

Software Program Design and Implementation of Voice Robots

Zhixin Hu

Nanchang Institute of Science & Technology

Keywords: Speech recognition; Robot; Signal Processing; Program design

Abstract. With the rapid development of modern science and technology, speech recognition technology is more and more concerned by people. As an important branch of intelligent robot research field, the speech recognition technology is to allow the machine to understand human language and to make convenient for human-machine communication. Therefore, applying the speech recognition technology in the field of robot control reflects the development level of the automation. Based on speech recognition technology, this paper takes the realization of simple motion control of robots as purpose, completes the recognition of isolated word speech signal of a specific person and mainly researches Feature extraction algorithm and speech recognition algorithm for speech signal. The speech recognition technology and the robot control technology are combined to analyze the software design process, as well as part of the function module division, which finally realizes the voice robot settings.

Introduction

As the robots are increasingly integrated into people's lives, the traditional man-machine exchange operation, such as buttons and joysticks, are relatively more trouble. To use voice recognition technology in the field of robots and to make robots perform specific tasks according to the voice commands of people can solve the barriers between man and machine, change the traditional way of human-computer interface and make the man-machine communication be more convenient. It is the valid proof of transferring voice recognition from theory into practice, is one of the important symbols of intelligent robot and is the embodiment of the development level of today's automation.

As one of the key technologies in the research of intelligent computer and human computer communication, voice recognition technology is drawing more and more attention. For example, the voice recognition technology can control the operation of large industrial machines, unmanned vehicle automatic driving, intelligent home and safe box. The intelligent robot controlled by voice recognition can carry out rescue activities in difficult environment, deep sea or space exploration, and can be used for family services, etc. In addition, voice recognition technology also has a wide range of uses in the military, communications, services, medical, banking, automotive electronics and other fields.

In this paper, the voice recognition technology and robot control technology are combined, the design process of the voice part of robots is described, which has important research significance to achieve the control of robots.

Basic Principles and Types of Voice Recognition

Voice is a kind of vibration sound wave which is produced by the vibration of a human voice organ. The purpose of voice recognition is to allow the machine to understand the human voice commands. The principle of voice recognition is shown in Fig. 1, the steps for voice recognition are generally as follows:

- (1) Signal collection: collect the voice of the people through the device or software, and conduct digital processing to voice signals.
- (2) Pretreatment: useful voice signals can be obtained through effective pretreatment.
- (3) Feature parameter extraction: extract low dimensional feature sequence which changes with time from sound wave to characterize the voice signal.

(4) Reference model library: produce different voice reference models after conducting special training to the extracted voice features.

(5) Pattern matching: The voice recognition process is essentially a pattern matching process. Pattern matching is to match and compare the unknown voice feature sequences with the reference template and to calculate the matching degree between them during the recognition process.

(6) Result analysis: Through the computer, analyze the results identified and determine the type of identification

Training

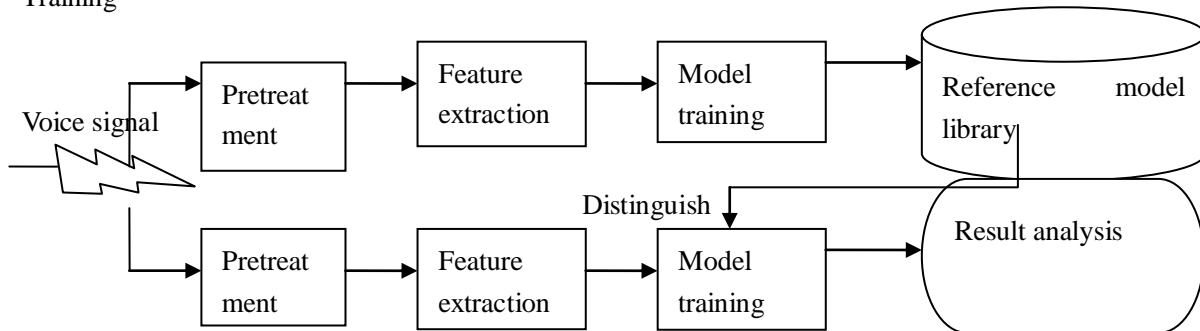


Figure 1. Principle of voice recognition

According to the different speaker, voice recognition can be divided into specific voice recognition and general voice recognition. Specific voice recognition is to recognize one or more specific voice, while general voice recognition is to recognize the voice of any person. Although the general voice recognition has better applicability and wider application, the voice recognition system of the general voice recognition is hard to achieve the satisfied recognition.

