

MODELING FOR RECOGNIZING RECITATION TYPE
OF THE HOLY QUR'ĀN

BY

BILAL YOUSFI

A thesis submitted in fulfilment of the requirement for the
degree of Doctor of Philosophy in Computer Science

Kulliyyah of Information and Communication Technology
International Islamic University Malaysia

SEPTEMBER 2021

ABSTRACT

Computer science and speech recognition have enjoyed a long and fruitful relationship for decades. Speech recognition has been beneficial for capturing and producing an accurate transcription of spoken words. In computer science, a prime challenge is to interpret these signals into meaningful data and to develop algorithms and applications to establish an interface between the human's voice signal and computer. Moreover, the major concerns of automatic speech recognition (ASR) are determining a set of classification features and finding a suitable recognition model for these features. Hidden Markov Models (HMMs) have been demonstrated to be powerful models for representing time-varying signals. The act of reading Qur'ān and pronouncing its sound dwells on the type of recitation. These are referring to the recitation of Warsh or the recitation of Hafss. According to the science of Qira'āt, it is essential to recognize the type of recitations, especially with the diversity and the spread of Qira'āt in the world. There are numerous efforts made by previous systems on the development of feasible guiding techniques to the act of reading the Holy Qur'ān (Tajweed rules). Unfortunately, liking the major control variables of the practices of both Usūl al Qira'āh (general principles) and Farsh al-huruf (specific variants) in those approaches were neglected. In order to fill this gap, this research thesis attempts to design and fabricate a speech recognition system that distinguishes the types of recitations (Qira'āt of Hafss An Assim and the Qira'āt of Warsh An Naafi') while reciting the Qur'ān. The proposed system is capable of recognizing, identifying, pointing out the mismatch and discriminate between two types recitations for Hafss and Warsh. An experiment among user's recitation for Hafss and Warsh with the recitation made by the expert Qur'ān reader stored in a database has been done. This thesis investigates acoustic models based on the Hidden Markov Models (HMM) classifier together with clustering algorithm for Qur'ān Speech Recognition. A significant improvement on the recognition performance was achieved when the HMM-clustering model was implemented compared to the baseline model's (single HMM model (conventional MFCC)) result. The results show that the proposed model has a faster ability for recognizing phonemes sequences than the (conventional MFCC model. Model. The adoption of the k-means algorithm for acoustic modeling is seen to be a more valid model for acoustic modeling speech recognition. However, our developed system shows a lower performance in some instances when it was compared to the other systems recently reported in the literature that used the same data. This due to the small size of training dataset used for this research, hardware availability and noise from the environment and from the speakers which can affect the improvement the results in this thesis as our aim is to investigate the proposed models for speech recognition and to make a direct comparison between these models.

