STAR LRT GUIDELINE USING SPEECH RECOGNITION

KWAN SWEE YEE

UNIVERSITI TEKNIKAL MALAYSIA MELAKA
BORANG PENGESAHAN STATUS TESIS

JUDUL: STAR LRT GUIDELINE USING SPEECH RECOGNITION

SESJ PENGAJIAN: 2009/2010

Saya Kwan Swee Yee mengaku membenarkan tesis (PSM/Sarjana/Doktor Falsafah) ini disimpan di Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dengan syarat-syarat kegunaan seperti berikut:

1. Tesis dan projek adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. ** Sila tandakan (/)

____ SULIT (Mengandungi maklumat yang berdaerah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)

____ TERHAD (Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

_/ TIDAK TERHAD

(TANDATANGAN PENULIS)
Alamat tetap: 35, Lorong 10/115c
Taman Kuchai Jaya, 58200 Kuala Lumpur.
Tarikh: 25/1/2010

(TANDATANGAN PENYELIA)
DR. SYARIFFANOR BINTI HISHAM

Tarikh: 25/6/2010

CATATAN: *
Tesis dimaksudkan sebagai Laporan Akhir Projek Sarjana Muda (PSM)

** Jika Tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa

© Universiti Teknikal Malaysia Melaka
STAR LRT GUIDELINE USING SPEECH RECOGNITION

KWAN SWEE YEE

This report is submitted in partial fulfillment of the requirements for the Bachelor of Computer Science (Computer Media Interactive)

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY
UNIVERSITI TEKNIKAL MALAYSIA MELAKA
DECLARATION

I hereby declare that this project report entitled

PUBLIC TRANSPORTATION GUIDANCE THROUGH SPEECH RECOGNITION

is written by me and is my own effort and that no part has been plagiarized without citations.

STUDENT: [Signature] Date: 25/6/2010

(KWAN SWEE YEE)

SUPERVISOR: [Signature] Date: 25/6/2010

(DR. SYARIFFANOR BINTI HISHAM)
DEDICATION

Hereby, I would like to show my dedication to my parents, Dr. Syariffanor binti Hisham (supervisor) and friends who have given me support and motivated me in accomplishing my PSM. Thank you for all throughout the completion of my project.
ACKNOWLEDGEMENTS

Upon accomplishment of the project, I would like to take this opportunity to express my appreciation and gratitude to my supervisor, Dr. Shariffanor binti Hisham, for her immeasurable advices and guidance throughout the project. With her advice and guidance, I managed to solve most of the problem that I had surfaced during the development of this project as well as completing before the deadline. Besides, I would also like to show my gratitude to my senior, Lau Kum Hoe, for spending his time on sharing his information and knowledge about speech recognition technology with me. The information and advices that he gave to me were very helpful in completing my final year project. Furthermore, I would also like to thanks my family for their encouragement and caring when I am proceeding with my final year project. Without their encouragement, I might not able to withstand and endure the stress level given by the project and this might lead to incompletion of the project. On the other hand, I would also like to apologize to those who I had interacted with if I had caused any unpleasant feeling or troubles during the process of the project. Last but not least, I would like to thank those who I did not mentioned above which have helped me in completing my project as well as readers who spend their time reading my report.
ABSTRACT

This project is focused on the technology of speech recognition and method to integrate the application with speech recognition software. Currently there are many websites provide the service of searching of location on map. However, it is lack of interactivity and inconvenient for user to search the specific location through the map. Thus, this project aimed to provide more interactive and convenient way for location searching. This project was developed using the System Development Life Cycle (SDLC) methodology which consists of planning, analysis, design, implementation and maintenance. Besides, hardware and software such as microphone, speaker, mouse, keyboard, workstation, Microsoft Speech Recognition, Microsoft SDK 5.1, Adobe Audition 1.5, Microsoft Visual Studio 2005, Microsoft Access 2007 and Adobe Photoshop CS3 have been chosen to develop the whole application. Furthermore, designs of the application such as architecture design, preliminary design and user interface design have also being sketched in chapter four. Next, the method of implementing and installing the application are being discussed in chapter five. In chapter six, target user for testing such as officer has been chosen in performing the test on application and it is found 80% of the users successfully control the application using speech. Lastly, the observation on strength and weakness of project and its proposition of improvement has been discussed in chapter seven as well as the contribution and overall conclusion of the project.
ABSTRAK

# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>SUBJECT</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DECLARATION</td>
<td>ii</td>
</tr>
<tr>
<td></td>
<td>DEDICATION</td>
<td>iii</td>
</tr>
<tr>
<td></td>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td></td>
<td>ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td></td>
<td>ABSTAK</td>
<td>vi</td>
</tr>
<tr>
<td></td>
<td>TABLE OF CONTENTS</td>
<td>vii</td>
</tr>
<tr>
<td></td>
<td>LIST OF TABLES</td>
<td>xii</td>
</tr>
<tr>
<td></td>
<td>LIST OF FIGURES</td>
<td>xiii</td>
</tr>
<tr>
<td></td>
<td>LIST OF ABBREVIATIONS</td>
<td>xv</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER I</th>
<th>INTRODUCTION</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.1 Project Background</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.2 Problem Statements</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1.3 Objective</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1.4 Scope</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1.5 Project Significance</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>1.6 Conclusion</td>
<td>3</td>
</tr>
</tbody>
</table>
## CHAPTER II  LITERATURE REVIEW & PROJECT METHODOLOGY

2.1 Introduction  
2.2 Domain  
  2.2.1 Speech Recognition  
2.3 Existing System  
  2.3.1 Comparison of Existing System  
  2.3.2 Comparison of Programming Languages  
2.4 Project Methodology  
  2.4.1 Planning Phase  
  2.4.2 Analysis Phase  
  2.4.3 Design Phase  
  2.4.4 Implementation Phase  
  2.4.5 Maintenance Phase  
2.5 Project Requirements  
  2.5.1 Software Requirements  
  2.5.2 Hardware Requirements  
2.6 Conclusion

## CHAPTER III  ANALYSIS

3.1 Current Scenario Analysis  
3.2 Requirements Analysis  
  3.2.1 Project Requirements  
    3.2.1.1 Need Analysis  
    3.2.1.2 User Analysis  
    3.2.1.3 Technical Analysis  
    3.2.1.4 Speech Recognition Software  
    3.2.1.5 Resource Analysis
3.2.1.6 Information Gathering 26
3.2.2 Software Requirement 26
3.2.3 Hardware Requirement 27
3.3 Project Schedule and Milestone 27
3.4 Conclusion 29

CHAPTER IV DESIGN
4.1 Introduction 30
4.2 System Architecture 30
4.3 Preliminary Design 33
   4.3.1 Storyboard Design 33
4.4 User Interface Design 38
   4.4.1 Navigation Design 38
   4.4.2 Input Design 38
   4.4.3 Output Design 39
4.4.4 Database Design 39
   4.4.5 Metaphor 40
4.5 Conclusion 40

CHAPTER V IMPLEMENTATION
5.1 Introduction 41
5.2 Media Creation 41
   5.2.1 Production of Texts 42
   5.2.2 Production of Graphic 43
   5.2.3 Production of Audio 44
5.3 Media Integration 44
5.4 Product Configuration Management 46
   5.4.1 Configuration Management Setup 47
5.4.2 Version Control Procedure
  5.4.2.1 Version 1: Test the database with application 50
  5.4.2.2 Version 2: Test the Microsoft Speech Recognition with application 50
  5.4.2.3 Version 3: Check for the deficiency of error of application 50

5.5 Implementation Status 51
  5.5.1 Create Database 51
  5.5.2 Record Sound 51
  5.5.3 Program Code 52
  5.5.4 Prepare Test Plan 52
  5.5.5 Perform Testing 53

