

A Speech Control System for Intelligent Industry

Jinlong Yao^{1, a}, Yuhuai Liu^{2, b} and Yuanhang Shi^{3, c}

^{1,2,3}School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China

^ayaojinlong2010@163.com, ^bieyhliu@zzu.edu.cn, ^cshiyuanhang1234@163.com

Keywords: Speech control. Intelligent industry. Spce061a MCU. Electric machinery.

Abstract. In recent years, the research of intelligent industry has become more and more prosperous. This paper designs a kind of speech control system based on the intelligent industry research background. In this study, the speech recognition theory summarized briefly. The speech recognition processing technology and the speech recognition model are simulated. SPCE061 and electric machinery realized the design of the system, users input control commands by speech and achieve the control of the motor. In this way, it realized the speech analog control industrial conveyor belt.

Introduction

The speech control system based on speech recognition. Speech recognition is not only used in speech communication system widely, it also applied in home life and industrial control, such as smart home devices in the intelligent home appliances, intelligent security, and hazardous areas under special environment of industrial speech control. It is almost involved in every aspect of people's lives. The study of speech recognition can provide core technology for speech applications growing market. Meanwhile, it meets the people's growing user experience, which has far-reaching enterprise and social values. Although speech recognition has developed for decades and great breakthroughs have appeared. However, it is still a challenge for the realizing of the machine like a person with complete speech recognition capability. Therefore, further research is necessary, such as the speech processing technology in acoustics and phonetics.

In recent years, the research of speech control system mainly covers the following aspects: The researchers build speaker independent speech model on the windows platform (Jing Lei, 2002), and realized the simulation of breakpoint detection and speech training template (Jun Ma, 2004). Meanwhile they analyzed the difficulty of hidden Markov model and improved the confidence values (Zhihui Wang, 2005). Based on HMM (hidden Markov model) and ANN (Artificial Neural Networks), Chinese speech recognition was studied (Liwei Chen, 2005). The scholars used Matlab simulating the speech recognition system, and put forward a better initialization algorithm, at the same time, they put forward the new IMFCC (Inverted Mel-scale Frequency Cepstral Coefficients) feature parameters (Yaqiang Gu, 2009). Not long afterward, researchers established the DNN-HMM acoustic mode (Gang Liu, 2014), and the mainstream technology of speech recognition system has been gradually replaced by DNN-HMM model, no longer the hidden Markov model mixed Gauss model HMM-GMM (Shaofei Xue, 2015).

Research progress of speech recognition promotes the application of speech control system. This paper simulates the models of speech recognition and demonstrates the design of design of a speech control system with SPCE061A chip and unSP IDE2.0.0 as software environment. Proteus 8 Professional and Altium Designer Summer 09 are used for the design of the hardware circuit.

The theoretical derivation and Simulation of speech recognition

Speech is composed of a series of sound. In fact, these speech signals are sound waves and carry language information. Meanwhile the sound waves can be converted into electrical signals or light signals. Before simulating the speech recognition theory and model, the procedure of classic hidden Markov model introduces in Fig.1.

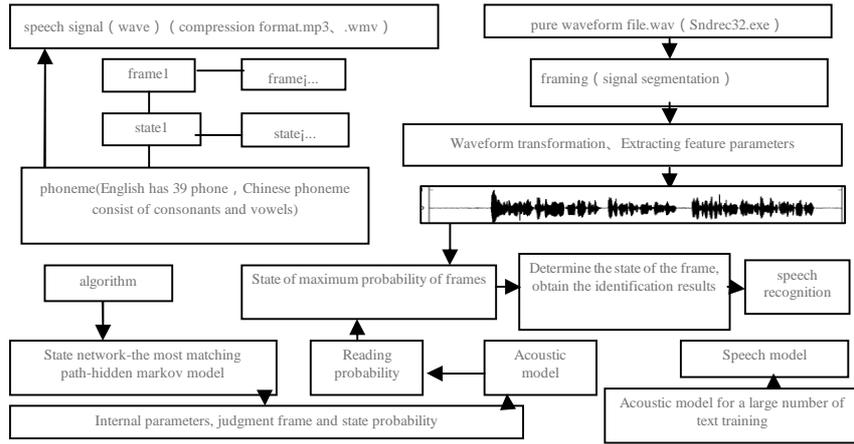


Fig. 1 HMM speech recognition process

It can be seen from the Fig.1 that the speech recognition principle of classical HMM. Before speech recognition process, it needs the processing of speech signal. The word pronunciation is made up of a phoneme, while one phoneme is composed of 3 states, and a state including a number of frames, as long as we identify the state of the frame, we can get the results of the recognition. First, the speech signal is divided into many short frames. Then we transformed the frame and extracted the characteristic parameters. In order to find the state of the maximum frame correspondence probability, it should use acoustic model and HMM. In the state of the network, the search algorithm of the optimal path is the algorithm that we need to study.

