Frequency Domain Processing for Artificial Synthesis of Swiftlet's Sound Waves

Mohammad Ali Nematollahi[1], S.A.R Al-Haddad[1], Abdul Rahman Ramli[1], Azhar Kassim[2], Shaiful Jahari Hashim[1]
[1]*Department of Computer & Communication Systems Engineering,*
[2]*Department of Animal Science,*
University Putra Malaysia, Selangor, Malaysia.
sar@upm.my

***Abstract*—Swiftlet is a valuable farming industry in South-East Asia for earning foreign exchange. Many countries at this region are competing with each other to localize this industry. However, inviting the swiftlets into the farms based on traditional bird-call playing involves a trial and error process. This paper proposes a new technique to mechanize this process using a spectrogram processing approach. A novel model has been designed to recognize the bird-call and construct the enrollment database. Then, a frequency-based processing of swiftlet's sound waves based on the database was built with real world swiftlet's sound waves. The proposed prototype can be applied successfully to improve this industry.**

***Index Terms*—Bird Call Recognition; Swiftlet Sound; Synthesis Swiftlet Call; Frequency Domain Processing.**

I. INTRODUCTION

Considering the main habitat of the swiftlets is at the South-East of Asia, Malaysia located in this region has a big potential for swiftlet farming. According to the Ministry of Housing and Local Government, which provides guidelines on swiftlet farming industry. Malaysia was the main contributor of this industry in 2004, although some countries such as Vietnam, China, Cambodia, Philippines and Thailand are actively working on the swiftlet industry.

Some species of swiftlets that grow in a nest located near the urban areas have a commercial value. To invite the swiftlets to these farms, factors, such as such as humidity, darkness, and sound need to be considered. Further, two main sounds that should be played to attract the birds to the farm include internal and external sounds. DWNP and Bird Nest Association (BNA) provide a guideline for farmers to give advice on the best time for playing the internal and external sounds

The main aim of this project is to develop a new algorithm based on three processing stages: First, identifying the bird calls through frequency domain using signal processing methods, then detecting the direction of the song, and finally synthesizing (generating) the signal of the song by computer. The whole process is shown in Figure 1.

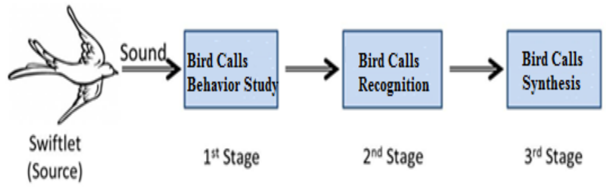

Figure 1: General system block diagram

Currently, to attract the swiftlets, their chirping and mating sound is captured from the cave, and this sound is then played using an acoustic device in the swiftlet house. This process is noisy because one sound may disrupt other sounds. Moreover, this type of swiftlet sound production lacks of any analysis as it is mainly based on trial-and-error process. Figure 2(A) presents the swiftlet’s man-made house in 3-D with a specific size, while Figure 2(B) shows the name of each room in 2-D view. The swiftlets enter from the open roof to the roving room, and finally aggregate in the VIP rooms. Nowadays, swiftlet researchers investigate the control of humidity and temperature of nesting house [1], swiftlet hearing system [2], swiftlet habits [3] and swiftlet attraction by sound [4]. However, sound must be studied significantly due to following reasons: It is cheap and the most attractive technique and this technique is partly unknown from different viewpoints.

