

Research on Speech Recognition Technology Applied for Intelligent Mobile Navigation System

Mi WANG, Bingxuan GUO, Deren LI, Jianya GONG
National Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing
Wuhan University, P.R. China

wangmi@rcgis.wtusm.edu.cn, gbx@rcgis.wtusm.edu.cn,
dli@wtusm.edu.cn, jgong@rcgis.wtusm.edu.cn

KEY WORDS: Navigation, Speech Recognition Technology, Intelligent

ABSTRACT

The capability of human-computer interaction reflects the intelligent degree of Mobile Navigation System. This paper applies Speech Recognition Technology to Intelligent Mobile Navigation System and does some deep research on the integration of Speech Recognition Technology with Mobile Navigation System, which makes human-computer interaction easy during Navigation. After processing the data and functions of mobile navigation system, a group of speech commands are abstracted. Speech Command Interface of Navigation System is implemented by Dutty ++ Software, which is based on Speech Recognition System —Via Voice of IBM. Through navigation experiments, it proved that human-computer interaction is very convenient by Speech Commands and the reliability is also better.

1. INTRODUCTION

At present, human-computer interaction is generally by using keyboard, mouse and screen in many application systems. The efficiency of this interaction mode is not high and the operation is inconvenient in many occasions. Mobile navigation System is a good example for this, as the space in car is very limited, it is very inconvenient and insecure to implement human-computer interaction by keyboard or mouse during navigation. It goes without saying intelligence. Well then, how can we reduce or not use keyboard or mouse during navigation. Speech is a better solution.

Since 1960s, many research institutes begin to pursue the research of Speech Recognition Technology. Up to now, it has achieved great progress in speech recognition technology. Many Company have developed speech recognition software, for example Via Voice4.0 of IBM Company, VoiceExpress of Microsoft and some other Company and DP1000 of Olympus. Among these systems, ViaVoice is a successful product of IBM Company. After speech training, the correct rate of Chinese nature language recognition can be up to 90 percent in suitable environment.

But if we apply Speech Recognition Technology to Intelligent Mobile Navigation System, two problems must be solved. One is that it must satisfy real-time and reliable need, the other is resistibility of disturbance of the car. As the rate of Chinese nature language recognition is not very well in the car noise environment and it must perform speech training before recognition. So, from point of view of data processing and system functions, this paper get a series of speech commands. Using these speech commands we develop our navigation system on the basis of the software Dutty ++ based on Speech Recognition System —Via Voice which is developed by IBM Company. After experiments, it proves to achieve very well Human-Computer Interaction results.

2. Introduction of Via Voice and Dutty++

Via Voice is a Chinese Speech Recognition developed by IBM Company, which does deep research on speech recognition for twenty-seven years and have invested over two hundred million dollars. It is Chinese mandarin recognition and dictation system and developing tools. It is put forward after American English, British English, French, German, Italian, Spanish and Japanese.

The main functions of Via Voice4.0 are (1)Chinese Character input, edit and print etc; (2)speech dictation; (3)speech command; (4)the core interface with other application software. This system can translate user dictation into text instead of keyboard input. It has the functions of accent adaptation according to different speech characteristics. Accent adaptation can make Speech Recognition System record and compile user speech information, which can improve recognition rate of dictation. It is generally perform accent adaptation for a long time before nature language recognition. It has a basic vocabulary which includes thirty thousand Chinese phrases and many computer commands. In addition, user can add phrases to the system. Via Voice uses these basic vocabulary and user vocabulary to process voice information during dictation.

Some people said that Via Voice is a revolution of Chinese character input, make a great improvement in computer toward human and it is an important landmark in Chinese information processing technology.

Dutty++ speech recognition system is developed by Ruian Company, Beijing. It is a second develop on the basis of Via Voice and improve speech recognition functions. It implements command control and dictation input function in Windows 95 and Windows NT. At the same time, it does not affect inhere functions of Windows 95 and Windows NT. It supports dictation input in all text editors of Windows 95 and Windows NT and speech control command operation. As it is independent speech platform, it is very convenient for the need of word process.

Dutty++ speech recognition system has two state during running: command control and speech input. The switch is implemented by the command of start dictation and stop dictation. Usually, the usage of computer has a few steps as follow: 1.Turn on power; 2.startup operate system; 3.Run application; 4.Input, typeset, save and exit; 5.shutdown operate system; 6.turn off power. The whole processes need use keyboard or mouse. If you use DUTTY ++, many operations can be implemented by speech command except turn on and turn off power.