5.6 Conclusion 54

CHAPTER VI TESTING AND EVALUATION

6.1 Introduction 55

6.2 Test Plan 55
  6.2.1 Test User 56
  6.2.2 Test Environment 56
  6.2.3 Test Schedule 57
  6.2.4 Test Strategy 57

6.3 Test Implementation 57
  6.3.1 Test Description 58
  6.3.2 Test Data 58
  6.3.3 Test Result and Analysis 59
  6.3.4 Analysis Testing 60

6.4 Conclusion 64
CHAPTER VII  PROJECT CONCLUSION

7.1 Observation on Strengths and Weaknesses  65
  7.1.1 Strengths  65
  7.1.2 Weaknesses  66
7.2 Proposition for Improvement  66
7.3 Contribution  68
7.4 Conclusion  68

REFERENCES  69

BIBLIOGRAPHY  70

APPENDIX A: EXAMPLE OF QUESTIONNAIRE

APPENDIX B: USER MANUAL
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Comparison of Existing Services</td>
<td>7</td>
</tr>
<tr>
<td>2.2</td>
<td>Comparison between Java and VB.net</td>
<td>9</td>
</tr>
<tr>
<td>2.3</td>
<td>List of Software</td>
<td>14</td>
</tr>
<tr>
<td>2.4</td>
<td>List of Hardware</td>
<td>15</td>
</tr>
<tr>
<td>3.1</td>
<td>Comparisons of Speech Recognition Software</td>
<td>21</td>
</tr>
<tr>
<td>4.1</td>
<td>Table of Input Design</td>
<td>38</td>
</tr>
<tr>
<td>4.2</td>
<td>Table of Output Design</td>
<td>39</td>
</tr>
<tr>
<td>4.3</td>
<td>Database design of Station Table</td>
<td>39</td>
</tr>
<tr>
<td>4.4</td>
<td>Database design of Place Table</td>
<td>39</td>
</tr>
<tr>
<td>5.1</td>
<td>Text Used in Project</td>
<td>42</td>
</tr>
<tr>
<td>5.2</td>
<td>Description of status in create database</td>
<td>51</td>
</tr>
<tr>
<td>5.3</td>
<td>Description of status in audio production</td>
<td>52</td>
</tr>
<tr>
<td>5.4</td>
<td>Description of status in programming application</td>
<td>52</td>
</tr>
<tr>
<td>5.5</td>
<td>Description of status in preparing test plan</td>
<td>53</td>
</tr>
<tr>
<td>5.6</td>
<td>Description of status in performing testing</td>
<td>53</td>
</tr>
<tr>
<td>6.1</td>
<td>Degree of Agree</td>
<td>59</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>DIAGRAM</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Model of SDLC</td>
<td>11</td>
</tr>
<tr>
<td>3.1</td>
<td>Flow Chart of Searching Location through Google Map</td>
<td>17</td>
</tr>
<tr>
<td>3.2</td>
<td>Flow Chart of Calculate Distance Between Location A and Location B</td>
<td>18</td>
</tr>
<tr>
<td>3.3</td>
<td>Gantt Chart for Project Development</td>
<td>28</td>
</tr>
<tr>
<td>4.1</td>
<td>System Architecture of Star LRT Guideline Using Speech Recognition</td>
<td>31</td>
</tr>
<tr>
<td>4.2</td>
<td>Flow Chart of Star LRT Guideline Using Speech Recognition</td>
<td>32</td>
</tr>
<tr>
<td>4.3</td>
<td>Splash Screen for Starline LRT Guidance</td>
<td>34</td>
</tr>
<tr>
<td>4.4</td>
<td>frmSearch for Starline LRT Guidance</td>
<td>35</td>
</tr>
<tr>
<td>4.5</td>
<td>Help for Starline LRT Guidance</td>
<td>36</td>
</tr>
<tr>
<td>4.6</td>
<td>frmResult for Starline LRT Guidance</td>
<td>37</td>
</tr>
<tr>
<td>4.7</td>
<td>Navigation flow of Starline LRT Guidance</td>
<td>38</td>
</tr>
<tr>
<td>5.1</td>
<td>Process of Audio Production</td>
<td>44</td>
</tr>
<tr>
<td>5.2</td>
<td>Process of Using Microsoft Voice</td>
<td>44</td>
</tr>
<tr>
<td>5.3</td>
<td>Summarized of the process of media integration</td>
<td>46</td>
</tr>
<tr>
<td>5.4</td>
<td>Choose ‘Open the Speech Dictionary’</td>
<td>48</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>DIAGRAM</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.5</td>
<td>Select ‘Add a new word’</td>
<td>48</td>
</tr>
<tr>
<td>5.6</td>
<td>Insert the place name</td>
<td>48</td>
</tr>
<tr>
<td>5.7</td>
<td>Check ‘Record a pronunciation upon Finish’</td>
<td>49</td>
</tr>
<tr>
<td>5.8</td>
<td>Record pronunciation of word</td>
<td>49</td>
</tr>
<tr>
<td>6.1</td>
<td>Percentage of Success through Speech Control</td>
<td>60</td>
</tr>
<tr>
<td>6.2</td>
<td>Average rating for ease of users by each tester</td>
<td>61</td>
</tr>
<tr>
<td>6.3</td>
<td>Average rating for efficiency by each tester</td>
<td>61</td>
</tr>
<tr>
<td>6.4</td>
<td>Average rating for interface design by each tester</td>
<td>62</td>
</tr>
<tr>
<td>6.5</td>
<td>Average rating for satisfaction by each tester</td>
<td>63</td>
</tr>
<tr>
<td>6.6</td>
<td>Choice of preferable input</td>
<td>64</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

