BIO ACOUSTIC SIGNAL IDENTIFICATION BASED ON SPARSE REPRESENTATION CLASSIFIER

FROG SPECIES VOICE IDENTIFICATION

WAN ZHI XUAN

UNIVERSITI SAINS MALAYSIA
2018

BIO ACOUSTIC SIGNAL IDENTIFICATION BASED ON SPARSE

REPRESENTATION CLASSIFIER

FROG SPECIES VOICE IDENTIFICATION

by

WAN ZHI XUAN

Thesis submitted in partial fulfilment of the
requirements for the degree of
Bachelor of Engineering (Mechatronics Engineering)

JUNE 2018

ACKNOWLEDGEMENT

Table of Contents

iv

PENGENALAN SIGNAL BIO AKUSTIK BERDASARKAN KLASIFIER PERWAKILAN SPARSE - PENGENALAN SPESIS KATAK MELALUI BUNYI

ABSTRAK

Kebanyakan serangga dan haiwan menghasilkan bunyi sebagai cara komunikasi dalam spesies mereka atau sebagai bunyi yang dikeluarkan semasa makan atau perjalanan. Pengiktirafan automatik isyarat bio-akustik menjadi penting dalam aspek penyelidikan biologi atau pemantauan alam sekitar. Dengan peningkatan teknologi, para saintis hari ini dapat mengklasifikasikan jenis dan spesies haiwan dengan suara mereka tanpa perlu melihat haiwan atau serangga dengan mata kasar. Oleh itu, pengenalan spesies berdasarkan bunyi mereka adalah topik penting untuk meningkatkan aspek penyelidikan ekologi. Projek ini bertujuan untuk membangunkan sistem pengenalan suara spesies katak, mengenali spesies katak yang berlainan dengan menganalisis panggilan mereka. Dalam peringkat pemerolehan data, pangkalan data dari *Intelligent Biometric Research Group* (IBG), Pusat Pengajian Kejuruteraan Electrik dan Electronik Universiti Sains Malaysia dan Pusat Pengajian Farmasi Universiti Sains Malaysia telah digunakan untuk menilai prestasi sistem. Fail-fail panggilan katak mentah diproses dengan menggunakan teknik *Mel- Frequency Cepstrum Coefficient* (MFCC) untuk mengekstrak ciri-ciri yang diperlukan dalam menguji dan melatih sistem. Dalam projek ini, pengelas yang digunakan adalah *Sparse Representation Classifier* (SRC) dan *Kernel Sparse Representation Classifier* (KSRC). Prestasi SRC and KSRC akan dibincangkan dan dibandingkan dalam projek ini. Selain itu, antara muka pengguna grafik (GUI) juga dibangunkan untuk memudahkan pengguna semasa berinteraksi dengan sistem. Pendek kata, KSRC (96.6667%) mempunyai prestasi yang lebih tinggi berbanding dengan SRC (95.6667%). Walau bagaimanapun, KSRC mengambil masa pengiraan yang lebih panjang berbanding dengan SRC. GUI yang melaksanakan KSRC telah diprogramkan dengan dimensi ciri 64-64 sebagai produk akhir.

BIO ACOUSTIC SIGNAL IDENTIFICATION BASED ON SPARSE REPRESENTATION CLASSIFIER – FROG SPECIES VOICE IDENTIFICATION

ABSTRACT

Most insects and animal produce sounds as a way of communication within their species or as noises resulting from feeding or travelling. Automated recognition of bio-acoustic signals is becoming vital in the aspect of biological research or environmental monitoring. With the improvement of technology, scientists today are able to classify types and species of animals by their vocalizations without even need to see the animal or insects with naked eye. Hence, species identification based on their calls or vocalization is an important topic to enhance in the aspect of ecological research. This project aims to develop a frog species voice identification system, recognizing different frog species through analyzing their calls. In the data acquisition stage, databases from Intelligent Biometric Research Group (IBG), School of Electrical and Electronics Engineering, Universiti Sains Malaysia in collaboration with School of Pharmacy, Universiti Sains Malaysia have been used to evaluate the performance of the system. Raw frog call files are processed using Mel-Frequency Cepstral Coefficient (MFCC) technique to extract features that will be needed in testing and training the system. In this project, the classifier used is Sparse Representation Classifier (SRC) and Kernel Sparse Representation Classifier (KSRC). Performance between SRC and KSRC is compared and discussed in this project. Besides, a graphic user interface (GUI) is also developed to facilitate the user while interacting with the system. Two experiments were done in this project, both using SRC and KSRC. In short, KSRC (96.6667%) has a higher performance in accuracy compared to SRC (95.6667%). However, KSRC takes a longer computation time compared to SRC. A GUI was developed implementing KSRC with feature dimension of 64-by-64 as an outcome of this project.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

dB   Decibel

DCT   Discrete Cosine Transform

DFT   Discrete Fourier Transform

GPS   Global Positioning System

GUI   Graphic User Interface

KSRC   Kernel Sparse Representation Classifier

MFCC   Mel Frequency Cepstral Coefficients

MP3   MPEG Audio Layer III

SNR   Signal to Noise Ratio

SRC   Sparse Representation Classifier

SVM   Support Vector Machine

WAV   Waveform Audio Format

WMA   Window Media Audio

CHAPTER 1

INTRODUCTION

1.1. Research Background

In this section, background of research which are closely related to our project title will be introduced. Researches about animal species recognition, sound recognition in animals and also an overview of frog call recognition system will be revealed.

