ARDUINO BASED VOICE CONTROLLED ROBOT

K. Kannan[1], Dr. J. Selvakumar[2]

[1]*PG Scholar, Embedded System Technology, SRM University, Tamilnadu, India*

[2]*AP (S.G), ECE Department, SRM University, Tamilnadu, India*

---***---

*Abstract --- Voice Controlled Robot (VCR) is a mobile robot whose motions can be controlled by the user by giving specific voice commands. The speech is received by a microphone and processed by the voice module. When a command for the robot is recognized, then voice module sends a command message to the r**obot's** microcontroller. The microcontroller analyzes the message and takes appropriate actions. The objective is to design a walking robot which is controlled by servo motors. When any commands are given on the transmitter, the EasyVR module will take the voice commands and convert the voice commands into digital signals. Then these digital signals are transmitted via ZIGBEE module to the robot. On the receiver side the other ZIGBEE module receives the command from the transmitter side and then performs the respective operations. The Hardware Development board used here is ATmega 2560 development board. In ATmega 2560 there are 15 PWM channels which are needed to drive the servo motors. Addition to this there is camera which is mounted in the head of the robot will give live transmission and recording of the area. The speech-recognition circuit 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 pro**cessing 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. The software part is done in Arduino IDE using Embedded C. Hardware is implemented and software porting is done.*

Key Words: Arduino, ATmega 2560, EadyVR, Servo

1. INTRODUCTION:

1.1 Voice Control:

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 form the robot's main intelligence [central processing unit (CPU)]. This a good thing because it doesn't takes any of the robots 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. You can program and train the SRC to recognize the unique words you want recognized. The SRC can be easily interfaced to the robot's CPU.

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. By using a menu all we are doing is limiting the input domain space.

1.2 Use of Building Robots:

Robots are indispensable in many manufacturing industries. The reason is that the cost per hour to operate a robot is a fraction of the cost of the human labor needed to perform the same function. More than this, once programmed, robots repeatedly perform functions with a high accuracy that surpasses that of the most experienced human operator. Human operators are, however, far more versatile. Humans can switch job tasks easily. Robots are built and programmed to be job specific. You wouldn't be able to program a welding robot to start counting parts in a bin. Today's most advanced industrial robots will soon become "dinosaurs." Robots are in the infancy stage of their evolution. As robots evolve, they will become more versatile, emulating the human capacity and ability to switch job tasks easily. While the personal computer has made an indelible mark on

society, the personal robot hasn't made an appearance. Obviously there's more to a personal robot than a personal computer. Robots require a combination of elements to be effective: sophistication of intelligence, movement, mobility, navigation, and purpose.

Without risking human life or limb, robots can replace humans in some hazardous duty service. Robots can work in all types of polluted environments, chemical as well as nuclear. They can work in environments so hazardous that an unprotected human would quickly die.

1.3 THE TASK:

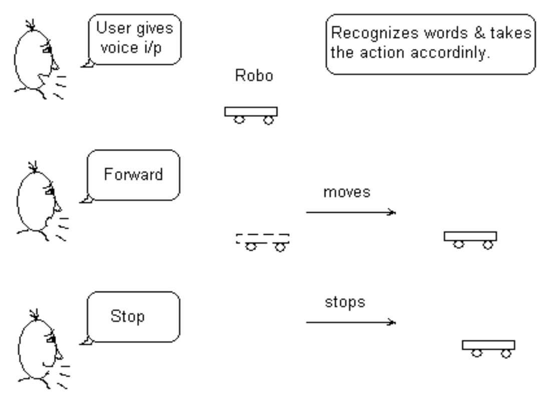


Fig 1: Basic Tasks of Robot

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.

The robot 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)

INPUT(Speaker speaks)	OUTPUT (Robot does)
Forward	moves forward
Back	moves back
Right	turns right
Left	turns left
Load	Lifts the load
Release	Releases the load
Stop	stops doing current task

Table 1: Tasks

2. SYSTEM DESCRIPTION:

2.1 Existing System:

The Existing system is a speech recognizing system. Speech recognition is the process of capturing spoken words using a microphone or telephone and converting them into a digitally stored set of words.

The quality of a speech recognition systems are assessed according to two factors: its accuracy (error rate in converting spoken words to digital data) and speed (how well the software can keep up with a human speaker).