خلاصة البحث

يتمتع علم الحاسوب وتقنية التعرف على الكلام بعلاقة وطيدة منذ عقود. حيث تعتبر هذه التقنية مفيدة في التقاط وإنتاج نسخة دقيقة من الكلمات والمقاطع والأصوات المنطوقة. حيث يكمن التحدي الرئيسي في علم الحاسوب في ترجمة هذه الإشارات وتحويلها إلى بيانات ذات معنى ومفهومة للحاسب لتطوير خوارزميات وتطبيقات لإنشاء حلقة وصل بين الإشارة الصوتية للإنسان ولغة الكمبيوتر. علاوة على ذلك، فإن الاهداف الرئيسية للتعرف التلقائي على الكلام (ASR) هي تحديد مجموعة من ميزات التصنيف وإيجاد نموذج للتعرف عليها، حيث أثبتت نماذج ماركوف المخفية (HMM) أنها نماذج قوية لتمثيل إشارات الصوتية وتصنيف ميزات. تعتمد طريقة نطق آيات القرآن الكريم وإخراج حروفه ولفظه الصوتي على نوع القراءة. القراءات القرآنية من أهم علوم القرآن، حيث يعتبر من الموضوعات الشديدة الصلة بنص القرآن الكريم، لأنه يعنى بكيفية النطق بألفاظ القرآن، وتحقيق الروايات المنقولة في ذلك عن أئمة القراءة. من الضروري التعرف على نوع القراءة المتبعة عند قراءة القرآن، لا سيما مع تنوع وانتشار القراءات في العالم. هناك العديد من الجهود التي بذلتها النظم السابقة في تطوير تقنيات توجيه قراءة القرآن الكريم (قواعد التجويد). لسوء الحظ، لم تعالج هذه الابحاث أحكام علم القراءات من أصول القراءات (المبادئ العامة) وفرش الحروف (المتغيرات المحددة). من أجل سد هذه الفجوة، تحاول هذه الدراسة البحثية تصميم نظام للتعرف وتمييز نوع القراءة (قراءة حفص عن عاصم وقراءة ورش عن نافع) من خلال قراءة القرآن عن طريق استعمال تقنية التعرف على الكلام. النظام المقترح قادر على التعرف على عدم التطابق وتحديد التمييز بين نوعين القراءة، حيث تم إجراء تجارب بين قراءات المشتركين مع اخرى لقراء مشاهير مخزنة في قاعدة البيانات. كما تبحث هذه الأطروحة عن مدى جدوى استعمال النماذج الصوتية المستندة إلى مصنف نماذج ماركوف المخفية (HMM) جنبًا إلى جنب مع خوارزمية التجميع clustering للتعرف على الكلام في القرآن. اشارت النتائج ان النظام حقق تحسن كبير في أداء التعرف باستعمال نموذج HMM-Clustering مقارنة بنموذج MFCC - HMM التقليدي. علاوة على ذلك، اشارت النتائج الى أن النموذج المقترح لديه قدرة أسرع على التعرف على عينات الكلمات مقارنة بنموذج MFCC التقليدي. كما بينت الدراسة ان اعتماد خوارزمية k-means للنمذجة الصوتية يُعتبر نموذجًا أكثر فعالية للتعرف على الكلام في النمذجة الصوتية. غير ان النظام المطور اظهر أداءً أقل جودة في بعض الحالات عند مقارنته بالأنظمة المنشورة سابقا والتي تستخدم نفس البيانات. هذا نظرًا لصغر حجم مجموعة البيانات التدريبية المستخدمة في هذا البحث ولوجود الضوضاء من البيئة ومن مكبرات الصوت والتي يمكن أن تؤثر سلبا على تحسين النتائج في هذه الرسالة .


ASST. PROF. DR. RAWAD ABDULKHALEQ ABDULGHAFOR
Assistant Professor
Department of Computer Science
Kulliyah of Information and Communication Technology
International Islamic University Malaysia

APPROVAL PAGE

The thesis of Student's Name has been approved by the following:

Akram M Z M Khedher
Supervisor

Abdul Wahab Abdul Rahman
Internal Examiner

Syed Ahmad Sheikh Aljunid
External Examiner

Shahrul Na'im Sidek
Chairman

DECLARATION

I hereby declare that this dissertation is the result of my own investigations, except where otherwise stated. I also declare that it has not been previously or concurrently submitted as a whole for any other degrees at IIUM or other institutions.

Bilal Yousfi

Signature

Date

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

**DECLARATION OF COPYRIGHT AND AFFIRMATION OF
FAIR USE OF UNPUBLISHED RESEARCH**

**MODELING FOR RECOGNIZING RECITATION TYPE OF THE
HOLY QUR'ĀN**

I declare that the copyright holder of this thesis is jointly owned by the student and IIUM.

Copyright © 2021 Bilal Yousfi and International Islamic University Malaysia. All rights reserved.

No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the copyright holder except as provided below

1. Any material contained in or derived from this unpublished research may be used by others in their writing with due acknowledgement.
2. IIUM or its library will have the right to make and transmit copies (print or electronic) for institutional and academic purposes.
3. The IIUM library will have the right to make, store in a retrieved system and supply copies of this unpublished research if requested by other universities and research libraries.

By signing this form, I acknowledged that I have read and understand the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Bilal Yousfi

.....
Signature

.....
Date

This thesis is dedicated to my beloved parents (Abdullah and Sassia).

To

My dearest brothers (Mohammed Ridha and Alaeddine), sisters (Nadjah, Amira and Chaima), nephews (Abdelbaset Abdisamad, Ayoub and Jawad), and nieces (Bayan, Maria and Rosline)

For their love, prayers, encouragement, perseverance, and faith in me.

ACKNOWLEDGEMENTS

In the name of Allah, the Most Gracious, the Most Merciful. He, who has given me the strength to continue doing this research. Alhamdulillah, All Praise to Allah that he has answered my prayers to submit this thesis with the hope to pass with excellence

All the languages of the world cannot produce words that thank my parents for their great support to my education, eventually resulting in successful completion of this work. They were the cause, and the inspiration without which this work would have never been completed. I wish a happy long life for my parents; without them I simply would not be where I am now.

