

# Master Control Design of Intelligent Homes Based on Voice Recognition Technology

Chen Hong

School of Computer Science, Wuhan Donghu University, Wuhan, Hubei, 430212 China

fofochen@163.com

**Keywords:** Voice Recognition Technology; Intelligent Homes; Master Control Design

**Abstract:** This study proposes the design architecture of a smart home service robot, which takes the TMS320F2812 as the main control chip, is embedded with the monitor module, and utilizes the delay estimation method to realize the real-time detection of the sound source orientation. The voice recognition module, the motor drive module and the obstacle avoidance module are embedded so as to realize the control of voice to the movement of the robot. And the ultrasonic sensor is used to make the robot avoid obstacles during its movement. Moreover, the programming interface is provided to enable the robot to take corresponding actions according to the family members' voice commands.

## 1. Introduction

With the development of computer network technology, "Intelligent Homes" has been increasingly valued. Based on the development trends, the increasing popularity of intelligent homes will be a necessity. However, the intelligent home in China is still at the initial development stage. It is necessary to continuously improve the platform, the application compatibility and user experience. Android, as an excellent open source software solution, is widely applied into mobile phones, tablets and other devices. It is becoming a hot spot to develop intelligent homes with complete functions, simple operations and stable systems for the development of intelligent homes in the world.

## 2. Software Realization of Master Control Design of Intelligent Homes Based on Voice Recognition Technology

### 2.1. Voice recognition engine library call

Voice Recognition JNI encapsulates the API provided by the voice recognition engine, providing the Voice Control Service with the Java API under the voice control. The encapsulated interface mainly includes the initialization of recognition engines, the creation of recognizers, and the callback processing of recognition results and so on. After initializing the recognition engine, create a local recognizer according to the application requirements, start the recognition process, callback function through the Java, obtain the voice recognition results by notifying the Java layer, and process them based on the application situations

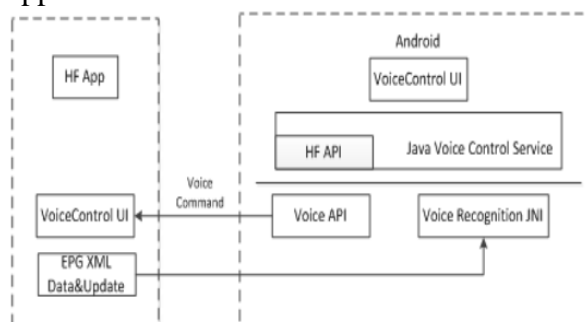


Fig.1 Voice recognition hierarchy diagram

## 2.2. Realization of Voice Control Service

The voice recognition engine obtains the character string of the voice recognition results by matching. And it's necessary to create the voice recognition local service at the Framework layer to process the recognition results.

## 2.3. Design of Voice Control Service UI

The component Voice UI is the information display of the voice recognition process and results. Enabling/disabling the Voice UI keeps synchronized with the Service. When the user enables or disables the Voice function on the menu, the event that the Android sends information to the intelligent home module will enable and disable the Voice Service and UI. UI includes the Android application software and the UI design.

## 2.4. Endpoint detection

Experience has shown that the zero-crossing rate of the unvoiced segment is usually the largest. Endpoint detection is first to judge whether there is a sound. If there is a sound, then judge whether it is unvoiced or voiced. In order to carry out the endpoint detection correctly, the two characteristics of short-time energy and zero-crossing rate are generally used and a double threshold detection method is adopted.

(1) Short-term energy

Set the short-time energy of the n-th frame speech signal  $X_n(m)$  as  $E_n$ , then the formula is:

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

Here,  $N$  is the frame length of the signal.

(2) Zero-crossing rate

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

Here,  $\text{sgn}[]$  is a symbolic function, that is:

$$\text{sgn}[x] = \begin{cases} 1, & (x > 0) \\ -1, & (x < 0) \end{cases}$$

According to the signal, set three thresholds: the energy thresholds  $TL$  and  $TH$ ; the zero-crossing rate threshold  $ZCR$ . When a frame signal is greater than  $TL$  or  $ZCR$ , it is regarded as the start of the signal; when it is more than  $TH$ , the formal speech signal is considered. If the considered formal speech signal keeps for a while, it is confirmed that this signal is the desired signal.

