

**SPEECH-ACTIVATED TELEPHONE DIRECTORY**

**ASSISTANCE**

**MALI H. HAKEM ALAMEADY**

**UNIVERSITI UTARA MALAYSIA**

**2007**

# **Speech-Activated Telephone Directory Assistance**

A thesis submitted to the Graduate School in partial fulfilment of the requirements  
for the degree Master of Science (Intelligent System)

Universiti Utara Malaysia

By

MALI H. HAKEM ALAMEADY

© MALI H. HAKEM ALAMEADY, 2007

All rights reserved.



**PUSAT PENGAJIAN SISWAZAH**  
**(Centre For Graduate Studies)**  
**Universiti Utara Malaysia**

**PERAKUAN KERJA KERTAS PROJEK**  
**(Certificate of Project Paper)**

Saya, yang bertandatangan, memperakukan bahawa  
*(I, the undersigned, certify that)*

**MALI H. HAKEM**

calon untuk Ijazah  
*(candidate for the degree of)* **MSc. (Intelligent Systems)**

telah mengemukakan kertas projek yang bertajuk  
*(has presented his/her project paper of the following title)*

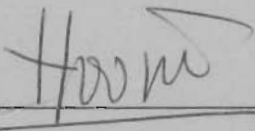
**SPEECH ACTIVATED TELEPHONE DIRECTORY ASSISTANCE**

seperti yang tercatat di muka surat tajuk dan kulit kertas projek  
*(as it appears on the title page and front cover of project paper)*

bahawa kertas projek tersebut boleh diterima dari segi bentuk serta kandungan  
dan meliputi bidang ilmu dengan memuaskan.  
*(that the project paper acceptable in form and content, and that a satisfactory  
knowledge of the field is covered by the project paper).*

Nama Penyelia Utama  
*(Name of Main Supervisor):* **MISS NOORAINI YUSOFF**

Tandatangan  
*(Signature)*

:   
\_\_\_\_\_

Tarikh  
*(Date)*

: 31/5/2007  
\_\_\_\_\_

## **PERMISSION TO USE**

In presenting this thesis in partial fulfilment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the University Library may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor(s) or, in their absence by the Dean of the Graduate School. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part, should be addressed to

**Dean of Faculty Information Technology**

**Universiti Utara Malaysia**

**06010 UUM Sintok**

**Kedah Darul Aman.**

## ABSTRACT

*In everyday life people are all liable to find themselves in an emergency situation and, more commonly, will require to be alerted to signals that give a warning or indication of action to be taken. The means to access services and equipment to ensure safety and comfort are well known, readily available and usually at little or no cost. Majority of public emergency services are accessed by telephone for example the access to the Fire, Police and Ambulance services by picking up the telephone and dialing 994, 999 etc. However, less-able people such as elderly, disabled, children and lower-educated may not be able to use the services that are available because of disability or find equipment that will ensure personal safety and comfort. Hence, this study aims to propose a speech activated telephone assistance K-Nearest Neighbours. User only needs to give command through voice, and then the system will assist the caller to search the telephone directory and dial the required phone number. This can fasten the dialing as the user no need search or recall the number. The system include three main module namely, Speech Features Extraction Module; that extracts the speech features, Speech Recognition Module; that classifies spoken word ("Fire", "Ambulance" and "Police") and Automated Dialing; that searches and dials the number that matched with the spoken word.*

## ACKNOWLEDGEMENT

*By the Name of Allah, the Most Gracious and the Most Merciful*

First, I would like to express my appreciation to Allah, the Most Merciful and, the most compassionate who has granted me the ability and willing to start and complete this study. I do pray to his greatness to inspire and enable me to continue the work for the benefits of humanity.

My most profound thankfulness goes to my supervisor Miss. Nooraini Yusoff for his scientifically proven and creativity encouraging guidance.

Last but not least, i wish to thank all my dearest family members, especially Dad, Mum, and my husband Mohmoud for being by my side since my daughters Zainab, Mariam, Safa, my son Mohamed, my Sisters, and my brothers left Home. To my lecturers and friends who have given me emotional support during my study.