In addition, users can define their own speech commands. Users can connect speech commands with their application.

3. System Design and Constitue of Speech Navigation System

In mobile navigation system, the spatial data and attribute data are organized by GIS and road network topology is built. GPS receives the car position information, which matches with electronic map.

3.1 System Constitue

The Navigation System includes industry computer with multimedia equipment, GPS receiver, Electronic compass, wheel counter, GPS signal process software, the integration software of GPS and GIS, Speech recognition software, Electronic map.

3.2 System Functions

1. Electronic map display, zoom in, zoom out, zoom pan and map rotation.
2. Query function. Attribute and spatial inter-query
3. Best route computation. Route selection and computation.
4. Two Navigation Mode: Best road Navigation and Random Navigation
5. Real-time display car position and track.
6. The Error of GPS Correct
7. The process of GPS signal loss.
8. GPS positioning signal match with road
9. Speech Navigation without accent adaptation.
10. Car State real-time display

4、 The Abstract and Classification of Speech Commands

Speech Commands of Navigation System is crucial for the results of Speech Navigation. In Mobile Navigation System, the principles of the abstract and classification of speech commands are as few as possible and as definite as possible. According to these principles, as every speech command has definite meanings, the navigation can achieve better results of speed, nicety and reliability. According to Navigation functions, speech commands can be classified two types as follows.

1. System Commands. These include electronic map zoom in, zoom out, zoom pan left, zoom pan right, zoom pan up, zoom pan down, start navigation, compute best road, select point by name etc. The characteristics of these commands are when they are executed without parameter. They can guarantee nicety and reliability.
2. None System Commands. During navigation experiments, we mainly do research on the problem of best road selection by speech commands. Best road selection is the most complex operation and is also common function in mobile navigation. It needs many parameters while computing best road and there are many human-computer interactions. If we finish best road operation in the car with keyboard and mouse, it is very inconvenient. Best Road selection begin with start place and pass or not pass other places, at last get to the destination place. There are many candidate places in a city. If every place has a speech command, thus there will be thousands of speech commands. The speed and reliability of speech recognition will be affected. So we divide the candidate places into eight kinds: hotel, education, amusement, hospital, bank, enterprise, government and culture establishment. Each kind has his own kind speech command: The First Kind, The Second Kind, The Third Kind ...The Eighth Kind. There are many candidate places in each kind. So the candidate places in each kind is divided into many pages. We can use next page and previous page as page speech command to turn over pages. In each page, we use place speech command: One, Two, Three, Four, ...Seventeen to identify each place. Thus, if we use kind commands, page commands and place commands, each place will be identified exclusively in the candidate places. The detail is shown as below.

Table 1 Speech Command Table

| Kind Command Table | | Page Command Table | |
|--------------------|-----------------------|---------------------|---------------------|
| Speech Command | Command Description | Speech Command | Command Description |
| The First Kind | Hotel | The Next Page | Change Pages |
| The Second Kind | Education | The Previous Page | Change Pages |
| The Third Kind | Amusement | Place Command Table | |
| The Fourth Kind | Hospital | | |
| The Fifth Kind | Bank | | |
| The Sixth Kind | Enterprise | | |
| The Seventh Kind | Government | Speech Command | Command Description |
| The Eighth Kind. | Culture establishment | One | Place A |
| | | Two | Place B |
| | | Three | Place C |
| | | ... | ... |

Now take a example to illustrate the process of best road compute. If we want to compute a best road from Wuhan University to Hankou Cinema and pass ZhongNan bookstore, this function is finished with speech Commands as follows: Select Point By Name-The Second Kind-The Second Page-Ten-Add-The Eighth Kind-Five-Add- The Third Kind-The Third Page-Two-Add-Compute. Through above speech commands, we will finish best road computation.

5、 The Implement of Speech Command Interface

It is very simple to implement speech Command interface. For program developer, the process of speech commands is the same as menu command message. It illustrates as follows:

First, Speech commands are store as text into the file Dutty.ini. In addition, application title and user define message also store in the file Dutty.ini. DUTTY ++ can connect speech recognition system with

the application. When user sends speech command by microphone, speech recognition system translate and recognize speech command, then convert the speech command into user defined message. DUTTY++ send the message to the main window of the application. The functions of speech commands are implemented after User responds the message in the application. The detail processes as below:

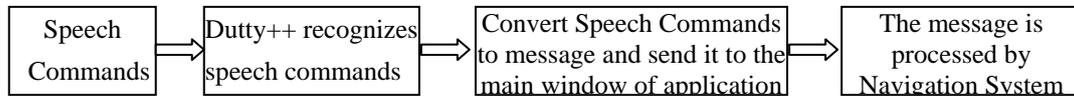


Figure 1.The Process of Speech Command Recognition

6. Conclusions

Figure 2 is the computer screen of navigation system. There are two electronic map windows on the screen. The left window is electronic map window in driver reference frame; the right window is electronic map window in absolute reference frame. Figure 3 shows the Experiment Environment of Navigation System. Figure 4 shows the interface of best road selection by speech commands.



