Multi-function robots with speech interaction and emotion feedback

Hongyu Wang, Guanting Lou and Mengchao Ma
School of Instrument Science and Opto-electronic Engineering, Hefei University of Technology, Hefei, Anhui 230009, P.R. China

Corresponding author’s e-mail: louguanting@gmail.com

Abstract. Nowadays, the service robots have been applied in many public circumstances; however, most of them still don't have the function of speech interaction, especially the function of speech-emotion interaction feedback. To make the robot more humanoid, Arduino microcontroller was used in this study for the speech recognition module and servo motor control module to achieve the functions of the robot's speech interaction and emotion feedback. In addition, W5100 was adopted for network connection to achieve information transmission via Internet, providing broad application prospects for the robot in the area of Internet of Things (IoT).

1. Introduction

With the rapid development of China in recent years, robots in the manufacturing and industrial areas have been applied in large quantities, but their applications in the service industry are not particularly extensive yet. In recent years, there have been cleaning robots and cooking robots emerging in the market, indicating that it has been realized that service-oriented robots will be a very important application in the future. However, most of the robots, which has been put into use in the current domestic services, can only complete a single function and they do not have the ability to communicate with people and cannot provide emotional feedback either. At the international market, the research on service-oriented robots in Japan has reached a very high level. The "HRP-2" developed by the Japanese Industrial Technology Research Institute and the Shimizu Construction Company and the "Small IF" developed by the Japanese Human Robot Foundation have not only the ability to communicate with humans, but also the ability to speak, they all represent the current highest level in the development of service-oriented robots [1].

Based on these realities, under the premise of widely used Internet technology, this investigation designed and implemented a multi-function robot with speech interaction and emotion feedback. Different from the current majority of service-oriented robot, the robot can do speech recognition and face expression change. Those would not only meet the user's emotional and operational requirements, but also set the dialogue content and feedback mechanism in a specific occasion and realize some of the specialized services in public places. It is another prominent feature of its internal work instructions are achieved by the open source development platform. Therefore, customers can set their own parameters according to their needs and achieve its function and application at a great convenience. The diversity of the scene also allows users to personally participate in their own robot partners "recycling" process.
2. Realization of speech interaction function
Automatic speech recognition (ASR) is a technology that transforms a speech signal into a corresponding text file or command by identifying and understanding the process [2]. Speech recognition is the key technology for speech interaction robots, while the dialogue between human being and robot in natural language are realized. In order to achieve the speech interaction function, ASR M08-A non-specific speech recognition module was used in this study.

2.1. Work principle for ASR M08-A module
LD3320 speech recognition chip functions as a core of ASR M08-A non-specific speech recognition module. The chip is designed and manufactured by ICRoute company, integrates speech recognition processor and some external circuits including AD, DA converter, microphone interface, sound output interface, and so on. The all set-up do not have any external auxiliary chip, whereas speech recognition/speech/man-machine dialogue functions can be realized [3-4]. Also, the identified list of key words can be edited dynamically, which provides a basis for its wide application in the field of robots.

![Figure 1. The connection details between speech modules](image)

2.2. Hardware design of speech recognition function
ASR M08-A non-specific speech recognition module has total of six external serial port, namely SPK-, SPK +, 3.3V, TXD, RXD and GND. Wherein SPK is the interface of the external speakers, 3.3V and GND are the positive and negative terminals of the module power supply, TXD and RXD are the communication interface of the modules. In normal operation, the module receives the external voice signal input, through the LD3320 chip into a serial output signal, and sent to the Arduino microcontroller. The specific implementation steps are shown in Figure 1.

3. Realization of emotional feedback function
In order to achieve the emotional feedback function, a steering gear is connected to the robot's eyebrows, eyes, mouth, eyelids as shown in Figure 2. With the coordination of the steering control, the speech interaction is achieved and at the same time this unique configuration allows the robot to demonstrate emotion feedback as well.
3.1. Steering gear control

The control signal of the servo is a pulse width modulation (PWM) signal with a period of 20ms, where the pulse width is in the range of 0.5-2.5ms and the corresponding rudder position with the freedom of linear change can be 0-180 degrees. In other words, at a giving pulse width, its output shaft will remain at the corresponding angle, no matter how the external torque changes. The rudder will change the output angle to a new correspondence position until another pulse is provided.

The control of the steering gear generally requires a time base pulse of about 20ms, the time base pulse is in the range of 0.5ms-2.5ms. The corresponding relationship of the steering gear and time base pulse width is shown in Table 1:

| Time Base Pulse Width (ms) | Steering Gear Position (degree) |
|----------------------------|---------------------------------|
| 0.5                        | 0                               |
| 1.0                        | 45                              |
| 1.5                        | 90                              |
| 2.0                        | 135                             |
| 2.5                        | 180                             |

The steering gear has a reference circuit inside to produce a reference signal with a period of 20ms and a width of 1.5ms. There is a comparator that compares the applied signal with the reference signal to determine the direction and size to produce the motor's rotation signal. It can be seen that the steering gear is a position servo drive, the rotation range cannot exceed 180 degrees. The steering gear is the key device to realize the change of expression and maintain multi-functions for the investigated robot.