## 2.5. Voice recognition

The voice recognition will be realized by comparing the similarity degree between the sound pressure levels of the template and the input speech. The fundamental idea is to find the best function  $\varnothing(i_n)$  so that the function can satisfy the shortest cumulative distance of the template's sound pressure level and the input speech's sound pressure level. The formula is:

$$D = \min_{\varphi(i_n)} \sum_{i_n}^n (d(T(i_n)), R(\varphi(i_n)))$$

Here,  $(d(T(i_n)), R(\varphi(i_n)))$  represents the distance between the  $i_n^{\text{th}}$  frame of the vector to be tested and the  $i_n^{\text{th}}$  frame of the training template.



### 3.2. Zig Bee module design

Based on Smart RF technology, the process of CC2530 chip adopts QFN40 encapsulation with few external components and stable operation. The chip integrates the CPU of an 80\_51 core, including three different types of memory access buses (SFR, DATA and CODE/XDATA), up to 2\_56KB programmable flash and 8KB SRAM. The chip has the power management function and various different working modes. In the sleeping mode, the chip's current consumption is only 0.4\_SuA. It includes eight independent 12-bit ADCs with configurable resolution and independent channel, several timers, AES 128 coprocessor, watchdog timer, sleep timer, random number generator, battery monitor, temperature sensor, voltage regulator, a RF transceiver with high sensitivity and high resistance to interference, USART supporting multiple communication protocols, and 21 general-purpose I/O pins. Since the internal working voltage of the CC2530 chip is 1.8V, the power consumption is small. The internal voltage regulator can directly convert the 3.3V voltage into 1.8V voltage, thus improving the chip's compatibility.

In the hardware circuit design of ZigBee module, XOSC\_Q1 and XOSC\_Q2 are connected to 32MHz crystal oscillators, p1 and p2 respectively. RF\_N and RF\_P input and output wireless signals. The CC2530 ports P0.2 and P0.3 are respectively linked with USART 1\_ RXD and USART1\_ TXD of STM32 chip so as to realize serial port communication. This circuit design is a typical ZigBee wireless communication circuit design with good reliability and low loss rate of data packet. The specific hardware circuit design is presented in Figure 4.

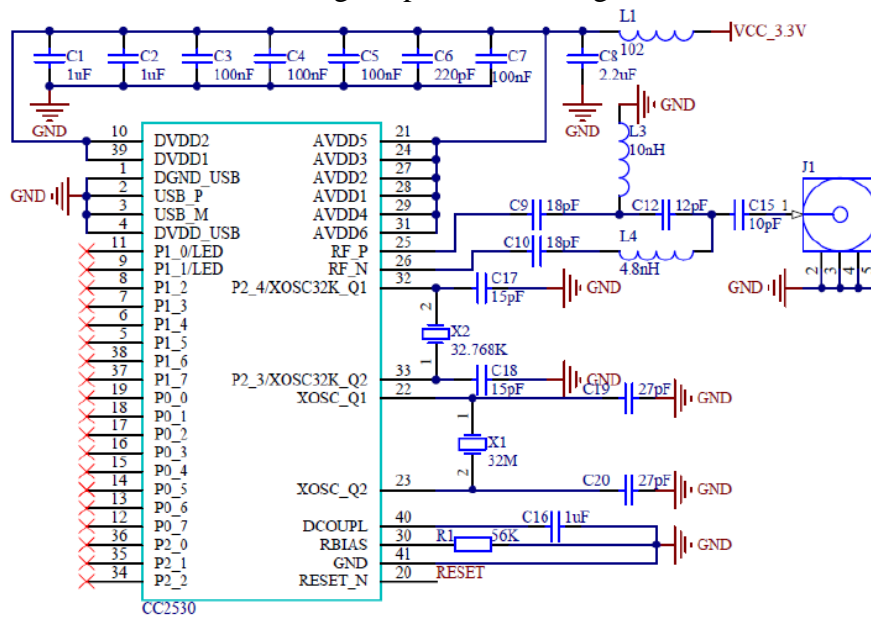


Fig. 4 Circuit diagram of ZigBee module

### 3.3. Motor drive module

The motor drive module drives the motor's operation (forward, turn, retreat, etc.) based on the control signal of the DSP controller. The motor's drive circuit uses an H-bridge drive circuit, as shown in Figure 5. The four triodes Q1, Q2, Q3 and Q4 form 4 bridge arms with Q1 and Q4 one set and Q2 and Q3 one set. Q5 controls the connection and disconnection of Q2 and Q3, and Q6 controls the connection and disconnection of Q1 and Q4. Since Q5 and Q6 are controlled by the GPIO port of TMS320F2807, the high and low levels of the I/O output can control the connection and disconnection of four bridge arms, further controlling the running state of the motor (forward or reverse rotation) and controlling the movement direction of the trash can. The design utilizes the L298N chip driving component, which can effectively avoid too complicated circuit board caused by the discrete component.