Speech recognition technology has endless applications. Commonly, such software is used for automatic translations, dictation, hands-free computing, medical transcription, robotics, automated customer service, and much more. If you have ever paid a bill over the phone using an automated system, you have probably benefited from speech recognition software.

Speech recognition technology has made huge strides within the last decade. However, speech recognition has its weaknesses and nagging problems. Current technology is a long way away from recognizing

conversational speech. Despite its shortcomings, speech recognition is quickly growing in popularity. Within the next few years, experts say that speech recognition will be the norm in phone networks the world over. Its spread will be aided by the fact that voice is the only option for controlling automated services in places where touch tone phones are uncommon.

2.2 PROPOSED WORK:

Voice Controlled System

While speech recognition is the process of converting speech to digital data, voice recognition is aimed toward identifying the person who is speaking.

Voice recognition works by analyzing the features of speech that differ between individuals. Everyone has a unique pattern of speech stemming from their anatomy (the size and shape of the mouth and throat) and behavioral patterns (their voice's pitch, their speaking style, accent, and so on).

The applications of voice recognition are markedly different from those of speech recognition. Most commonly, voice recognition technology is used to verify a speaker's identity or determine an unknown speaker's identity. Speaker verification and speaker identification are both common types of voice recognition.

Speaker verification is the process of using a person's voice to verify that they are who they say they are. Essentially, a person's voice is used like a fingerprint. Once a sample of their speech is recorded, a person's speech patterns are tested against a database to see if their voice matches their claimed identity. Most commonly, speaker verification is applied to situations where secure access is needed. Such systems operate with the user's knowledge and cooperation.

Speaker identification is the process of determining an unknown speaker's identity. Unlike speaker verification, speaker identification is usually convert and done without the user's knowledge.

For example, speaker identification can be used to identify a criminal solely by their voice. In this situation, a sample of their voice would be checked against a database of criminals' voices until a match is found. Recently, this technique was used to identify a South American drug kingpin who had obscured his physical identity by undergoing extensive plastic surgery.

3. DETAILED DESCRIPTION OF PROPOSED WORK:

3.1 Transmitter Section:

On the transmitter section, voice commands are given to the EasyVR module. The EasyVR module will then take the voice commands convert it into digital values by using inbuilt analog to digital converter (ADC) and compare it with the predefined voice commands (for eg: 11 – forward, 12 – backward) and transmits those values according to the voice commands in the form of binary. This binary information is then received by the Microcontroller (ATmega 2560) and enters into the switch case. It will compare the value with the cases and according to it the string with the command is transmitted via ZIGBEE module.

3.2 Receiver Section:

On the receiver section, the digital signals are received by the ZIGBEE receiver module, and it sends the binary values to the microcontroller (ATmega 2560). The micro controller enters into the switch case and compares those string values with the values in switch case. Then according to the string value it will drive the servo motors in a continuous loop.

Here simultaneously three or more servos work, because the robot should walk for that at least three servos has to work together.

4. ALGORITHM:

1. The voice commands should be trained to the EasyVR module.
2. Then the stored voice commands are represented in the form of binary numbers such as move forward – 001, move backward – 010 etc.
3. These binary values are transmitted via zigbee module which is a transceiver.
4. The transmitted binary values are then received by another zigbee module which is present on the receiver side.
5. Microcontroller will take those binary values and performs action(servo motors) according to the binary values.

5. DISADVANTAGES OF EXISTING SYSTEM:

1. Even the best speech recognition systems sometimes make errors. If there is noise or some other sound in the room (e.g. the television or a kettle boiling), the number of errors will increase.
2. Speech Recognition works best if the microphone is close to the user (e.g. in a phone, or if the user is wearing a microphone). More distant microphones (e.g. on a table or wall) will tend to increase the number of errors.
3. In Speech recognition system, there is a possibility of unauthorized usage. Since this doesn't depends upon which person is speaking.
4. No password protection.

6. RESULTS:

6.1 Training of Voice Module:

For the first-time use, we need to do some configuration:

1. Select the serial baud rate (default 9600)
2. Select the communication mode: Common Mode or Compact Mode
3. Recording five instructions of the first group(or 2nd or 3rd as required)
4. Import the group you need to use (only recognize 5 instructions within one group at the same time)

After all the setting above, you can speak or send voice instruction to it. If identified successfully, result will be returned via serial port in the format: group number + command number. For example, return Result: 11 (Compact mode returns 0x11) means identified the first command of group 1.

