

LB / DON / 106 / 2016

IT 01 / 26

Speech Recognition Research for Recognize Digits in Sinhala

LIBRARY
UNIVERSITY OF MORATUWA, SRI LANKA
MORATUWA

R.M.D.T. Asiri

139154A

004 "16"

004 (013)

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

University of Moratuwa

April 2016



TH3161

TH 3161

+ 1 DVD ROM

(TH 3160 - TH 3180)

TH3161

Declaration

I declare that this thesis is my 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 Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.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 the student: Asiri RMDT

Student Number: 139154A

Signature of the student: Asiri

Date: 29/04/2016

Supervised By

Name of the supervisor: Mrs. Indika Karunaratne

Signature of the supervisor: **UOM Verified Signature**

Date: 29/04/2016

Dedication

This Dissertation is dedicated to my loving Family for being part of me and encouraging me always being by my side.

Acknowledgement

First I express my heartfelt appreciation and gratitude to my supervisor Mrs. Indika Karunaratne for his most valued guidance, commitment and kind support to make this research success

Also I express my kind appreciation to our MSc Coordinator Mr. Saminda Premaratne to for his guidance and support and given flexibility with some timelines in order to complete the research activities

Also sincere appreciation is extended to Eng. Mr. Kanishka Jayasekara, Deputy General Manager (Information Technology) for his valuable opinions and encouragement given to me in this endeavor

It is my great pleasure to thank all the other Senior lecturers, Lecturers, Instructors, and staff members who helped us in many ways to make this research. The guidance and support received from all the members who contributed and who are contributing to this project, was vital for the success of the Research.

I would also like to thank my all the batch mates of the MSc in Information Technology batch 7 in faculty of Information Technology for their various help and support. And also other friends of the faculty as well as friends for outside who gave me support and encourage me with their best wishes.

Abstract

There is growing tendency in the human computer interaction activities especially within last decade with smart world. Day by day lot of smart options are coming to the market which has more powerful human computer interactive options. With this kind of an environment speech recognition is also became very vast and interested research area among modern computer researches. When we considering speech recognition there are lot of smart applications even in smart phones are available all around the world right now, but major problem of those application are the accuracy and the localization. Especially considering a language like Sinhala, current industry doesn't have much accurate of efficient recognition system to cater with. This research is basically focuses on how to identify user audio signals in Sinhala language and how to convert them in to text. Speech recognition becomes very popular research topic with the highly incremental human computer interactions today. Most of the popular or vastly using languages are already have well developed speech recognition systems, but as mentioned earlier languages like Sinhala it is very rare, but for Sri Lanka it is very useful to have a recognition system.

Ceylon Electricity Board (CEB) is the sole agent to generate and distribute electricity power within the Sri Lanka. So it has several call centers running around the country to get customer feedbacks especially regarding power failures. Recently CEB management came with an idea to employ some disable (blind) people to those call centers to collect customer complains thru telephone line and log them into a call center web application and then that application will process those complains to forward to correct maintenance party to attend to the problem quickly as possible. In order to get the input first idea was to use a brail keyboard; same time management has an idea to get the input thru a microphone. But that audio input needs to be in Sinhala language. So now our problem is to develop good speech recognition for Sinhala language to identify those vocal signals. As part of solution to the above mention problem, this research is conducting to create well trained Sinhala identification system. In this phase research is focus only to identify digit's vocal inputs created using Sinhala language. Than can be used to identify customer according to his/her electricity account number. Once that is done call center application can validate the customer using customer billing database.

Table of Contents

Declaration	ii
Dedication	iii
Acknowledgement	iv
Abstract.....	v
List of Figures.....	viii
List of Tables	viii
1. Introduction.....	1
2. Review of others work – Literature Review	2
2.1. Introduction.....	2
2.2. Classification of Speech Recognition Systems.....	2
2.3. Speech Recognition in MATLAB [1].....	2
2.4. Speech Detection Algorithm	4
2.5. Acoustic Model Development	5
2.6. Google Speech Recognition Researches [2].....	7
2.7. Automatic Pronunciation Verification for Speech Recognition [3].....	8
2.8. Open Source Speech Recognition Toolkit –Source forge Project [4].....	8
3. Technology Adapted: Sphinx 4.....	10
3.1. Introduction.....	10
3.2. Why Sphinx 4?.....	10
3.3. Overview [5]	10
3.4. Acoustic Model.....	12
3.6. Language Model	13
3.6.1. JSGF (Java Speech Grammar Format) [8].....	13
3.6.2. Dictionary	13
3.7. Summary.....	14
4. Research Approach.....	15
4.1. Introduction.....	15
4.2. Approach in Steps	15
4.3. Obstacles faced and how they overcome with this Approach?.....	16
5. Analysis and Design	17
5.1. Introduction.....	17
5.2. Abstract of the System Design.....	17

5.3.	Functional overview of the system	17
5.4.	System Architecture and structure	18
5.4.1.	Architecture.....	18
5.5.	Detail System Software Design	19
5.5.1.	Use Cases	19
6.	Implementation	22
6.1.	Introduction.....	22
6.2.	Major Parts of the Implementation	22
6.6.	Outline Architecture in Brief	23
6.7.	Logical View of the Implementation	23
6.8.	Class Description	24
6.9.	Activity Diagram	26
6.9.1.	User Training Activity	26
6.9.2.	Recognizing speech Activity	26
6.10.	Sequence Diagram for Decoding Speech.....	27
6.11.	Deployment Architecture.....	28
7.	Evaluation	29
8.	Discussion	29
9.	Conclusion	30
10.	Future Work	30
11.	References	31
12.	Appendixes	32

List of Figures

Figure 2-1 - MATLAB Code Sample	3
Figure 2-2- Zero Crossing	4
Figure 2-3- PSD estimate of three different utterances of the word "ONE."	5
Figure 2-4 PSD estimate of three different utterances of the word "TWO."	6
Figure 3-1 - Sphinx 4 - Architecture	12
Figure 4-1- Research Approach in brief	16
Figure 5-1- Frontend Behavior of the System	18
Figure 5-2- Level 0 Use case	19
Figure 5-3- Level 1: Use Case for "Initiate New User Profile"	20
Figure 5-4 -Level 1: Use Case for "Recognize Digits"	21
Figure 6-1- Main Class Diagram	24
Figure 6-2 - User Training Activity Diagram	26
Figure 6-3 - Recognizing Speech Activity Diagram	27
Figure 6-4 - Decoding Speech Sequence Diagram	28

List of Tables

Table 3-1- Sample Dictionary	14
------------------------------	----