Investigation of Robust Speech Feature Extraction Techniques for Accents Classification of Malaysian English Speakers

by

Yusnita Mohd Ali
(1040610467)

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

School of Mechatronic Engineering
UNIVERSITI MALAYSIA PERLIS

2014

UNIVERSITI MALAYSIA PERLIS

DECLARATION OF THESIS

Author's full name : YUSNITA BINTI MOHD ALI

Date of birth : 09 MARCH 1974

Title : INVESTIGATION OF ROBUST SPEECH FEATURES EXTRACTION TECHNIQUES FOR ACCENTS CLASSIFICATION OF MALAYSIAN ENGLISH SPEAKERS

Academic Session : 2013/2014

I hereby declare that the thesis becomes the property of Universiti Malaysia Perlis (UniMAP) and to be placed at the library of UniMAP. This thesis is classified as :

- [ ] **CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)*
- [ ] **RESTRICTED** (Contains restricted information as specified by the organization where research was done)*
- [x] **OPEN ACCESS** I agree that my thesis is to be made immediately available as hard copy or on-line open access (full text)

I, the author, give permission to the UniMAP to reproduce this thesis in whole or in part for the purpose of research or academic exchange only (except during a period of _____ years, if so requested above).

Certified by:

SIGNATURE

SIGNATURE OF SUPERVISOR

740309035636
(NEW IC NO. / PASSPORT NO.)

Prof. Dr. PAULRAJ M P
NAME OF SUPERVISOR

Date : 23/7/2014

Date : 23/7/2014

DR. PAULRAJ M P
Professor
School of Mechatronic Engineering
University Malaysia Perlis
02600 Ulu Pauh, Perlis, Malaysia.
Email: paul@unimap.edu.my

NOTES : * If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization with period and reasons for confidentially or restriction.

THESIS DECLARATION

ACKNOWLEDGMENT

First and foremost, I thank God the Almighty, Allah Subhana Wata'ala, for giving His blessing to me to complete this fruitful journey of doctorate study. This research is never solitary and thus, I have many people to thank for making the writing of this thesis bearable and to reach the end of my battle successfully.

I should thank the Ministry of Higher Education Malaysia and my employer, Universiti Teknologi MARA (UiTM) for giving this golden opportunity and financial support, without which I may have not pursued this study diligently. Also, I would like to acknowledge the Vice Chancellor of University Malaysia Perlis (UniMAP), Brig. Jeneral Dato' Prof. Dr. Kamaruddin Hussin, the Dean and staffs of Mechatronic Engineering of UniMAP and the Dean and staffs of Centre for Graduate Studies of UniMAP for their support and encouragement throughout my candidature.

Especially, I would like to thank my main supervisor, Prof. Dr. Paulraj Murugesa Pandiyan for the three years and nine months of supervising my doctorate study. He has been very patient and dedicated with the writing and eventual completion of this thesis. His inputs and comments were tremendous and constructive and his insight proved to be very valuable. Towards the end, I voiced out to him my tiredness of the thesis writing, and it was him who reminded me of the effort and time I had invested in this research and the meaning of holding the doctorate title is just a license to start doing any research afterwards. And for that, I shall always be indebted to him.

My deepest gratitude also goes out to Prof. Dr. Sazali Yaacob, my co-supervisor especially during the first year of my candidature, who, despite his hectic schedule as prominent lecturer and deputy vice chancellor, graciously set aside time to entertain my fearful emotion of dark ages and guide me through the literature review of this study. I deeply appreciate his genuine concern for my well-being, his truly belief in me, his steady support and his kind words of encouragement.

I wish to also thank my second co-supervisor, Dr. Shahriman Abu Bakar for his advice and support especially in my data collection process and made the Malaysian English (MalE) accent database initiation most successful. Also, many thanks go to fellow members at the Intelligent Signal Processing Cluster of UniMAP, especially to my kind friends, Sathees Kumar Nataraj and Lim Sin Chee, who, closely share their knowledge and guide me very well through real examples.

It is no denying that I went through a very stressful time during my study. During these times of enormous difficulty, my family was the most that I sacrificed. My sincerest and utmost gratitude goes to my beloved husband, Syahreen bin Alim for his kindness, allegiance and unconditional love and support by taking finest care of our children, Harith Dzakwan and Ameera Mardhiyyah and for helping me with the household chores. I am also grateful to my beloved parents, Mohd Ali bin Yusof and Nik Jah binti Abdullah, for their patience, love, understanding and endless prayers towards my success.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AAC	Automatic Accent Classification
AE	American English
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
ASR	Automatic Speech Recognition
BrE	British English
BruE	Brunei English
BS	Baseline
CALL	Computer-assisted Language Learning
CRD	Completely Randomized Design
CRs	Classification rates
dB	Decibels
DOE	Design of Experiments
DWT	Discrete Wavelet Transform
FF-MLP	Feed forward Multilayer Perceptron
FIR	Finite Impulse Response
FIS	Fuzzy Inference System
Formants	Formant frequencies
GMM	Gaussian Mixture Model
GSTs	Global Statistical Thresholds
HMM	Hidden Markov Model
Hz	Hertz
ICA	Independent Component Analysis
IWs	Isolated Words
KNN	K-nearest Neighbors
LDA	Linear Discriminant Analysis
LPC	Linear Prediction Coefficients
MalE	Malaysian English
MAP	Maximum-a-Posteori
MBSE	Mel-band Spectral Energy
MFCC	Mel-frequency Cepstral Coefficients
MLLR	Maximum Likelihood Linear Regression

