. Due to the man-made noise introduces into their environment, some of species lack advertisement calls altogether. It is believed by some biologists that man-made noise is a major contributing to the falling numbers of suitable habitat [6].

In order to overcome the problem of identification frog species in noisy environment, an automated frog identification system that robust in noisy environment has been proposed in this study. The database consists of 675 audio data obtained from 15 species of frogs in the Malaysian forest and recorded in an outdoor environment. A variety of natural background noises and man-made noises have been added in the frog calls and corrupted in 5dB and 10dB of signal noise to ration (SNR). The system consists of three important processes i.e. syllable segmentation, feature extraction and classification where the Short Time Energy (STE) and Short Time Average Zero Crossing Rate (STAZCR) [7], Mel-Frequency Cepstrum Coefficients (MFCC) [8] and local means kNN with fuzzy distance-weighting (LMKNN-FDW) [9] are employed in every process, respectively. This paper is outlined as follows. The architecture of frog identification system contains the background of the study, the syllables segmentation, feature extraction and the classification process is explained. Consequently, the experimental results are presented. Finally, conclusions are summarized in final section.

2. Materials and Methods

2.1. The Architecture of Frog Identification System, Data Acquisition and Noise Recording

The frog calls were collected from 2 localities in the Malaysian forest in the state of Kedah between February 2012 and July 2013. The first locality is in Sungai Sedim, Kulim and the sounds were recorded next to a running stream from 8.00 pm to 12.00 pm. The second locality is in Baling where the frog sounds were recorded in a swampy area from 6.00 pm to 10 pm. The recordings were made using a Sony Stereo IC Recorder ICD-AX412F supported with a Sony electric condenser microphone 32kHz sampling frequency with WAV format. The sounds samples are then converted to 16-bit mono. The recording dataset consists of 15 species where the scientific name, common name and images are tabulated as in Table 1.

In order to investigate the effect of frog calls in noisy environment, a variety of natural background noises i.e. rain, running stream and insects have been added in the frog calls. The noises were recorded in the study sites with the microphone position placed at ground level, protected from direct impact of the other elements such as wind and raindrops. For the purpose of evaluating the frog calls in man-made noisy environment, different sound of transportation i.e. car, train, and airplane has been downloaded from internet sound database, www.soundjay.com is employed in this research.

2.2. Syllables Segmentation

A syllable is basically a sound that frog produces with a side blow of air from the lungs. In order to segment the syllables, the STE and STAZCR have been applied in this study. The STE function is defined by the following expression;

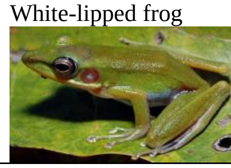
$$E_k = \frac{1}{N} \sum_{k=1}^{N} \left[x(k) w(k-m) \right] \tag{1}$$

where E_k is the energy of the sample k of the signal, and $w[k]$ is a window of size N. On the other hand, the STAZCR is employed as a part of the front-end processing in the frog identification system. As the amplitude of the unvoiced part normally has much higher value and vice- versa for frog sounds, the STAZCR is defined as;

$$Z_k = \frac{1}{2N} \sum_{k=1}^{N} |\operatorname{sgn} x(k) - \operatorname{sgn}[x(k-1)]|\, w(k-m) \tag{2}$$

ISSN Number: 2289-3946

Table 1: List of frog call samples.

Family, scientific name, common name and image				
Microhylidae *Hylarana glandulosa* Rough sided frog 	Ranidae *Odorrana hosii* Poisonous rock frog 	Microhylidae *Microhyla heymonsi* Taiwan rice frog 	Rhacophoridae *Philautus petersi* Kerangas bush frog 	Rhacophoridae *Polypedates leucomystax* Common tree frog 
Bufonidae *Phrynoidis aspera* River toad 	Microhylidae *Kaloula baleata* Flower pot toad 	Dicroglossidae *Fejervarya limnocharis* Grass frog 	Microhylidae *Microhyla butleri* Painted chorus frog 	Bufonidae *Duttaphrynus melanostictus* Black-spectacled toad 
Microhylidae *Kaloula pulchra* Asian painted bullfrog 	Rhacophoridae *Philautus mjobergi* Bubble-nest frog 	Ranidae *Hylarana labialis* White-lipped frog 	Rhacophoridae *Rhacophorus appendiculatus* Frilled tree frog 	Bufonidae *Genus ansonia* Stream toad 

In order to ensure that the start and end point detection of the syllable performs well, threshold levels need to be set. If the values of the STE is continuously lower than a certain set of threshold, or if the most values of the STAZCR in the segment are lower than certain set of threshold, then the segment is indexed as syllable. The threshold level is given as;

$$T_h = \begin{cases} 1, E_k \geq T_E \text{ and } \quad Z_k \geq T_z \\ 0, \text{otherwise} \end{cases} \tag{3}$$

where T_E and T_Z are the thresholds for STE and STAZCR, respectively and they are defined as;

$$T_E = \frac{W(E_{k\,\max}) + E_{k\,\max}}{W+1} \tag{4}$$

$$T_Z = \frac{W(Z_{k\,\max}) + Z_{k\,\max}}{W+1} \tag{5}$$

where W is the weight parameter, $E_{k\max}$ and $Z_{k\max}$ the maximum values of frequency distribution that has been determine by the tabular data organizes by histogram.