Figure 2 The Screen of Navigation System



Figure 3 The Experiment Environment of Navigation System

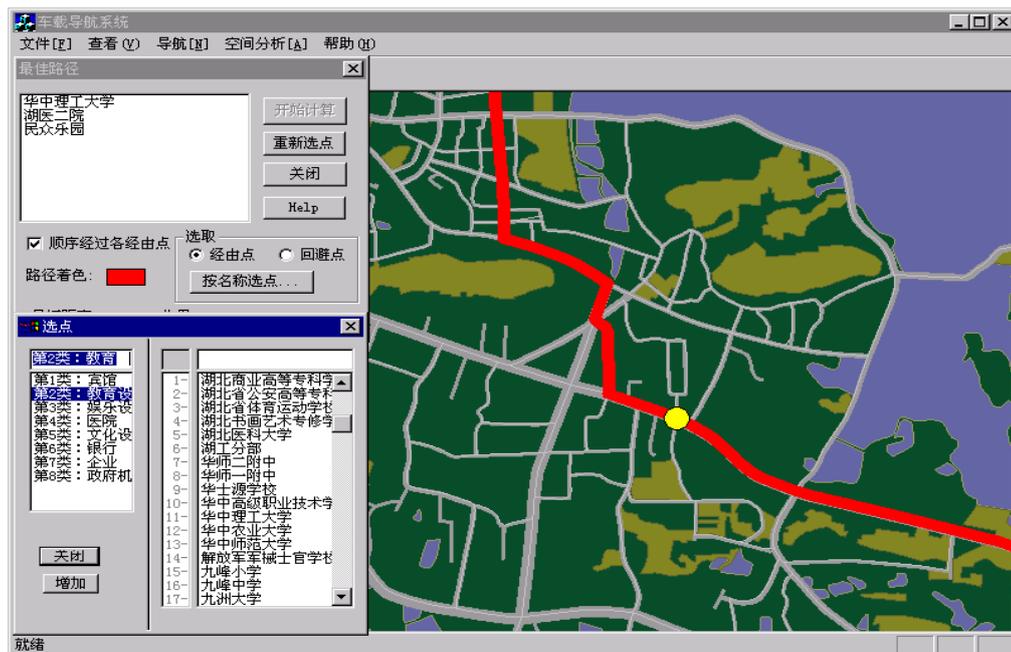


Figure 4 Best Road Calculation by Speech Commands

After experiments, it proved that speech command is indeed good way to operate navigation system and can make navigation system intelligent.

6. ACKNOWLEDGMENT

The research described in this paper was funded by the Natural Science Foundation project of China (No. 49631050)

7. REFERENCES

FANG Zhigang, WU Xiaobo, Ma Weijuan. The Progress on the Study of Human-Computer Interaction Technology. Computer Engineering and Design. 1998 Vol.19 No.1

LI Deren, GUO Bingxuan, WANG Mi, LEI Ting. Vehicle Navigation System Design and Implement Based on Integration of GPS and GIS. Journal of Wuhan Technical University of Surveying and Mapping. 2000 Vol. 25 No.3

LIU Deping, LI, Xingsheng, LIU Hongxia, LIU Wufa. Research on Speech Recognition Technology. Journal of Zhengzhou University of Technology. 1998 Vol.19 No.1

KONG Qigui. Introduction of IBM Via Voice Speech Recognition System. Electronic Technology. 1998 Vol.9.

WANG Mi, GUO Bingxuan, LEI Ting , LI Deren. Research on GPS Positioning Matching with Road In Mobile Navigation System. Journal of Wuhan Technical University of Surveying and Mapping. 2000 Vol. 25 No.3

WANG Yingzhi, ZONG Chengqing, CHEN Zhaoxiong, HUANG Heyan. Nature Language Human-Computer Interface design and implement In ITS. Computer Research and Development. 1998 Vol. 3 No.9