1.1.1. Animal Species Recognition

Recognition of animal species have always been common. However, with the improvement of technologies, there are numerous ways to recognize animal species. The most common animal recognition techniques are through the appearance of the animal itself. Animal species can also be recognized through DNA sequence variation [1]. Wildlife monitoring of animals is also one of the ways to identify species of animal. Various modern technologies have been developed for wild animal monitoring, including radio tracking, wireless sensor network tracking, satellite and global positioning system (GPS) tracking, and monitoring by motion-sensitive camera traps [2]. Species recognition of animals by using their calls or sound is also one of the advanced ways to identify their species, as the animals may not be disturbed and harmed.

1.1.2. Sound Recognition in Animals

With the continuous improvement of technology and the precision of researches, sound recognition in animals have become common for numerous purposes. For animals, initiation of sound can be means of information transfer or can be the noise they made while travelling or feeding. Most animals make produces sound to communicate with

their own species. However, some sound produced may not falls in the hearing frequency of human being.

In the recent years, numbers of researches regarding automated animal call recognition published have increased with various purposes. Animal sound recognition has been carried out at different environment to analyze the habits and distribution of animals. The purpose for this is to monitor and later improve the survivability of the animals [3].

Not only land animals are in the researchers' interest, marine mammals sound classification was also done by [4]. A feature extraction was perform using 1/6 octave analysis to capture various sounds in the ocean that can allow marine scientist to detect, identify and locate endangered marine species.

The recognition of types of animals, or more precisely, the species of the same animal are also becoming common. A research on bird species identification via transfer learning from music genres was reported by [5]. In the research of [5], transfer learning is proposed to transfer knowledge existing in music genre classification to identify bird species, in conjunction with the existing acoustic similarities.

However, frog species recognition by frog call is still to be detailly explored. An intelligent system that can facilitate the effort to estimate frog community calling activity and species richness has been developed by [6]. A frog call biometric identification system for recognizing frog species has also been developed by [7]. In this study, frog calls were processed into signals and classified using support vector machine (SVM) technique.

1.1.3. Overview of Frog Call Recognition System

This project will be focusing on development of frog calls biometric identification system using automated classifiers (Sparse Representation Classifiers). Classification is a fundamental and quintessential task in pattern recognition and machine learning [8]. For the past decades, many classification approaches based on the Gaussian mixture model (GMM), the hidden Markov model, the support vector machine (SVM) etc. have been successfully applied [9], [10].

Existing works in literature had applied Kernel Sparse Representation Classifiers (KSRC) in automated emotion recognition from speech [11], stated that KSRC has higher effectiveness. In the research, group sparsity constraint in KSRC is proposed to improve the performance by estimating more discriminative and accurate weights.

Study and development of SRC has become a popular title to be studied in recent years. Sparse representation of signal can be expressed by a linear combination of atoms in an over-complete dictionary [12], in which some of the entries are non-zero. In mathematical terms, a linear combination of SRC can be written as Equation 1.1 below:

$$y_{m \times 1} = A_{m \times n} \, x_{n \times 1} \qquad\qquad (Eq\ 1.1)$$

where y is the input signal which is in $R^{m \times 1}$ space, A is the dictionary which is in $R^{m \times n}$ space and x is the sparse solution which is in $R^{n \times 1}$ space.

In this project, we will be going to train and develop the dictionary A to identify the input signal to provide an accurate sparse solution. Frog calls in signal wave form will be taken as data (Data Collection) to undergo feature extraction module before further classified by SRC to identify their species. The accuracy of the output or sparse solution is what we will be putting effort in enhancing.

1.2. Motivation

The motivation of developing a Frog Species Identifier is because the amphibians plays an important role in the environment and ecosystem. Certain species of amphibians are useful as indicators of ecosystem stress. Normally, environmental stress is defined as the biological, chemical, and physical, constraints on the productivity of species and on the development of ecosystems. When the exposure to environmental stressors increases or decreases in intensity, it indicates the ecological responses Populations of stream amphibians can be particularly sensitive to increased siltation because they frequent interstitial spaces among the loose, coarse substrates that comprise the matrix of most natural streambeds [13].

Besides that, frogs also play an important role in pharmaceutical industry. Over the years, researchers found two proteins produced from the skin of frogs that could help treat cancer and other diseases. The proteins disrupt the development of blood vessels: one turn on the process of "angiogenesis" while the other switches it off. The scientists say this discovery has the potential to transform cancer from a terminal illness to a chronic condition [14].