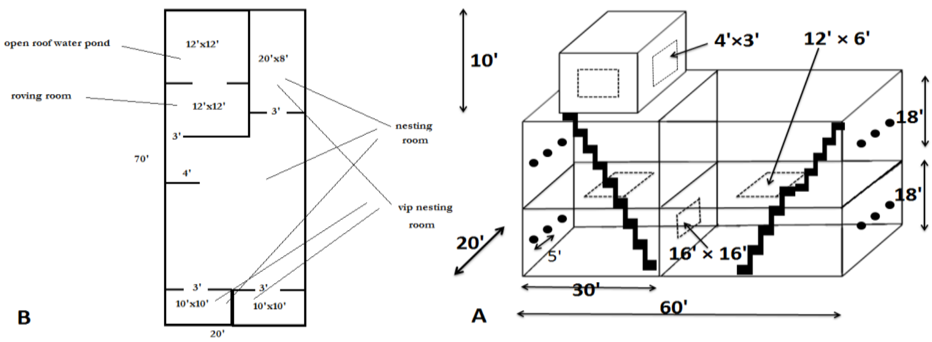

Figure 2: 2-D and 3-D man-made swiftlet house with size and name of each room

This paper studied various acoustic factors of swiftlet signal to synthesis high quality sound that has attraction ability. This paper is organized into six sections. The first section discusses the swiftlet sound production. Due to synthesis process of the swiftlet sound, the second part discusses the hearing system in the swiftlet. The third section reviews related works in this field. The fourth section presents the methodology of the study and proposes a swiftlet acoustic system. The fifth section discusses the results, followed by the conclusion and future work, which are drawn for the recombination future research approach.

II. METHODOLOGY

Previously, T.A.Parker recorded the bird songs in Amazon to identify the species, which reached to 85% of recognition accuracy [5]. Another study has been done to propose a learning technique for bio-acoustical as a tool for species recognition [2]. For this purpose, the sounds of birds, frogs, bats, and orthopnoea were studied. A Support Vector Machine (SVM) approach was developed for bird sound recognition, especially for the species recognition [6]. Other studies have developed Dynamic Time Warping (DTW), Hidden Markova Model, and sinusoidal technique [4-7]. Cornell Macaulay Sound Library is used as a database to label and evaluate the learning systems. However, in reality, there is always noise, which can degrade the recognition rate.

In case of sound direction detection, researchers have proposed a variety of methods based on time difference of arrival (TDOA) Using microphone arrays, which are spatially distributed microphones as the underlying principle in sound localization, Byoungho, Youngjin and Youn-sik proposed the sound source localization using the spatially mapped GCC (generalized cross correlation) functions based on TDOA [8]. Walworth et. al. proposed a linear equation formulation for the estimation of the three-dimensional (3-D) position of a wave source based on the time delay values [9-12]. These studies aim to avoid some effects on direction recognition that naturally comes from topology of the array the direction estimation. Further, front-back confusion effect can be solved by applying more recorders for capturing the sounds, which help to determine the source of the sounds [9].

Two major steps are developed in the proposed algorithm. In the first step, a de-noising is applied on the swiftlet sound to provide a significant frame of the swiftle signal for processing. In the second step, these frames are processed for recognition. For this purpose, Linear Predictive Analysis (LPA) is applied on each frame to compute the LPCs as the main features of the sound. Then, a SRNFN recognizers system is built by applying fuzzy logics and If-then-else in order to capture all of the temporal features. SRNFN can successfully capture the temporal relations among the frames, which improve the recognition among the swiftlet species. The main criterion for evaluating the recognition algorithm is the error of the SRNFN. To improve the recognition, back propagation artificial neural network and Takagi-Sugeno-Kang (TSK)-type recurrent fuzzy network (TRFN) can be used. The recognition process is shown in Figure 3.