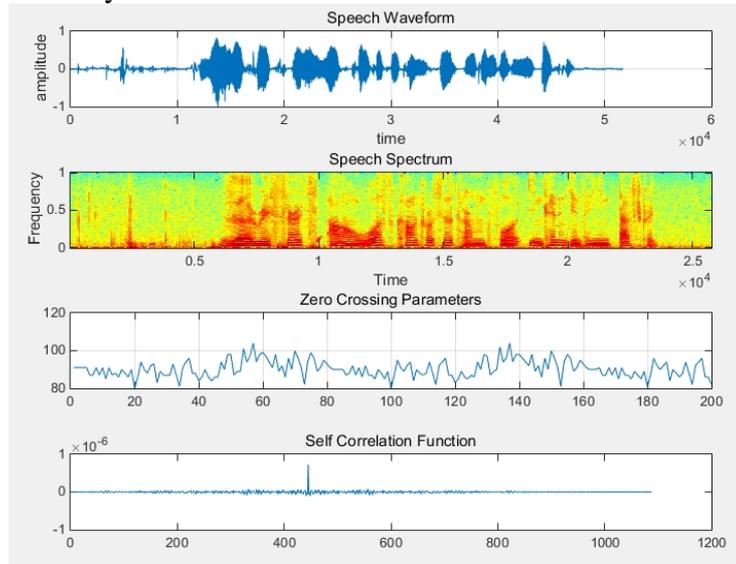


Fig. 2 Speech signal processing simulation

Speech signal is non-stationary signal, but in a very short period time, ranged from 10ms to 30ms, the speech signal can be regarded as the steady state signal. Before the speech signal processing, it needed to sample and quantify the speech signal, then digitize the speech signal. The process of pretreatment usually uses a first order digital filter such as Eq. 1.

$$H(z) = 1 - \mu \cdot z^{-1} \tag{1}$$

Defines the number of the n frame's average power is E_n .

$$E_n = \sum_{m=-\infty}^{\infty} [x(m)w(n-m)]^2 \tag{2}$$

The mean amplitude of the speech signal is M_n .

$$M_n = \sum_{m=-\infty}^{\infty} |x(m)|\omega(n-m) = |x(n)| * \omega(n) \tag{3}$$

The average zero crossing rate of x(n) in short time speech sequences is z_n .

$$z_n = \sum_{m=-\infty}^{\infty} |sgn[x(m)] - sgn[x(m-1)]|\omega(n-m) \tag{4}$$

The autocorrelation function of short-time speech signal in Eq. 5.

$$R_n(k) = \sum_{m=-\infty}^{\infty} x(m)\omega(n-m)x(m+k)\omega(n-m-k) \quad (5)$$

For simulation, first read the title speech file of the paper, then normalized treatment of the speech signals. We selected frame length 200 and frame shift 80, and selected the Hanning window function. The speech signal waveform, spectrum, zero crossing parameters and self-correlation function are shown in Fig2. By analyzing the results of simulation, it can find that the normalized processing of speech signal is useful in waveform display. Average zero parameters in speech recognition can distinguish between speeches and speechless. Auto correlation function is widely used in endpoint detection and pitch extraction. Certainly, there are many ways to deal with the speech signal, but here no more narrative. After the basic processing of speech signal, it needs to establish a model of speech recognition. In order to identify the speech more effectively, a optimized algorithm will be very important.

The design of speech control system hardware circuit

SPCE061A is a 16-bit processing chip with digital signal processing function; it has a strong processing capacity for speech information. SPCE061A has 2K static RAM and 32K flash memory, its clock operating frequency of CPU between 0.32MHz-49.152MHz, its real-time clock is 32768Hz. The operating voltage is between 2.4-3.6V. The chip has an audio programming capabilities and a special microphone interface. The chip has 32 programmable input and output ports. Also, it has 14 forms of timer B interrupt sources and two external clock source, 7 channel 10 bit ADC voltage and single channel speech ADC. The speech ADC is for specific speech signal input, there is a mic passage with automatic gain amplifier in it, meanwhile the MIC input channel and peripheral circuit has two-stage operational amplifier.