Hence, the identification of different frog species must be precise and accurate as different species of amphibians have different functions and niche in the ecosystem.

1.3. Problem Statement

There are approximately 4,740 species of frogs around the entire world [15] and it is almost impossible for the wildlife researchers to identify every species of frogs through bare ears by their sounds. The current manually method of identifying frog species carried out by the experts may not be efficient and might not be effectively used by the non-expertise. Some species of frogs may be hard to be spotted, for example in secluded area under the leaves or up in the trees. Hence, there is always visual limitation for researchers to identify frog species without looking at it. Therefore, instead of visual recognition, sound recognition is surely a better solution for species identification. The frog species identification by their call through Sparse Representative Classifier (SRC) strongly contributes to the wildlife researchers. The proposed project will be able to detect the samples accurately thus more samples can be collected compared to manual research. By that, wildlife field work can be done efficiently and effectively, reducing time and manpower.

1.4. Objectives of the Project

The aim of this Final Year Project is to develop an automated system to detect and recognize frog sound hence further identify frog species based on the recordings. As below are the main objectives:

1. To develop a frog sound recognition module using Sparse Representative Classifier (SRC) and Kernel Sparse Representation Classifier (KSRC).
2. To compare the difference in performance between SRC and KSRC.
3. To develop a graphic user interface (GUI) to show the results of frog species identification.

1.5. Scope of Project

This project will be focusing on the developing the frog species identifier using SRC and KSRC algorithm. Database of frogs are made up of 15 different frog species that are from Malaysia. Besides that, the performance of SRC and KSRC will be evaluated through experiments by comparing their accuracy and computation time by manipulating the feature sizes and training samples. A graphic user interface (GUI) for the frog species identifier is programmed by using KSRC algorithm to classifier the unknown frog species. The system enables users to identify unknown frog species by uploading the frog calls they recorded. The frog species identifier only accepts wav. files for feature extraction to obtain its features for classification processes. As a result, the frog species common name, scientific name and also a sample image will be shown. Apart of that, a short description about the frog species and a syllable of the frog will be shown as well.

1.6. Thesis Outline

This thesis is divided into 5 chapters. Chapter One introduces about the overall concept and background of the frog sound recognition system. The problem statement, project objectives and the scope of research is included in this chapter as well.

Chapter Two is the chapter for literature review, where past researches and projects of bio-acoustic recognition system, speech recognition and analysis, sound feature extraction using Mel-frequency cepstral coefficient (MFCC) are reviewed and studied as part of the project reference.

Chapter Three is the chapter for project methodology. In this chapter steps and methods to carry out the project is stated and explain in detail. Procedures from data acquisition, feature extraction, training and testing using SRC will be elaborated in this chapter

Chapter Four is the chapter for experimental results and discussion. Results and performance of the system is tested and analyzed. Besides, the development of GUI is also explained in this chapter.

Chapter Five is the conclusion of the project. Future improvements and suggestions of ideas are also included in this chapter.

CHAPTER 2

LITERATURE REVIEW

2.1. Research Background

This chapter will revise related literatures of researches and projects for the development of frog call recognition system. Studies regarding bio-acoustic recognition system, Mel-frequency Cepstral Coefficients (MFCC) signal processing and Sparse Representation Classifier (SRC) will be reviewed in this chapter.

2.2. Bio-acoustic Recognition System

Bio-acoustics normally refers to zoology and closely related to ethology. These branches of bioacoustics study sound production and reception in human and animals. Also, how animals communicate by using sound also is part of their investigation. Apart of that, bioacoustics also investigates the organs of listening and production of sounds as well as the physiological and neurological processes, where sounds of animal are developed and received for communication and echolocation purposes. This field also worked to clarify the relationship between features of sound an animal produces and the environment where the animals are used and the functions their acoustics organ are designed to bring out the function. The field of bioacoustics researches are effectively developed in the 1950s. Since then, recording techniques and analyzing ways became readily available to the research community [16].

Methods to identify or locate bird, animal and insect species by recognizing their calls had been implemented since then [17]. Undoubtedly, these techniques are time consuming, slow and strongly depends on the wildlife researchers' expertise of surveyor's knowledge under investigation procedures. The survey to locate targeted species

normally take places at infrequent intervals mainly because of the time required, which in turn causing problems in analyzing long-term trends. However, high speed advances in electronics and computation are causing the development of automated recognition systems able to handle long-term and continuous monitoring in inhospitable regions without man power. These systems can be designed for hand-held use and applications range from rapid biodiversity assessment especially in acoustically rich habitats [18], electronic identification guides, acoustic autecology and the detection and recognition of pest species. Investigations and studies that involve automated bioacoustics species identification is highly efficient that manual surveying.