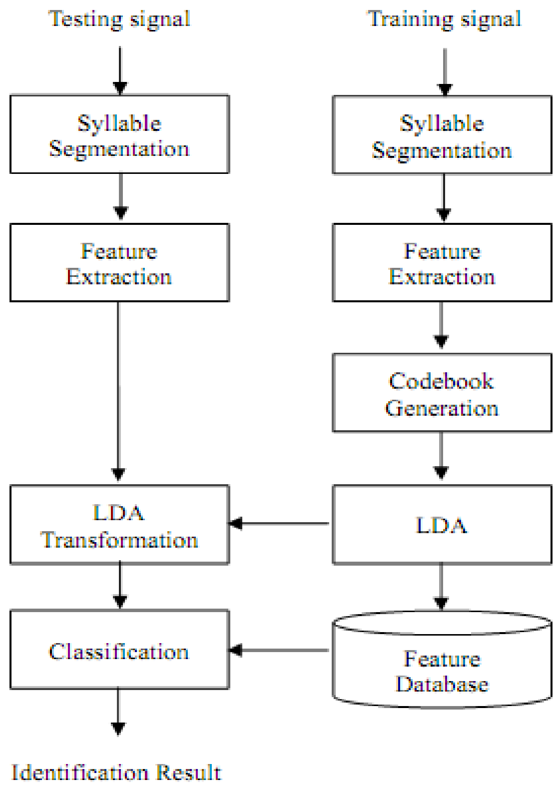

Figure 3: The block diagram of the proposed bird calls Recognition system

In Figure 3, two phases of identification are shown. In each phase, when the sound is recorded, it is divided into a syllable, followed by the extraction of the form each syllable feature. In the training phase, the features are used for generating a codebook. Linear discrimination analysis is also applied to select and reduce the size of the codebooks for efficiency. Finally, the feature of the database is constructed. However, in the testing phases, the extracted features are then transformed by LDA to a proper space by using a database, which is constructed in the training phase before the result is identified.

The recognition of the direction of the Swiftlet’s sound can be done using TDOA method. Estimating the time difference between two or more microphones used to get the sound of the birds will be based on Generalized Cross Correlation (GCC) and a linear equation is used to estimate the direction of the sound source. First, we get the result of recognition system and use Matlab to perform our algorithms of time difference between the microphones as mentioned above. Secondly, the direction of the sound is derived using a triangulation formula. Finally, we implement the designed and simulated algorithms and use them as part of the whole system.

The synthesis part will be based on the samples analyzed in the recognition stage by extracting the pattern of the song signal and applying some of digital processing techniques (DSP) to add or subtract or manipulate many sound waves to produce the desired song signal. In this paper, subtractive synthesis technique will be tested to generate the song by processing the amplitude value (a1) and the frequency (f1) as illustrated in the block diagram as shown in Figure 4. The subtractive synthesis technique can be applied to some additive partials by using some GUI software, like the Harmor.

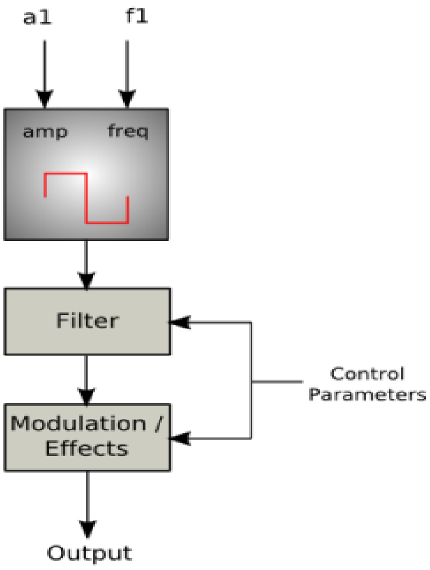

Figure 4: System block diagrams

III. Proposed System

Figure 5 presents the proposed system for the recognition and synthesis of swiftlet call. The first sound of the swiftlet is recognized based on acoustic features, and then it is synthesized based on the recognition step. Finally, the CCTV cameras are used to supervise the behavior of the swiftlets.

Fourier and Linear Predictive (LP) transform are a strong tool for analyzing the frequency domain of any signal. In this research, Fast Fourier Transform (FFT) MATLAB function is applied for analyzing the swiftlet signal. The internal and external swiftlet sounds were recorded by Song Meter (SM2), which can record the ultrasonic sound till 48 kHz. Figure 6 shows SM2 devices at the internal and external position of the swiftlet bird-house. In this experiment, every 2 seconds of swiflet calls were analyzed to investigate the ultrasonic features in it. As shown in Figure 7, the frequency is between 0-48KHhz. However, there are two distinct peaks in the LP spectrum that is beyond the human hearing, which are more than 20 kHz.

