Speech Enabled System for Robot

Rishabh Anand
Computer Science Department
Lingaya's University
Faridabad, Haryana, India
rishabhanand1911@gmail.com

Siddharth
Computer Science Department
Lingaya's University
Faridabad, Haryana, India

Dr. Kamlesh Sharma
Associate Professor, CSE
Lingaya's University
Faridabad, Haryana, India
Kamlesh0581@gmail.com

***Abstract*—** The science of speech recognition has been advanced to the state where it is now feasible to communicate reliably with a computer or machine by speaking command to it in a disciplined manner using a vocabulary of the moderate size. It is the purpose of this paper to outline the fundamental concept of voice operated machine, Robot in this case. Now-a-days, Scientists are trying to develop Human-like or Humanoid Robots to enable them for work which either a simple human can't perform or takes more time. Voice System which will be used for controlling the Robot via speech input from any user with the help of an Android application using Bluetooth technology. Voice Recognition system allow people to control a machine by speaking to it through a microphone, entering text, or issuing commands to the system.

***Keywords*—** Voice recognition, Voice Operated, Speech Recognition, Motor driven circuit, Bluetooth device.

1. INTRODUCTION

1.1 Speech Recognition in simple terms is considered as, making the system understand Human voice. It is a technology where the machine is able to understand the basic words and not the meaning of the spoken words given through speech or voice. Speech is an ideal method for controlling robots or to communicate with them. The speech recognition protocol we will maintain, functions independently from the robot's main intelligence i.e. the Central Processing Unit (CPU). This results in less consumption of the robot's main CPU processing power for word recognition protocol. The CPU only has to check the voice circuit's recognition lines periodically to check if a command has been issued to the robot or not. We can even improve this system by connecting the recognition line to one of the CPU interrupt lines of the machine. By performing this function, a recognized word would cause an interrupt, making the CPU aware of a recognized word that has been spoken. The basic advantage of using an interrupt is that polling the circuit's recognition line periodically would not be necessary, further reducing any CPU overhead.

1.2 One more advantage to this, Stand-Alone Speech Recognition Circuit (SRC) is its simple programmability. We can easily program and train the SRC to recognize the unique words we want to be recognized by the machine. This SRC can easily be interfaced to the machine's CPU. To command as well as control an appliance (such as a computer, a VCR, TV security system, etc.) by speaking to it, will make it easier, while increasing the efficiency & effectiveness of working with the machine. At the most fundamental level of speech recognition, it allows the user to perform parallel tasks, for example; Even if the hands and eyes are busy elsewhere you can continue to work with the appliance or machine simultaneously.

1.3 **Robotics** is an ever evolving technology and here are numerous ways to build robots, and no one can be sure of which technology will be used 100 years from now. Like the biological systems, robotics is also evolving by following the Darwinian model of **"Survival of the Fittest".**

1.4 For instance, we want to control a menu driven system. What is the most essential property that we can think of? Well, the initial thought that came to our mind is the range of inputs in a menu driven system are limited. As a matter of fact, by using a menu we are just limiting the input domain space. But this is one characteristic could be very useful in implementing the menu in the stand-alone systems. For example think of a washing machine menu. How many different commands do they need? Something to think upon! [1], [2], [13], [14]

2. THE TASK

2.1 The aim of this project is to build a basic robot which could be controlled using voice commands of humans which in our case will be through a mobile device using Bluetooth technology. Basically these kinds of systems are known as Speech Controlled Automation Systems (SCAS). Hence, our system can be considered as a prototype of SCAS.

2.2 We are certainly not aiming to build a high tech robot which can recognize a lot of words. Our simple idea is to develop some sort of menu driven control for our robot, where the menu is going to be a voice driven one. What we are aiming to control the robot using the following voice commands. [2], [14]

2.3 Robot which can do these simple tasks:-

1. Move forward
2. Move back(reverse)
3. Turn left
4. Turn right
5. Light on
6. Light off

2.4 Sample of Input & Output

INPUT (Voice Commands)	OUTPUT (Robot performs)
Forward	Moves forward
Reverse	Moves back
Left	Turns left

Right	Turns right
Light On	Light turns on
Light Off	Light turns off
Stop	Robot stops

(Words are chosen in a way that they sound least familiar to one another)

3. SPEECH RECOGNITION

3.1 The voice enabled devices basically use the principal of speech recognition system. It is the process of electronically converting a speech waveform into words.

