ACOUSTIC PROFILES BASED ON QURANIC MAQAMAT AUDIO FEATURES

Farah Hanim bt Seman @ Abd Jabar

Thesis submitted in partial fulfilment for the degree of
DOCTOR OF PHILOSOPHY IN
SCIENCE AND TECHNOLOGY

UNIVERSITI SAINS ISLAM MALAYSIA

November 2022

AUTHOR DECLARATION

I hereby declare that the work in this thesis is my own unless specified and duly acknowledged by quotation.

Date:

Signature:
Name:
Matric No:
Address:

ACKNOWLEDGEMENTS

In the name of Allah, The Most Beneficent, The Most Merciful.

Grateful to Allah SWT and His Messenger Muhammad SAW. Firstly, I would like to express my deeply gratitude to my late supervisors, Allahyarham Prof Madya Dr Janatul Islah binti Mohammad who had extremely supported me throughout my first 3 years of my research with her utmost patience and knowledge. Then I would also like to convey my gratitude and appreciation to my current supervisor, Ts. Dr. M Norazizi Sham Mohd Sayuti and my co-supervisor Prof. Dr. Kamaruzzaman Seman for their great supervision and guidance toward the end of my thesis journey.

I owe my gratitude to the Institute of Higher Education of Malaysia and Universiti Sains Islam Malaysia (USIM) for providing me this opportunity to gain experience and broaden my knowledge.

Finally, thanks to my parents for their prayers and believe in my capabilities, and lastly yet most importantly, to my husband and children, for their unconditional love, patience and support throughout my struggle over the years.

ABSTRAK

Pengekstrakan Sifat Audio memberikan peranan besar dalam pemprosesan pertuturan terutama dalam analisis audio bersemantik yang memberikan maklumat sifat audio berasaskan intonasi, emosi dan rentak. Ramai penyelidik telah mengutarakan cabaran yang dihadapi oleh penutur Bahasa Arab terutama di dalam Al-Quran, disebabkan oleh perbezaan dalam teknik penulisan dan bacaan. Intonasi perkataan di dalam Al-Quran akan memberikan maksud terjemahan atau persepsi kepada objektif surah secara amnya atau ayat secara khususnya. Antara kaedah untuk mengekstrak sifat jelas yang mencirikan keaslian signal pertuturan bermelodi yang kompleks ini adalah analisis *kepstral*. Kini, kajian lepas telah menunjukkan kebanyakan penyelidikan tentang ciri-ciri akustik dan retorik di dalam Al-Quran memfokuskan analisis teks berbanding bunyi bacaan. Dalam kajian ini, analisis spektral dipertingkatkan telah dicadangkan bagi mengekstrak sifat audio yang berasaskan ontologi yang terkandung dalam rakaman audio Bacaan Al-Quran Bermaqamat. Sistem yang dicadangkan ini dimulakan dengan mengekstrak sifat audio yang terkandung dalam 242 buah rakaman audio menggunakan analisis berspektral sedia ada dan yang dipertingkatkan berasaskan fungsi meledingkan. Sebanyak sembilan set sifat audio diekstrak dari algoritma deskriptor berspektral dan dua set adalah dari algoritma pengektrakan sifat audio yang dipertingkatkan dengan fungsi yang diledingkan. Kerangka ini berupaya mengesan sampul mean spektral dan deskriptor spektral sebagai sifat audio yang jelas yang digunakan sebagai entiti bagi penandaan atribut dan padanan peraturan. Peringkat akhir adalah analisis semantik berasaskan ontologi Maqamat dengan penandaan atribut dan padanan peraturan bagi binaan pengkalan data. Hasil penyelidikan dan prestasi dianalisis dengan semantik audio berdasarkan bentukan frekuensi dari sifat spektral. Kajian ini diharapkan dapat membantu dalam pemahaman kandungan maqamat dan menyumbang kepada meningkatkan inisiatif dalam memprofail sifat akustik dalam analisis penuturan kompleks bagi pemprosesan penuturan masa hadapan.

ABSTRACT

Audio feature extraction underpins a massive proportion of speech processing, mainly in semantic audio analysis which retrieves sound features information based on intonation, emotion and rhythm. Many researchers have addressed various challenges that most speakers faced when dealing with Arabic language especially in the Quran, due to its differences in written and recital technique. The intonation of words in the Quran will bring different translation or perception to the objective of the chapter generally or to the exact verse particularly. Among other methods to extract strong features that characterise the complex nature of complex and melodious speech signals is cepstral analysis. Nowadays, existing literature have shown that most of the study on acoustical and rhetorical element of the Quran is focusing on the textual analysis rather than recitation of the Quran. In this research an enhanced audio feature extraction has been presented to extract significant ontological audio features contained in Quranic Maqamat Recitation audio recording. The proposed system is initiated by extracting the maqamat features from 242 sets of audio files using existing and enhanced cepstral analysis based on warping function. Nine sets are extracted from spectral descriptors algorithm and the other 2 sets are extracted from the audio feature extraction techniques enhanced with warping function. The framework managed to detect the mean of spectral envelope and spectral descriptors as significant audio features which used as entities for attributes tagging and rule matching. The final stage is performing the semantic analysis based on the proposed ontology with attribute tagging and rule matching for the insight knowledge base construction. The results is analysed in semantic audio analysis and their performance based on the formants frequencies extracted from the spectral properties. This study will initiate an understanding on the characteristics of the maqamat content and hopefully contribute to raising initiatives in profiling the acoustical elements of complex speech analysis for further speech processing analysis.

