

Design of speech recognition system based on LD3320 chip

Lei Wang^{1, a}

¹School of Jianqiao University, Shanghai201316, China;

^agench_wl009@126.com

Keywords: Speech recognition, Embedded technology, Pattern matching

Abstract. Speech recognition technology is a key to voice control technology, through speech recognition enable computer can read the speech command vocabulary, and can be through voice control realize control on all kinds of home appliances in the home, the system combines ld3320 chip ICRoute company, voice control system design, completed the software and hardware system, realized the non - specific human speech recognition, through tests the average recognition rate reached 90%, and the system is easy to build, with good practical value.

Introduction

Speech Recognition Automatic (ASR) refers to the machine through the recognition and understanding of the process of converting the signal to the corresponding text file or the command of the technology. The ultimate goal of speech recognition is to enable the machine to understand human language, and the language can be translated into the corresponding command and control equipment to achieve the specified function^[1].

The essence of speech recognition is the process of pattern matching recognition, speech recognition process includes four parts of speech signal preprocessing, speech feature extraction, training and recognition^[2]. In speech recognition, you must first preprocessing and feature extraction of speech, through these two steps can user speech feature parameters, and according to the parameters of reference model established, and the content of the training is mainly through the user multiple and continuous speech input, speech reference model was established, recognition is the input feature parameters and model reference library reference model are compared, similarity of the highest the as output results, so as to achieve the purpose of speech recognition.

The concept of speech recognition in foreign countries have already put forward, until in the late 1960s and early 1970s, speech recognition technology only substantive breakthrough, the representative events is linear predictive coding technology and dynamic time warping technology is put forward. At that time, the research field focused on small vocabulary, specific and isolated word speech recognition research. After 1980s, the focus of speech recognition research has gradually shifted to a large vocabulary, not a specific person. To the 90s of 20 century, speech recognition technology start from the laboratory to the society, many manufacturers began to use speech recognition technology in their products, and with the promotion of the networking technology, speech recognition technology in intelligent control, intelligent home application has been widely promoted and applied.

At present, there are two main types of speech recognition technology, one is the large vocabulary continuous speech recognition technology based on the computer platform, and the other is the small vocabulary speech recognition technology based on embedded technology. In the use of smart home system is embedded words sink isolated word speech recognition technology, the whole recognition process is as follows, firstly in the environment of the smart home system through a microphone for voice input, voice chip embedded system through the monitoring process receives input speech to determine whether it is the correct pronunciation and format, if not continue to monitor until it receives correct voice format of speech signal, and start the recognition, the input speech and the reference model of data for comparison, the command with the highest similarity, the command is transmitted to the corresponding equipment operation, and through voice reporting the implementation. Speech recognition technology in the smart home system can be more

convenient for users to use all kinds of household equipment, to understand the environment, to provide users with a more convenient and comfortable environment.

The hardware design

The hardware circuit of the system mainly comprises a core control components and speech recognition components in two parts, the core control unit uses is STC company produced the STC89LE52RC main chip, using the speech recognition component is Id3320 chip ICRout company produced, the chip embedded voice module, that is, the user does not need training, you can directly use, manufacturers have used the large amounts of speech data for voice training, and establish the model, the chip is a kind of non specific recognition, the main purpose of the chip just do identification. Chip supports up to 50 identification entries, each entry is standard Mandarin Chinese Pinyin, each with a space between the two Chinese Pinyin^[3-4].The related circuit diagram is shown in figure 1.