Second, in order to be able to coordinate the control of multiple steering gears, a 32-way servo controller is used to achieve the corresponding function. The 32-way servo controller connected to the steering gear control chip functions as the core control unit. Therefore, 32 external steering gears can be integrated for simultaneous control with conveniently adding the program according to the requirements in the corresponding control software. Meanwhile, the user can customize the instruction based on such configuration. The 32-way servo controller provides a variety of instructions, which can be distributed into the following three categories as shown in Table 2:

| Component Name                  | Commands                                      |
|--------------------------------|-----------------------------------------------|
| Single Steering Gear           | #1P1500T1000\n                                 |
| Multiple Steering Gear         | #1P1500#2P1500T1000\n                          |
| Action Chains                  | G1F3\r                                       |

For a single commander and a control command for multiple servos, the numbers 1 and 2 after "#" are the number of the steering gears. The number 1500 after "P" is the steering angle value in the range of
500-2500. The number 1000 after "T" is the speed of the steering gear in the range of 100-9999. For action group instructions, the number 1 after "G" represents the action group number, and the number 3 after "F" represents the cycle of executions of the action group.

In this study, the action group instructions are applied the steering gear number to adjust the position of the steering gear, and then add the action group. The dynamic process is achieved by the cycle number of operations. The specific program codes see the section 4.2 control system software design in this article.

3.2. Hardware design of emotional feedback function
Arduino receives the data from the ASR M08-A module and converts them to the 32-way steering controller through the serial port. The TXD and RXD serial ports on the 32-way servo controller functions as the serial port for data exchange with the Arduino microcontroller, and control the steering gear for the corresponding action.

4. Application of network module
W5100 network expansion module [5] directly uses the Ethernet library file in the Arduino IDE development environment and makes it as a simple Web server. Read and writing the digital and analog serial port and other network applications can be achieved through the network control as well. In this study, the W5100 module is used as a simple self-service hotel service system. The robot will automatically help to realize the corresponding function by saying the corresponding instructions such as "room reservation" and "check room" as well as the corresponding room on the webpage is listed as "subscribed", "can be booked" and so on (web page program to rely on JavaScript language to complete).

5. Design of control system
The core of the control system used in this study is the Arduino microcontroller. Arduino is an open source electronic design platform. The hardware part consists of an Atmel AVR microcontroller, I/O interface and related circuits. The software part includes the standard program compiler and program downloader, with similar Java and C language Processing / Wiring Development environment [6-7]. The main controller used in this study is Arduino Mega2560. Arduino Mega2560 core is ATmega2560 microcontroller, which has 54 digital input / output interface (16 channels can be used as PWM output), 16 analog input, 4 UART interface, a 16 MHz crystal oscillator, a USB port, and a power outlet, an ICSP header and a reset button. The control board is connected to a computer via standard USB cable for easy program debugging [8].

5.1. Design of control system hardware
Arduino Mega2560 as the core of the control system, including the ASR M08-A non-specific speech recognition module, 32-way servo controller, a number of steering gear, and W5100 network module. The circuit wiring diagram is shown in Figure 3.

Figure 3. The connection details of hardware components
5.2. Design of control system software

This study makes use of the Arduino IDE development environment combined with the JavaScript language written by the web program to achieve the simple application of the robot environment. Because the actual program is way too long, only an example to achieve a smile action is demonstrated here:

```cpp
#include"SoftwareSerial.h"

void setup()
{
    Serial.begin(9600);
    Serial1.begin(9600);
}

void loop()
{
    switch(Serial.read())
    {
        case 20:
        {
            Serial1.print("G21F3\n\n");
            delay(800);
            }break;
        default:  break;
    }
}
```

In case 20, 20 is the return value for the voice module receiving a smile command transferred to the microcontroller. 21 in the serial output of G21F3 is the action group label of smile and 3 is the number of executions. If you want to achieve more actions, it is only needed to add action group and set the corresponding return value.

6. Conclusions

Through the study of multi-function robots with speech interaction and emotion feedback, combined with the Internet of things technology, not only the basic functions of the case are realized, but also the self-help hotel applications are completed with the aid of the robots. In conjunction with the increasing demand for service-oriented and networking products in the today's society, we believe that this robot will certainly have a broader development prospects in the future.
7. Acknowledgement
The project is supported in part by the Undergraduate Innovative Discipline Project in China (201610359048) and the National Science Foundation of China (51405126).

8. References
[1] Qin D, 2003 Alive robots, News Lover, April 30(46)
[2] Liu R, Peng S and Liu G, 2014 Speech recognition embedded in intelligent house-holding system, Journal of Guangdong University of Technology, 2, pp. 50-53
[3] Li H, Mao X and Song J, 2016 Intelligent bionic two-way sign language translation system, Application of Electronic Technology, 42(7), pp. 83-86
[4] Su B, 2012 Design of speech recognition system based on AVR microcontroller, Modern Electronic Technique, 35(11), pp. 136-138
[5] Chen S, Wang Q and Wang T, 2013 A data transmission system based on W5100, Application of Electronic Technology, 39(9), pp. 54-60
[6] Ji X, 2012 Design of intelligent light-seeking car based on Arduino development environment, Modern Electronic Technique, 35(15), pp. 161-163
[7] Wu Y, Li L, Chen S and Pen W, 2014 Design of photoelectric encoder detector based on Arduino development environment, Modern Electronic Technique, 37(2), pp. 124-125
[8] Cui Y, Zhang W and Bai Y, 2013 An intelligent home control system based on Arduino, Application of Electronic Technology, 39(9), pp. 54-60