If voice instruction is recorded, each time after you power it on, you need to import the group before letting it identify voice instructions.

6.2RECORDING STAGE:

1. Record indication: D1 (RED) flashes 3 times within the 600ms, then off for400ms, and then flashes quickly for 4 times within 600ms. Now the recording indication is over.
2. Begin to speak: D1 (RED) is off for 400ms, and then is on. Voice during the time while D1 (RED) is on will be recorded by this module.
3. Recording a voice instruction successfully for the first time: D1 (RED) off, D2 (ORANGE) on for 300ms.
4. Recording a voice instruction successfully for the first time: D1 (RED) off, D2 (ORANGE) on for 700ms.
5. Recording failure: D2 (ORANGE) flashes 4 times within the 600ms. In cases that voice instructions detected twice don't match, or the sound is too large, or there is no sound, recording will fail. You need to start over the recording process for that instruction.

6.3 Waiting Mode:

In waiting mode, D2 (ORANGE) is off, and D1 (RED) is on for 80ms every other 200ms, fast flashing. In this mode, it doesn't recognize voice command, only waiting for serial commands.

6.4 Recognition Stage:

In identification stage, D2 (ORANGE) is off, and D1 (RED) is on for 100ms every other 1500ms, slow flashing. In this stage, this module is processing received voice signal, and if matching, it will send the result immediately via serial port.

6.5 Recording:

Before using it, we have train it by recording voice instructions. Each voice instruction has the maximum length of 1300ms, which ensures that most words can be recorded. Once you start recording, you can't stop the recording process until you finish all the 5 voice instructions recording of one group. Also, once you start recording, the previous voice instructions in that group will be erased.

First, you need a serial tool AccessPort

Serial port setting:

Baud rate: 9600

Parity bit: None

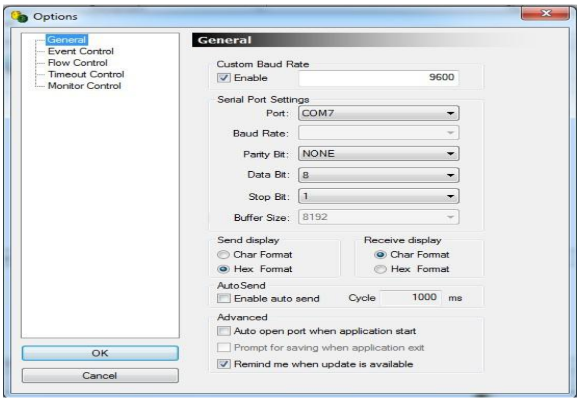
Data bit: 8

Stop bit: 1

Send format: Hex

Receive format: Char

Fig 2: Serial Port Setting



Send commands:

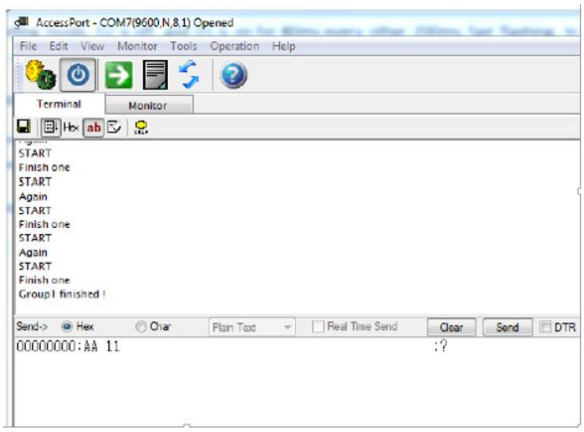


Fig 3: Commands

6.6 Recognition

We added another way to import the voice instruction group on V2. For V1, the only way is to send command to it through serial port, for example:

- Send command 0xAA21 to import group 1. Send command 0xAA22 to import group 2.
- Send command 0xAA23 to import group 3.