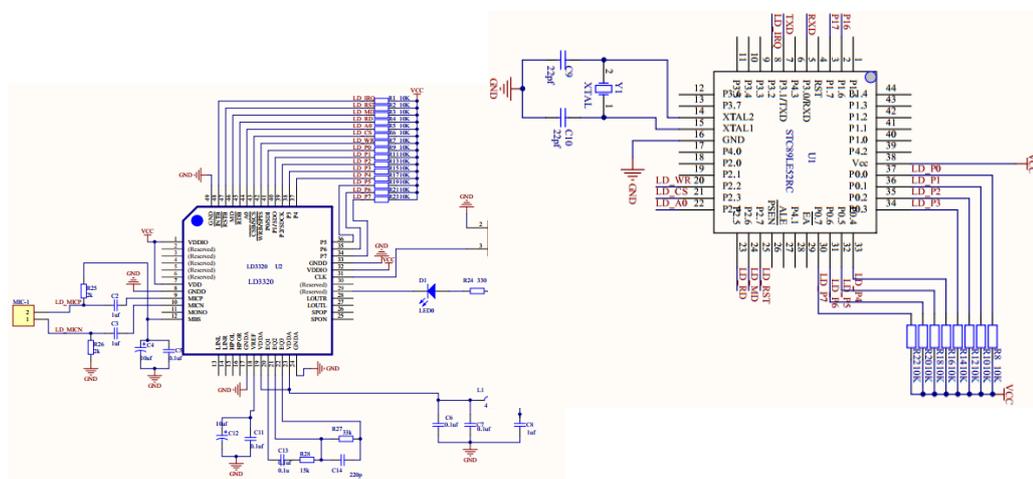


Fig. 1 Chip circuit diagram

Software design

Software interrupt mode is adopted in the system to work, basic work flow, including general initialization, speech recognition with initialization, write list recognition, start a voice recognition and response interrupt; in the initialization procedure, mainly to complete the soft reset, set mode, clock frequency setting and FIFO setting. Write identification list refers to each identification entry corresponding to a specific number, number can be the same, can not be continuous, but the value is less than 256. Through a microphone for voice recognition, when received by the microphone to send voice data, whether or not it can be recognition, the system will generate an interrupt signal, and the read register values to determine whether the correct identification of the signals.

Introduction of speech recognition algorithm

Language is unique to humans used to transmit information in an important way, by means of language can be realized between person and person communication, understanding what the other needs, if can make the machine understand "the language of the people, and in accordance with the language of the people to perform related operations will is human-computer interaction in the coordination of communication mode. The generation of language recognition system provides the possibility of the realization of this goal^[5].

Speech recognition system, its essence is a kind of pattern matching process, that is, through the voice signal acquisition to obtain specific parameters of speech signal and compare the specific parameters and phonetic reference model library model identified that matching rate was the highest

results [6]. Speech recognition system mainly includes signal preprocessing, speech endpoint detection, feature extraction, reference template library, pattern matching, and so on. Pretreatment is the first of the speech signal sampling, and then the speech signal pre-processing, such as pre emphasis processing, plus window processing, sub frame processing, etc.. Speech endpoint detection is to be reedited by preprocessing the speech signal, and to extract the effective speech signal. Feature extraction is to extract the most critical speech feature parameters from the speech signal, which is used to prepare the model. The establishment of reference model base is the key characteristic parameters of speech signal obtained by training, and establish the reference model database of speech recognition according to the parameters. Recognition is the last step of the speech recognition system, through the establishment of early speech model library, the recognition is in the input speech signal and model base model were compared with matched and the highest similarity outputs the recognition result.

Currently using the recognition algorithm, a total of four, one is the dynamic time warping (DTW, DTW algorithm, and the second is hidden Markov model (HMM, HMM), the third is the VQ (Vector Quantization VQ), and the fourth is to artificial neural network algorithm of artificial neural networks (ANN).

Performance testing

The speech recognition module and intelligent home control board are connected and through a microphone were the voice command input, by selecting the help from different regions of the number of students, each student test 50 times, has completed the speech recognition test, some test results are shown in Table 1 below.

Table 1 results of speech test

content	Non - specific person1	Non - specific person2	Non - specific person3	Non - specific person4	Non - specific person5
Turn on the lights	48/50	44/50	46/50	44/50	43/50
Turn off the lights	41/50	42/50	46/50	47/50	45/50
Open the curtain	47/50	46/50	44/50	47/50	44/50
Close the curtain	46/50	42/50	45/50	41/50	46/50

According to the test results, the average speech recognition rate can reach about 90%, which can meet the control requirements of all kinds of household devices in the smart home system.

Conclusion

This system uses the LD3320 speech recognition chip to realize the non - specific personnel's speech recognition, has the good practical application value, and has achieved the anticipated design goal through the performance test.

Reference

- [1]Zhang Zhixia, Han Huilian, XueHongwei. Analysis of speech recognition technology [J].computer development and application.2008, 21:33-35
- [2]Li Dan, Ming Yong, Lu Hanyu. Research on speech recognition program behavior detection method based on [J]. CDHMM Journal of Wuhan University of Technology, 2012,34 (9): 1-5
- [3] Application of embedded speech recognition system of Hong Jiaping [J].LD3320 new technology 2012, 2:47-49

[4] LD3320 Data sheet[EB/OL]. (2010-06-08) [2015-1-5]

[5] Yoo Oh, Jae Yoon, Ji Park A name recognition based call-and-come service for home robots[J].IEEE Transactions on Consumer Electronics, vol.54, no.2: 247-251, 2008.

[6] Jia Jing. Design of embedded speech recognition module based on [J]. STM32 digital technology and application 2012,6:152-153