**Thank you UUM.**

## TABLE OF CONTENT

<b>PERMISSION TO USE .....</b>	<b>i</b>
<b>ABSTRACT.....</b>	<b>ii</b>
<b>ACKNOWLEDGEMENT .....</b>	<b>iii</b>
<b>TABLE OF CONTENTS.....</b>	<b>iv</b>
<b>LIST OF TABLES.....</b>	<b>v</b>
<b>LIST OF FIGURES.....</b>	<b>viii</b>
<b>LIST OF ABBREVIATIONS .....</b>	<b>ix</b>
<b>CHAPTER ONE: INTRODUCTION.....</b>	<b>1</b>
1.1 Background.....	1
1.2 Problem Statement.....	3
1.3 Research Objectives and Scope of The Study .....	4
1.4 Significance of the Study.....	5
1.5 Organization of Report .....	6
1.6 Summary.....	7
<b>CHAPTER TWO: LITERATURE REVIEW.....</b>	<b>8</b>
2.1 Speech Recognition .....	8
2.2 Applications Of Speech Recognition.....	10
2.3 Speech Feature Extraction .....	12
2.4 Speech Recognition Techniques .....	14
2.5 Summary.....	19
<b>CHAPTER THREE: METHODOLOGY .....</b>	<b>20</b>
3.1 Overview Of The Methodology.....	20

3.1.1	Phase 1: Awareness Of The Problem .....	21
3.1.2	Phase 2: Suggestion .....	22
3.1.3	Phase 3: Development .....	25
3.1.4	Phase 4 and 5: Evaluation And Conclusion.....	27
3.2	Summary.....	28
<b>CHAPTER FOUR: FINDINGS AND DISCUSSION .....</b>		<b>29</b>
4.1	System Design .....	29
4.2	System Development .....	31
4.2.1	Data Acquisition .....	32
4.2.2	Data Description .....	32
4.3	Screenshots:Speech Recognition System .....	33
4.3.1	Speech Feature Extraction Module.....	33
4.3.2	Speech Recognition Module.....	36
4.3.3	Automated Dialing.....	37
4.4	Experimental.....	39
4.5	Summary.....	43
<b>CHAPTER FIVE: CONCLUSION.....</b>		<b>44</b>
5.1	Findings .....	44
5.2	Contribution Of Study .....	46
5.3	Future Work.....	46
<b>REFERENCES .....</b>		<b>47</b>
<b>APPENDIX</b>		
A:	Source Code.....	51



## LIST OF TABLES

No	Title	Page
4.1	The Description of main attributes	32
4.2	Results of spoken word "Fire"	40
4.3	Results of spoken word "Police"	41
4.4	Results of spoken word "Ambulance"	42

## LIST OF FIGURES

No	Title	Page
1.1	Overview of system	5
2.1	Typical speech recognition processes	9
3.1	The General Methodology of Design Research	21
3.2	Use Case Diagram View	23
3.3 A	General Idea of Speech-activated Telephone Directory Assistance	24
3.3 B	Feature Extraction	25
3.4	K-Nearest Neighbours Algorithm	26
4.1	Use Case Diagram View	34

4.2	Interface for Feature Extraction	34
4.3	Extracted Speech Features	35
4.4	Sample of words	37
4.5	Interface for Automated Dialing Facility	38
4.6	Automated Dialing	39

## LIST OF ABBREVIATIONS

<b>AI</b>	Artificial Intelligence
<b>ANN</b>	Artificial Neural Network
<b>CBR</b>	Case -Based Reasoning
<b>KNN</b>	K-Nearest Neighbours
<b>SR</b>	Speech Recognition

# **CHAPTER 1**

## **INTRODUCTION**

This chapter briefly explains the background of the study that mainly involves the integration between K-Nearest Neighbours (KNN), Speech Feature Extraction to help user which needs to give command through voice, and then the system will assist the caller to search the telephone directory and dial the required phone number. The problem statement, objectives, significance of the project and scopes will also be introduced.