MMS	Maximum-mean subtraction normalization
mse	Mean-squared errors
msec	milliseconds
MVN	Mean and variance normalization
PCA	Principal Component Analysis
PD	Pronunciation Dictionary
PPRLM	Parallel Phone Recognition Language Modeling
PRLM	Phone Recognition Language Modeling
QMF	Quadrature Mirror Filters
RFE	Recursive Feature Elimination
RP	Received Pronunciation
SBS	Statistical Band Selection
SFFs	Spectral Feature Fusions
SgE	Singapore English
SNR	Signal-to-Noise Ratio
STE	Short-time Energy
STFT	Short-time Fourier Transform
STs	Sentences
SVM	Support Vector Machine
SVM-RFE	Support Vector Machine-Recursive Feature Elimination
TTS	Text-to-Speech
V-UV	Voiced-Unvoiced
ZCR	Zero-crossing Rate

LIST OF SYMBOLS

α	Learning rate of ANN
β	Momentum rate of ANN

Kajian terhadap Teknik Penyarian Sifat Pertuturan Lasak untuk Pengelasan Loghat Bagi Penutur Bahasa Inggeris Berbangsa Malaysia

ABSTRAK

Sistem pengecaman pertuturan automatik (ASR) bukanlah suatu topik baru dalam pemprosesan pertuturan dan interaksi manusia-mesin. Ianya telah dikaji lebih daripada lima dekad lepas. Walau bagaimanapun, loghat kekal sebagai cabaran besar berkait rapat dengan kepelbagaian bahasa dalam isu-isu ASR masakini yang menggambarkan perbezaan pertuturan dalam sebutan dan intonasi seseorang yang mempunyai pelbagai perbezaan latar-belakang dari segi sosiolinguistik. Terdapat banyak dan pelbagai literatur yang telah mendedahkan kesan negatif daripada pelbagai loghat sebagai penyebab kemerosotan prestasi ASR. Walaupun loghat Bahasa Inggeris telah menjadi jenis loghat paling banyak dikaji kerana diangkat sebagai bahasa yang paling penting dan berprestij, MalE yang merupakan versi baru didalam 'New Englishes' dikalangan penutur bukan ibunda masih belum diterokai. Dalam produk pasaran ASR pada masa kini, MalE dianggap sebagai versi yang seragam secara konvensional walaupun tanggapan ini dipertikaikan oleh ramai sarjana dan penyelidik yang menganggap MalE sebagai penuturan yang terhasil daripada implikasi setempat kepelbagaian etnik. Kajian persepsi yang lepas telah melaporkan kemungkinan tinggi mengesan identiti etnik daripada penuturan Singapore English (SgE) dan Brunei English (BruE) yang boleh dijadikan perbandingan yang sesuai dengan MalE melalui ujian pendengaran. Pada masa ini, tiada kajian yang telah dilakukan untuk mengenal pasti asal usul etnik dari sampel penuturan MalE menggunakan pelbagai teknik analisis pertuturan dan algoritma pembelajaran mesin untuk pengelasan automatik yang lebih dapat dipercayai, standard dan tepat melalui kaedah eksperimen. Kajian ini merupakan satu cubaan untuk mengisi jurang tersebut dan untuk tujuan ini, satu pangkalan data baru loghat MalE telah dibina. Kajian ini merangsang sebutan jenis IWs dan STs daripada para pelajar universiti (lelaki dan perempuan) yang terdiri daripada tiga etnik utama di negara ini iaitu Melayu, Cina dan India yang mewakili para penutur berpendidikan tinggi menggunakan perkataan yang sensitif kepada loghat, dipilih daripada kajian yang lepas. Reka bentuk sistem yang dicadangkan terdiri daripada pra- pemprosesan, penyarian sifat dan pengelasan. Selain daripada pra-pemprosesan asas, kajian ini mencadangkan integrasi dengan sistem inferens kabur untuk segmentasi asas frem suara kepada bergetar-tidak bergetar (FIS V-UV) turut menyumbang kepada pelaksanaan sistem keseluruhan yang lebih baik berbanding sistem pengelasan loghat (AAC) konvensional. Satu kaedah baru yang dicadangkan yang dinamakan sebagai ambang statistik global (GSTs) untuk membina fungsi keahlian masukan-masukan tenaga pendek masa (STE) dan kadar lintasan sifar (ZCR) dalam segmentasi FIS V-UV telah mengurangkan jumlah frem yang perlu diproses di peringkat penyarian ciri. Keputusan eksperimen menunjukkan keberkesanan AAC-terbantu FIS V-UV yang dicadangkan menggunakan GSTs dengan peningkatan tertinggi dalam kadar ketepatan sebanyak 7.70% dan pengurangan frem sebanyak 24.26% berbanding AAC konvensional. Pada peringkat kedua, ciri-ciri akustik berkait rapat dengan loghat daripada tiga etnik dibangunkan melalui beberapa kaedah analisis bank-penuras, model saluran vokal, analisis hibrid dan analisis paduan. Daripada lapan vektor sifat yang diuji ke atas pangkalan data MalE, perihalan statistik tenaga spektrum jalur-Mel (MBSE), analisis komponen utama-berubah MBSE (disingkatkan sebagai PCA-MBSE), dua teknik hibrid ombak kecil diskret berubah diperolehi pekali ramalan