Furthermore, from the spectrogram, it is seen that there are some frequencies, which are more than the human hearing range (20 Hz to 20 kHz).

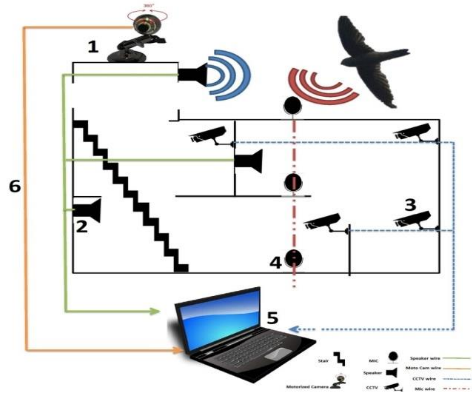

Figure 5: The proposed system for swiftlet call recognition and synthesis

Figure 6: Internal and external swiftlet sound collection by SM2

IV. Swiftlet Frequency Spectrum

Figure 8 shows the different frequency spectrum for the different internal and external swiftlet sounds. As seen, the higher frequency in the internal sound is more than the external. These results demonstrate the results in [4] because it seems all external frequency are limited between the ranges of 1 KHz to 16 KHz. However, for internal sound due to echo and more background noise, it is extended more than 42 KHz.

Figure 9 presents the spectrogram for the internal and external sound of the different swiftlet species. Further, specific pattern is available for each sound. Many syllables shapes, such as U, /, \, - and n may be extracted from the spectrogram for the identification and synthesis process. Moreover, the noise appears more in the internal sound than the external sound. Furthermore, other information such as average silence time interval between syllables, clear or noisy frequency structure and number of rattles (short clicks) could be found from the spectrogram processing.