PSM - Projek Sarjana Muda
SDLC - System Development Life Cycle
LRT - Light Rail Transit
SDK - Software Development Kit
GUI - Graphical User Interface
API - Application Program Interface
SRGS - Speech Recognition Grammar Specification
SSML - Speech Synthesis Markup Language
CHAPTER I

INTRODUCTION

1.1 Project Background

This project focused on the domain of speech recognition technology. Currently, there are many websites provide services to search for the location on map and the transportation that able to travel user to the destination. User can find the route on railway station on the map surrounded with the location nearby station through map. However, it is not convenient and user friendly as in user is required to search for the accurate information themselves through the area of map that as an output of the search. Furthermore, the interactive between user and the system is at minimum. Thus, Star LRT Guideline Using Speech Recognition System that focused on Starline LRT which have been utilize the speech recognition technology will be able to solve problem of lack of interactive of current system as well as provide a more convenient way for user to search for their desired information.
1.2 Problem Statements

Nowadays, as the current LRT information retrieval system LRT does not provide the support for technology of speech recognition and hence it is said to be lack of the choice of interactive. Besides, passenger does not know the specific information about transportation they should take in order to travel from one place to another. Furthermore, it is time consuming for users to queue up to make enquires about transportation information. Lastly, the current system is said to be not disabled-friendly especially for those who have difficulty in writing.

1.3 Objective

- To study on the technology of speech recognition.
- To design and develop an application that uses the technology of speech recognition.
- To enhance the technology of speech recognition on Starline LRT station in order to guide user on the LRT transportation.

1.4 Scope

- Focus on teenagers and those have difficulty in writing.
- Standalone application.
- Limitation on accuracy of voice recognition due to the cloudy environment and different regional accents.
- Limitation in controlling the application as not much people used to command of windows speech recognition software.
1.5 Project Significant

In the end of the project, developer will gain the knowledge in the area of speech recognition technology. Besides, developer will be able to design and build a simple speech recognition application based on the knowledge obtained from the research done in speech recognition. Furthermore, users of Starline LRT will be able to obtain relevant information in order to reach their desired destination. Indirectly, this project boosts up the ticketing counter efficiency of LRT station because it reduces the time of guiding the user when there is a request. Lastly, both users who are able to read and unable to read will benefit from the new system.