I would like to express my deep gratefulness, admiration, and honor to my supervisor Dr. Akram M Z M Khedher. With such richness in experience, knowledge, and dedication, he supported and provided me with the direction to complete this work; always having time to listen, always prepared to help and always open to discussion. I must present my thanks to him for his advice, valuable feedback and suggestions and support during the research journey.

Likewise, I am greatly indebted to my brothers and sisters, Mohammed Radha, Nadjah, Asma, Chaima, for their continuous support and words of encouragement throughout the course of this Ph.D. study. This thesis would not have been completed without their help and support.

I would also like to thank my nephew and nieces, Bayan, Abdelbaset Abdessamed, Ayoub, Djouad, Maria, and Rosline.

I would also like to thank KICT Post Graduate Office Members, Kyra, Narita, Izzati, and Sarah for their support and help.

In the same vein, I am equally grateful to all my colleagues and friends at the IIUM for providing soundboards and light relief.

TABLE OF CONTENTS

Abstract	ii
Abstract In Arabic	iii
Approval Page	iv
Declaration	iv
Copyright	vi
Dedication	vii
Acknowledgements	ii
Table Of Contents	iii
List Of Tables	v
List Of Figures	vii
List Of Abbreviations	xi
CHAPTER ONE: INTRODUCTION	1
1.1 Overview	1
1.2 Background About Speech Recognition	5
1.3 Problem Statement	7
1.4 Research Objectives	10
1.5 Research Questions	11
1.6 Research Design	11
1.7 Scope and Constraints	13
1.8 Research Significance	15
1.9 Thesis Organisation	16
CHAPTER TWO: LITERATURE REVIEW	17
2.1 INTRODUCTION	17
2.2 Overview	17
2.3 Theoretical Aspects of Speech	18
2.4 Features Extraction Techniques	23
2.4.1 Linear Predictive Coding and Linear Predictive Cepstral Coefficients	24
2.4.2 Perceptual Linear Prediction (PLP)	24
2.4.3 Mel-Frequency Cepstral Coefficient (MFCC)	25
2.5 Features Classification	28
2.6 Review on speech recognition methods for the Holy Qur’ān	30
2.7 Qira’āt Background	40
2.7.1 The Classification of the Various Qira’āt	44
2.7.1.1 The Authentic Qira’āt	45
2.7.1.2 The Exceptional (Shadh) or Unauthentic Qira’āt	46
2.7.2 The Qira’āt of Warsh A’n Naafi’	47
2.8 The Qira’āt of Hafss An Assim	48
2.8.1 The Qira’āt Today	49
2.9 The discipline of Qira’āt	52
2.9.1 Differences between Hafss and Warsh	54
2.10 The Qur’ān and the Arabic language	56
2.10.1 Arabic Phonemes	59

CHAPTER THREE: RESEARCH METHODOLOGY	61
3.1 Introduction	61
3.2 Data Collection	63
3.3 Pre-processing	72
3.4 Features Extraction	74
3.5 Features Selection Procedure.....	79
3.5.1 Clustering Classifier Generation Procedure.....	80
3.5.2 Identify Target Data.....	82
3.6 Features Classification Process	83
3.6.1 Training Dataset.....	85
3.6.2 Testing Dataset.....	87
3.7 Conclusion	87
CHAPTER FOUR: DATA ANALYSIS AND RESULTS.....	89
4.1 Introduction	89
4.2 Acoustic Analysis	90
4.2.1 The Acoustic Analysis of Madd (Lengthening) Rule	90
4.2.1.1 The Acoustic Analysis of Exchange Prolongation Rule ..	91
4.2.2 The Acoustic Analysis for The Straight and The Imālah Rules ...	96
4.2.3 The Acoustic Analysis for The Idghaam and Al Ith-Haar	
Rules	100
4.2.4 The rules of the letter lam (L) and the letter raa (R) (Tafkheem	
and tarqeeq)	103
4.2.4.1 The Acoustic Analysis for The Rules of the Letter Lam..	103
4.2.4.2 The Rules of The Letter Raa	108
4.3 Presentation of The Results	112
4.3.1 Results of The Experiment.....	113
4.3.1.1 Results of Experiment I.....	115
4.3.1.2 Results of Experiment II.....	117
4.3.1.3 Results of Experiment III	120
4.4 A comparison between the proposed model and mfcc-mlp model	125
4.4.1 MFCC-MLP MODEL IMPLEMENTATION	125
4.4.1.1 Data collection and feature extraction	125
4.4.1.2 CLASSIFICATION.....	128
4.4.2 MFCC-MLP model Experiments and results	132
4.5 External Works Comparison	142
4.6 Conclusion	144
CHAPTER FIVE: CONCLUSION AND FUTURE WORK.....	145
5.1 Introduction	145
5.1.1 Research findings.....	146
5.1.2 Research Limitation	150
5.1.3 Future work.....	150
REFERENCES.....	150
PUBLICATIONS	155