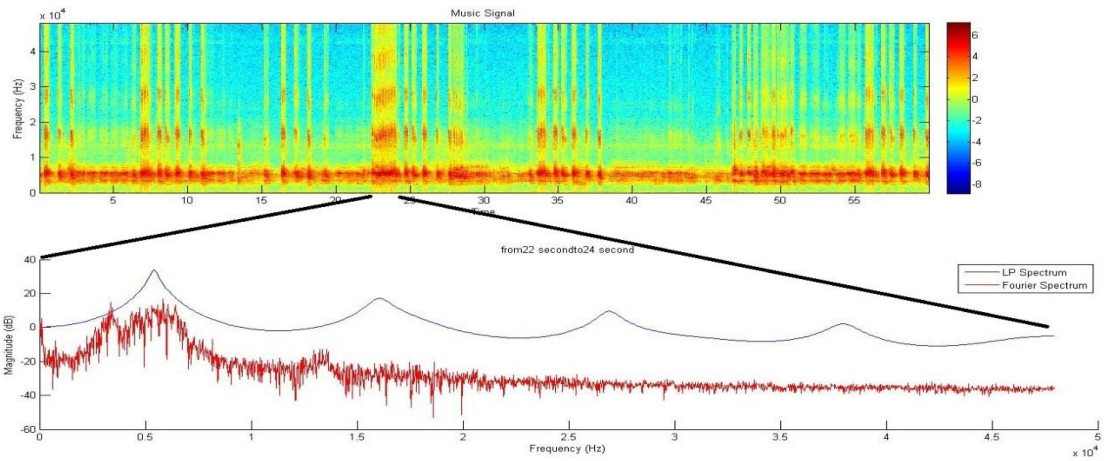

Figure 7: Swiftlet spectrogram, Fourier and LP spectral from 0 Hz to 48 kHz

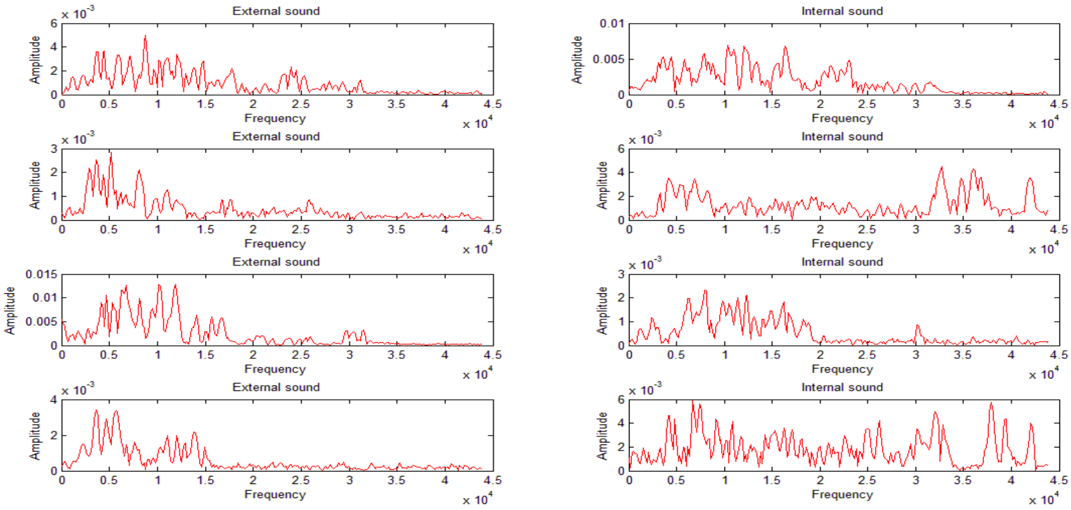

Figure 8: Different internal and external swiftlet sound in frequency domain

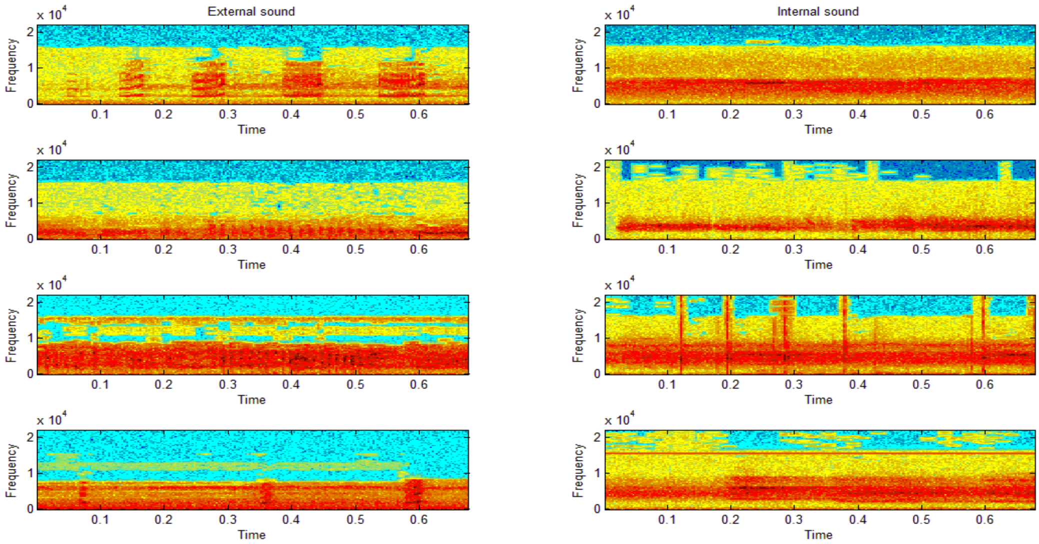

Figure 9: Spectrogram shapes for different swiftlet internal and external sound

V. CONCLUSION AND FUTURE WORK

Swiftlet industry has a good potential to be invested by governments and organization. The main limitation is the traditional process could not be invested in high scale. This paper proposed a new system to improve the traditional swiftlet calling system by frequency-based processing of the swiftlet's sound waves based on the database built with the real swiftlet's sound waves. Further, the success of the proposed system will give big impact to the swiftlets industry, especially in Malaysia and Asia's country by eliminating the noisy voice produced by current systems (in the market). Therefore, the detection, recognition, and synthesis of the swiftlet sound can be performed by the proposed system. In addition, these processes must be done through a spectrogram analysis.