1.6 Conclusion

This project highlights utilization of the technology of speech recognition. This new system provides effective guidance to Starline LRT users as well as increases the efficiency of Starline LRT station’s operation. In the next chapter, some literature reviews such as comparison of existing system, comparison of programming language, methodology, instructional design, software specification and hardware specification are discussed to assist the development of Starline LRT guidance system.
CHAPTER II

LITERATURE REVIEW & PROJECT METHODOLOGY

2.1 Introduction

In this chapter, the domain of the project will be identified as well as study on existing system that has similarity with this project. By gone through this process, the project able to refine from the disadvantages of existing system and make the developing system more effective and attractive. Next, software and hardware technology that will be used are to be determine in order to improve the efficiency of the system.

2.2 Domain

According to Wikipedia and proven by few articles, the first speech recognizer appeared in 1952 and consisted of a device for the recognition of single spoken digits. Although the technology has been developed for some decades, it is not widely utilized in Malaysia. Thus, the domain of this project is to introduce the utilization of the speech recognition technology in providing guidance for Starline LRT user in Malaysia.
2.2.1 Speech Recognition

Speech recognition can be categorized into speaker dependent and speaker independent. For speaker independent speech recognition, the speech recognition software will not rely on the voice of user who speaks so that everybody can interact with the system with the same rate of accuracy. On the other hand, speaker dependent will recognize the voice of the specific user to make the system work more accurately.

The advantage of using speaker independent software is the software provides vast vocabulary that can be use for dictation. However, users are required to train their voice into a standard accent in order to allow the software to provide a more accurate output. In contrast, the speaker dependent software does not consist of large numbers of vocabulary provided. However, it is suitable to use on application that require much specific word such as telephony, call center and so on.

In conjunction with the development of application, it enhances the use of speech recognition technology. Grammar is required to recognize the specific word that will be used for the application which does not available in the speech recognition software. For example, the word *Pudu Raya* needed to store inside the grammar file in order to enable the use of word inside the application.

Similarly with grammar, vocabulary is also needed to allow the speech recognition to identify word that is not being stored inside the speech recognition engine. Unfortunately, the word that is being saved as vocabulary does not enable the feature of word reservation for event-handling.
In conclusion, developers who wish to implement the technology of speech recognition on the application are required to choose the type of speech recognition software that is being use as well as identify the grammar and vocabulary being used in the application.

2.3 Existing System

Currently, there are a few websites that provide user information of the route of LRT which are Google Map, streetdirectory.com, maps.701panduan.com and etc. Those examples provided above are time-consuming as it requires users to search for it on the map manually.

2.3.1 Comparison of Existing System

In this section, analysis for several existing resources that provide the service for checking the nearby Starline LRT station for desired destination will be selected to perform comparison. The 3 existing application selected to perform comparison are generally web-based application which are Google Map, streetdirectory.com and maps.701panduan.com. The table below is the summary for the services provided by the application.
<table>
<thead>
<tr>
<th>Existing Service</th>
<th>Google Map</th>
<th>Streetdirectory.com</th>
<th>Maps.701panduan.com</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Service</strong></td>
<td>Search location on map, search for the path to the destination, search the time duration from one place to another according to method of transportation, view satellite map</td>
<td>Search location on the map, provide path of driving from one place to another</td>
<td>Search spot of tourist</td>
</tr>
<tr>
<td><strong>Purpose</strong></td>
<td>Provide location, direction and duration to the destination</td>
<td>Provide location and direction to the destination</td>
<td>Provide spot of tourist</td>
</tr>
<tr>
<td><strong>Method of searching</strong></td>
<td>Destination address, street name, company name, drag on map</td>
<td>Drag on map, street name, company name</td>
<td>Company name</td>
</tr>
<tr>
<td><strong>Degree of zoom</strong></td>
<td>18</td>
<td>Ten</td>
<td>Seven</td>
</tr>
<tr>
<td><strong>Satellite Map</strong></td>
<td>Yes. Provided with the freedom of zooming at certain degree</td>
<td>Yes. Without functionality of zooming</td>
<td>None</td>
</tr>
<tr>
<td><strong>Coverage of Map</strong></td>
<td>World-wide</td>
<td>Singapore and Malaysia only</td>
<td>Malaysia only</td>
</tr>
<tr>
<td><strong>Information on Transportation</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>