Voice controlled robot

Vishnu Theja Reddy MJ
vishnutheja1998@gmail.com
SRM Institute of Science and Technology,
Chennai, Tamil Nadu

Ravi Shankar Reddy
ravishankarreddychenchu666@gmail.com
SRM Institute of Science and Technology,
Chennai, Tamil Nadu

ABSTRACT

When we say voice control, the first term to be considered is Speech Recognition i.e. making the system to understand human voice. Speech recognition is a technology where the system understands the words (not its meaning) given through speech. Speech is an ideal method for robotic control and communication. The speech-recognition circuit we will outline, functions independently from the robot’s main intelligence [central processing unit (CPU)]. This is a good thing because it doesn’t take any of the robot’s main CPU processing power for word recognition. The CPU must merely poll the speech circuit’s recognition lines occasionally to check if a command has been issued to the robot. We can even improve upon this by connecting the recognition line to one of the robot’s CPU interrupt lines. By doing this, a recognized word would cause an interrupt, letting the CPU know a recognized word had been spoken. The advantage of using an interrupt is that polling the circuit’s recognition line occasionally would no longer be necessary, further reducing any CPU overhead. Another advantage to this stand-alone speech-recognition circuit (SRC) is its programmability.

Keywords— Arduino, Voice commands, CPU, Bluetooth module

1. INTRODUCTION

To control and command an appliance (computer, VCR, TV security system, etc.) by speaking to it, will make it easier while increasing the efficiency and effectiveness of working with that device. At its most basic level speech recognition allows the user to perform parallel tasks, (i.e. hands and eyes are busy elsewhere) while continuing to work with the computer or appliance. Robotics is an evolving technology. There are many approaches to building robots, and no one can be sure which method or technology will be used 100 years from now. Like biological systems, robotics is evolving following the Darwinian model of survival of the fittest. Suppose you want to control a menu driven system. What is the most striking property that you can think of? Well, the first thought that came to our mind is that the range of inputs in a menu-driven system is limited. In fact, by using a menu all we are doing is limiting the input domain space. Now, this is one characteristic which can be very useful in implementing the menu in standalone systems. For example, think of the pine menu or a washing machine menu. How many distinct commands do they require? Ease of Use

2. LITERATURE SURVEY

The purpose of this project is to build a robotic car which could be controlled using voice commands. Generally, these kinds of systems are known as Speech Controlled Automation Systems (SCAS). Our system will be a prototype of the same.

We are not aiming to build a robot which can recognize a lot of words. Our basic idea is to develop some sort of menu driven control for our robot, where the menu is going to be voice driven. What we are aiming at is to control the robot using following voice commands. Robot which can do these basic tasks:

1. Move forward
2. Move back
3. Turn right
4. Turn left
5. Load
6. Release
7. Stop ( stops doing the current job )

Table 1: Input and output actions

<table>
<thead>
<tr>
<th>INPUT (Speaker speaks)</th>
<th>OUTPUT (Robot does)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward</td>
<td>Moves forward</td>
</tr>
<tr>
<td>Back</td>
<td>Moves back</td>
</tr>
<tr>
<td>Right</td>
<td>Turns right</td>
</tr>
<tr>
<td>Left</td>
<td>Turns left</td>
</tr>
<tr>
<td>Stop</td>
<td>Stops doing current task</td>
</tr>
<tr>
<td>Load</td>
<td>Lifts the load</td>
</tr>
<tr>
<td>Release</td>
<td>Releases the load</td>
</tr>
</tbody>
</table>

A. Speech recognition is to determine which speaker is present based on the individual’s characterization. Several techniques have been proposed for compensating the mismatch occurred between the testing and training sessions. The communication among human-computer interaction is called a human-computer interface.
B. Since 1960s computer scientists have been researching ways and means to make computers able to record, interpret and understand human speech. In computer science, speech recognition (SR) is the translation of spoken words into text. It
is also known as "automatic speech recognition", "ASR", "computer speech recognition", "speech to text", or just "STT".

3. DESIGN APPROACH
The most challenging part of the entire system is designing and interfacing various stages together. Our approach was to get the analog voice signal digitized. The frequency and pitch of words be stored in a memory. These stored words will be used for matching with the words spoken. When the match is found, the system outputs the address of stored words. Hence we have to decode the address and according to the address sensed, the car will perform the required task. Since we wanted the car to be wireless, we used RF module. The address was decoded using a microcontroller and then applied to the RF module. This together with driver circuit at receivers end made complete intelligent systems.

3.1 System design
(a) Voice Recognition Module
(b) Microcontroller and Decoder
(c) RF module
(d) Motor Driver Circuit
(e) Buffer

4. BLOCK DIAGRAM
The speech recognition module basically consists of:

4.1 Voice Recognition Chip
It is the heart of the entire system. HM2007 is a voice recognition chip with the on-chip analog front end, voice analysis, recognition process, and system control functions. The input voice command is analyzed, processed, recognized and then obtained at one of its output port which is then decoded, amplified and given to motors of robot car. We initially used an Indian manufactured voice recognition chip AP7003. It is a monolithic user dependence speech recognition IC designed for toy application. AP7003 consist of microphone amplifier, A/D converter, speech processor and I/O controller. After pre-recording, AP7003 can recognize up to 12 different sentences each with 1.5-sec length with highly I/O programmability. However, it was not much accurate and reliable. So we started looking for another alternative. We found HM 2007 as a right choice. The chip provides the options of recognizing either forty 96 second words or twenty 1.92 second words. This circuit allows the user to choose either the .96 second word length (40-word vocabulary) or the 1.92-second word.
4.3 Microcontroller circuit

It is the second most important part of the project. The output from the chip is given to decoder (micro-controller) which acts as a DMC i.e. a Digital Motor Controller. DMC senses the output ports of the HM2007 chip and produces proper o/p as per the commands forward, backward, left, right, load, release, stop. The proper functionality of the system depends on the proper decoding logic. We use port0 as an input port and port1 as output port. P0.0 to P0.6 are given inputs from 7 output pins of the voice recognition module while P0.7 is kept grounded.

5. FUTURE SCOPE

(1) This research work has been narrowed down to a short-range ZigBee module. Using long-range modules will result in connectivity with the robot for long distances.
(2) Power Optimization such as sleep and wake-up schedules can be incorporated.
(3) Image processing can be implemented in the robot to detect the color and the objects.
(4) For more accurate working servo motors can be deployed.
(5) Automatic Targeting System can be implemented in the robot for tracking the target.

6. REFERENCES

[8] The 8051 microcontroller –Kenneth Ayala, 3rd reprint, 2005; Thomson Asia Ltd., Singapore; Chapter 3,6,7,8. For programming 89S51