LIST OF TABLES

Table 2. 1	Critical Analysis of Features Extraction Techniques	28
Table 2. 2	Taxonomic Table for Research Studies Related to the Qur'ān Speech Recognition	39
Table 2. 3	The Most Authentic Qira'āt.	47
Table 2. 4	Well-Known Qira'āt	48
Table 2. 5	The areas in the world where other Qira'āt are prevalent(Qadhi, 2003)	53
Table 2. 6	Some differences between Hafss and Warsh according the Usūl al Qira'āh.	56
Table 2. 7	Some Differences between Hafss and Warsh according to Farsh al-huruf.	56
Table 2. 8	Arabic–Roman letters mapping table.	59
Table 2. 9	Phonetic transcription for diacritics.	60
Table 2. 10	Arabic phonemes	61
Table 3. 1	List of some rules used in the system	68
Table 4. 1	The difference between the exchange prolongation according to Hafss and Warsh in term of time and prolongation of vowels.	96
Table 4. 2	The Overall average time of Madd types according to Hafss and Warsh.	96
Table 4. 3	The acoustics characteristics of the spoken word الْفُؤَى .	100
Table 4. 4	The acoustics characteristics of the spoken word فَقْدُ ظَلَمَ .	104
Table 4. 5	The acoustics characteristics of the spoken word مَطَّلَع .	107
Table 4. 6	The Overall average for acoustics characteristics of the spoken word الأَبْرَار	112
Table 4. 7	Recognition time and recognition rate or the word مَا تَشَاءُونَ	116
Table 4. 8	Recognition rates and the recognition time for experiment 1.	117
Table 4. 9	Recognition rates and the recognition time for experiment II.	120

Table 4. 10	Recognition rates and the recognition time for experiment III.	122
Table 4. 11	Recognition time and recognition rate under the number of Different ks clusters for Experiment I and Experiment II	124
Table 4. 12	Recognition time and recognition rate for the system under the number of different Ks cluster	126
Table 4. 13	MLP Network Parameters.	130
Table 4. 14	Confusion matrix for testing the Qira'āt of Hafss and the Qira'āt of Warsh accuracy.	134
Table 4. 15	Confusion Matrix for recognising the type of Qira'āt accuracy.	135
Table 4. 16	A comparison between MFCC-MLP and HMM-Clustering model.	140
Table 4. 17	The performance of most published work.	141
Table 4. 18	A comparison between Single HMM and HMM-Clustering model	142

LIST OF FIGURES

Figure 1. 1	The Stages of Research Methodology.	13
Figure 1. 2	Research Diagram.	14
Figure 1. 3	The illustration of the research scope	15
Figure 2. 1	Block diagram of Automatic Speech Recognition System	21
Figure 2. 2	Block diagram for a conventional MFCC implementation.	28
Figure 2. 3	Classification of speech recognition approaches(Anusuya & Katti, 2010)	31
Figure 2. 4	The ten imams (reciters) of Qira'āt (Sal, 2014)	45
Figure 2. 5	The narration chain of The Qira'āt of Warsh An Naafi'(Sal, 2014)	49
Figure 2. 6	The narration chain of The Qira'āt of Hafss An Assim (Sal, 2014)	50
Figure 2. 7	Arabic language classification (Elmahdy et al., 2012)	58
Figure 3. 1	Proposed Methodology for Data Analysis	64
Figure 3. 2	Data Collection Procedure	65
Figure 3. 3	The flow chart of the data collection steps.	66
Figure 3. 4	The recitation rules used in data collection	67
Figure 3. 5	Percentage of recitation's type used in data collection.	70
Figure 3. 6	Percentage of speaker group used in data collection.	70
Figure 3. 7	Percentage of Qari recitation used in data collection.	71
Figure 3. 8	Percentage of recitation rules used in data collection.	72
Figure 3. 9	Percentage of each rule of usūl al Qira'āt rules used in data collection.	72
Figure 3. 10	Speech pre-processing steps	75