Taking [19] as an example, the researcher did an experiment to track the stages of insects' activity in grain bulks by means of acoustic sensing and automated identification using noise spectra processing. According to [19], insects produce noises in audible range inside a grain bulk which the wavelength of noises can be sensed by high performance acoustic detector. Figure 2.1 shows a portable probe of 1.4m length. The portable probe
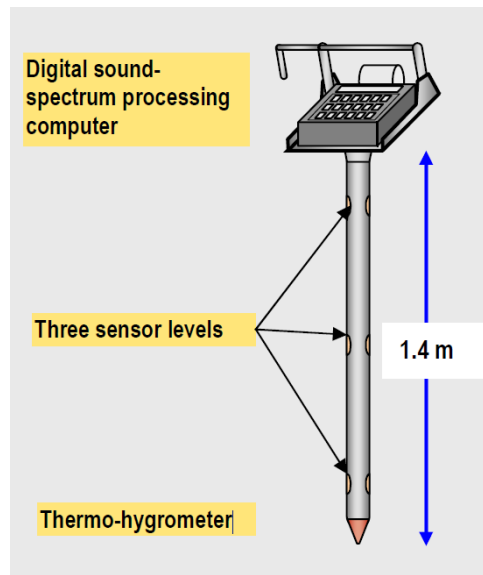

Figure 2. 1 Schematic representation of the portable acoustical probe [19]

was built with three levels of acoustical sensor connected to a computer assisted processing signals. Before storing into the database, the recorded data will be digitized. An automated recognition system programmed with classification algorithm was built to identify insect noise signals. This classification system allows the user to sort the stages of insect, either adult of larval stage.

Studies in environmental sound recognition (ESR) has drastically increased in recent years as the problems in ESR has risen in the past decades. Recent works are more prioritizing in appraisal of non-stationary aspect of sounds as well as developing new predicated features in non-stationary characteristics. These features aim to increase information content pertaining to signal's temporal and spectral characteristics. Moreover, sequential learning methods have also been implemented to capture long-term variation of environmental sounds. In the study of [20], a survey was conducted to offer a qualitative and elucidatory survey on recent developments. The study consists of 3 main parts: basic environmental sound processing schemes, stationary ESR techniques and non-stationary ESR techniques.

Although spectral features that are mainly comprised in stationary ESR techniques were easy to compute, there are also some limitations in the modelling of non-stationary sounds. The non-stationary ESR techniques obtain features derived from the wavelet transform, the sparse representation and the spectrogram. Wavelet based methods give results comparable to stationary methods. Sparse representation and spectrogram-based methods in general perform better. To increase the precision of classifier, MFCC features are often used with one or more features. However, this method of computation is often costlier [20].

2.3. Mel-Frequency Cepstral Coefficients (MFCC)

Mel-Frequency Cepstral Coefficients (MFCC) processing is the most commonly used feature extraction method in automatic speech recognition (ASR) [21]. MFCC is able to copy region of interest as part of human speech production and extract the features vector containing all information about the linguistic message. MFCC mimics the logarithmic perception of loudness and pitch of human auditory system and tries to eliminate speaker dependent characteristics by excluding the fundamental frequency and their harmonics. Figure 2.2 shows the standard implementation of computing the MFCC [22].
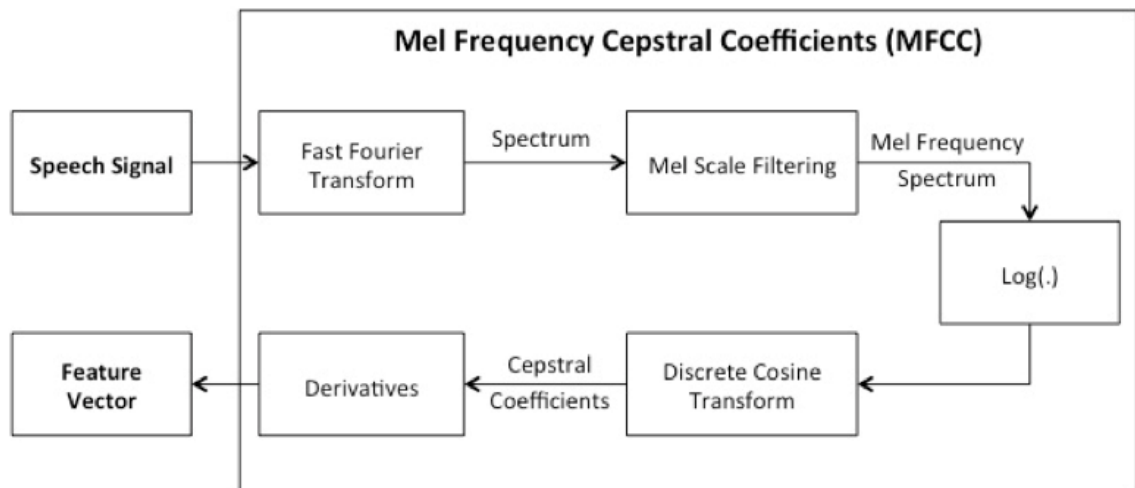
Figure 2. 2 Block Diagram of MFCC Algorithm

Besides, there were also experiments on MFCC application in speaker recognition done using MATLAB. In the research, it was stated that speaker recognition is a new challenge for technologies, where a lot of algorithm have been suggested and developed for feature extraction. This paper evaluates experiments conducted along each step of MFCC process. Apart of that hamming window and rectangular window technique were

11

also compared, taking number of filters for accuracy and efficiency of results as manipulating and responding variable respectively. From the research, it can be concluded that using a 32-filter hamming window has higher accuracy compared to using windowing techniques and number of filters [23]. In this paper, [23] also explained that the process of speaker identification is divided into two main phases, Training (enrolment) Phase and Testing (identification) phase, as shown in Figure 2.3 and Figure 2.4 below.

