LB/DON/106/2016
IT 01/128

Speaker independent Sinhala Speech to Text SMS application for Mobile Phones

B.G.D. Anuradha Bopagama

139157K

LIBRARY
UNIVERSITY OF MORATUWA, SRI LANKA
MORATUWA

Dissertation submitted to the Faculty of Information Technology, University of Moratuwa, Sri Lanka for the partial fulfillment of the requirements of the Master of Science in Information Technology.

004"16"
004(043)

April 2016

University of Moratuwa
TH3163

TH 3163
+
1 DVD-ROM
(TH 3160 - TH3180)

TH 3163

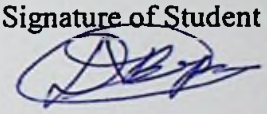
Declaration

We declare that this thesis is our own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Name of Student (s)

B.G.D.A.Bopagama

Signature of Student (s)



Date: 29/04/2016

Supervised by

Name of Supervisor(s)

G.T.I. Karunaratne

Signature of Supervisor(s)

UOM Verified Signature

Date: 29/04/2016

Dedication

To my parents and my wife for their constant support and encouragement over the years.

Acknowledgement

I would first like to thank my project supervisor Lecturer Mrs. Indika Karunaratne of the Faculty of Information Technology at University of Moratuwa Sri Lanka. She was there for me whenever I ran into a trouble or had a question about my research or writing. She consistently allowed this paper to be my own work, and steered me in the right direction whenever she thought I needed it.

I would also like to thank the Students in University of Moratuwa Information Technology Faculty who were involved in the survey and Mr. Amal Wijewardane who did a great job by coordinate with other fellow students and get the required voice recordings for this research project. Without their passionate participation and input, this project could not have been successfully conducted.

Finally, I must express my very profound gratitude to my parents, parents in law, sister in law, brother in law and to my loving wife for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Author

Anuradha Bopagama

Abstract

Speech recognition is one of the most discussed topics by researchers in recent years. Because of the limitless applications and the competition of making more user-friendly systems, lots of researchers put their effort on speech recognition system developments. There are lot of applications have already developed for English language. But for the languages like Sinhala, Hindi, Tamil are still at their preliminary stages.

The main purpose of this study was to develop a speaker independent automatic speech recognition android application for Sinhala language. Sinhala is the native language for Sri Lankans and they are the only people who speak Sinhala Language. So this study will create a great opportunity for the Sri Lankans to build their own speech recognition applications using and enhancing this language model.

The products of the study include frequently used SMS phrases speech corpus in Sinhala language which can be used in sending SMS. A survey has conducted among university students in order to collect frequently used SMS phrases. Those data will be used to create and implement the speech recognition system. Since the speech recognition training task is a time consuming one and the time limitation for the project restrict the size of the SMS phrases corpus to a limited one. At the moment audio recordings were only taken from one female and a one male and more recordings needed to build more accurate speech model.

The System was implemented using a HMM toolkit call CMUsphinx. CMUSphinx toolkit is a leading open source speech recognition toolkit with various tools used to build speech applications. CMU Sphinx toolkit has a number of packages for different tasks and applications. Pocketsphinx is one of the tools that support Android operating system which comes under CMUSphinx. Pocketsphinx tool used to create a speech model that can be used in various applications. To build the speech model it needs audio recordings of text and corresponding text. Once the model created it can be used in various applications. The main aim of this project is to build an accurate speech recognition model for Sinhala language that can be used in Android operating system.

UNIVERSITY OF MORATUWA LIBRARY

Contents

		Page
Chapter 01 - Introduction		1
1.1	Background and Motivation	2
1.2	Aim and Objectives	3
1.3	Structure of dissertation	3
1.4	Summary	3
Chapter 02 - Literature Review: Current tendency in Speech recognition systems		4
2.1	Introduction	4
2.2	Comparison	4
2.2.1	Problem derivation/statement - Research questions	7
2.3	Speech Recognition system approach	8
2.4	Hidden Markov model (HMM)-based speech recognition	9
2.5	Dynamic time warping (DTW)-based speech recognition	11
2.6	Deep Neural networking	11
2.7	Similar projects carried out	12
2.8	Summary	12
Chapter 03 - Technology enrichment of speech recognition on mobile devices		14
3.1	Introduction	14
3.2	Pocketsphinx process	14
3.3	Sinhala Speech Recognition with Pocketsphinx	16
3.4	Summary	17
Chapter 04 - Approach to Sinhala speech recognition Using Pocketsphinx		18
4.1	Introduction	18
4.2	Proposed setup	18
4.3	Survey	19
4.4	Android Testing Environment	20
4.5	Android interface development	20
4.6	Recording	20
4.7	Summary	21
Chapter 05 - Design of Sinhala speech recognition system for Android OS		22
5.1	Introduction	22
5.2	Speech recognition design	22

5.3 Data preparation 23
5.4 Recordings 24
5.5 Training 25
5.6 Design assumptions and dependencies 26
5.7 Summary 26
Chapter 06 - Implementation process of Speech Recognition System 27
6.1 Introduction 27
6.2 Collecting data set 27
6.3 Creating Language Model File 27
6.4 Creating Phonetic Dictionary File 28
6.5 Creating other files 28
6.6 Training 28
6.7 Creating Android Environment 29
6.8 Summary 29
Chapter 07 - Evaluation 31
7.1 Introduction 31
7.2 During training process 31
7.3 Testing on Real Application 31
7.4 Testing Conditions 32
7.5 Assumptions 33
7.6 Results 33
Chapter 08 - Conclusion and Future Work 34
8.1 Introduction 34
8.2 Completed work and future work 34
8.3 Summary 36
References 37
Appendix A 40
Appendix B 41
Appendix C 42
Appendix D 43
Appendix E 45
Appendix F 46
Appendix G 48
Appendix H 49
Appendix I 50

Appendix J 51
Appendix K 52

List of Tables

Table 2.1: Summarization 08

Table 4.1: Android test environment details 20

Table 5.1: CMUSphinx data file details 24

Table 7.1: Group of people use to test the system 32

Table 7.2: Results 33

List of Figures

Figure 2.1: Search Graph 10

Figure 3.1: Pocketsphinx Architecture 16

Figure 5.1: Speech recognition system design 22

Figure 5.2: CMUSphinx Training Process 25