UNIVERSITI PUTRA MALAYSIA

DESIGN OF INTELLIGENT QIRA'AT IDENTIFICATION ALGORITHM

NORAZIAHTULHIDAYU BINTI KAMARUDIN

FK 2018 100

DESIGN OF INTELLIGENT QIRA'AT IDENTIFICATION ALGORITHM

By

NORAZIAHTULHIDAYU BINTI KAMARUDIN

Thesis submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfillment of the Requirements for the Degree of Doctor of Philosophy

December 2017

DEDICATION

Firstly, everything would not be possible without the support of the higher education ministry of Malaysia(MOE) for the financial support of my PhD study in Universiti Putra Malaysia(UPM)so, great thanks to both the MOE and UPM.

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirement for the degree of Doctor of Philosophy

DESIGN OF INTELLIGENT QIRA'AT IDENTIFICATION ALGORITHM

By

NORAZIAHTULHIDAYU BINTI KAMARUDIN

December 2017

Chairman : Associate Professor Syed Abdul Rahman Al Haddad, PhD
Faculty : Engineering

The speaker's native dialect, accents and the socioeconomic background are few factors that influence the speaking style. The mixing of Qira'at is considered forbidden in Islam, especially during salat prayer. The identification of threats that could influence the accuracy of voice recognition and influence decisions in recitation recognition performance of accents recognition. On the other hand, only few studies focus on research of the performance factor or accuracy in the reverberant environment and none yet focusing on the factors that would affect Qira'at speech signals and identification.

The main objective of this thesis is to propose the identification process of Quranic recitation but oriented to the identification of various Qira'at with the emphasis on recognition without being affected by echo or noise during live recitation or in recordings. Sequential Windowing Parameterizing of Affine Projection Algorithm (SPAP) is proposed to improve windowing parameterizing during echo cancellation, while recognition accuracy factors are taken into account for further improvement. The process of the SPAP Algorithm is to extend parameters of the Affine Projection Block with two different selections of windowing length that affect the final accuracy on pattern classification.

The usage of Feature Selection (FS) contributes to simplify and enhance the quality of the dataset used by selecting significant features. Qira'at audio files for Surah Ad-Dhuha are used in this study to re-sample an audio sample. Clean audio signals from AEC are used with proposed feature selection technique called as X-Ant Colony Optimization, that utilizes the concept of Ant Colony Optimization, and can enhance feature extraction. For the feature vectors that are collected from feature extraction (MFCC) and feature selection (X-ACO), the feature vectors are used as input for the classification phase. A combination of Principal Component Analysis (PPCA) and Gaussian Mixture Model (GMM) is proposedly in used for the classification phase as it is able to reduce any redundancy from the latent variables and carries only the most important information through dispersion of entropy.

To evaluate the algorithm, 350 samples for 10 types of Qira'at recitation are in used, and for justifying the best pattern classification, few algorithms are tested in the early preliminary evaluation with K-Nearest Neighbour, GMM and PPCA. And in the final evaluation for PPCA, it achieved high accuracy with 95.15%, while WER and EER are around 7.63%. The current evaluation for SPAP tests another echo database of a Lecture Room that presents a reduction in the accuracy rate of around 92.1% while the WER and EER are around 7.53%. But both of the results are significant compared to achieve results for MFCC without SPAP feature selection technique that just acquired 89.1% in early preliminary test. It proves that the current proposed algorithm achieves better results in Echo Greathall comparable to Echo Lecture Room and finally, these results will be used as foundation for any upcoming related research that may improve the understanding of Qira'at among the Muslim.

Abstrak tesis dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

REKABENTUK ALGORITMA KECERDIKAN IDENTIFIKASI QIRA'AT

Oleh

NORAZIAHTULHIDAYU BINTI KAMARUDIN

Disember 2017

Pengerusi : Profesor Madya Syed Abdul Rahman Al-Haddad, PhD
Fakulti : Kejuruteraaan

Dialek bahasa asli, cara percakapan dan latar belakang sosioekonomi adalah beberapa faktor yang mempengaruhi gaya bercakap. Pencampuran Qira'at adalah dilarang dalam Islam, terutamanya semasa pembacaan sewaktu solat. Pengenalpastian ancaman yang boleh mempengaruhi ketepatan pengenalan suara dan keputusan mempengaruhi pengenalan bacaan itu sendiri belum lagi dibincangkan secara meluas, terutamanya dengan pelbagai kesan yang berlaku dalam bunyi bising dan kekeliruan latar belakang, prestasi pengecaman cara percakapan dalam persekitaran yang mempunyai bunyi bising dan gema akan menurunkan kualiti audio secara dramatik juga. Namun begitu, hanya sedikit kajian yang menumpukan pada penyelidikan faktor prestasi atau ketepatan dalam persekitaran gema dan masih lagi belum ada yang mengkaji faktor-faktor tersebut pada kesannya di dalam sampel Qira'at dan pengelasannya.

Tujuan utama tesis ini adalah untuk mencadangkan proses pengenalan bacaan Al-Quran tetapi berorientasikan identifikasi pelbagai Qira'at dengan penekanan atas pengiktirafan tanpa dipengaruhi oleh gema atau bunyi semasa bacaan secara langsung atau dalam rakaman. Melalui teknik Turutan Tetingkap Parameter melalui teknik Affine Projection (SPAP) dicadangkan untuk meningkatkan tetingkap parameter semasa pembatalan gema, manakala faktor ketepatan pengiktirafan diambil kira untuk penambahbaikan selanjutnya. Proses Algoritma SPAP adalah untuk mendapatkan parameter Blok Perangkaan Affin dengan dua pilihan panjang tetingkap yang mempengaruhi ketepatan corak pengelasan akhir identifikasi.

Penggunaan Pemilihan Ciri (FS) menyumbang untuk memudahkan dan meningkatkan kualiti dataset yang digunakan dengan memilih ciri-ciri penting. Fail-fail audio Qira'at untuk Surah Ad-Dhuha digunakan dalam kajian ini untuk menyusun semula sampel

audio. Isyarat audio bersih dari AEC digunakan dengan teknik pemilihan ciri yang dicadangkan dipanggil sebagai Pengoptimuman Koloni X-Semut, yang menggunakan konsep Pengoptimuman Koloni Semut, dan dapat meningkatkan pengekstrakan ciri.Untuk vektor ciri yang dikumpulkan dari pengekstrakan ciri (MFCC) dan pemilihan ciri (X-ACO), vektor ciri digunakan sebagai input untuk fasa klasifikasi. Gabungan Analisis Komponen Utama (PPCA) dan *Gaussian Mixture Model* dicadangkan untuk digunakan untuk fasa klasifikasi kerana ia dapat mengurangkan sebarang redundansi daripada pembolehubah *laten* dan hanya membawa maklumat yang paling penting melalui penyebaran entropi.

Untuk menilai algoritma, 350 sampel untuk 10 jenis bacaan Qira'at digunakan, dan untuk membenarkan klasifikasi corak yang terbaik, beberapa algoritma diuji pada awal penilaian awal dengan *K-Nearest Neighbor*, GMM dan PPCA. Dan dalam penilaian akhir untuk PPCA, ia mencapai ketepatan yang tinggi dengan 95.15%, manakala WER dan EER adalah sekitar 7.63%. Penilaian semasa untuk SPAP menguji satu lagi pangkalan data gema Bilik Kuliah yang menunjukkan pengurangan kadar ketepatan sekitar 92.1% manakala WER dan EER adalah sekitar 7.53%. Tetapi kedua-dua keputusan adalah signifikan berbanding dengan mencapai keputusan untuk MFCC tanpa teknik pemilihan ciri SPAP yang hanya memperoleh 89.1% pada awal ujian awal. Ia membuktikan bahawa algoritma yang dicadangkan semasa mencapai hasil yang lebih baik dalam Echo Greathall setanding dengan Bilik Kuliah Echo dan akhirnya, keputusan ini akan digunakan sebagai asas bagi sebarang penyelidikan berkaitan yang akan meningkatkan pemahaman Qira'at di kalangan umat Islam.

ACKNOWLEDGEMENTS

In the Name of Allah, the Most Merciful, the Most Compassionate all praises to Allah swt, the Lord of the world;and prayers and peace be upon Mohamed His servant and messenger. First and foremost, I must acknowledge my limitless thanks to Allah, the Ever Magnificent; the Ever Thankful, for His help and blessings. I am totally sure that this work would have never come to the fore, without His guidance. Sincere gratitude goes to my advisor Assoc. Prof. Dr. Syed Abdul Rahman Al Haddad for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study. Furthermore, I would like to thank the rest of the thesis committee: Assoc. Prof. Shaiful Jahari Hashim, and Assoc Prof. Dr. Abd Rauf Hassan for their insightful comments and encouragement, but also for the hard questions which incentivised me to widen my research from various perspectives and also my fellow examiners during my PhD verification and viva session whom giving their helpful comments and suggestions.