Figure 2. 3 Enrolment Phase
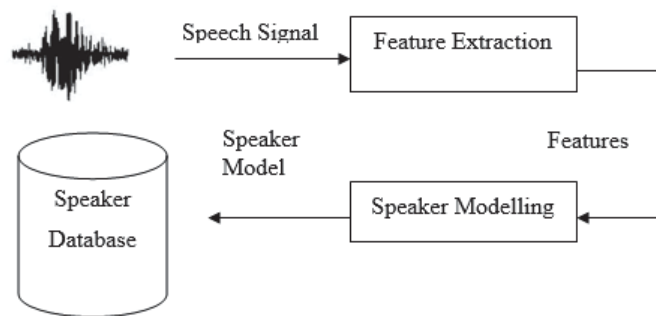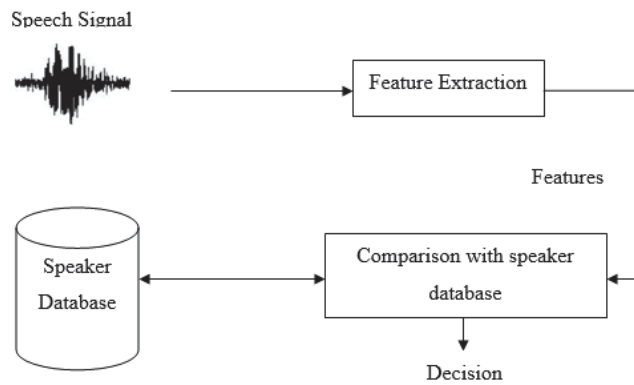

Figure 2. 4 Identification Phase

Apart from that, pitch prediction also can be done from MFCC by using Sparse Spectrum Recovery. This study proposed a three-step method to estimate pitch from MFCC vectors. Firstly, the Mel-filterbank energies (MF-Bes) are estimated from MFCC vectors. Next, a novel method was proposed to estimate the spectrum from MFBE that exploits the sparse

12

nature of the voiced speech spectrum. Lastly, the pitch is estimated from the recovered spectrum. Furthermore, the effect of different levels of truncation of the discrete cosine transformation (DCT) coefficients in MFCC was also explored. [24].

In short, Mel Frequency Cepstral Coefficient processing consists of 6 major procedures. The signal pre-emphasis, windowing, spectral analysis, filter bank processing, log energy computation and Mel frequency cepstrum computation. [25]

MFCC is the most prevalent and dominant method used to extract spectral features according to [26]. In this study on human speech recognition, it was found that frequency domain using the Mel scale is based on human ear scale. MFCC is a representation of real cepstral of a windowed short time signal derived from the Fast Fourier Transform (FFT) of that signal. MFCC is also and audio feature extraction technique that extracts parameter from speech similar to ones that are used by humans in hearing speech. In the study of [26], a summary of MFCC process is also discussed. A basic idea of acoustic feature extraction includes the following algorithmic blocks: Fast Fourier Transformation (FFT), calculation of logarithm (LOG), the Discrete Cosine Transformation (DCT). Figure 2.5 below shows a short summary of MFCC derivation.
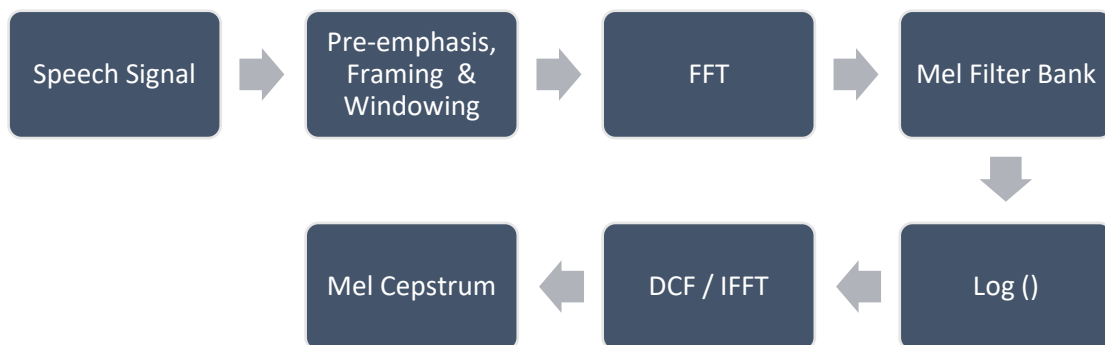
Figure 2. 5 MFCC Derivation