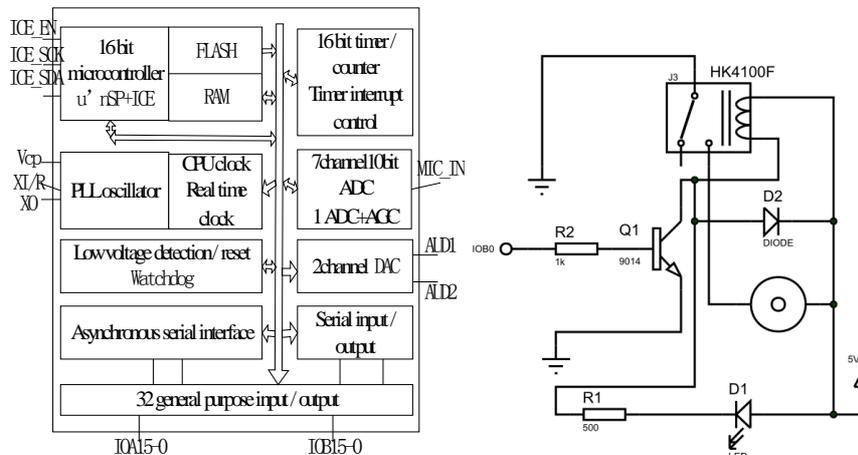


Fig. 3 Spce061a internal circuit structure and output circuit

The internal circuit structure diagram of Spcen61a and the speech output circuit are shown in Fig.3. When input the speech control command, the speech recognition results will be output from the IOB0 port in a level way. The IOB0 port external connect 9014 transistor, and the collector of the transistor connected with the relay port. The normally closed port of the relay is disconnected. In addition, another port connected to the motor. When SPCE061a recognizes speech commands, IOB0 output high level. At the same time, the transistor turns on, the relay turns off normally state, and the motor starts to rotate. The process of closing the relay is just the opposite.

The design of speech control system software programming

SPCE061A system software development based on IDE2.0.0 unSP integrated development environment and Tool_CHS.exe ISP program download tool. Program design procedure list as following: (1) Open the IDE2.0.0 unSP integrated development environment, file the new project New Project File and the new .C file/.Asm assembly file. (2) Add the speech recognition function library

bsrv222SDL.lib, the speech recognition header file bsrSD.h/bsrSD.inc, Hardware.asm/ Hardware.inc and Sacmv26e.lib to the project folder. (3) Write functional programs and add speech resources index table. (4) Compiler and downloaded the .S37 files to the chip. The design of the program flow chart is shown in Fig.4.

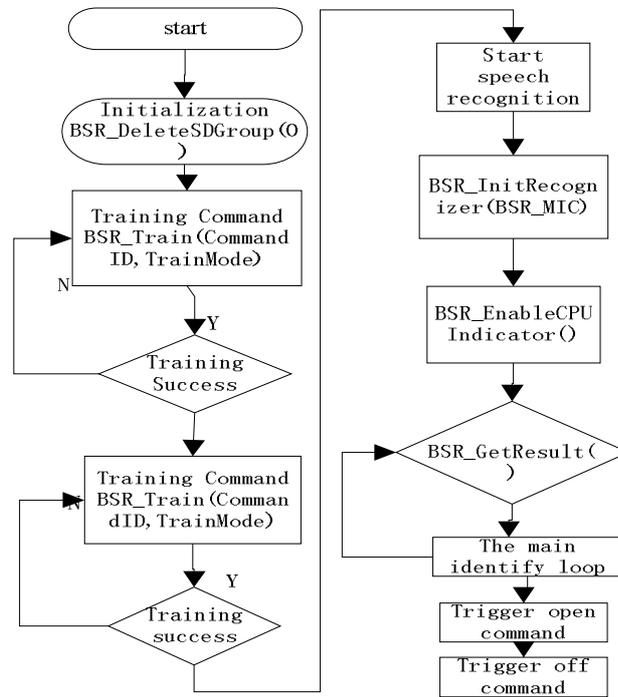


Fig. 4 Main program flow chart

Before the speech recognition, it needs the sound playing function PlayRespond and the training function TrainWord. For starting the speech training, it needs to initialize the IO port and the RAM memory. The initialization identification function is BSR_InitRecognizer, and the real time monitoring function is BSR_EnableCPUIndicator. In these programs, the IOB0 port is the output control program. After the speech recognition completed, the turn on control procedure of the motor is as follows: *P_IOB_Attrib = 0xffff; *P_IOB_Dir = 0xffff; *P_IOB_Data = 0xffff. In this way, IOB0 port outputs high level and drives the motor rotation.

Conclusions

This paper has achieved the following results: The paper discussed and simulated the models of speech recognition based on the classical hidden Markov model as an example. A speech control system was designed and it used an electric motor to simulate a conveyor belt for applications in the intelligent industry. The improvement of speech recognition algorithm will be further studied with more functions that are practical for industrial applications on the existing basis.

References

- [1] Information on http://www.gov.cn/zhengce/content/2015-05/19/content_9784.htm
- [2] Zhuo Xiang, Huihuang Zhao and Sixin Tang, in: Journal of Hengyang Normal University, No.6 Vol.35 (2014), p.66-68
- [4] Jinjian Lin, Runjie Chen and Jiayi Deng, in: Journal of Zhongshan University graduate(2014)
- [5] Zhihui Wang: The research on speech recognition algorithm based on HMM modeling (2005)
- [6] Jing Lei: The research and implementation of speech recognition technology (2009)

- [7] Chun Liang Hsu. Constructing, in: Expert Systems with Applications(2009), p.9308-9318
- [8] Zhenghua Tan, Paul Dalsgaard, Borge Lindberg: Borge Lindberg. Automatic speech recognition over error-prone wireless networks(2005)