Last but not least, I give sincere appreciation to my fellow labmates (Dr. Asem Khmag, Nogol Memari, Dr. Merhdad Moghbel) for the stimulating discussions, for the sleepless nights working together before deadlines, and for all the fun we have had in the last four years. I really do appreciate all the time that we had together and all the pain that we suffered too, this will always be in our memories. Also I thank all my nice colleagues within institution (Dr. Faraneh Zarafshan, Dr Sura, Abdul Mawla, and Dr. Maryam Hejazi). My fellow friends;Fizah and Husband, Doctorate Support Groups,sis Khyrina,Bro Samir, Riki, Ija, Maryam, Salziana, Saleha and Noraini and all my colleague that I may not able to include here too.

The deepest appreciation goes to my family: my parents(Hj. Kamarudin Sanusi and Hjh. Hamidah Abdul Malik) and to my late brother Zainal Fitri Kamarudin and sister in law;Norhana Azlin Binti Ab. Hamid, Mohd Farid Adam, my lovely kids (Putri & Ateh), my aunty Hasnah Zakaria and her families, and my beloved cousin whom always be there throughout my journey and concern on my parents matter while I'm away, making generous doa and for supporting me spiritually throughout writing this thesis and my life in general.

I would like to convey my appreciation to our fellow lab technicians En.Fathullah Hakim and En Azlan, thanks for providing a perfect laboratory environment, great support and assistance and fellow UTeM staff and collegues. I would like to present my special appreciations to Universiti Putra Malaysia for providing a beautiful, peaceful and calm academic environment for research and study.

Noraziahtulhidayu Kamarudin

I certify that a Thesis Examination Committee has met on 8 December 2017 to conduct the final examination of Noraziahtulhidayu binti Kamarudin on her thesis entitled "Design of Intelligent Qira'at Identification Algorithm" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Thesis Examination Committee were as follows:

Sharifah Mumtazah Syed Ahmad Abdul Rahman, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Wan Azizun binti Wan Adnan, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Internal Examiner)

Md Nasir bin Sulaiman, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

Akhtar Kalam, PhD
Professor
Victoria University
Australia
(External Examiner)

NOR AINI AB. SHUKOR, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 26 April 2018

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfillment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Syed Abdul Rahman bin Syed Mohammad, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Shaiful Jahari bin Hashim, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Abd. Rauf Hassan, PhD
Associate Professor
Faculty of Modern Language
Universiti Putra Malaysia
(Member)

____

ROBIAH BINTI YUNUS, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: ________________________ Date: ______________________

Name and Matric No : Noraziahtulhidayu binti Kamarudin, GS37053

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: _____________________________________
Name of Chairman of Supervisory Committee: Associate Professor Dr. Syed Abdul Rahman bin Syed Mohammad

Signature: _____________________________________
Name of Member of Supervisory Committee: Associate Professor Dr. Shaiful Jahari bin Hashim

Signature: _____________________________________
Name of Member of Supervisory Committee: Associate Professor Dr. Abd. Rauf Hassan

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

LIST OF ABBREVIATIONS

AC	Accents Conversion
AEC	Acoustic Echo Cancellation
AP	Affine Projection(AP)
ALGASD	ALGerian Arabic Speech Database
ADC	analog to digital conversion
ACO	Ant Colony Optimization
ANN	Artificial Neural Network
ASR	Automatic Speech Recognition
DFT	Discrete Fourier Transform
DTW	Dynamic Time Warping
ERLE	Echo Return Loss Enhancement
EER	Equal Error Rate
FFT	Fast Fourier Transform
FRLS	Fast Recursive Least Squares
GMM	Gaussian Mixture Model
HMM	Hidden Markov Model
LI	Language Identification
LMS	Least Mean Squares
LS	Least Squares
LDA	Linear Discriminant Analysis
LP	Linear Predictive
LPCC	Linear Predictive Cepstral Coefficients
LPC	Linear Predictive Coding
MLLR	Maximum Likelihood Linear Regression
MFCC	Mel Frequency Cepstral Coefficients
MSA	Modern Standard Arabic
MSE	Mean Square Error
MLP	Multilayer Perceptron
NLMS	Normalized Least Mean Square
PLP	Perceptual Linear Prediction
PSD	Power Spectral Density
PSNR	Peak to Signal Noise Ratio
PCA	Principal Component Analysis
PPCA	Probabilities Principal Component Analysis
RANSAC	RANdom Sample And Consensus
RASTA	Relative Spectra Filtering
RNN	Recurrent Neural Network (RNN)
RLS	Recursive Least Squares (RLS)
RASTA-PLP	Relative Spectral Transform - Perceptual Linear Prediction
RIR	Room Impulse Response (RIR)
SVM	Support Vector Machine (SVM)
VQ	Vector Quantization (VQ)
WER	Word Error Rate (WER)
ZCR	Zero Crossing Rate(ZCR)

والضحى	Wadh-dhuha
سجى	Saja
فترضى	Fa-tardha
فاوى	Fa-aawa
فهدى	Fa-hada
فاغنى	Fa-aghna

Imam Abu Amru al Basri	Al Douri from Abu Amru Al Susi Abu Amru
Imam Asim al Kufi	Hafs an Asim Shoba from Asim
al Kisai Al Kufi	Abu al Hareth al Kisaie Ad-Duri al Kisaie
Imam Nafi' al Madani	Warsh, Qaloon
Abu Ja'far al Madani	Wardan, Ibn Jammaz
Hamzah al-Kufi	Khallad from Al Hamzah Khallaf from al Hamzah
Ibnu Kathir al-Makki	Qunbul Ibnu Katheer, Albazzi Ibnu Kathir
Khalaf al-bazar	Edrees Al Haddad, Khalf Ishak Al Bazaar
Imam Ibnu Amir As-Syami	Hesham from Ibn Aamer, Ibnzakwan
Imam Yaaqub Al Basri	Rauh from Yaacob, Ruwais from Yaacob

CHAPTER 1

INTRODUCTION

1.1 Overview

The understanding of accents relates to a person's language while dialects mostly apply to the varieties of vocabularies, idioms, grammars and pronunciation of a specific language within certain community, often due to geographical differences. Accents for a selected language vary between speakers (Han-ping Shen, et al,2015), but they are still understandable within the same scope of speakers. But, if it is applied within speech procesings field; it is a requirement to have certain standard of accent identification;in terms of improving the execution of speech, speaker recognition systems, human machine interaction and developing regional dialect spoken query systems(Borah and Sharma, 2016). This will create difficulties to model input speech for the development of speaker-independent systems.Factors of rate recognition accuracy is affected due to the spectrum chanes for Quranic signals that prones to additive noise or echoes from recordings.

Mixing different types of Qira'at recitation considers forbidden in Islam, especially during Solat. As Quran is written in Arabic Language and Arabic is one of the languages that are often described as morphologically complex and the problem of language for Arabic are multipart by the variation of dialectal (Almeman K, 2015;N. Zainon et al,2012 ;Noor Jamaliah,2011).

Al Quran is the Holy Scripture given by Allah swt to Muslims worldwide and recited with the Arabic Language, the only language it was revealed in. All Muslims believe that the Quran's content is neither corrupted nor changed and hence the value, and the original text is protected and maintained. Therefore, it is forbidden to recite Holy Quran in other languages except Arabic with neither additions nor subtractions(Mohammed, Shahrizal, and Salam, 2015).

Although recited the Quran is not used in communication, it is extremely important to teach the pronunciation of classical Arabic sounds (language of the holy Quran). Moreover, teaching how to pronounce Quranic sounds is indispensable to correctly read the holy Quran, either for memorizing it or for reciting it in Islamic worship, such as prayers(Elhadj, Alghamdi, and Alkanhal, 2013). Basically, there are 10 differentf Qira'at recitations (7 types known as Mutawatirrah and 3 Mashuur)(Rauf,2014), which use the same Mushaf Uthmani and the differences between those Qira'at happened due to the location of where the Holy Quran was revealed (eg; Nafi' IbnKathir, Abu 'Amr ibn al-'Ala', Ibn 'amir, 'Asim, Hamza and al-Kisa'i). The sensitivity of recitation and pronunciation of the same vowel and different words by a particular speaker may affect the sensitivity or accuracy of automatic identification for language accents and this may happen with Quranic recitations too. Figure 1.0 shows the coverage of different Qiraat based on geographical location. The most dominant Qira'at is Hafs An Asim, which is widely recited by most Muslims worldwide, then followed by the Qira'aat of Warsh, Ad-Douri and Qaloon.