2.4. Sparse Representation Classifier

Sparse Representation is commonly used for classification. A novel supervised matrix factorization method, which can also be used as a classifier with multiple classes was proposed in [27]. In the study, sparse l_1-norm regularization is used as the coefficient matrix of the factorization. The coefficient matrix is mainly formed by combining atom dictionaries of various classes that are trained by penalizing inhomogeneous representation in a joint supervised manner. The samples are also labelled according to classes. The data of interest is modelled as a combination of discriminative linear subspace by projection of sparse. The model that proposed by [27] is based on the observation that many high-dimensional natural signals lie in a much lower dimensional subspaces or union of subspaces. In this paper, the high performance of this representation model for classification is proven by the conducted experiments. The author also suggested that a tight reconstructive representation model can be useful for further improve the effectiveness of the classifier.

Face recognition is also often done by using sparse representation. Automated human faces recognition has always been a challenging field to work on, especially from different viewing angles with varying expression and illumination. In [28], a general classification algorithm for image-based object recognition was proposed. The classification fundamental is based on sparse representation computed using l_1-normalization. This new framework provides new insights into two crucial issues in face recognition: feature extraction and robustness to occlusion. Using feature extraction, the study proved that the selection of feature is no longer critical if the sparsity of the recognition problem is properly harnessed. But, if the feature size is too large or the sparse representation is wrongly computed, the recognition might be critical. Even if the features

are unconventional, down sampled images and random projections, for instance, the prediction perform just as good as conventional features such as Eigenfaces and Laplacian faces, provided the feature dimension is higher than a certain threshold. Errors caused by occlusion or corruption can be solved by making full use of this framework. However, these errors must be sparse with respect to the standard (pixel) basis. The theory of sparse representation is responsible to predict how much occlusion the recognition algorithm can handle and ways to maximize robustness to occlusion by choosing training samples.

Sparse representation has developed into a very basic tool in numerous learning algorithms and received satisfying improvements and outstanding results. Computer vision, pattern recognition, image and signal processing using sparse representation are getting famous among researchers. Dictionary learning is a very famous topic in Industry 4.0 and closely related to sparse representation. A study was done to provide comprehensive investigation and an up to date summary on sparse representation as a guidance for researchers [29]. This paper helps those who are keen in sparse representation by providing fundamentals of researches on sparse representation besides giving a head start for the freshies in computer visioning and pattern recognition segments. In this study, different sparse representation investigations were done by manipulating the norm regularizations. Based on [29], up to now, Sparse based dictionary learning, robustness and performances of sparse representation have become the main branches of investigations in the field of sparse.

Sparse representation has also been implemented in pattern recognition and computer vision. In the field of computer vision, sparse signal representation had shown to be a very convincing tool in the process of acquisition, representation and compression of high dimension signals. To successfully apply sparse representation to a computer

vision tasks, the basis for representing the data have to be addressed. The training dictionary often have to be done by learning from given sample images to a task-specific dictionary. This procedure can maximize the existing algorithms and theory in sparse representation based on new scenarios. In [30], some examples of sparse representation in computer vision is discussed. These examples are to verify that the sparsity is a powerful prior for visual inference. Besides that, solution on improving problems faced in computer vision that can be enhanced by sparse representation theory are also proposed in [30]. In the conclusion of the paper, Sparse representation is indeed a strong tool prior for inference with high-dimensional visual data that have intricate low-dimensional structures. Also, the key to realizing the power of sparse representation achieving state-of-the-art performance is by picking the dictionary that sparse representations correctly reveal the semantics of the data to the dictionary.

However, if sparse representation classifier were to be implemented in real time application, the elapsed time for solving classification will be a big disadvantage. This problem is mainly caused by the sparse signal solver, which is based on l_1 minimization or Basis Pursuit [12]. Hence, researches had been done on to improve the efficiency of sparse signal recovery solver. In this work, a smooth l_0 norm solver is modified and implemented to improved accuracy of classification apart of reducing computation time. Apart of that, kernel sparse representation as a modified version to this solver is also described in this paper.

Since sparse classification based on Basis Pursuit is slower [31], a new algorithm for Sparse Component Analysis (SCA) or atomic decomposition on over-complete dictionaries is presented in [32]. The algorithm is essentially a method for obtaining sufficiently sparse solutions of underdetermined system of linear equations. The solution

obtained by the proposed algorithm is compared with the minimum L1-Norm solution achieved by Linear Programming (LP). It is experimentally shown that the proposed algorithm is about two orders of magnitude faster than the state-of-art L1 Magic, besides giving almost similar or higher accuracy. The authors concluded that sparse decomposition problem is not computationally as tough as suggested by LP approach.