Analysis and Pretreatment of Voice Signals

From an overall point of view, the properties and characterization of its essential characteristic parameters of the voice signal are time-varying, it is a nonstationary process with time-varying. But in a short time range (usually 10~30 ms), its characteristics remain unchanged, which can be considered as a quasi stationary process, namely, short-time stationary. Therefore, the analysis and treatment of the voice signals must be built on the basis of "short time".

The voice signal analysis includes time domain analysis and frequency domain analysis, etc. Time domain analysis is to analyze the system directly in the time domain and from the perspective of statistics, which can obtain amplitude, short-time energy and short-time zero crossing rate and other basic parameters of voice signal. The characteristics of time domain analysis is intuitive, accurate with simple calculation and clear physical meaning.

The signal intensity obtained from time domain analysis changes with the time. However, signal also relates with frequency, phase and so on, which requires further the signal frequency analysis. Mainly by means of Fourier series and Fourier transform, change signal from time domain to frequency domain to obtain a description of the frequency spectrum, that is, the size of component of signals in different frequency.

(1) Short-time energy

Set the time domain signal of the voice as $x(m)$, after the process of window frame, the N frame signal is $x_n(m)$, then

$$\begin{cases} x_n(m) = x(n+m)\omega(m) \\ \omega(m) = \begin{cases} 1, & m = 0 \sim N-1 \\ 0, & m = \text{other values} \end{cases} \end{cases}$$

Other values: other values

In which, $n = 0, 1T, 2T, \dots$

N—Frame length

T—Frame shift length

Set the short-term energy E_n of the n frame $x_n(m)$, then:

$$E_n = \sum_{m=0}^{n-1} x_n^2(m)$$

E_n is used to characterize the energy size of a voice signal, commonly used in voice endpoint detection.

(2) Short-time zero crossing rate

The short-time zero crossing rate represents the times of a frame signal waveform crossing lateral axis (zero level), which is used to distinguish the sound of the voice. The short-time zero crossing rate is a simple and effective method to analyze the signal. The flow to achieve the short-time zero crossing rate is shown in Fig. 2.

The short-time zero crossing rate is a simple and effective method for the analysis of the signal. The flow chart is shown in Fig. 2.

Define the short-time zero crossing rate Z_n of a frame signal $x_n(m)$

$$Z_n = \frac{1}{2} \sum_{m=0}^{N-1} |\text{sgn}[x_n(m)] - \text{sgn}[x_n(m-1)]|$$

In the formula, $\text{sgn}[x]$ is the symbolic function, which is $\text{sgn}[x] = \begin{cases} 1, (x \geq 0) \\ -1, (x < 0) \end{cases}$

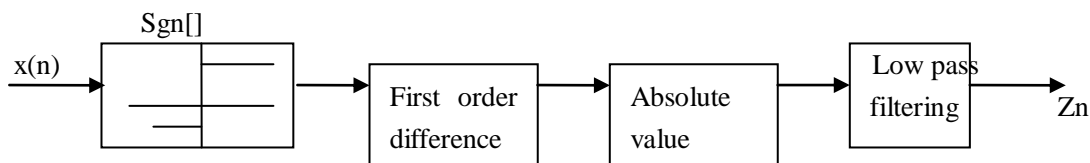


Figure 2. Realization flow chart of short time zero crossing rate

Software Program Design of Voice Robot

Program Design Process. To achieve human-computer interaction of the voice robots, it is necessary to design a set of reasonable man-machine dialogue process. Process of dialogue between people can be described as: people begin a dialogue → begin to ask a question → the other side reaction → start the second round of dialogue. According to the mode of human dialogue, man-machine dialogue process is shown in Fig. 3.

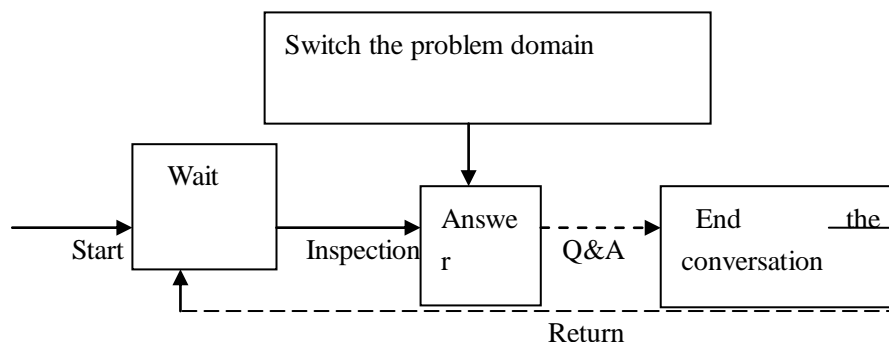


Figure 3. Voice robot software program running process

Function Module Design. To achieve the above procedure, the above functions are divided into several modules to complete procedures. Program module is as shown in Fig. 4, voice robot accepts a conversation, it will find examples with similar voice band in speech recognition library to narrow the range of understanding the issues. Voice robot extracts the keywords from the sentence and understand the sentence based on the keywords. According to the comprehension of the keywords, the voice robot gets the judgment of credibility: if the credibility is low, the robot will give a sign of speech error; if the credibility is high, it will give corresponding reply according to the keywords problem understanding. Voice robot generates sentences and completes the first round of the dialogue.