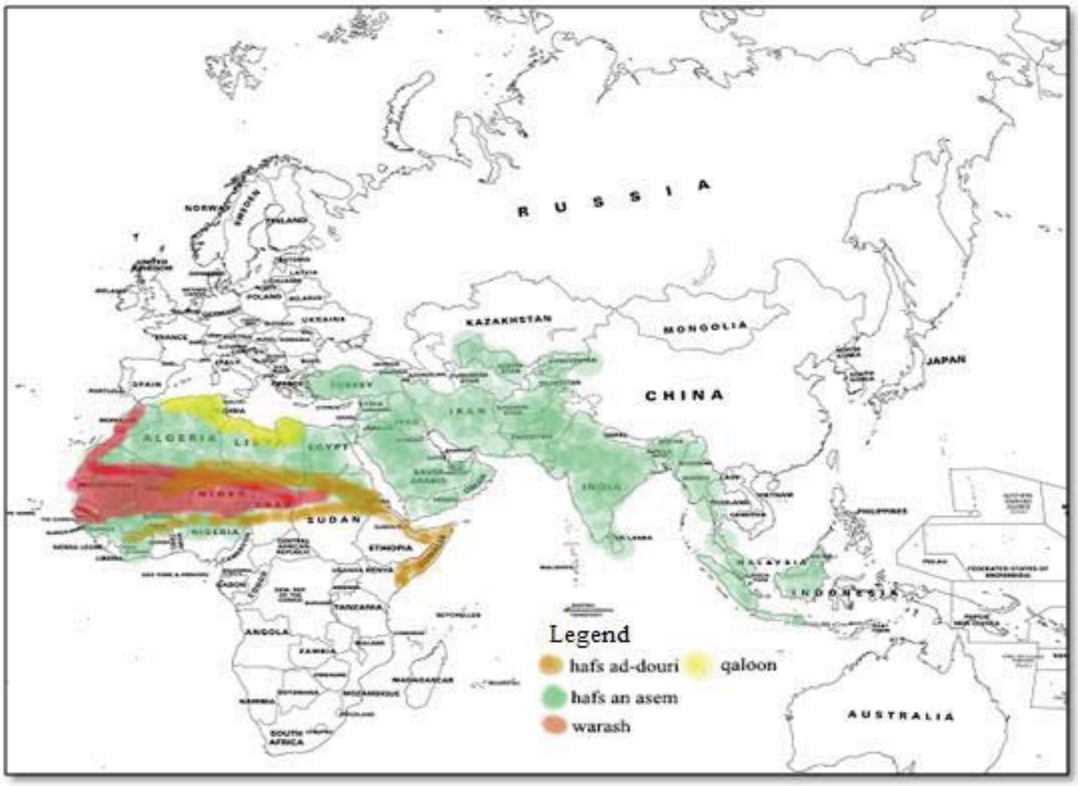

Figure 1.0 : World map with different Qira'at Recitations
(Ibnamin,2017)

As the Holy Quran was revealed to Muslims in the Arabic Language, it is an obligation for Muslims to learn and mastered it. Arabic is the official dialect for more than 22 nations. This study, the focus is to identify the Qira'at, which therefore would help to identify the Qira'at signals based on the stored recorded signal. Throughout this study, some preliminary discovery has been done on related scholars that focus on similar topics like Tajweed Automatic Identification, Language with Accents Automatic Identification and Music Genre Identification. The previous scholars have successfully created an application, which is able to detect errors in recitation by using Mel Frequency Cepstral Coefficients(MFCC) for feature extraction and also Hidden Markov Model (HMM)as pattern classification for identifying error in Quranic Recitation(Tajweed) especially during recitation. The best results achieved by them were within 86% to 92 % for both different researchers (Noor Jamaliah et al,2013;Elhadj et al, 2014). The rate of recognition accuracy is considered to be less due to the spectrum of the Quranic signals, which may be prone to additive noise or echoes from the recordings. In any case, the latest speech databases have representation of distinctive levels and range from the phonemes to sounds.

Furthermore, it will interpret the speech; and the interpretation is adjusted to the speech acoustic signal(Auran et al, 2004). But the process of transcribing Quran verse surely will take a long time and should be done carefully as it involves a Holy Surah. Therefore, this research embarks new process in identifying the Qira'at only not to be used for a specific Surah but can be applied to the whole Holy Quran. Errors can be missing words, verse, misreading Harakat (articulation, accentuations and diacriticization) (WaqarMirza et al, 2010).

From Waqar Mirza Muhammad et al (2010), E-Hafiz system is developed for answering these problems that happen due to noise in the environment for detecting speech recognition while reciting a Surah. The application is produced using Mel-Frequency Cepstral Coefficient (MFCC) strategy to concentrate voice features from Quranic verse recitation and maps them with the information gathered amid the preparation stage. Any dissimilarity of Tajweed recognition is pointed out for the recitors. Testing results of short verses of the Quran utilizing the E-Hafiz framework are exceptionally promising as what's being done by the researcher of the E - Hafiz system (Waqar Mirza Muhammad et al, 2010). For those Quranic research done, they have used multiple wave files based on collected databases, which varies by gender and age. In this specific research of Qira'at identification, only expert scholars shall be collected for the Qira'at database in terms of feature extraction and pattern classification purpose.

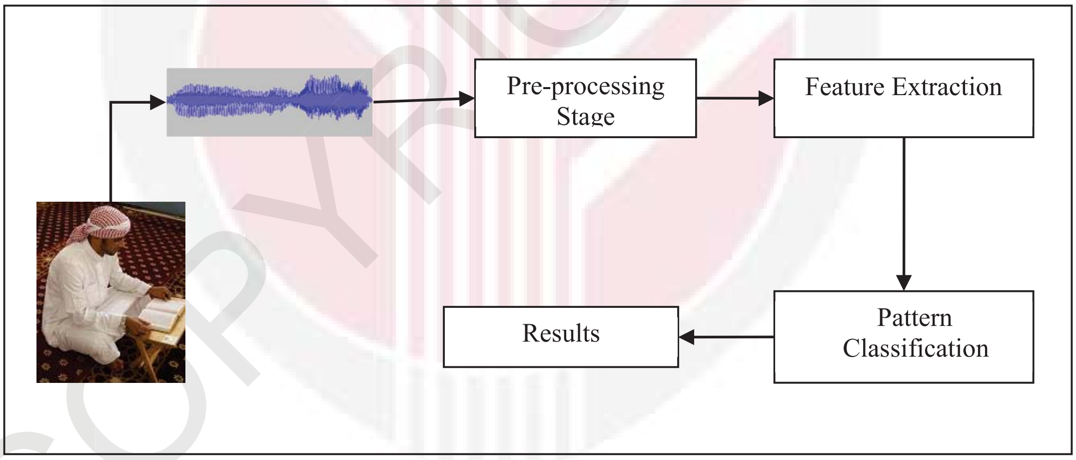

Figure 1.1 : Basic flow of Qiraat Identification

Quranic signal is an audio signal under study to investigate the Qira'at classification. Therefore, Quranic signals are closely related to music signals in terms of pitch and skewness. For the targeted Surah of Ad-Dhuha, which also contains *imalah, taqlil, naqal* and *badal*; these attributes contribute to the parameters of Qira'at identification. The interruption of echo that always occur in pre-processing stages as shown in Figure 1.1, may reduce the accuracy rate achieve during recording, or in the live environment. Therefore, methods of applying suitable adaptive algorithms are investigated with the robust feature selection algorithms that can give a higher value of accuracy during the feature extraction process. The processes focused in this study will involve the main stages of Pre-Processing and Feature Extraction, while the whole processes of identification will still cover till pattern classification.

Figure 1.2 illustrates the accuracy results achieved by different implementation of research done by previous researchers, those implementations vary with the accuracy of Arabic speech recognition, automatic error (Tajweed) identification for Quranic recitation, songs and singer voice classification. The investigation done by those researchers was based on identifying on the accuracies factor and Word Error Rate (WER) and the results were presented in percentage values. Among that research the highest accuracy of 94.82% on songs classification accuracy was done by Ghosal et al (2012). Feature extraction techniques used were solely using MFCC for the feature vector while wavelet was used in songs identification.