Another study published also proposed a smooth approximation L0-norm constrained affine projection algorithm (SL0-APA) to improve the convergence speed and the steady-state error of affine projection algorithm (APA) for sparse channel estimation [33]. By merging smooth approximation L0-norm (SL0) into APA cost function, the algorithm ensures improved performance in terms of the convergence speed and the steady-state error, which gives rise to a zero attractor that promotes the sparsity of the channel taps in the channel estimation and hence accelerates the convergence speed and reduces the steady-state error when the channel is sparse. The simulation results demonstrate that our proposed SL0-APA is superior to the standard APA and its sparsity-aware algorithms in terms of both the convergence speed and the steady-state behavior in a designated sparse channel. Furthermore, SL0-APA is shown to have smaller steady-state error than the previously proposed sparsity-aware algorithms when the number of nonzero taps in the sparse channel increases.

It is believed that kernel sparse method is an improvement from the normal basic pursuit of sparse representation classification. Performance of redundant representation and sparse coding against classical kernel methods for classifying histological sections are compared in the study of [34]. Since sparse algorithm has been convinced to be a higher performance method for restoration, its function has been widely used in classification [34]. However, biological and technical fluctuations have led to inherent

heterogeneity in classification histology. For instance, technical variations come from sample preparation, fixation, and staining from multiple laboratories, where biological variations are caused by issue content. Image patches are represented with invariant features at local and global scales, where local refers to responses measured with Laplacian of Gaussians, and global refers to measurements in the color space. Experiments are designed to learn dictionaries, through sparse coding, and to train classifiers through kernel methods with normal, necrotic, apoptotic, and tumor with characteristics of high cellularity. The kernel classification results are compared with two different kernel methods of support vector machine (SVM) and kernel discriminant analysis (KDA). Preliminary investigation on histological samples of Glioblastoma multiforme (GBM) indicates that kernel methods perform as good if not better than sparse coding with redundant representation.

Besides, two powerful algorithms are proposed to investigate the sparse representation on high-dimensional Hilbert space, stated in the study of [35]. By proving that all the calculations in Orthogonal Match Pursuit (OMP) are essentially inner-product combinations, the OMP algorithm is improved by implementing the kernel-trick to become Kernel OMP (KOMP). KOMP is has shorter computation time, besides providing results with higher accuracy. A rigid group-sparsity constraint was applied to KOMP, leading to a noniterative variation. The constrained cousin of KOMP, dubbed as Single-Step KOMP (S-KOMP), performs better in sparse coefficients. S-KOMP is proven to achieve an improvement (up to 2,750 times) in its performance, with almost zero loss of accuracy.

2.5. Summary

In chapter 2, a simple outline of topics of literature review was outlined in section 2.1. Results of researches on bio-acoustic recognition are discussed and reviewed in section 2.2. Besides, research regarding method used for signal processing – Mel-Frequency Cepstral Coefficient (MFCC) was also review and written in section 2.3. Lastly, classification method of Sparse Representation and list of researches and studies done using this classifier were also mentioned in section 2.4.

CHAPTER 3

PROJECT METHODOLOGY

3.1. Introduction

In this chapter, all methodologies of the process involved in the system will be described. The frog sound identification system module consists of 3 sub-modules, they are the data acquisition module, feature extraction module and lastly the classification module.

In the data acquisition module, raw frog call data will be acquired from Intelligent Biometric Research Group (IBG), School of Electrical and Electronics Engineering, Universiti Sains Malaysia (USM) in collaboration with USM School of Pharmacy. Noise reduction by using band pass filter and syllable segmentation that are used to enhance the quality of the data will be include in this module as well.

In this system, the feature extraction method that is used is MFCC to obtain the desired frog features to be processed before undergoing classification. In the feature extraction module, pre-processing of signals will be done which includes pre-emphasis, framing and windowing of signals. The pre-processing procedures will be done before MFCC.

Lastly, the classification module of this system consists of two classifier, Sparse Representation Classifier (SRC) and Kernel Sparse Representation Classifier (KSRC). The performance of both classifier will be evaluated and discussed.

As a product of this project, a Frog Species Identifier GUI will be developed using the highest performance classifier.