Figure 3. 11	The single for the word مشاؤون before and after pre_emphasis.	76
Figure 3. 12	Steps of features vector extraction using MFCC Algorithm	77
Figure 3. 13	Block Diagram for the MFCC algorithm (Jurafsky, 2009)	77
Figure 3. 14	Framing Signal for the word ما تشاؤون	78
Figure 3. 15	Filter bank on a Mel-Scale	79
Figure 3. 16	The result of MFCC Cepstral Coefficients for the word وما تشاؤون	80
Figure 3. 17	MFCCs in frames for the word وما تشاؤون	81
Figure 3. 18	Steps of features clustering using K-means Algorithm	82
Figure 3. 19	Steps of features selection for target data group using DTW Algorithm	84
Figure 3. 20	Steps of templates generation using HMM	85
Figure 3. 21	The framework of speech classification	86
Figure 3. 22	The framework of speech Training dataset	89
Figure 3. 23	System Architecture	90
Figure 4. 1	The different types of Madd in Tajweed rules	93
Figure 4. 2	Part of the waveform of the spoken word مَا تَشَاءُونَ According to Hafss and Warsh respectively.	94
Figure 4. 3	The 2D Plot of the codebook for the spoken word مَا تَشَاءُونَ According to Hafss and Warsh	95
Figure 4. 4	2D plot of acoustic vectors of the spoken word مَا تَشَاءُونَ According to Hafss and Warsh	95
Figure 4. 5	The time average for different Madd rules for both Hafss and Warsh	97
Figure 4. 6	Part of the waveform of the spoken word الْقُوَى According to Warsh and Hafss.	98
Figure 4. 7	The 2D Plot of the codebook of spoken word الْقُوَى (The Lord of Mighty Power).	99
Figure 4. 8	2D Plot of the acoustic vector of spoken word الْقُوَى (The Lord of Mighty Power)	99

Figure 4. 9	The overall average of frequency for the straight and the Imālah rules	101
Figure 4. 10	The overall average of intensity for the straight and the Imālah rules	101
Figure 4. 11	The overall average time for the straight and the Imālah rules	102
Figure 4. 12	Part of the waveform of the spoken word فَفَدُّ ظَلَمَ According to Warsh and Hafss	103
Figure 4. 13	The 2D Plot of the codebook for spoken word فَفَدُّ ظَلَمَ According to Warsh and Hafss	103
Figure 4. 14	2D Plot of the acoustic vector of spoken word فَفَدُّ ظَلَمَ According to Warsh and Hafss	104
Figure 4. 15	The Overall average for acoustics characteristics for the idghaam and al ith-haar rules	105
Figure 4. 16	Part of the waveform of the spoken word مَطَّلَعِ According to Warsh and Hafss.	106
Figure 4. 17	The 2D Plot of the codebook for spoken word مَطَّلَعِ According to Warsh and Hafss.	106
Figure 4. 18	2D Plot of the acoustic vector of spoken word مَطَّلَعِ According to Warsh and Hafss.	107
Figure 4. 19	The overall average of frequency for the rules of the letter lam	108
Figure 4. 20	The overall average of intensity for the rules of the letter lam	109
Figure 4. 21	The overall average time for the rules of the letter lam	109
Figure 4. 22	Part of the waveform of the spoken word الأبرار According to Warsh and Hafss.	110
Figure 4. 23	2D Plot of the acoustic vector of spoken word الأبرار According to Warsh and Hafss.	111
Figure 4. 24	The 2D Plot of the codebook for spoken word الأبرار	111
Figure 4. 25	The overall average of frequency for the rules of the letter Raa	112
Figure 4. 26	The overall average of intensity for the rules of the letter Raa	113