The frequency-based processing of the swiftlet's sound waves has been presented. For future trend, it is potentially applicable to the study of swarm dynamics in a population. They have similar mechanisms, like the "influence of the number of topologically interacting neighbors on swarm dynamics". Providing a comment/remark on this possible extension could also be another trend.

In recent years, many companies and government are investing on various swiftlet research areas. Currently, many swiftlet's research are continuing and undergoing in University Putra Malaysia (UPM). This research is funded by UPM.

ACKNOWLEDGMENT

This study has been conducted under Ministry of Agriculture Malaysia, Swiftlet Excellence Centre UPM project number 6371400-10301 (Group 3).

REFERENCES

[1] H arma, A. "Automatic recognition of bird species based on sinusoidal modeling of syllables", *IEEE Int. Conf. Acoust. Speech and Signal Processing (ICASSP'2003), Hong Kong*, 2003.

[2] SitiNurzalikhaZainiBintiHusniZaini, Sunardi, "Attraction Sound in the Swiftlets House", *Proc. of the Second Intl. Conf. on Advances in Electronic Devices and Circuits - EDC*, 2013.

[3] H arma, A. and Somervuo,P. "Classification of the harmonic structure in bird vocalization", submitted to this conference, 2004.

[4] Catchpole C. K. and Slater P. J. B., "Bird Song: Biological Themes and Variations". *Cambridge, UK: Cambridge University Press*, 1995.

[5] Harma , A. " Automatic identication of bird species based on sinusoidal modeling of syllables*", International Conference on Acoustic Speech Signal Processing*, vol. 5, pp. 545-548, 2003.

[6] Kogan, J. A. and Margoliash , D.. "Automated recognition of bird song elements from continuous recordings using dynamic time warping and Hidden Markov Mod-els: A comparative study". *The Journal of the Acoustical Society of America*, vol. 103, no. 4, pp. 2185-2196, 1998.

[7] Anderson, S. E. , Dave, A. S., and Margoliash,D. "Template-based automatic recognition of birdsong syllables from continuous recordings," *The Journal of the Acoustical Society of America*, vol. 100, pp. 1209–1219, 1996.

[8] Yong-Eun Kim, Dong-hyun SUI ,Jin-Gyun Chung, Xinming Huang and Chul-dong Lee. "Efficient Sound Source Localization Method Using Region Selection" *IDiv IEEE International Symposium on Industrial Electronics (ISlE 2009)* Seoul Olympic Parktel, Seoul, Korea July 5-8, 2009 .

[9] Thorburn, Craig Carpenter. "The Edible Nest Swiftlet Industry in Southeast Asia: Capitalism Meets Commensalism." *Human Ecology* vol. 43, no. 1, pp. 179-184, 2015.

[10] Sureshkumar, A., S. Muruganand, and P. Balakrishnan. "Design and Development of a Virtual Instrument for Hazardous Environment Monitoring and Control Using Lab View." *Intelligent Automation & Soft Computing ahead of print*, pp. 1-12, 2015.

[11] Au, Whitlow WL, and Roderick A. Suthers. "Production of Biosonar Signals: Structure and Form." *Biosonar. Springer* New York, pp. 61-105, 2014.

[12] Sia, Yik Hei. "Swiftlets farming for production of edible bird's nests." *U.S. Patent Applicatio*n 14/328,285, filed July 10, 2014.