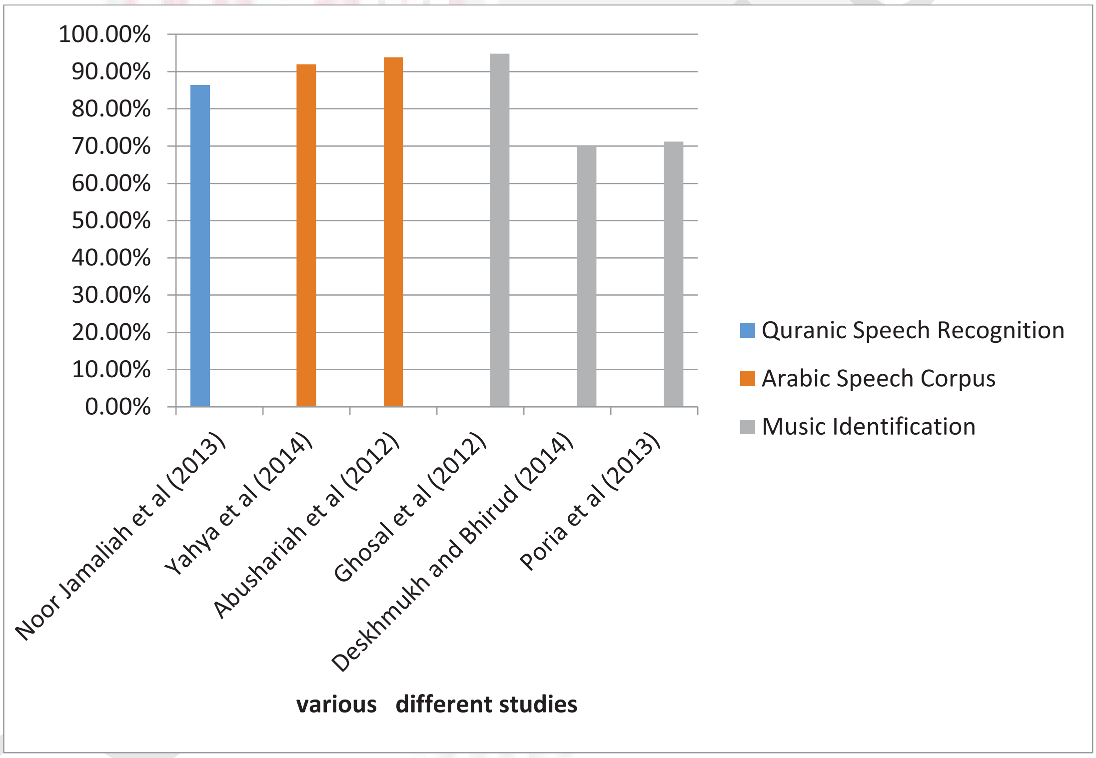

Figure 1.2 : Accuracies achieved by Music Identification, Arabic Speech and Quranic Speech

Throughout this studies, preliminary test are already done to verify the suitable algorithms to be used for the current research as presented in Literature Review.

1.2 Problem Statement

In solving the problems of Qira'at identification and to avoid from wrong recitation during salat; this study of Qira'at identification considered new in the current research area. It is important to understand the correct concept and flows of identification that may involved by first extracting the significant cleaned signal as an input for feature vector while the second one is the deep emphasis on the feature selection process.

Acoustic noise from ambience and sound systems distorted the real signals. It happens from frequency response and masking produces unusually large peaks as a result music recording are prone to distraction and the identification would be a problem. Noise is omnipresent in almost all acoustical environments(Utpal et al, 2016;Hom K,2013; Gil-Cacho, J,2013). Amongst popular solution are Least Mean Square(Zhang et al, 2015;Deepika M,2013;Longbiao et al,2012) and Recursive Least Square(Rakesh P, Kumar T, 2015;Adapa N,Bollu S,2012;PSR Diniz, 1984). But the techniques used to have slow convergence rate(Gonzalez A et al, 2012), high computational complexity taken (Roshahliza et al, 2012;Ranbeer T, Dheeraj A,2012; Nielsen JK, Jensen SH, 2012)based on Mean Square Error (Weiyang Cai,2013;Nurul et al,2013, Arablouei R, Do K,2012)and Peak to Signal Noise Ratio(Verma A, Verma N,2016;Santosh D et al,2012;Kun Han et al,2014).

Interruptions by echoes from sound systems, their frequency response and masking produces unusually large peaks. Noise appears in music recordings such as country music; and few studies with different individual have shown that they may be prone to distraction and for gender difference especially during indexing process(Jing, et al, 2012), and this may happen also to Quranic recordings (Adapa and Bollu, 2012; Razak et al,2008). The performance of the accuracy prediction model are severely influenced by noise sources from a computer, Air Conditioner (AC) duct, fan etc and always depends on number model data, noisiness and target classes(Julian Teryzk, 2014; Lerch A,2012). While, for hands free telephony inside a car, the noise source could be from car engine noise, outside vehicle and ambient wind. It is even possible that surrounding noise sources are captured at a higher level than the speech signal (Ganesan and Manoharan, 2015). The detection may drop for some of Qira'at due to appearance of noise/echo, the identification shall be robust enough to face possible environment variability and shall be able to achieve a high accuracy for short input speech samples (Safavi, 2015).

Secondly, in finding significant features vector it has to find the best feature coefficient frame selections and also the proper technique of feature extraction. The results later can be used in pattern classification for Qira'at identification but on the other hand it may results over fitting of training irrelevant data (Tabakhi, et al, 2014; Rashno A et al,2013;Unler et al,2011).

Thirdly, the identification shall be robust enough to face possible, environment variability and shall be able to achieve a high accuracy using selected pattern classification (Safavi, 2015; Gholiput A et al,2012; Rao KS, Koolagudi SG,2013) of Probabilistic Principal Component Analysis(Xiao-chun L et. al,2012; Xiaodong Cui MA and BZ,2012; Ismail MNI, Muse ME-HM,2014).

1.3 Objectives

The main objectives of this research is to propose an intelligent Qira'at identification algorithm. The specific objectives are as follow;

a. To design dynamical sequential windowing selection for Acoustic Echo Cancellation using Affine Projection (SPAP) from the recording sound on Qira'at speech signals, within best parameters and experimental setup.
b. To optimize new feature selection using X-Ant Colony Optimization(X-ACO) on current feature extraction technique with Mel Frequency Cepstral Coefficients(MFCC),
c. To design Qira'at Identification from speech signals using Probabilistic Principal Component Analysis and Gaussian Mixture Model.

1.4 Thesis Scope

This research presents the analyzation of Qira'at Identification based on digital speech audio recordings to improve the recognition perfomance and verification on the accents classifications and finally would solve the problems of Qira'at identification. The perfomance of recognition for accuracies is easily degraded when noise from surroundings and echoes (reverberant) corrupt the live or recorded Qira'at speech. Time is wasted to find the significant data for feature vectors to be used in the classification phase.

In this study, the effectiveness of applying accents identification is investigated for a recorded sample database of Quranic speech recitation. For the accuracy factor of the current investigation, three methods of performance. First using the technique of accuracy in percentage values, Word Error Rate and Equal Error Rate. On the segmentation part, current studies use 16kbit with Surah Ad-Dhuha verse segmentation that is done by manually for 10 Qira'at with a variety of recitation from selected trusted scholars with 2 faces of Qira'at recitation that by total covers 20 types of Qira'at. There are 350 speech segmented verses collected for Quranic speech audio, Sequential Parameterization Affine Projection for AEC. For feature extraction Mel Frequency Cepstral Coefficient (MFCC) together with X-Ant Colony Optimisation (X-ACO) on feature selection and Probabilities Principal Component Analysis (PPCA) is used for pattern classification. The effects of echo or reverberation in this study are also investigated using Aachen Impulse Response Database and the metrics that are used in evaluating for the whole signals are Mean Square Error and Peak to

Signal Noise Ratio. The data used are from Islamway.net since the audio samples are from trusted scholar (Mohamed, Hassanin, Taher, and Othman, 2014).

Sample data process based on 10 Qira'at includes; Riwayat Imam Abu Amru Al Basri (Al Douri from Abu Amru and Al Susi Abu Amru), Riwayat Imam Asim al Kufi (Hafs an Asim and Shoba from Asim), Riwayat al Kisai Al Kufi (Abu al Hareth al Kisaie, Ad-Duri al Kisaie), Riwayat Imam Nafi' al Madani (Warsh, Qaloon), Riwayat Abu Ja'far al Madani (Wardan, Ibn Jammaz), Riwayat Hamzah al-Kufi (Khallad from Al Hamzah, Khallaf from al Hamzah), Riwayat Ibnu Kathir al-Makki (Qunbul Ibnu Katheer, Albazzi Ibnu Kathir), Riwayat Khalaf al-bazar (Edrees Al Haddad, Khalf Ishak Al Bazaar), Riwayat Imam Ibnu Amir As-Syami (Hesham from Ibn Aamer, Ibn Zakwan) and Riwayat Imam Yaaqub Al Basri (Rauh from Yaacob, Ruwais from Yaacob).