Figure 4. 27	The overall average time for the rules of the letter Raa	113
Figure 4. 28	Recognition rate under the number of Different ks clusters for Experiment I and Experiment II	125
Figure 4. 29	Recognition time under the number of different ks clusters for Experiment I and Experiment II	125
Figure 4. 30	Recognition time for the system under the number of Different ks clusters.	127
Figure 4. 31	Block Diagram for Analysing Raw Signal.	126
Figure 4. 32	Schematic Representation of fivefold cross validation.	127
Figure 4. 33	The MLP Network with one hidden layer.	129
Figure 4. 34	Five required stages of data to obtain results.	131
Figure 4. 35	A histogram used for a comparison between the average results of HMM-Clustering Model with MFCC-MLP Model for the Qira'āt of Warsh training dataset.	137
Figure 4. 36	A histogram used for a comparison between the recognition time results of HMM-Clustering Model with MFCC-MLP Model for the Qira'āt of Warsh training dataset.	138
Figure 4. 37	A histogram used for a comparison between the average results of HMM-Clustering Model with MFCC-MLP Model for the Qira'āt of Hafss training dataset.	139
Figure 4. 38	A histogram used for a comparison between the recognition time results of HMM-Clustering Model with MFCC-MLP Model for the Qira'āt of Hafss training dataset.	139

List of Abbreviations

AI	Artificial intelligence
ASR	Automatic Speech Recognition
DFT	Discrete Fourier transform
DTW	Dynamic Time Warping
FIR	Finite Impulse Response
HMM	Hidden Markov model
LPC	Linear Prediction Coding
MFCC	Mel-Frequency Cepstral Coefficients
MSA	Modern Standard Arabic
WRA	Word Recognition Accuracy
ZCR	Zero Crossing Rate

CHAPTER ONE

INTRODUCTION

1.1 OVERVIEW

Allah Almighty said: "And make loose the knot (the defect) from my tongue, (i.e., Remove the incorrectness from my speech) " [20:27].

﴿وَأَحْلَلْ عُقْدَةً مِّن لِّسَانِي ۚ﴾

Language can be described as a system of speech sounds, words, and patterns used by humans to communicate their thoughts and feelings (Goldstein, 2014). In other words, speech is a natural communication method between humans and is a meaningful, useful and the fastest form of data input between humans and technology. That is why there are a different variety of languages used in verbal communication. However, it is estimated that there are almost 16 languages which are spoken by more than 500 million people, i.e., English, Arabic, Chinese, Italian, Spanish, Russian, French, Hindi, Portuguese, Bengali, German, Japanese and Urdu.

The Arabic language is considered one of the most ancient living and Semitic languages in the world and it is also the Fifth most generally used language. The Arabic language is the mother tongue for roughly 240 million people (Garcia, 2018), and the Arabic language is linked with Islam since it is the language of the Qur'ān, and as well Arabic is the language that is used in most of the Islamic literature.

The Qur'ān is the Speech (Kalaam) of Allah Almighty, which he revealed to The Prophet (ﷺ) in wording and meaning, and which has been preserved in the Mushafs, and has reached us by Mutawātir transmissions.

As mentioned in the Holy Qur'ān itself, the Arabic language is the language of the Qur'ān, and with it, the Qur'ān was revealed to the Prophet (ﷺ). The knowledge of the Arabic language is essential for every Muslim so that they can perform their religious acts of worship and they can be proficient in the recitation of the Qur'ān. Allah Almighty says in His Book:

إِنَّا أَنْزَلْنَاهُ قُرْآنًا عَرَبِيًّا لَعَلَّكُمْ تَعْقِلُونَ (2)

“Verily. We have revealed this as an Arabic Qur'ān” [12:2]

وَهَذَا لِسَانٌ عَرَبِيٌّ مُبِينٌ (103)

... this (the Qur'ān) is in a clear Arabic tongue” [16:103]

وَكَذَلِكَ أَوْحَيْنَا إِلَيْكَ قُرْآنًا عَرَبِيًّا ... (7)

«And thus, we have inspired you with an Arabic Qur'ān... [142:7].

The practice of reciting the Qur'ān differs from the normal reading of Arabic text, and it is performed according to a set of guidelines known as Tajweed. Therefore, learning how to recite the Qur'ān correctly is an obligatory act upon every Muslims, as in Sahih al-Bukhari Hadith number 5027: The Prophet (ﷺ) said, "The best among you (Muslims) are those who learn the Qur'ān and teach it." (Al-Bukhārī & Ḥān, 1997). Many people, especially beginners or new learners have difficulties in reading the Qur'ān base on the designed rules of recitation. The Majority of them mismatch their act of recitations, from the basic prescribed ones and this is most likely caused by their lack of oral practices and monitoring. A research finding from (Mohd, 2006) states that "most Muslims are weak in reading the Qur'ān in the field of Islamic education for some time". Furthermore, some Muslims only practices reading the Qur'ān during their lesson in school as this could be attributed to the lack of resources that can ease their recitations

skills and encourage them to learn the recitation types according to the rules governing those styles.