الملخص

يرتكز استخراج الميزة الصوتية بشكل كبير على معالجة الكلام خاصة حين يتعلق الأمر ب تحليل الصوت
وقد تناول مجموعة .الدلالي الذي يستخلص معطيات الميزات الصوتية بناء على التنغيم والعاطفة والإيقاع
من الباحثين بالدراسة التحديات التي يواجهها غالبية المتحدثين مع اللغة العربية وخاصة القرآن لاختلاف
ومن ثم كان لتنغيم أو ترتيل الكلمات في القرآن أن يضيف للسورة بصفة عامة .أسلوب الكتابة والتلاوة
ومن بين طرق استخراج الميزات القوية الأخرى التي تميز طبيعة .والآية بصفة خاصة ترجمة أو تصورا مختلفا
Cepstral إشارات الكلام المعقد والمنغوم المعقدة التحليل الطيفي الذي سأعتمد عليه في هذا البحث
. وتكشف لنا المؤلفات اليوم أن معظم الدراسات حول عنصري الصوت والبلاغة تهتم بالتحليل *Analysis*
لهذا أقدم في هذا البحث خاصية لاستخراج الميزات الصوتية المحسنة من .النصي بدلا من تلاوة القرآن
أجل استخراج الميزات الصوتية الوجودية المهمة التي يتضمنها التسجيل الصوتي لتلاوة المقامات القرآنية.
وقد خلصت إلى هذا النظام المقترح باستخراج ميزات المقامات من 242 مجموعة من الملفات الصوتية
باستعمال التحليل الطيفي الراهن والمحسن المبني على وظيفة التزييف و9 مجموعات من الميزات الصوتية
التزييف. من خوارزمية الواصفات الطيفية ومجموعتين من تقنيات استخراج الميزات الصوتية المحسنة بوظيفة
وقد نجح هذا الإطار في رصد متوسط الغلاف الطيفي والواصفات الطيفية باعتبارها مميزات صوتية هامة
ومطابقة القواعد. وتتمثل المرحلة النهائية في *Attributes Tagging* تستخدم ككيانات لوسم الصفات
التحليل الدلالي بناء على الأنطولوجيا المقترحة مع وسم الصفات ومطابقة القواعد لبناء قاعدة معرفة
عميقة. أما النتائج فقد تم تحليلها في التحليل الصوتي الدلالي بناء على ترددات صفات أصوات الكلام
المستخرجة من الخصائص الطيفية. إن هذه الدراسة تقدم فهما جديدا لخصائص محتوى المقامات على
أمل أن تساهم في رفع مبادرات لتحديد خصائص عناصر التحليل الكلامي المعقد الصوتية لأجل المزيد
.من تحاليل معالجة الكلام

TABLE OF CONTENTS

CONTENT **PAGE**

CHAPTER 5 : MPS CLASSIFICATION FOR QMR KNOWLEDGEBASE CONSTRUCTIONS

CHAPTER 6 : PERFORMANCE EVALUATION WITH CROSS VALIDATION TECHNIQUES

CHAPTER 7 : CONCLUSIONS AND FUTURE WORKS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF APPENDICES

LIST OF ABREVIATIONS

Notation	Descriptions
c_r	Real spectrum
c_j	Imaginary part
n	Signal sequence
e	Excitation sequence
h	Vocal tract filter sequence
s	Speech sequence
E	Excitation sequence in frequency domain
H	Vocal tract filter sequence in frequency domain
S	Speech sequence in frequency domain
ω	Spectrum sequnce
α	pre-emphasis parameter
a	coefficient
H	Transfer function
Z	z-domain
f_{mel}	Mel-frequency
H_m	Weight of given spectrum relative to energy spectrum bin
M	Total number of triangular mel weighting filters
c	Cepstral coefficients
C	Total number of MFCCs
T	Total number of successive frames
N	Number of points used to compute the DFT
T_w	Time frame
T_s	Shift frame

LIST OF ACRONYMS

DSP	Digital Signal Processing
QMR	Quranic Maqamat Recitation
CCA	Cepstral Coefficient Analysis
IDFT	Inverse Discrete Fourier Transform
DFT	Discrete Fourier Transform
IDCT	Inverse Discrete Cosine Transform
DCT	Discrete Cosine Transform
FIR	Finite Impulse Response
SD	Spectral Descriptors
ASR	Automatic Speech Recognition
HMM	Hidden Markov Models
HTK	HMM Toolkit
SAA	Semantic Audio Analysis
W-DFT	Warped DFT
MFCC	Mel-Frequency Cepstral Coefficient