Speech Recognition Technology Applied to Intelligent Mobile Navigation System

1 Introduction

At present, human-computer interaction is performed 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 on many occasions. Mobile navigation system is a good example for this case, 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. Well then, how can we reduce or quit using keyboard or mouse during navigation? Speech is a better solution.

Since the 1960s, many research institutes have begun to pursue the research of speech recognition technology. Up to now, great progress in speech recognition technology has been achieved. Many companies have developed speech recognition software, for example Via Voice4.0 of IBM Company, VoiceExpress of Microsoft and some other companies and DP1000 of Olympus. Among these systems, Via Voice is a successful product. After speech training, the correct rate of Chinese natural 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 be real-time and reliable, and the other is resistibility to disturbance of the car. As the rate of Chinese natural language recognition is not very satisfactory in the car noise environment, and the speech training before recognition must be performed. From point of view of data processing and system functions, this paper presents a series of speech commands. Using these speech commands we develop our navigation sys -
System on the basis of the software Dutty + + based on the speech recognition system —Via Voice. After experiments, it is proved to achieve very good human-computer interaction results.

2 Introduction of Via Voice and Dutty + +

Via Voice is a Chinese speech recognition, which does deep research on speech recognition for 27 years and have invested over 200 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 Voice 4.0 are:
1) Chinese character input, edit and print,
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 the speech recognition system record and compile user speech information, which can improve recognition rate of dictation. It is generally to perform accent adaptation for a long time before natural 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 and make a great improvement in computer toward human, and that it is an important landmark in Chinese information processing technology.

Dutty + + speech recognition system is a second development on the basis of Via Voice and improves speech recognition functions. It implements command control and dictation input function in Windows 95 (NT) and speech control command operation. As it is an independent speech platform, it is very convenient for the need of word processes.

Dutty + + speech recognition system has two states 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) start operate system; 3) run application; 4) input, typeset, save and exit; 5) shut operate system; 6) turn off power. The whole process needs using keyboard or mouse. If you use Dutty + + , many operations can be implemented by speech command except turning on and turning off power.

In addition, users can define their own speech commands.

3 System design and constitution 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 maps.

3.1 System constitution

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

The functions of navigation system include:
1) Electronic map display, zooming in, zooming out, zooming 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 displaying 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 Abstract and classification of speech commands

Speech commands of navigation system is crucial for the results of speech navigation. In the mobile navigation system, the principles of the abstract and classification of speech commands should be as few as possible and as definite as possible. According to these principles, as every speech command has a definite meaning, the navigation can achieve better results of speed, nicety and reliability. According to navigation functions, speech commands can be classified into two types.

1) System commands. They include electronic map zooming in, zooming out, zooming pan left, zooming pan right, zooming pan up, zooming 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 begins with starting place and passes other places or not, at last gets 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 decrease. 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 of each kind. So the candidate places of each kind are 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 commands, 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 in Fig. 1.

| Kind command |
|--------------|
| Speech command | Command description |
| The first kind | Hotel |
| The second kind | Education |
| The third kind | Amusement |
| The fourth kind | Hospital |
| The fifth kind | Bank |
| The sixth kind | Enterprise |
| The seventh kind | Government |
| The eighth kind | Culture establishment |

| Page command |
|--------------|
| Speech command | Command description |
| The next page | Change pages |
| The previous page | Change pages |

| Place command |
|---------------|
| Speech command | Command description |
| One | Place A |
| Two | Place B |
| Three | Place C |
| ... | ... |

Fig. 1 Speech command

Now take an example to illustrate the process of best road computation. If we want to compute a best road from Wuhan University to Hankou Cinema and pass Zhongnan bookstore, this function is finished by 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 Implementation of Speech Command Interface

It is very simple to implement speech command interface. The Dutty ++ software can convert
speech commands into window messages on windows operating system. So, for program developer, the process of speech commands is the same as menu command message. It is illustrated as follows.

First, speech commands are stored as text into the file Dutty.ini. In addition, application title and user defined message also stored in the file Dutty.ini. DUTTY ++ can connect speech recognition system with the application. When user sends speech command by microphone, the speech recognition system translates and recognizes the speech command, then converts it into user defined message. Dutty ++ sends the message to the main window of the application. The functions of speech commands are implemented after the user responds the message in the application. The detail process is shown in Fig. 2.

![Fig. 2 Process of speech command recognition](image)

**Speech commands**

Dutty ++ recognizes speech commands

Convert speech commands to message and send it to the main window of application.

The message is processed by the navigation system.

**Fig. 2 Process of speech command recognition**

6 Conclusion

Fig. 3 is the computer screen of the navigation system. There are two electronic map windows on the screen. The left window is electronic map window in driver reference frame and the right window is electronic map window in absolute reference frame. Fig. 4 shows the interface of best road selection by speech commands.

![Fig. 4 Best road calculation by speech commands](image)

The experiments prove that speech command is indeed a good way to operate navigation system and can make navigation system intelligent.

References

1 Fang Z G, Wu X B, Ma W J (1998) The progress on the study of human-computer interaction technology. *Computer Engineering and Design*, 19(1) (in Chinese)

2 Li D R, Quo B X, Wang M, Lei T (1998) Vehicle navigation system design and implement based on integration of GPS and GIS. *Journal of Wuhan Technical University of Surveying and Mapping*, 25(3) (in Chinese)

3 Liu D P, Li X S, Liu H X, et al. (1998) Research on speech recognition technology. *Journal of Zhengzhou University of Technology*, 19(1)

4 Kong Q G (1998) Introduction of IBM Via Voice speech recognition system. *Electronic Technology*. (in Chinese)

5 Wang M, Quo B X, Lei T, et al. (2000) Research on GPS positioning matching with road in mobile navigation system. *Journal of Wuhan Technical University of Surveying and Mapping*, 25(3) (in Chinese)

6 Wang Y Z, Zong C Q, Chen Z X, et al. (1998) Natural language human-computer interface design and implement in ITS. *Computer Research and Development*, 3(9) (in Chinese)