3.2 Converting a speech waveform into a sequence of words involves several necessary steps:

1. A microphone (mobile's mic in this case) picks up the signal of the speech to be recognized and converts it into an electrical signal (done by an application). A modern speech recognition system also requires the electrical signal to be represented digitally by means of an analog-to-digital conversion process, so that it can be processed with a digital computer or a microprocessor.
2. Then the speech signal is analyzed to produce a representation consisting of salient features of the speech module. The most general feature of speech is derived from its short-time spectrum, measured consecutively over short-time window of length 20 to 30 milliseconds overlapping at intervals of 10 to 20 milliseconds. Each short-time spectrum is converted into a feature vector, and the temporal sequence of such feature vectors in turn forms a speech pattern.
3. The speech pattern is further compared to a store of phoneme patterns or models through a dynamic programming process in order to produce a hypothesis (or number of hypotheses) of the phonemic unit sequence. (A phoneme is a basic unit of speech and a phoneme model is a concise representation of the signal that corresponds to a phoneme, usually embedded in a statement.) The speech signal has built-in substantial variations along many dimensions.

3.3 Before we understand the blueprint of the project, let's understand speech recognition types and styles first. Speech recognition is classified into two basic categories, *Speaker dependent* and *Speaker independent*.

1. ***Speaker dependent*** systems are trained by an individual who will be using the system by himself. Such systems are capable of achieving a high command count and more than 95% accuracy for speech/word recognition. The drawback of this approach is that the system will only respond accurately to the individual who trained the system. Hence, this is the most common approach enrolled in software for personal computers.
2. ***Speaker independent*** are those systems which are trained to respond to a word regardless of who speaks it. Therefore the system responds to a very large variety of speech patterns, inflections and enunciation's of the targeted word. The command word count is hence lower than the speaker dependent however high accuracy can still be maintained within the processing limits. The industrial requirements more often need speaker independent voice systems, such as the AT&T system used in the telephones.

3.4 A more established form of voice recognition is available through feature analysis and this technique usually leads to the "*speaker-independent*" voice recognition type. Instead of trying to find an exact match between the actual voice input and a previously stored voice template, this method initially processes the voice input using **"Fourier transforms"** or **"Linear Predictive Coding (LPC)"**, and then attempts to find typical similarities between the expected inputs and the actual digitized voice input. These similarities are present for a wide range of speakers, and so the system does not need to be trained by every new user. The types of speech differences that the speaker-independent methodology can deal with, but which pattern matching would fail to handle, include accents, and different speeds of delivery, volume, inflection, and pitch. Speaker-independent speech recognition system has proven to be quite difficult, with some of the huge hurdles being the variety of accents and inflections used by speakers of different origins. The recognition accuracy for the speaker independent systems is somewhat less than of that of the speaker-dependent systems, usually lying between 90 to 95 percent. Speaker independent systems does not ask to train the system as an advantage, but to perform with lower quality. Such systems find applications in telephonic communications such as dictating a number or a word where many people are in concern. However, there is a certain need for a well training database in the speaker independent systems. [1], [2], [13]

4. RECOGNITION STYLE

4.1 ***Speech recognition*** systems have another limitation concerning the style of speech they could recognize. They are three fundamental styles of speech: *Isolated*, *Connected* and *Continuous*.

a) ***Isolated*** speech recognition systems can only handle words that are spoken separately and is the most common speech recognition systems available at present. The user must take a pause between each word or command that is spoken. The S-R circuit is set up to identify isolated words of 0.96 seconds of length.

b) ***Connected*** is a half way point between isolated and continuous speech recognition. It allows users to speak multiple words at same time. The HM2007 can be set up to identify words or phrases up to 1.92 seconds in length which reduces the word recognition vocabulary number to just 20.

c) ***Continuous*** comes in the natural conversational speech of which we are used to in everyday life. It is exceptionally difficult for a recognizer to shift through the texts as the words tends to merge together. For example; "Hi, how are you doing?" sounds like "Hi,.howyadoin" in everyday life.