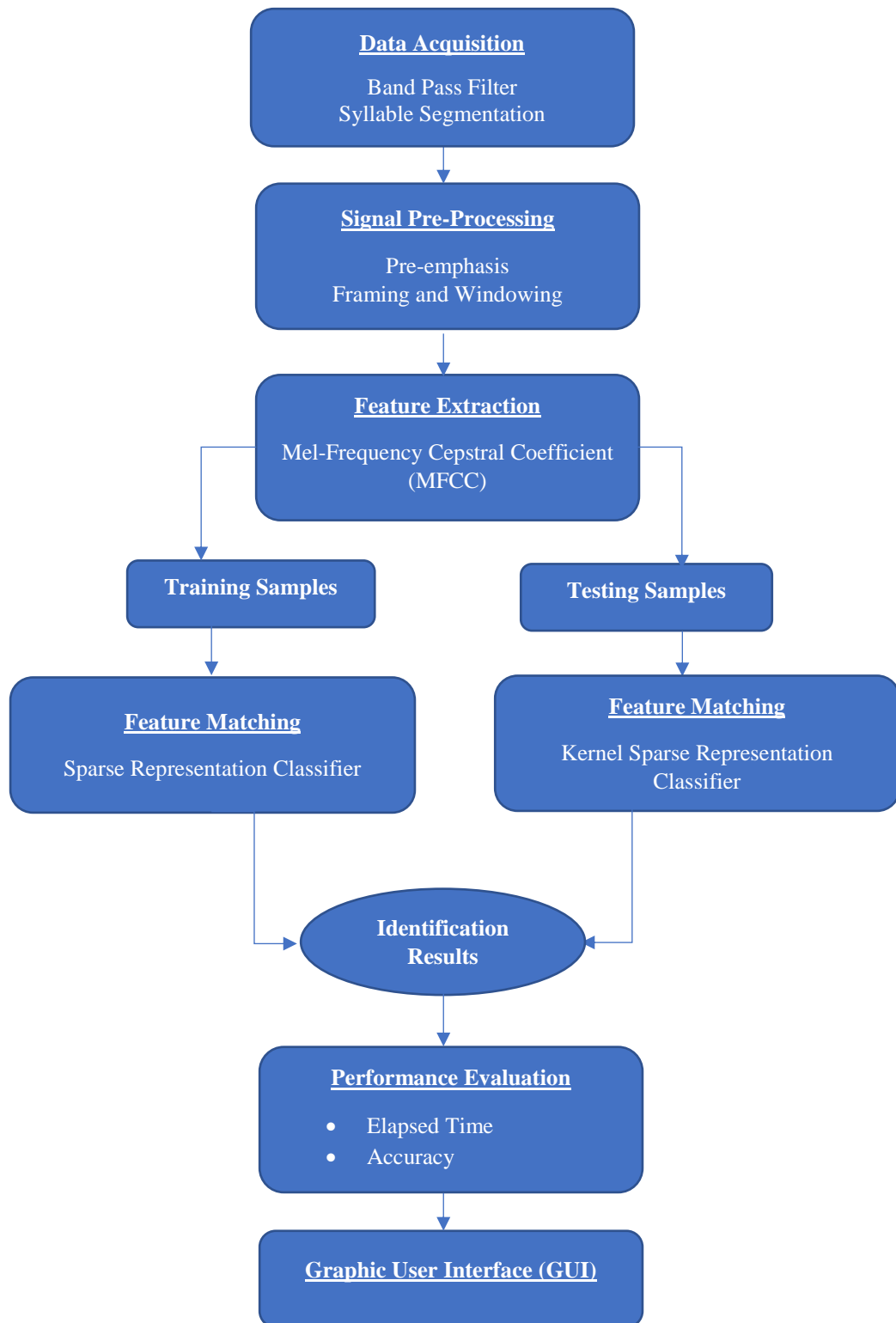Figure 3.1 below shows the overview of the system.

Figure 3. 1 Overview of frog call identification system

3.2. Data Acquisition

For this project, the raw data that have to acquire are the digital frog call samples. The samples are obtained from Intelligent Biometric Research Group (IBG), School of Electrical and Electronics Engineering, Universiti Sains Malaysia (USM) in collaboration with USM School of Pharmacy.

Based on IBG, frog calls were collected from two different segments of the forest in Kedah, Malaysia. The time frame to obtain the raw data of frog calls is between February 2012 and July 2013. The first segment of forest to collect data was occurred at Sungai Sedim, Kulim. The sounds were recorded beside a river from 8.00pm to 12.00am. On the other hand, there were also frog calls collected from Baling, Kedah at a swamp area between 6.00pm to 10.00pm.

The frog calls from the woods were recorded by using a Sony Stereo IC Recorder ICD-AX412F together with an electric condenser microphone of 32kHz sampling frequency with WAV format. The sound samples were then converted to 16-bit mono. Finally, the frog call database was formed, comprising of 15 known frog species. Their scientific names, common names and images are shown in Table 3.1 as follows.

Table 3. 1 Frog Species in database

Family, scientific name, common name and image		
Microhylidae *Hylarana glandulosa* Rough sided frog 	Rhacophoridae *Polypedates leucomystax* Common tree frog 	Microhylidae *Microhyla heymonsi* Taiwan rice frog
Bufonidae *Phrynoidis aspera* River toad 	Microhylidae *Kaloula baleata* Flower pot toad 	Dicroglossidae *Fejervarya limnocharis* Grass frog
Microhylidae *Kaloula pulchra* Asian painted bullfrog 	Rhacophoridae *Philautus mjobergi* Bubble-nest frog 	Ranidae *Hylarana labialis* White-lipped frog

Table 3.1 Continued

Ranidae	Bufonidae	Bufonidae
Odorrana hosii	*Duttaphrynus melanostictus*	*Genus ansonia*
Poisonous rock frog	Black-spectacled toad	Stream toad
Rhacophoridae	Microhylidae	Rhacophoridae
Philautus petersi	*Microhyla butleri*	*Rhacophorus appendiculatus*
Kerangas bush frog	Painted chorus frog	Frilled tree frog