There are seven accepted readings in the system of Qira'āt. The number seven is based on a well-known hadith of several variants, in the Hadith that was narrated by al-Bukhari, 3047; Muslim, 819: The Prophet (ﷺ) said: "The Qur'ān has been revealed in seven different ways (modes (Ahruf)), so recite it whichever is easiest for you ". (Al-Bukhārī & Ḥān, 1997) , which means that the Holy Qur'ān was revealed in seven variants Ahruf (dialect), the Ahruf which refers to different dialects among the Arabs at the time of the revelation of the Qur'ān and the term Qira'āt (recitations) usually denotes the accepted variant Ahruf of the Qur'ān. Moreover, the word Qira'āt is the variations in words and pronunciations of the Qur'ānic verses (Qadhi, 2003). In another definition, Qira'āt refers to the changes which occur in the words of the Qur'ān with regard to the prolongation (Muddud), shortening (Kassar), the punctuation of the written text and the pronunciation of the Qur'ānic words. (Dogan, 2014)

In the 4th century of Hijra, Abū Bakr b. Mujāhid (d. 324/936) standardized the number of recitations to seven acts of Qira'āt and these are recognised as the seven well-known reading modes of the Holy Qur'ān. The seven readings that were accepted in Ibn Mujāhid's time as the accepted Qira'āt includes: Ibn Kathīr (Mecca, d. 120/738), Nāfi' (Medina, d. 169/785), Ibn `Āmir (Damascus, d. 118/736), Abū `Amr (Baṣra, d. 154/770), `Assim (Kūfa, d. 127/745), Ḥamza (Kūfa, d. 156/773), and al-Kisā'ī (Kūfa, d. 189/804). In the Qira'āt science today, the most popular readings are those transmitted by Hafss (d. 180/796) on the authority of `Assim and Warsh (d. 197/812) on the authority of Nafi. (Dogan, 2014).

Computer science and speech recognition have enjoyed a long and fruitful relationship for decades. Automatic Speech Recognition (ASR) is a field of computer

science which deals with designing computer systems that can recognise spoken words or take dictation which involves the ability to match a voice pattern against a provider or acquired vocabulary (Abushariah & Gunawan, 2011). According to Rabiner (Rabiner & Schafer, 1978), Speech recognition is the process of converting or translating an acoustic signal, captured by the microphone to a set of textual words. A word is recognised by extracting features from the captured signal and classifies the features with the given voice sample in the database. One of the earliest researches of ASR is in Human-Computer Interaction (HCI) such as speech to speech translation (Lavie et al., 1997; Wahlster, 2000), Computer-aided language learning (Witt and Young, 1997; Xu et al., 2009; Peabody, 2011), dictation (Murveit et al., 1993; Lee, 1997; Lee et al., 2009), and voiced based information retrieval (Franz and Milch, 2002; Zue et al., 2000) etc.

Currently, there have been numerous researches related to Al-Qur'ān and Arabic speech; Such as speech recognition for Tajweed rules has been found to focus on Automated Tajweed checking rules engine for Qur'ānic learning (Raja-Jamilah Raja-Yusof and Fadila Grine, 2013), as well as Tajweed checking system to support recitation (Ahsiah, Noor, & Idris, 2013). There are other studies which also streamline Qur'ānic Verse recitation recognition module for support in J-QAF learning (abbreviation to Jawi, Qur'ān, Arabic, Fardu Ain, which is a religious learning approach in Malaysian schools), and Makhraj recognition using speech processing (Wahidah et al., 2012) in order to isolate the right sound speech for the right Qur'ān words.

In Consistent with the previous studies, the goal of this research was to investigate the relationship between reciting the Qur'ān and understanding the type of recitation using spectrogram voice analysis. To accomplish this goal, new speech recognition model for Qur'ān have been developed based on the two famous Qira'āt

which are the Qira'āt of Hafss An Assim (Hafss from Assim) and the Qira'āt of Warsh An Naafi' (Warsh from Naafi').

This study aims to deliver types of recitation identification system that uses pattern recognition and classification methods as this would assist people who wish to practice reading the Qur'ān following one of the two famous Qira'āt, while constituting a step toward developing a machine interface that can respond to the human type of recitation following during the learning of the Qur'ān.