### **1.1 Background**

Speech recognition (SR) has developed very rapidly over the last ten years. This has largely been the result of advances in computer processing power combined with correspondingly rapid advances in speech recognition algorithms. As computer processing power increased and became more widely available and less costly, the

## REFERENCES

- AILIA Inc. (2004). The Canadian Speech Processing Industry. Report of Technology Roadmap (2003-2007). Available at:  
[http://www.ailia.ca/documentVault/Rpts/The\\_Canadian\\_Speech\\_Processing\\_Industry.pdf](http://www.ailia.ca/documentVault/Rpts/The_Canadian_Speech_Processing_Industry.pdf)
- Belkin, N., & Croft, B. (1992). "Information filtering and informatioun retrieval". *Communications of the ACM*,35, pp. 29-37.
- Bailey,j. & Bakos, y.,(1997). "An Exploratory Study of the Emerging Role of Electronic Intermediaries." *InternationalJournal of Electronic Commerce* , 1(3), Spring 1997.
- Franco, H., Weiniruub, M. & Cohen, M.,(1997). "Context Modeling in a Hybrid HMM-Neural Net Speech recognition System", *Proceedings of the International Conference on Neural Networks, Houston, TX, 1997*

- Fukuda, S. and V. Kostov, (1999)“Extracting Emotion from Voice,” *Proceedings of IEEE International Workshop on Systems, Man, and Cybernetics*, vol. 4, 1999, pp.299-304.
- Kjell E. & Gyorgy T. ,(2000). Acoustic-phonetic recognition of continuous speech by artificial neural networks. Technical Report STL-QPSR 2-3/1990, Institutionen för tal, musik och hörsel, KTH, Stockholm, Sweden, 2000.
- Gosztolya,G A. Kocsor, L. & Felf,l.,(2003). *Various Robust Search Methods in a Hungarian Speech Recognition System*, Acta Cybernetica 16., pp. 229- 240., 2003.
- Golbeck, J. & Hendler, J. ,(2004). *Reputation Network Analysis for Email Filtering* .University of Maryland, College Park . from Available at: <http://www.ceas.cc/papers-2004/177.pdf>
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). *Design Science in Information Systems Research*. *MIS Quarterly*, 28(1), pp. 75-105. Available at: <http://www.misq.org/PLS2.Call.pdf>
- Jelinek ,F,(2001). *Statistical Methods for Speech Recognition*, The MIT Press, 1997.7.
- X. Huang, A. Acero and H.-W. Hon, Spoken Language Processing, Prentice Hall PTR, 2001.

- Robert J. R., Kelly, R. , Abu-Amer, T. ,& Walsh, M. (2001). *Speaking Autonomous ntelligent ,Devices* Department of Computer Science, University College Dublin, Ireland,Retrieved on junuary-2007.
- Sadeh, N. ,Eriksson, J. , Finne, N. & Janson, S. ,(2003). *TAC-03: A supply-chain trading competition*. *AI Magazine*, 24(1):92-94.
- Kocsor, L. T\_oth, A., Kuba Jr., K., Kov\_acs, M., Jelasity, T., Gyim\_othy & Csirik,,J. (2000). *A Comparative Study of Several Feature Space Transformation and Learning Methods for Phoneme Classi\_cation*, *International Journal of Speech Technology*, Vol. 3, Number 3/4, pp. 263-276, 2000.
- Ishikawa, Sh. & Ikeda, T.,(2004). *Speech-Activated text retrieval system for multimodal cellular phones*, *Proceedings of IEEE* . PP.1-453 -1-456.
- Lesser, V. , Horling, V., Klassner, V. , Raja, A. Wagner, T. & Zhang, S. , X. ,(2000). *Big: an agent for resource-bounded information gathering and decision making*. *Artif. Intell.*, 118(1-2):197-244.
- Nouza, J., Nouza, T.,(2004). *A Voice Dictation System for a Million-Word Czech Vocabulary*. Department of Electronics and Signal Processing Technical University of Liberec, pp.149-152.



Wu ,Ch .& Chen , J,(1997). *Speech Activated Telephony Email Reader based on speaker verification and text-to-speech conversion*, *Proceedings of IEEE Transactions on Consumer Electronics*, pp. 707-716 , Vol. 43, No. 3, AUGUST 1997.