2.3. Feature Extraction

The segmented syllables then were extracted to represent the audio signal in the compact digital form. In this paper, MFCC is selected due to the feature is more robust to noise compared to other feature extraction such as linear predictive coding (LPC). MFCC processing consists of signal preemphasis, windowing, spectral analysis, filter bank processing, log energy computation and mel frequency cepstrum computation as shown in Fig.1. There are 15 melcepstrum coefficients, one log energy coefficient and three delta coefficients per frame have been set in the experiments.

ISSN Number: 2289-3946

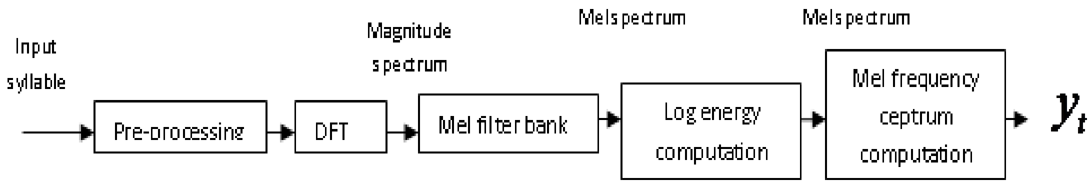


Figure 1: Typical MFCC process

2.4. Classification

Basically, LMkNN-FDW is the improvement of k-nearest neighbor (kNN) where the distance between query pattern or testing sample and local means vector is assigned using fuzzy algorithm [11]. Let $X_i = \{x_n \in R^m\}_{n=1}^{N}$ be a training sample where *m* is the number of dimensional in feature space, *N* is the total number of training sample and $y_n \in \{c_1, c_2, ..., c_M\}$ denotes the class label for x_n. Then, the kNN is determined from the set X_i for each class y_n by Euclidean distance as given as;

$$d\left(x, x_{ij}^{N}\right) = \sqrt{\left(x - x_{ij}^{N}\right)^{T}\left(x - x_{ij}^{N}\right)} \tag{6}$$

The local mean vector Y_{ik} is obtained by applying k nearest neighbor of training sample such that;

$$Y_{ik} = \frac{1}{k}\sum_{j=1}^{k} x_{ij}^{N} \tag{7}$$

By applying fuzzy with distance weighting, the fuzzy membership $U_{ij}(k)$ can be computed as follows;

$$u_{ij} = \frac{\sum_{j=1}^{k} u_{ij}\left[\frac{1}{\left\|x - Y_{ik}\right\|^{2/(m-1)}}\right]}{\sum_{j=1}^{k}\left[\frac{1}{\left\|x - Y_{ik}\right\|^{2/(m-1)}}\right]} \tag{8}$$

Note the notation *m* denotes the fuzzy weight of the distance or fuzzy relationship. If value *m* increases, the neighbors are more evenly weighted. This causes the distance between training and query pattern to have less effect on each other and vice-versa.

Table 2: Performance results obtained in natural background noises.

Species	Rain			Running stream			Insect		
	Clean	10dB	5dB	Clean	10dB	5dB	Clean	10dB	5dB
Hylarana glandulosa	25	20	25	25	24	25	24	23	25
Kaloula pulchra	25	25	24	25	22	19	24	23	25
Odorrana hossi	19	17	13	18	16	13	20	13	9
Polypedates leucomystax	25	22	14	25	21	10	25	21	19
Kaloula baleata	25	24	20	25	8	8	25	25	17
Philautus mjobergi	25	25	20	25	24	20	23	25	25
Duttaphrynus melanostictus	24	24	24	24	23	21	24	25	24
Phrynoidis aspera	25	25	24	25	25	23	25	25	23
Microhyla heymonsi	18	12	11	17	15	13	19	12	9
Fejervarya limnocharis	25	22	16	25	16	13	25	25	18
Genus ansonia	21	24	22	25	25	21	25	25	25
Rhacophorus appendiculatus	25	24	14	25	22	13	25	25	23
Hylarana labialis	25	25	24	25	22	23	25	25	24
Philautus petersi	25	24	25	23	18	20	25	24	25
Microhyla butleri	23	25	24	21	22	21	23	25	24
Total	355	338	300	353	303	263	357	341	315
C_A (%)	94.67	90.13	80	94.13	80.8	70.13	95.2	90.93	84

ISSN Number: 2289-3946

3. Results and Discussion

The experiments are implemented using Matlab R2010 (b) and have been tested in Intel Core i5, 2.1GHz CPU, 6G RAM and Windows 7 operating system. In this experiment, the data of 45 syllables are extracted by a MFCC in which 20 syllables are used for training and 25 for testing. Classification done using LMkNN-FDW where k is set to 3 and the classification accuracy (C_A) is defined as;

$$C_A = \frac{N_C}{N_T} \times 100\% \tag{9}$$

where N_c is the number of syllables which are recognized correctly and N_T is the total number of test syllables.