And for this research, Surah Ad- Dhuha is in used for Quranic samples and only male offline recordings are collected and audio database are collected from Islamway.web while echo database from Aachen Impulse Response database.

1.5 Thesis Outlines

This thesis comprises of 5 chapters as follows:

Chapter 1 present the general overview of the research, which reflects on the significance of the study, problem statement, research objective, aim, and the background of study. This chapter gives general views of the continuous topics that are covered in the upcoming chapter.

In Literature Review for Chapter 2, the chapter explains a comprehensive understanding of recent achievement of previous research which relates to Quranic speech signals, the techniques and algorithms used for the feature extraction, and pattern classifications in identifying Quranic verse especially, performance of other research related, topics and issues, which arises for the automatic identification that are currently available. The issues of acoustic echo cancellation and parameters involved and the requirement of feature selection algorithms and the important parameters for feature extraction are also discussed in this chapter.

For Research Methodology in Chapter 3, the first section describes and discusses algorithms and all methods, including techniques and tools, which are involved in the sampling and pre-processing (basic feature extraction and pattern classification). For the second section an overview on sequential windowing length parameterization for acoustic echo cancellation is presented and the third section focuses on feature selection algorithms for the Quranic speech signals accents identification. Any parameters and metrics for evaluation, which supports this study are also investigated.

Furthermore, the Word Error Rates, Accuracies, Equal Error Rates given for recorded recitations from total signals are also evaluated in this study.

In Chapter 4, the overall results, which are implemented as described in Research Methodology, are presented. Therefore, the effectiveness based on accuracy achievement and Word Error Rate (WER) is evaluated too in this chapter. Overall discussion on the proposed algorithm is discussed in this chapter. The review results for a recorded recitation, is also covered within this chapter.

The final chapter 5, discusses the whole contributions of the currents studies, which includes; conclusions and a proposals for future work. In this chapter too, the trade off of this accents identification techniques with various language also presented together with music genre identification.

REFERENCES

A.K. Verma and N. Verma,.(2016). A Comparative Performance Analysis of Wavelets in Denoising of Speech Signals Int. J. Comput. Appl.

A.Lerch. (2012). Audio Content Analysis: Applications in Signal Processing and Music Informatics. John Wiley & Sons.

Mousa, M. Qados, and S. Bader. (2011). Adaptive Noise Cancellation Algorithms Sensitivity to Parameters in 2011 International Conference on Multimedia Computing and Systems (ICMCS), pp. 1–4.

Abd. Rauf Hassan,(2014) Rahsia Qiraat Mutawatirah Majalah Al- Ustaz Isu No-21, 2014,pp-96-97

Abushariah MAM, Ainon RN, Zainuddin R.(2012). Impact of a Newly Developed Modern Standard Arabic Speech Corpus on Implementing and Evaluating. pp:1–12.

Adami AG, Hermansky H.(2003). Segmentation of Speech for Speaker and Language Recognition.

Adapa NS, Bollu S.(2012). Performance Analysis of different Adaptive Algorithms based on Acoustic Echo Cancellation. Blekinge Institute of Technology, 371 79 Karlskrona Sweden; 2012.

Ahmad AM, Ismail S, Samaonl DF, Fay T.(2004). Recurrent Neural Network with Backpropagation through Time for Speech Recognition.

Ahmad Khan AF, Mourad O, Dahan HBAM, Abushariah M a. M. (2013),Automatic Arabic pronunciation scoring for computer aided language learning. 1st Int Conf Commun Signal Process their Appl [Internet]. Ieee; 2013 Feb;(1):1–6.

Ahsiah I, Noor NM, Idris MYI. (2013).Tajweed checking system to support recitation. 2013 Int Conf Adv Comput Sci Inf Syst [Internet]. Ieee; 2013 Sep;189–93.

Almeman KA. (2015).Reducing Out-of-Vocabulary in Morphology to Improve the Accuracy in Arabic Dialects Speech Recognition. 2015.

Al-Haddad S.A.R., Samad S.A., Hussain A. KAI .(2008) Isolated Malay Digit Recognition Using Pattern Recognition Fusion of Dynamic Time Warping and Hidden Markov Models. Am J Appl Sci 5 714-720, 2008. 2008;5(6):714–20.

Amudha V, Venkataramani B, Vinoth R, Ravishankar S.(2007). SOC Implementation of HMM Based Speaker Independent Isolated Digit Recognition System. In: 20th International Conference on VLSI Design, 2007. 2007. p. 1–6.

Arablouei, R., and Do, K. (2012). Affine projection algorithm with selective projections. An International Journal Signal Processing, 92, 2253–2263. http://doi.org/10.1016/j.sigpro.2012.02.018

Asem Khmag (2016). Denoising of Digital Images using Second Generation Wavelet Transforms-Hidden Markov Model, PhD theses, Universiti Putra Malaysia.

Ashikin, N., Kadir, A., Sudirman, R., Mahmood, N. H., & Ahmad, A. H. (2013). Jurnal Teknologi Cascade-forward Neural Networks for Arabic Phonemes Based on k-Fold Cross Validation, 2(1), 13–18.

Attarian, A., Danis, G., Gronsbell, J., Iervolino, G., & Member, H. T. (2013). A Comparison of Feature Selection and Classification Algorithms in Identifying Baseball Pitches. In International MultiConference of Engineers and Computer Scientists (Vol. I, pp. 263–268).

Auran, Cyril, Caroline Bouzon and Daniel Hirst. (2004).The Aix- MARSEC project: an evolutive database of spoken British English. International Conference: Speech Prosody 2004. Nara, Japan.

Balen J Van. (2011).Automatic Recognition of Samples in Musical Audio. 2011.

Banos, O., Galvez, J.-M., Damas, M., Pomares, H., & Rojas, I. (2014). Window size impact in human activity recognition. Sensors (Basel, Switzerland), 14(4), 6474–99.

Barbosa, J. D. R. (2012). DATORRO'S Natural Reverberator (Vol. 311275087).

Bhattacharjee, U., & Sharma, R. (2016). A Statistical Analysis on the Impact of Noise on MFCC Features for Speech Recognition.

Biadsy F. (2011) Automatic Dialect and Accent Recognition and its,"

Bobbert D., Wolska M. (2007) .11th Workshop on the Semantics and Pragmatics of Dialogue, 159-160.Trento, Italy.

Borah, J., & Sharma, U. (2016). Automatic Identification of the Dialects of Assamese Language in the District of Nagaon. In 3rd International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 2827–2831).

Byrne W, Doermann D, Franz M, Member S, Gustman S, Oard D.(2004). Automatic Recognition of Spontaneous Speech for Access to Multilingual Oral History Archives. 2004;12(4):420–35.

C. Vimala.(2012)."A Review on Speech Recognition Challenges and Approaches," vol. 2, no. 1, pp. 1–7, 2012.

Chia Ai O, Hariharan M, Yaacob S, Sin Chee L. (2012).Classification of speech dysfluencies with MFCC and LPCC features. Expert Syst Appl. 2012;39(2):2157–65.

Chin, A. J., Mirzal, A., Haron, H., Member, S., Nuzly, H., & Hamed, A. (2015). Supervised , Unsupervised and Semi - supervised Feature Selection : A Review on Gene Selection. IEEE/ACM Transaction on Computational Biology and Bioinformatics, 5963(d), 1–20.

Costa YMG, Oliveira LS, Koerich AL, Gouyon F. (2012).Comparing textural features for music genre classification. 2012 Int Jt Conf Neural Networks [Internet]. 2012

Daniel Pullella. (2011). Model-based Time-Frequency Mask Enhancement for Robust Missing Data Speaker Identification.

Daniel, J. , J.B. Rault, J.D. Polack. (1998). Ambisonics encoding of other audio formats for multiple listening conditions" - Pre-prints of the 105th AES Convention, S.Francisco

Dadaneh, B. Z., Markid, H. Y., & Zakerolhosseini, A. (2016). Unsupervised probabilistic feature selection using ant colony optimization. Expert Systems with Applications, 53(October), 27–42.