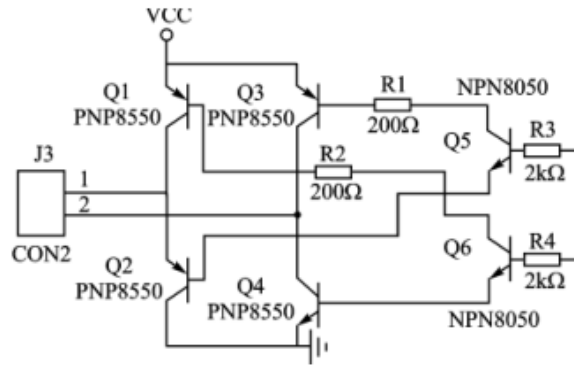


Fig. 5 H-bridge drive circuit

#### 4. System Software Design of Intelligent Homes Based on Voice Recognition Technology

The system software mainly includes the specific sound source localization algorithm module, obstacle avoidance software module, data acquisition module, and SCI communication module. The overall flow chart of system software is shown.

##### 4.1. Specific sound source localization algorithm module

The sound source localization algorithm in this system is mainly based on the microphone array time delay estimation method, the A/D sampling module is used to detect the sound signals received by different pickups, and the general cross-correlation method is adopted to help implement the algorithm. The waveform sampled by the DSP is close to a sine wave, and the sampled values are stored in an array. The correlation operation is firstly performed to obtain a cross-correlation spectrum. In order to avoid errors and noise interference, multiple calculations are performed to remove those larger error values. For the sake of avoiding the robot locating any received sound, the voice recognition is added into the system during positioning, which means only the desired sound is positioned.

##### 4.2. Obstacle avoidance software module

The obstacle avoidance software module adopts a counter to sample time. After the robot is powered on, firstly power on the two ultrasonic waves along the forward direction and start the counter while triggering one of the ultrasonic waves. The return of each echo will trigger an external interrupt, and the external interrupt processing program will record the return time of the ultrasonic. At the same time, activate the other ultrasonic sensor and turn off the currently working sensor. In fact, only one ultrasonic sensor is working at each moment and the data is collected. When the data collected by any one of the two ultrasonic waves is smaller than the constant value (the value that the obstacle is considered), the other three pairs of ultrasonic sensors are turned on successively.

#### 5. Conclusion

Voice recognition and processing are a key and difficult point in the field of voice signal processing. This design combines STM32C8T6 microcontroller with LD3320 non-specific human voice recognition chip to realize the overall design of voice control system of intelligent air conditioning. The test results show that this system can produce a 90% success rate of voice recognition of non-specific human, realizes the control of voice to air conditioner, which has made good experimental results. The system is characterized by strong versatility, high portability and promising development prospects.

#### Acknowledgements

This paper is supported by Scientific Research Project of Education Department of Hubei

Province (B2018295).

## References

- [1] Wei Gaowu, Feng Zuyong. Design of DSP Voice Recognition System Based on Denoising Technology [J]. Transducer and Microsystems, 2017, 36(01):108-110+118.
- [2] Luo Guizhou, Xiong Xiaodong. Research on Intelligent LED Lighting System Based on LD3320 [J]. Electronic World, 2017(02):182-183+186.
- [3] Sun Yongjie. Behind the Smart Speaker War: The Application of Voice recognition Technology Still Faces Challenges [J]. Communication World, 2017 (14): 54-55.
- [4] Liu Jia, Zhang Weiqiang. Research Progresses on Several Key Technologies of Low-resource Voice Recognition [J]. Journal of Data Acquisition & Processing, 2017, 32(02):205-220.
- [5] Dai Lirong, Zhang Shiliang, Huang Zhiying. Current Status and Prospect of Voice Recognition Technology Based on Deep Learning [J]. Data Acquisition & Processing, 2017, 32(02): 221-231.
- [6] Han Yuan, Jiang Chenghui, Li Sheng, Wang Binwei, Wang Yuchen, Qiu Qiu, Shi Xinghui. Preliminary Study on the Effect of Mobile phone Voice Recognition Technology on the Children's Voice Pronunciation in Mandarin [J]. Stomatology, 2017, 37(06): 515-519.