1.2 BACKGROUND ABOUT SPEECH RECOGNITION

Speech recognition systems, which come under the natural language processing umbrella, in particular, have witnessed a significant breakthrough based on the advances in Machine Learning (ML) and it can be categorized into two modes. The first is isolated word recognition which is the simplest speech recognition mode, where each word is surrounded by silence, so that word boundaries are well known. The second is the continuous speech recognition which is more natural than the isolated mode. This mode is more difficult than the isolated word recognition because the word boundaries are difficult to recognise (detecting the beginning and the end of each word). In speech conversation, a word may be uttered in a different way due to the variations in dialects, pronunciations, accents, volumes, gender and age. (Tebelskis, 1995). Also, the speaker variability such as emotion and illness is combined with the attention to the environment. Speech recognition systems are also influenced by variations between the environments which can introduce corruption into the speech signal as a result of background noise, transmission channels, and microphone characteristics. (Jou, 2008; Yuk, 1999)

The initial speech recognition stage is the pre-processing where recording, speech signals are passed through the pre-processing block for determining the correct boundary of the isolated words and rejecting the artifacts of speech such as noise and intra-word stops. With regards to large vocabulary continuous speech boundary detection, the problem gets to be much more difficult due to the intra-word silences and other artifacts. Generally, these problems are reduced by applying speech boundary detection algorithms. The most prevalent methods used for endpoint detection are energy profile and Zero Crossing Rate (ZCR). (Chen, 1988).

The most prominent phase in speech recognition is feature extraction which is designed to extract unique, robust, discriminative, and computationally efficient characteristic from the speech signals. The most common feature extraction techniques are Mel Frequency Cepstral Coefficients (MFCC) (S. Davis & Mermelstein, 1980), perceptual linear prediction coefficients (PLP) (Hermansky, 1990), and Fast Fourier Transform (FFT) (Van Loan, 1992). Typically, the ASR system represents the speech signal with state-of-the-art Mel Frequency Cepstral Coefficients (MFCCs).

The main concerns in the automatic speech recognition (ASR) are to Figure out a set of classification features and finding a suitable and appropriate model for these features. Hidden Markov Models (Rabiner, 1989), which used a special case of regular Markov models, have been demonstrated to be a powerful model for representing time-varying signals as a parametric random process (X.D. Huang et al., 2001, Rabiner, 1989). Artificial Neural Networks (ANNs) such as Multi-Layer Perceptron (MLP) (Bengio, 2009; Glorot & Bengio, 2010), Deep Neural Networks (DNN) (Glorot & Bengio, 2010), Recurrent Neural Networks (RNN), and Echo State Network (ESN) (Jaeger, 2002), have also been widely used for representing time-varying quasi-stationary signals.

Despite the use of the above techniques, current speech extraction and recognition techniques encounter errors in recognition accuracy rates and delay in processing time. An alternative is to perform the recognition process after using a clustering algorithm to classify the feature vectors, and the HMM is trained to recognise it in the same manner as the other patterns.

1.3 PROBLEM STATEMENT

The act of reading and pronouncing the Qur'ān dwells on the type of recitation as these refers to the recitation of Warsh or Hafss. It is very important to recognise the type of recitations, especially with the diversity and the spread of Qira'āt in the world (Qadhi, 2003). People are having difficulties in reading the Qur'ān with the correct rules of recitation type where they cannot recognise which recitation type they are following when reciting the Qur'ān due to lack of oral practices and monitoring.

Topics associated with the process of learning and memorization of the Qur'ān have been a subject of interest to several researchers. Al- Qur'ān words are believed by Muslims to be the words of Almighty Allah, which have distinctive characteristics and has been revealed in its precise meaning and wording through the Angel Gabriel. Moreover, The Qur'ān provides some instruction about how to perform its recitation, as in

﴿... وَرَتِّلِ الْقُرْآنَ تَرْتِيلًا﴾ [المزمل: 4]

«..., and recite the Qur'ān with measured recitation (4) » [Al-Muzzammil: 4]

In spite of the large amount of research on the Qur'ān and the act of learning and memorizing the Qur'ān in proper Qur'ānic recitation, an analysis of studies and sources conducted in literature (see section 03.06) indicate that—unfortunately—very little scientific research is available concerning the recitation types of the Qur'ān.