Continuous speech recognition systems are under continual development. [2], [8], [14]

4.2 *Base System Design:*

a. Voice Recognition Module

b. Microcontroller and Decoder(A mobile app in our case)
c. Motor Driver Circuit
d. Mobile robot

5 WORKING

This robot is the basic prototype of a Speech Enabled System. It houses a 8052 series Microcontroller consisting of 40pins & a clock frequency of 11.05MHz, and an automatic reset function, it also requires an external clock frequency for which a crystal oscillator is used. The code written in embedded C is burned inside it consisting of the commands and algorithms. There are two power supply modules; One of 12V DC Power supply for Relays, Light & Motors. Another of 5V DC Power supply for the Bluetooth module & Microcontroller respectively. A regulator of 7805 series is used for 12 to 5V step down as well as for constant power supply. Multiple capacitors and resistors are used for voltage fluctuations and filtering purposes. For connection, a Bluetooth module named HC-05 is added which through Serial communication, connects to the Microcontroller. The android application used, AMR_Voice is especially integerated with the code to perform uniquely and uses Google Speech-to-text modules through Internet. For the operations and command exchange between the Microcontroller & the Relays, a Driver, IC ULN 2003, is installed which transmits the operations one by one. Five Red LEDs are used to detect which relay is active and one Yellow LED is used to prompt if wrong command is given. Two heavy driven motors and axle system is used for driving the robot where two Relays are connected with one motor & two with another. And one Relay with the LED head light of the robot. The robot in the end works smoothly and efficiently after the complete system integration.

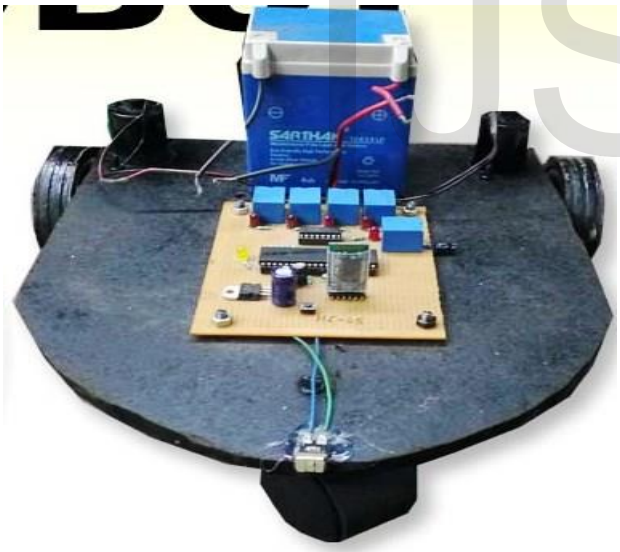

Fig. 5.1 Working Model of Robot

6 APPLICATIONS

Our belief is that such an interesting system will find wide variety of applications. The menu driven systems such as E-mail readers, common household appliances like washing machines, microwave ovens, and mobile phones etc. will completely become voice controlled in future time. [14]

• Voice enabled robots are useful in places where humans find difficult to reach but their voice & robots can reach easily. For example; In a fire-situations, in highly toxic areas, accident prone areas in mountains, etc.
• The robot can be used as a toy or even for household work.
• It can be used to bring and place large, heavy objects.
• This technology is the one of the important stage of constructing Humanoid robots.
• Command and control of appliances and equipment.
• Telephone assistance systems.
• Data entry systems.
• Speech and voice recognition security systems.

4 CONCLUSIONS AND FUTURE WORK

In this paper we present an idea to develop a Voice Operated Mobile Robot. The key factor in designing such system is to target the Future of Humanoid Robots. For example, Robots who will work simply upon the voice commands given to them and may as well interact with the Humans i.e. Humanoid Robots.