Dahan H, Hussin A, Razak Z, Odelha M. (2011).Automatic Arabic Pronunciation Scoring for Language Instruction. In: Proceedings of EDULEARN11 Conference 4-6. 2011. p. 145–50.

De Sena, E., Antonello, N., Moonen, M., & van Waterschoot, T. (2015). On the modeling of rectangular geometries in room acoustic simulations. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 23(4), 774–786.

De Veth J, Boves L.(1998). Channel normalization techniques for automatic speech recognition over the telephone. Speech Commun [Internet]. 1998 Aug;25(1–3):149–64.

Deepika M. (2013).Noise Cancellation In Speech Signal Processing Using Adaptive Algorithm. Int J Recent Innov Trends Comput Commun. 2013;1(September):743–6.

Deller, J. R. , J. H.L. Hansen and J. G. Proakis. (2000).“ Discrete –Time Processing o f Speech S ig nal s , ” The Institute of Electrical and Electronics Engineers, Inc., New York, 2000

Deskhmukh SH.(2014). North Indian Classical Music ’ s Singer Identification by Timbre Recognition using MIR Toolbox. Int J Comput Appl. 2014;91(4):1–5.

Diniz PSR.(2008). Adaptive filtering: Algorithms and Practical Implementations. 3rd ed. Boston, MA: Springer; 2008. ISBN: 978-0-387-31274-3

Droua-Hamdani G, Selouani S-A, Boudraa M. (2012).Speaker-independent ASR for Modern Standard Arabic: effect of regional accents. Int J Speech Technol [Internet]. 15(4):487–93.

E. C. Johnson, J. J. S. Norton, D. Jun, T. Bretl, and D. L. Jones .(2013). "Sequential selection of window length for improved SSVEP-based BCI classification," Conf. Proc. .Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf., vol. 2013, pp. 7060–7063, (2013).

Elhadj, Y. O. M., Alghamdi, M., & Alkanhal, M. (2013). Approach for Recognizing Allophonic Sounds of the Classical Arabic Based on Quran Recitations An In-house Developed Sound Database : Overview.

Elhadj, Y. O. M. (2010). E-HALAGAT : AnN E-Learning System for Teaching the Holy Quran. TOJET: The Turkish Online Journal of Educational Technology – January 2010, Volume 9 Issue 1, 9(1), 54–61.

Erguzel TT, Ozekes S, Gultekin S, Tarhan N. (2014).Ant Colony Optimization Based Feature Selection Method for QEEG Data Classification. Psychiatry Investig [Internet](2014);11(3):243–50.

Ernawan, F., Abu, N. A., & Rahmalan, H. (2012). Tchebichef moment transform on image dithering for mobile applications. Fourth International Conference on Digital Image Processing (ICDIP 2012) (Vol. 8334, p. 83340D–83340D–5).

Faris N. (2013).Speech Endpoint Detection : An Image Segmentation Approach.

Felps D, Geng C, Gutierrez-osuna R, Member S. (2012).Foreign Accent Conversion Through Concatenative Synthesis in the Articulatory Domain. IEEE Trans AUDIO, SPEECH Lang Process. 2012;20(8):2301–12.

Ferrer, M., Gonzalez, A., de Diego, M., & Piñero, G. (2012). Fast exact variable order affine projection algorithm. Signal Processing, 92(9), 2308–2314.

Fook CY, Hariharan M, Yaacob S, Ah A. (2012).Malay speech recognition in normal and noise condition. 2012 IEEE 8th Int Colloq Signal Process its Appl [Internet]. Ieee; 2012 Mar;409–12.

Fook CY, Muthusamy H, Chee LS, Yaacob S Bin, Adom AHB. (2013). Comparison of speech parameterization techniques for the classification of speech disfluencies. Turkish J Electr Eng Comput Sci. 2013;21(SUPPL. 1):1983–94.

Forsati R, Moayedikia A, Jensen R, Shamsfard M.(2014). Enriched ant colony optimization and its application in feature selection. Neurocomputing [Internet]. Elsevier; 2014;142:354–71.

Gales M, Young S. (2007). The Application of Hidden Markov Models in Speech Recognition. Found Trends® Signal Process [Internet]. Volume 1(3)pp:195–304.

Ganesan, V., & Manoharan, S. (2015). Surround Noise Cancellation and Speech Enhancement using Sub Band Filtering and Spectral Subtraction. Indian Journal of Science and Technology, 8(33).

Gannamaneni G. (2012). Master Thesis Electrical Engineering June 2012 Acoustic Echo cancellation inside a Conference Room using Adaptive Algorithms. Blekinge Institute of Technology, Karlskrona, Sweden; 2012

Gardner W.(1992). "The Virtual Acoustic Room," Master's Thesis, MIT Media Lab, 1992 August.

Gil-Cacho, J. M., Signoretto, M., Van Waterschoot, T., Moonen, M., & Jensen, S. H. (2013). Nonlinear acoustic echo cancellation based on a sliding-window leaky kernel affine projection algorithm. Audio, Speech, and Language Processing, IEEE Transactions on, 21(9), 1867–1878.

Ghosal A, Chakraborty R, Dhara BC, Saha SK. (2012). Music Classification based on MFCC Variants and Amplitude Variation Pattern : A Hierarchical Approach. Int J Signal Process Image Process Pattern Recognit. 2012;5(1):131–50.

Goldwater, S., Jurafsky, D., & Manning, C. D. (2010). Which words are hard to recognize? Prosodic, lexical, and disfluency factors that increase speech recognition error rates. Speech Communication, 52(3), pp;181–200.

Goodwin, G. C.; Payne R. L.,(1977) Dynamic System Identification: Experiment Design and Data Analysis. Academic Press, New York.

H. C. Lin.(2016). "Separation of adjacent interharmonics using maximum energy retrieving algorithm," IET Sci. Meas. Technol., vol. 10, no. 2, pp. 92–99, 2016.

Hermansky H. ,N. Morgan.(1994). RASTA processing of speech", IEEE Trans. on Speech and Audio Proc., vol. 2, no. 4, pp. 578-589, Oct. 1994.

Han, C.W, Kang, S.J., Kim, N.S. (2010). Implementation of hmm-based human activity recognition using single triaxial accelerometer IEEE Trans. Fundam. Electron. Commun. Comput. Sci., 93, 1379–1383, 2010.

Han, K., Wang, Y., & Wang, D. (2014). Learning spectral mapping for speech dereverberation. ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 23(6), 4628–4632.

Hanilci C, Ertas F.(2011). Impact of Voice Excitation Features on Speaker Verification. In: 7th International Conference on Electrical and Electronics Engineering. 2011. p. 157–60.

Han-Ping Shen, Chung-Hsien Wu* and Pei-Shan Tsai. (2015). Model Generation of Accented Speech using Model Transformation and Verification for Bilingual Speech Recognition." ACM Trans. Asian and Low-Resource Language Information Processing, Vol. 14, No 2, Article 6, March 2015.

Hasnat A. (2001). Isolated and Continuous Bangla Speech Recognition : Implementation, Performance and application perspective. In: International Symposium on Natural Language Processing (SNLP). 2001. p. 2–7.

Hassan T, Wassim A, Bassem M. (2006). Analysis and Implementation of an Automated Delimiter of " Quranic " Verses in Audio Files using Speech Recognition Techniques. In: Robust Speech Recognition and Understanding. pg. 352–62.

Haykin S. Adaptive Filter Theory. 4th ed. Prentice Hall, (2002)

Han, K., Wang, Y., & Wang, D. (2014). Learning spectral mapping for speech dereverberation. ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 23(6), 4628–4632. http://doi.org/10.1109/ICASSP.2014.6854479

Hermansky H. (1990). Perceptual linear predictive (PLP) analysis of speech. J Acoust Soc Am [Internet]. 1990;87(4):1738–52.

Hung, J. , H. Fan, and W. Tu (2012). Enhancing the magnitude spectrum of speech features for robust speech recognition," EURASIP J. Adv. Signal Process., pp. 1–20.

Hutson M. (2003). Acoustic Echo Cancellation Using Digital Signal Processing. The University of Queensland; 2003.

Ibnamin. (2017). http://www.ibnamin.com/recitations_current_places.htm

Ismail MNI, Muse ME-HM. (2014). Probabilistic PCA Mixture under Variance Preservation. World Comput Sci Inf Technol J. 2014;4(8):105–9.

Moorer. J. (1979) ."About this Reverberation Business," Computer Music Journal 3(2):13-28, 1979

Jing, Y., Jing, S., Huajian, C., Chuangang, S., & Yan, L. (2012). The gender difference in distraction of background music and noise on the cognitive task performance. 2012 8th International Conference on Natural Computation, (Icnc), pp. 584–587.