Table 2 lists the analytical results of the frog identification system in different types in natural background noises.

The C_A of frog sound for every species of frogs in different types in man-made noises is shown in Table 3. The results show that all of the propose identification frog calls are able to identify the frog species more than 70% in all conditions. It can be seen that the best noise condition that system can identify the frog calls when the frog calls was interrupted by insects noise. Although the CA in the man-made noise performance is lower than natural background noise, the frog identification system is still capable to achieve more than 80% in clean environment and more than 70% for corrupted SNR.

4. Conclusions

In this paper, a proposed frog calls identification system is proposed to overcome the problem of frog identification in noisy environment. The overall accuracy shows that the proposed system achieved CA more than 70% with the most outstanding result when the frog calls is added with insects noise. By using the STE and STAZCR techniques for syllable segmentation, MFCC for feature extraction and LMkNN-FDW as the classifier, their rates were further improved remarkable.

Table 3: Performance results obtained in man-made noises

Species	Airplane			Train			Car		
	Clean	10dB	5dB	Clean	10dB	5dB	Clean	10dB	5dB
Hylarana glandulosa	25	23	20	24	24	18	25	21	20
Kaloula pulchra	25	25	25	25	24	25	25	25	25
Odorrana hossi	15	13	9	13	10	10	15	13	9
Polypedates leucomystax	24	21	19	25	23	21	25	25	24
Kaloula baleata	25	25	25	25	25	25	25	25	25
Philautus mjobergi	25	16	13	19	17	15	24	19	15
Duttaphrynus melanostictus	25	25	25	25	25	25	25	25	25
Phrynoidis aspera	19	18	18	17	19	23	21	19	19
Microhyla heymonsi	12	9	7	18	14	12	13	13	10
Fejervarya limnocharis	24	22	20	17	20	14	25	23	20
Genus ansonia	21	21	12	17	15	13	24	14	18
Rhacophorus appendiculatus	25	24	23	25	22	20	25	18	21
Hylarana labialis	25	18	18	24	20	19	25	22	18
Philautus petersi	10	8	9	9	11	5	10	4	5
Microhyla butleri	22	21	20	22	22	21	24	25	24
Total	322	289	263	305	291	266	331	291	278
C_A (%)	85.86	77.07	70.13	81.33	77.60	70.93	88.27	77.60	74.13

Acknowledgment

This work was financially supported by Universiti Sains Malaysia Short Term Grant 60311048, Research University Grant 814161 and Research University-Post Graduate Grant Scheme 8046019 for this project.

References

[1] N.H. Shubin and F. A. Jenkins, in: *An Early Jurassic jumping frog*, Nature Vol. 377, No. 6544 (1992), p.49-52.

[2] J.W. Gibbons, C.T. Winne, D. E. Scott, J.D. Willson, X. Glaudas and K.M. Andrews, in: *Remarkable amphibian biomass and abundance in an isolated wetland: implications for wetland conservation*, Conservation Biology Vol. 20 (2006), p. 1457–1465.

ISSN Number: 2289-3946

[3] M.T. Southerland, R.E. Jung, D.P. Baxter, I.C. Chellman, G. Mercurio and J.H. Volstad, in: *Stream salamanders as indicators of stream quality in Maryland, USA*, Applied Herpetology Vol.2 (2004), p. 23–46.

[4] M.I. Herrera-Montes, and T. M.Aide, in: *Impacts of traffic noise on anuran and bird communities*, Urban Ecosystems Vol. 14 (2011), p. 415-427.

[5] M.K. Obrist, G. Pavan, J. Sueur, K. Riede, D. Llusia, and R. Márquez, in: *Bioacoustic approaches in biodiversity inventories*, In: Manual on Field Recording Techniques and Protocols for All Taxa Biodiversity Inventories, Abc Taxa, (2010), Vol. 8, p. 68-99.

[6] M. Penna, and C. Hamilton-West, in: *Susceptibility of evoked vocal responses to noise exposure in a frog of the temperate austral forest*, Anim. Behav. Vol.74, (2007), pp. 45-56.

[7] H. Jaafar, and D.A. Ramli, in: *Automatic syllables segmentation for frog identification system*, the 2013 9th IEEE Int. Colloquium on Signal Processing and its App, p. 90-95.

[8] H. Jaafar, D. A. Ramli, and S. Shahrudin, in: *MFCC based frog identification system in noisy environment*, 2013 IEEE International Conference on Signal and Image Processing Applications, p. 123-127.

[9] H. Jaafar, D. A. Ramli, B. A. Rosdi, and S. Shahrudin, in: Frog identification system based on local means k-nearest neighbors with fuzzy distance weighting, 8th International Conference on Robotic, Vision, Signal Processing and Power Application, Springer Singapore, p. 153-15

ISSN Number: 2289-3946