• Voice enabled robots are useful in places where humans find difficult to reach but their voice & robots can reach easily. For example; In a fire-situations, in highly toxic areas, accident prone areas in mountains, etc.
• The robot can be used as a toy or even for household work.
• It can be used to bring and place large, heavy objects.
• This technology is the one of the important stage of constructing Humanoid robots.
• Command and control of appliances and equipment.
• Telephone assistance systems.
• Data entry systems.
• Speech and voice recognition security systems.

5 CONCLUSIONS AND FUTURE WORK

In this paper we present an idea to develop a Voice Operated Mobile Robot. The key factor in designing such system is to target the Future of Humanoid Robots. For example, Robots who will work simply upon the voice commands given to them and may as well interact with the Humans i.e. Humanoid Robots.

6 REFERENCES

[1] Bojan Kulji, Simon János and Szakáll Tibor, Mobile robot controlled by voice, 5th International Symposium on Intelligent Systems and Informatics, 2007.

[2] Kamlesh Sharma, Dr. S. V. A. V. Prasad, Dr. T. V. Prasad, A Hindi Speech Actuated Computer Interface for Web Search, International Journal of Advanced Computer Science Applications, Oct 2012

[3] Kamlesh Sharma, Dr. T. V. Prasad, A Text-to-Hindi Speech Interface for Web Browsing, Proc. of National Conference on Advances in Knowledge Management, NCAKM 2010

[4] Dr. Kamlesh Sharma, Trilok Chand Sharma & Dr. T. V. Prasad , Voice Operated Computer Applicatio , Proc. of 2nd National Conference on Emerging Trends in Computer Science & Information Technology, ETCIT – 2009.

[5] Exploration of Speech Enabled System for English & Hindi by Dr. Kamlesh Sharma, T. Suryakanthi, Dr. T. V. Prasad, Proc. of the International Conference on System Modeling and Advancement in Research Trends (SMART), 2012.

[6] Article on Build a Speech Recognition Circuit. www.imagesco.com/articles/hm2007/SpeechRecognitionTutorial01.html

[7] For microphones types and properties. http://arts.ucsc.edu/ems/music/tech_background/TE-20/teces_20.html#I.

[8] For understanding microphone concepts, and other related concepts. www.Howstuffworks.com

[9] For multiple guiding video and research on voice operated robotic systems. www.youtube.com

[10] For multiple research work upon the voice integerated robots. http://ieeexplore.ieee.org

[11] Article on Voice operated Arduino for deep understanding. http://www.instructables.com/id/Voice-Activated-Arduino-Bluetooth-Android/

[12] Speech & Language Engineering by Martin Rajman. https://books.google.co.in/books?id=_ZTmDAAAQBAJ&pg=PA455&lpg=PA455&dq=voice+enabled+robot&source=bl&ots=4BV2Uz--Om&sig=w0Ix8CvlnCLKIJn3Cmgk-hHdm3M&hl=en&sa=X&ved=0ahUKEwjZr6SPlp7QAhUaTo8KHZ9eDXc4ChDoAQgkMAI#v=onepage&q=voice%20enabled%20robot&f=false

[13] Project report on Voice Controlled Camera enabled Robot. http://www.aiktcdspace.org:8080/jspui/handle/123456789/1561

[14] Article by Magesh Jayakumar. http://www.instructables.com/id/How-to-make-a-voice-control-direction-control-remo/

[15] Article & Tutorials on Controlling a robot using voice. https://www.generationrobots.com/en/content/59-speech-recognition-system-robot-parallax

[16] Society of Robots; Voice Recognition with Axon Tutorial. http://www.societyofrobots.com/sensors_voice_recognition_robot.shtml

[17] Thiang, Dhanny Wijaya, Limited Speech Recognition for controlling Movement of Mobile Robot Implemented on ATmega162 Microcontroller, International Conference on Computer and Automation Engineering, 2009.

[18] Report on Voice Controlled Robot. https://www.google.co.in/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&cad=rja&uact=8&ved=0ahUKEwiohcHXg7jQAhWJvY8KHU5sAd8QFggvMAQ&url=http%3A%2F%2Farfinfotech.com%2FPROJECTS%2Fvoice-controlled-robot.pdf&usg=AFQjCNHPLuKc7ZTyRhKBMUHjQ1Vx7kHzBw