Juraj Kacur, Radoslav Vargic PM. (2011).Speaker identification by K-nearest neighbors: Application of PCA and LDA prior to KNN. In: 18th International Conference on Systems, Signals and Image Processing (IWSSIP), 2011. 2011. p. 4.

Jurafsky D. and Martin. J.H. (2007).Automatic Speech Recognition,Speech And Language Processing: An Introduction to natural language processing, computational linguistics, and speech recognition.

K. Sreenivasa Rao, Sourjya Sarkar.(2014). Robust Speaker Recognition in Noisy Environments, Technology & Engineering, Springer.

Kashef S, Nezamabadi-pour H.(2013). A new feature selection algorithm based on binary ant colony optimization. 5th Conf Inf Knowl Technol [Internet]. 2013;50–4.

Khalifa, O., Khan, S., Islam, M. R., Faizal, M., & Dol, D. (2004).Text independent automatic speaker recognition. In 3rd International Conference on Electrical and Computer Engineering, Dhaka, Bangladesh, pp. 561–564.

Khan W and Holton. R(2015).Time warped continuous speech signal matching using Kalman filter. Int. J. Speech Technol., 2015.

Kim, Y. E. (2003). Singing Voice Analysis / Synthesis.

Kumar K. A.(2011) Spectro-Temporal Framework for Compensation of Reverberation for Speech Recognition. 2011.

Kumar P, Biswas A, Chandra M. (2010). Spoken Language Identification using Hybrid Feature Extraction Methods. J Telecommun. 2010;1(2):11–5.

L. R. Rabiner and R. W. Schafer(2009). "Theory and Application of Digital Speech Processing," in Theory and Application of Digital Speech Processing, 2009.

L. Singh, G. Chetty, and S. Singh. (2012). A Novel Algorithm Using Mfcc and Erb Gammatone Filters in Speech Recognition. J. Inf. Syst. Commun., vol. 3, no. 1, pp. 358–364, 2012.

Lee, D., & Cabrera, D. (2010). Effect of listening level and background noise on the subjective decay rate of room impulse responses: Using time-varying loudness to model reverberance. Applied Acoustics, 71(9), 801–811.

Li, X., Huang, X., & Jin, T. (2010). Estimation of wall parameters by exploiting correlation of echoes in time domain. Electronics Letters, 46(23), 1563. http://doi.org/10.1049/el.2010.1779

Lili Xiang,Linbo Xie, Yuwu Liao,Ruifeng Ding.(2010).Hierarchical least squares algorithms for single-input multiple-output systems based on the auxiliary model,Journal Mathematical and Computer Modelling: An International Journal Volume 52 Issue 5-6, September, 2010 Pages 918-924

Liu, F.H., Acero, A., and Stern, R.M.(1992). Efficient Joint Compensation of Speech for the Effects of Additive Noise and Linear Filtering", ICASSP-92 , pages 865-868, March 1992

Liu, H., Yu, L.(2005). Toward integrating feature selection algorithms for classification and clustering. IEEE Transactions on Knowledge and Data Engineering 17 (4), 491–502.

Loweimi E, Ahadi SM, Drugman T, Loveymi S. (2013) On the importance of pre-emphasis and window shape in phase-based speech recognition. Lect Notes

Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics). 2013;7911 LNAI:160–7.

Luu KY. (2010). Real Time Noise Robust Speech Detection. 2010.

Naji, M., Firoozabadi M., and Azadfallah. P. (2014). “Classification of Music-Induced Emotions Based on Information Fusion of Forehead Biosignals and Electrocardiogram,” Cognit. Comput., vol. 6, pp. 241–252, 2014.

Masikkaniemi, A. (2010). Acoustic Model and Language Model Adaptation for a Mobile Dictation Service.

McGregor, P. K., Leonard, M. L., Horn, A. G., & Thomsen, F. (2013). Anthropogenic Noise and Conservation. In H. Brumm (Ed.), Animal Communication and Noise (pp. 409–444). Berlin, Heidelberg, Germany: Springer Berlin Heidelberg.

Md. Rashedullslam , Md. Shafiul Azam SA.(2009). Speaker Identification System Using PCA & Eigenface. In: 12th International Conference on Computer and Information Technology (ICCIT 2009) 21-23 December, 2009, Dhaka, Bangladesh Speaker. 2009. p. 21–3.

Meseguer, N. A. (2009). Speech Analysis for Automatic Speech Recognition.

Mohamed Hassine, Lotfi Boussaid HM. (2015),Hybrid Techniques for Arabic Letter Recognition. Int J Intell Inf Syst [Internet]. 2015;4(1):27.

Mohammed, A., Shahrizal, M., & Hj, S. (2015). Jurnal Teknologi Quranic Verses Verification using Speech Recognition Techniques, 2, 99–106.

MonirulKabir M, Monirul Islam M, Murase K. (2009). A new wrapper feature selection approach using neural network. Neurocomputing [Internet]. Elsevier; 2010;73(16–18):3273–83.

Moradi P, Rostami M. (2015) Integration of graph clustering with ant colony optimization for feature selection. Knowledge-Based Syst [Internet]. Elsevier B.V.; 2015;84:144–61.

Mourtaga E, Sharieh A, Abdallah M. (2007).Speaker Independent Quranic Recognizer Based on Maximum Likelihood Linear Regression. World Acad Sci Eng Technol 36. 2007;

Mousa A, Qados M, Bader S. (2012).Speech Signal Enhancement Using Adaptive Noise Cancellation Techniques. Can J Electr Electron Eng. 2012;3(7):375–83.

Mohamed, S. A. E., Hassanin, A. S., Taher, M., & Othman, B. (2014). Virtual Learning System (Miqra ’ ah) for Quran Recitations for Sighted and Blind Students. Journal of Software Engineering and Applications, (April), 195–205.

Muhammad, W.M., Muhammad, R. & Muhammad, (2010) A., Voice Content Matching System for Quran Readers In Ninth Mexican International Conference on Artificial Intelligence. 2010.

Nakatani, R., Takiguchi, T., and Ariki, Y. (2013). Two-step correction of speech recognition errors based on n-gram and long contextual information. Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, (August), 3747–3750.

Narang S. and Gupta. M. D. (2015). Speech Feature Extraction Techniques : A Review.Int. J. Comput. Sci. Mob. Comput., vol. 4, no. 3, pp. 107–114, 2015.

Nelke C, Jax P, Vary P.(2016). Wind Noise Detection : Signal Processing Concepts for Speech Communication. Proc DAGA. 2016;972–5.

Nematollahi MA, Al-Haddad S a. R. (2015).Distant Speaker Recognition: An Overview. Int J Humanoid Robot [Internet]. 2015;12:1550032.

Nisar, S., Khan O. U., and Tariq M.(2016). “An Efficient Adaptive Window Size Selection Method for Improving Spectrogram Visualization,” Comput. Intell. Neurosci., vol. 2016, 2016.

Noor Jamaliah, I., Zaidi, R., Zulkifli, M.Y., M. Yamani I., and Emran, M.T. (2008), Quranic verse Recitation feature extraction using Mel- Frequency Cepstral Coefficients (MFCC) The 4th International Colloquium on Signal Processing and its Application (CSPA), Kuala Lumpur, Malaysia, pp. 13-18.

Ozeki K. and Umeda T. (1984). An adaptive filtering algorithm using an or- thogonal projection to an affine subspace and its properties Electron. Commun. Japan, vol. 67-A, (1984), pp. 19–27

Park Y, West T.(2010). TCS Innovation Labs - Mumbai, Tata Consultancy Services, Yantra Park, Thane (West), Maharastra, India. 2010;(Isspa):10–3.

Parvin S, Park JS.(2007). An Efficient Music Retrieval Using Noise Cancellation. In: Future Generation Communication and Networking (FGCN 2007) [Internet]. Ieee; 2007. p. 541–6.

Poria, S., Gelbukh, A., Hussain, A., Bandyopadhyay, S., & Howard, N. (2013). Music Genre Classification : A Semi-supervised Approach. Pattern Recognition.

Rakesh P, Kumar TK. (2015). A Novel RLS Based Adaptive Filtering Method for Speech Enhancement. Int J Electr , Comput Energ Electron Commun Eng. 2015;9(2):176–81.

Ramli RM, Noor AOA, Samad SA.(2012). Adaptive Line Enhancer Using Affine Projection Algorithm for Noise Cancelling in Speech. In: International Conference on Engineering and Built Environment. 2012. p. 1–6.