Once the group is important, it will output message through serial port. It could have 15 voice instructions in 3 groups. Each time you need to import the group before it could recognize instructions in that group. That means, this module could recognize 5 voice instructions at the same time. In recognition stage, this module could receive other serial commands. It will not exit the recognition stage until you send 0XAA00 or delete that group, or begin recording instruction

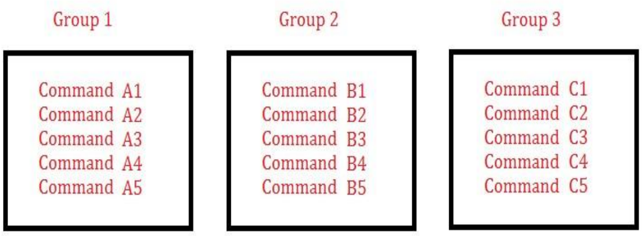
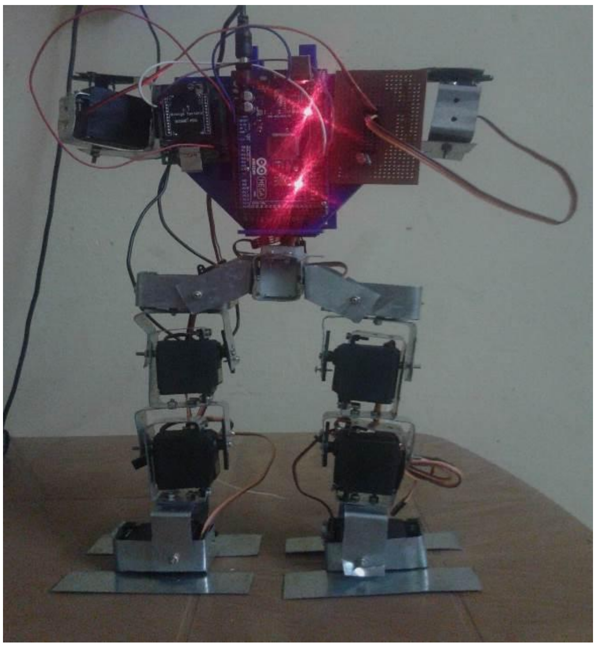


Fig 4: Importing the command

6.7 HARDWARE IMPLEMENTATION:



7. SCOPE FOR FUTURE WORK:

1. This research work has been narrowed down to short range zigbee module. Using a long range modules will result in connectivity with the robot for long distances.
2. Power Optimization such sleep and wakeup 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.

REFERENCES:

1. *A study on real-time control of mobile robot with based on voice command,* Byoung-Kyun Shim ; Yoo-Ki Cho ; Jong-Baem Won;Sung-HyunHan Control, Automation and Systems (ICCAS), 2011 11th International Conference on Publication Year: 2011

2. *Speech recognition and its application in voice-based robot control system*, Luo Zhizeng; ZhaoJingbing Intelligent Mechatronics and Automation, 2004. Proceedings. 2004 International Conference,

3. A study on precise control of autonomous driving robot by voice recognition, Sung-Won Jung ; Ki-Won Sung ; Moon-Youl Park ; Eon-Uck Kang ;Won-Jun Hwang ; Jong-Dae Won; Woo-Song Lee ; Sung-Hyun Han Robotics (ISR), 2013 44th International, Symposium on DOI:10.1109/ISR.2013.6695640 Publication Year: 2013

4. A mechanism for human - robot interaction through informal voice commands, D'Mello, S.; Lee McCauley, ; Markham, James Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop, DOI:10.1109/ROMAN.2005.1513777, Publication Year: 2005

BIBILOGRAPHY:

K. Kannan received his B.Tech.(Electronics and Communication Engineering)from Annamacharya Institute of Technology and sciences (AITS), India in 2013. He is currently pursuing his M.Tech. (Embedded System Technology) in SRM University, India. His interests include Robotics, Embedded Systems, Wireless Sensor Networks, Wireless Communication.

Dr. J. Selvakumar received his B. E (Electrical & Electronics) Madras University, India in 1999, M. E. degree in VLSI Design from ANNA University, Chennai in 2003. He received his Doctorate in the area of Low Power VLSI Design from SRM University, Chennai in 2013. Currently, he is an Assistant Professor (Selection Grade) in the Department of Electronics and Communication Engineering at S.R.M. University, Chennai, India, working for more than a decade, since 2003. He has more than 5 reputed publications in international journals with high citation (H-index), and 10 IEEE/SCOPUS cited international conference publication.