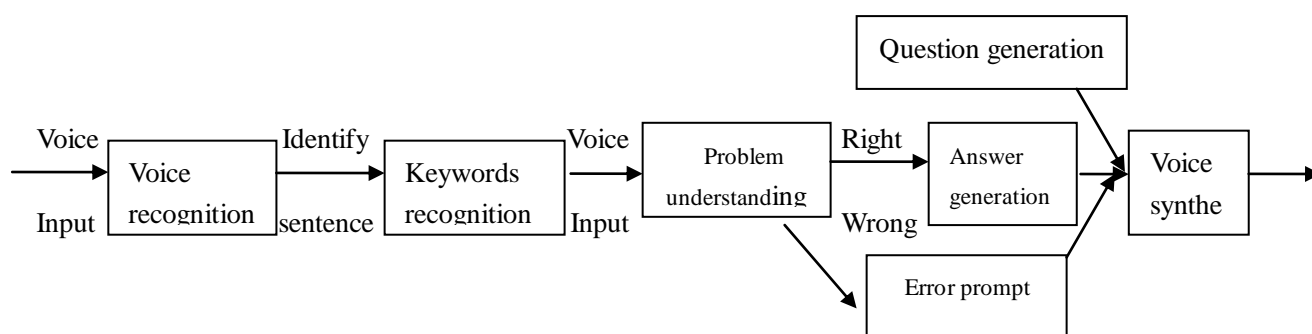


Figure 4. Voice robot software program module running process

Conclusion

With the development of society, people want to use voice to communicate with the robot. The existing technology can also achieve robot speech recognition technology both in the hardware and software. In order to verify the feasibility of the voice remote control system, this paper designs a voice remote control mobile robot experiment. Based on the research on speech recognition technology, this paper puts forward the voice recognition technology used by this topic, conducts a secondary development using Microsoft Speech SDK to construct a speech interaction system; establishes the basic model of remote control system and analyzes the remote control mode. Speech input is relatively flexible. In dictation mode, speech recognition rate is as high as 86.67%. In a limited range, the real-time, reliability and stability of the remote control system can be met. And client and server application program interface is friendly and the operation is easy, which forms the embryonic form of the remote control system of robot voice and builds a platform for the future system expansion and software development.

References

- [1] Luo Z, Zhao J. Speech recognition and its application in voice-based robot control system[C]// International Conference on Intelligent Mechatronics and Automation, 2004. Proceedings. IEEE, 2004:960-963.
- [2] Zhang Z Y, Song Y. Design and Realization of Voice-robot Based on Embedded[J]. Journal of Changchun Normal University, 2008.
- [3] Qi C. Application of Speech SDK in Exploiting Voice Robot[J]. Journal of Beijing Polytechnic College, 2008.
- [4] Okada, Shinichiro, Tanaba, et al. Single-surgeon thoracoscopic surgery with a voice-controlled robot.[J]. Lancet, 1998, 351(9111):1249.
- [5] Zhang Z Y, Song Y. Design and Realization of Voice-robot Based on Embedded[J]. Journal of Changchun Normal University, 2008.

- [6] Profio U D, Fujita M, Takagi T, et al. Robot apparatus, and behavior controlling method for robot apparatus: US, US7103447[P]. 2006.
- [7] Matsuda H, Shindo Y. Design and Implementation of Kana-Input Navigation System for Kids based on the Cyber Assistant [J]. Journal of Systemics Cybernetics & Informatics, 2004, 2(1):316.
- [8] Yan X U. WLAN based communication system design for fire searching robot[J]. Machinery, 2006.
- [9] Aleotti J, Micelli V, Caselli S. An Affordance Sensitive System for Robot to Human Object Handover[J]. International Journal of Social Robotics, 2014, 6(4):653-666.
- [10] Guan J S. Design of Voice Controlled System of Robot Voyager II[J]. Microcomputer Information, 2009.
- [11] Dan G. Robot Santa Discovers Meaning of Christmas at SantaCon | Underwire | Wired.com [J]. Bollettino Della Societa Italiana Della Scienza Del Suolo, 2003.
- [12] Pires J N. Robot-by-voice: experiments on commanding an industrial robot using the human voice[J]. Industrial Robot, 2004, 32(6):505-511.