Rashno, A., Sadeghiannejad, H., & Heshmati, A. (2013). Highly Efficient Dimension Reduction for Text-Independent Speaker Verification Based on Relief Algorithm and Support Vector Machines. International Journal of Signal Processing, Image Processing and Pattern Recognition, 6(1), 91–108.

Razak Z, Ibrahim NJ, Yamani M, Idris I, Tamil EM, Yakub M. (2008). Quranic Verse Recitation Recognition Module for Support in j-QAF Learning : A Review. IJCSNS Int J Comput Sci Netw Secur. 2008;8(8):207–16.

Rebecca Stewart and Mark Sandler. (2010). Database of Omnidirectional and B-Format Room Impulse Responses. In 2010 IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 165–168).

Regnier L, Peeters G. (2011).Singing voice detection in music tracks using direct voice vibrato detection. In: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings. 2009. p. 1685–8.

Rose RC, Arizmendi I, Parthasarathy S, Park F. (2003). An Efficient Framework For Robust Mobile Speech Recognition Sservices.

Roshahliza M Ramli, Ali O. Abid Noor, Salina Abd. Samad.(2014). "Modified Adaptive Line Enhancer in Variable Noise Environments using Set-Membership Adaptive Algorithm," International Review on Computers and Software (IRECOS),pp: 1468-1475, 2014

Safavi, S. (2015). Speaker Characterization using Adult and Children ' s Speech, (June).

Saha S. (1995). Biomedical Engineering Conference, Proceedings of the Fourteenth Southern. pp. 134-137.

Sandler RS and M. (2010). Database of Omnidirectional and B-Format Room Impulse Responses , Queen Mary , University of London Centre for Digital Music London , UK. Database. 2010;165–8.

San-segundo R, Hernandez LA. (2011). Robust Speech Detection for Noisy Environments. Aerospace and Electronic Systems Magazine, IEEE (Volume:26 , Issue: 11). 2011;(November):12.

Santosh, D. H. H., Aditya, S., Chandra, K. S., & Prasad, P. S. (2012). Performance Analysis of Noise Cancellation in SpeechSignals Using LMS, FT-LMS and RLS Algorithms. International Journal of Modeling and Optimization, 2(6), 667–671.

Saroha, R., Malik, S., & Anand, R. (2012). Echo Cancellation by Adaptive Combination of Nsafs Adaptedbystocastic Gradient Method. International

Schroeder. M. (1962). "Natural Sounding Artificial Reverberation," J.Audio Eng. Soc.47(2):219-223Journal of Computational Engineering Research, 2, 1079–1083.

Segura J.C., de la Torre A., Benitez M.C. and Peinado A.M..(2001). Model Based Compensation of the Additive Noise for Continuous Speech Recognition. EuroSpeech'01,Vol.I, Page(s):I – 941—944

Segura JC, Ben MC, Torre A De, Dupont S, Rubio AJ.(2002). Vts residual noise compensation. In: 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). 2002. p. 409–12.

Shrawankar U. (2011). Acoustic Echo Cancellation Postfilter Design Issues For Speech Recognition System. Int J Sci Adv Technol. 2011;1(5):38–43.

Singh N, Khan RA. (2012). MFCC and Prosodic Feature Extraction Techniques : A Comparative Study. Int J Comput Appl. 2012;54(1):9–13.

Sinha S, Jain A, Agrawal SS.(2015). Fusion of multi-stream speech features for dialect classification. CSI Trans ICT [Internet]. Springer India; 2015;2(4):243–52.

Sinha, S., Jain, A., & Agrawal, S. S. (2015). Acoustic-Phonetic Feature Based Dialect, 8(1), 235–254.

Soo, H., Man, Y., & Kim, H. (2007). Speech Enhancement via Mel-Scale Wiener Filtering with a Frequency-wise Voice Activity Detector. Journal of Mechanical Science and Technology, 708–722.

Srishti. (2012). REAL TIME FUZZY CLASSIFICATION OF CEPSTRAL COEFFICIENTS FOR SPEAKER IDENTIFICATION. 2012.

Stokes JW, Malvar HS. (2004). Acoustic Echo Cancellation with Arbitrary Playback Sampling Rate. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004 Proceedings (ICASSP '04)(Volume:4). 2004. p. 1–4.

Sukor. S. (2012). Speaker Identification System Using MFCC Procedure and Noise. 2012.

Sudhir V V, Murthy ASN, Rani DE.(2014). Acoustic Echo Cancellation using Adaptive Algorithms. Int J Adv Comput Sci Technol. 2014;3(4):248–52.

Tabakhi S, Moradi P, Akhlaghian F.(2014). An unsupervised feature selection algorithm based on ant colony optimization. Eng Appl Artif Intell. Elsevier; 2014;32:112–23.

Tabbaa HMA, Soudan B. (2015). Computer-Aided Training for Quranic Recitation. Procedia - Soc Behav Sci [Internet]. Elsevier B.V.; 2015;192:778–87.

Terzyk, J. (2014). Audio Feature Selection and Aggregation for Musical Genre Classification.

Tipping, M.E and C.M Bishop.(1999) Probabilistic Principal Component Analaysis, Journal of the Royal Statistical Society(Statistical Methdology)Vol 61, No 3.

Tyagi V, Wellekens C.(2005). On Desensitizing the Mel-Cepstrum to Spurious Spectral Components For Robust Speech Recognition

Tyagi, R., & Agrawal, D. (2012). Analysis the Results of Acoustic Echo Cancellation for Speech Processing using LMS Adaptive Filtering Algorithm. International Journal of Computer Applications, 56(15), 7–11.

Ueda Y, Wang L, Kai A, Xiao X, Chng ES, Li H.(2014). Single-channel dereverberation for distant-talking speech recognition by combining denoising autoencoder and temporal structure normalization. Proc 9th Int Symp Chinese Spok Lang Process ISCSLP 2014. 2014;379–83.

Unler, A., Murat, A., & Chinnam, R. B. (2011). Mr2PSO: A maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification. Information Sciences, 181(20), 4625–4641.

Välimäki V, Parker J.D., Savioja L., Smith J.O.,and Abel. J. S. (2012).“Fifty years of artificial reverberation,” IEEE Trans. Audio, Speech, Lang. Process., vol. 20, no. 5, pp. 1421–1448, Jul. 2012.

Venkateswarlu R.L.K., Kumari R. V. and Vani Jayasri G.(2011). “Speech Recognition Using Radial Basis Function Neural Network”, IEEE, 2011.

Veth J De, Boves L. (1996). Comparison of channel normalisation techniques for automatic speech recognition over the phone. Proceeding Fourth Int Conf Spok Lang Process ICSLP ’96. 1996;4(1):2332–5.

Wang, L., Odani, K., & Kai, A. (2012). Dereverberation and denoising based on generalized spectral subtraction by multi-channel LMS algorithm using a small-scale microphone array. EURASIP Journal on Advances in Signal Processing, 2012(1), 12.

Wei Yang Chai.(2013). University of Rochester. Department of Electrical and Computer Engineering, 2013 - 116 pages

Wu T. (2009). Feature Selection in Speech and Speaker Recognition.

Xiao-chun L, Jun-xun Y, Wei-ping H. (2012). A text-independent speaker recognition system based on Probabilistic Principle Component Analysis. 2012 3rd Int Conf Syst Sci Eng Des Manuf Informatiz [Internet]. Ieee; 2012 Oct;255–60.

Xiaodong Cui, M. A. and B. Z. (2012). Stereo-Based Stochastic Mapping With Context Using Probabilistic Pca for Noise Robust Automatic Speech Recognition. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012 (pp. 4705–4708).

Zainon, N. S. Z., Ahmad, Z. a., Romli, M. a., & Yaacob, S. (2012). Speech quality based on Arabic pronunciation using MFCC and LDA: Investigating the emphatic consonants. 2012 IEEE International Conference on Control System, Computing and Engineering, 398–403.

Zhang, Z., Wang, L., & Kai, A. (2014). Distant-talking speaker identification by generalized spectral subtraction-based dereverberation and its efficient computation, 1–12.

Zhang, Z., Wang, L., Kai, A., Yamada, T., Li, W., & Iwahashi, M. (2015). Deep neural network-based bottleneck feature and denoising autoencoder-based dereverberation for distant-talking speaker identification. EURASIP Journal on Audio, Speech, and Music Processing, 2015(1), 12.

Zheng, D. C., Dyke, D., Berryman, F., & Morgan, C. (2012). A new approach to acoustic analysis of two British regional accents-Birmingham and Liverpool accents. International Journal of Speech Technology, 15(2), 77–85. http://doi.org/10.1007/s10772-011-9123-3