

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Impaired Speech Recognition

A thesis presented in partial fulfilment of the
requirements for the degree of

Master of Information Sciences
in
Computer Science

Massey University, Albany, New Zealand.



Mohammed Nasser Almujiil

2015

Abstract

The purpose of this thesis is to present a novel mobile health application that can recognize impaired speech (using audio signals) and turn it into understandable speech. The system is developed to help Dysarthria of Speech patients communicate better with others in their everyday life. It will provide some background information about motor speech disorders, Dysarthria of Speech and the technical aspects of this application. It will then explain and test the algorithms to recognize impaired speech using audio fingerprinting technology. Finally it will discuss the test results and recommends some future work to improve the current algorithms.

Acknowledgements

I would like thank my project supervisor, Associate Professor Ruili Wang, for his advice and support throughout the course of this project, and for giving me the opportunity to write about this interesting topic.

Contents

Abstract	2
Acknowledgements	3
1 Introduction	7
1.1 Moto speech disorders	7
1.2 Dysarthria	8
1.3 Current situation	9
1.4 Objectives	9
1.5 Speech Assistant Introduction	10
2 Literature review:	12
2.1 Audio Fingerprinting	12
2.2 Audio Fingerprinting alternatives	14
2.2.1 Audio signal	15
2.2.2 Database	15
2.2.3 Robustness	16
2.3 Audio fingerprinting system properties	17
2.4 Audio fingerprint extraction and matching algorithms	18
2.4.1 Open source library: MusicG	19
2.4.2 Audio fingerprinting systems	19
2.5 Knowledge storage	21
2.6 Audio filters	22
2.7 Android	22
2.8 EEG	23
2.9 Human Computer Interaction	25
2.10 Summary	27
3 Methodology:	28
3.1 Training process	30

3.1.1	Phrase selection.....	30
3.1.2	Read selected phrase.....	30
3.1.3	Audio recording.....	31
3.1.4	Fingerprint extraction and organization.....	32
3.1.5	Fingerprint storage.....	35
3.2	Detection process	36
3.2.1	Start detection process	36
3.2.2	Recording and Fingerprint extraction.....	37
3.2.3	Detection algorithms.....	37
3.3	Summary	40
4	Application test	41
4.1	Quiet place.....	43
4.2	Slightly Noisy place	44
4.3	Summary	45
5	Discussion	46
5.1	Machine learning.....	46
5.2	Recording quality	47
5.3	False positive threshold.....	48
5.4	Scalability.....	48
5.5	Fingerprint extraction and audio analysis	49
5.6	Application Test	50
5.7	Human computer interaction.....	51
5.8	Summary	51
6	Conclusion and future work.....	52
6.1	Discussion	52
6.2	conclusion.....	53
	References	54

List of Figures

Figure 1: Neurologic communication disorders. Retrieved from: [3].....	8
Figure 2: Application framework.....	29
Figure 3: Phrase selection	30
Figure 4: Listening view.	31
Figure 5: Fingerprint extraction and organization.	34
Figure 6: Fingerprint storage.....	35
Figure 7: Start Detection Process.....	36
Figure 8: Audio capture in the detection process.....	37
Figure 9: Fingerprint comparison.....	39
Figure 10: Result view	40