RESEARCH ARTICLE   OPEN ACCESS

Speech Recognition for Robotic Control

Shraddha D. Gosavi*, Uday Pandit Khot**, Seema Shah***

*Post Graduate Student (M.E-EXTC), Vidyalankar Institute of Technology, Wadala, Mumbai-37
** Professor, Department of Electronics & Telecommunication, St. Francis Institute of Technology, Borivali, Mumbai.
*** Department of Information Technology, Vidyalankar Institute of Technology, Wadala-37

ABSTRACT
The main goal of this paper is to introduce "hearing sensor" and also the speech synthesis to the Mobile robot such that it is capable to interact with human through spoken natural language. The context of speech recognition refers to system where a person can speak via a microphone to a computer. The computer translates the spoken words into either text or commands to execute functions in the computer. The intelligent speech recognition system enables the robot to understand spoken instructions. The speech recognition system is trained in such a way that it recognizes defined commands and the designed robot will navigate based on the instruction through the speech commands. The complete system consists of three subsystems: the speech recognition system, a central controller, and the robot. The results prove that the proposed robot is capable of understanding the meaning of speech commands. They will act autonomously in a natural environment and will communicate in a natural way with those people they are supposed to support.
Keywords – human robot interface, robot, speech recognition,

I. INTRODUCTION

Since the first folding, tubular steel wheelchair was invented in 1932 and the first electric wheelchair was designed in the 1950's, the functionalities of the manual and electric wheelchair have been improved significantly. The electric wheelchairs have become more efficient, quieter and lower-maintenance in general. They also grant users more freedom with less assistance including in the control, styles, range or travel distance, maneuverability, seating and other user options. In contrast, the mobility of the manual wheelchair was limited by the user's physical condition and restricted his or her daily activities [1]. Over the past several decades a need has arisen to enable humans to communicate with machines in order to control their actions or to obtain information. Initial attempts at providing human-machine communications led to the development of the keyboard, the mouse, the trackball, the touch screen, and the joy stick. However, none of these communication devices provides the richness or the ease of use of speech which has been the most natural form of communication between humans for tens of centuries. As robots move from the factory into homes, offices, and other environments that are complex and unpredictable, it will be increasingly important for the software that controls those robots to be both adaptive and flexible. In the industrialized circumference, robotics is a step beyond mechanization. Robot plays an increasingly important role in the world economy and in daily experience. Engineers try to combine robots with mathematical and organizational tools to create complex systems for a rapidly expanding range of applications and human activities. It is one of the challenging areas in Human-Robot Interaction. Speech recognition technology is a great aid to admit the challenge and it is a prominent technology for Human-Computer Interaction and Human-Robot Interaction for the future.

Speech recognition allows you to provide input to an application with your voice. Just like clicking with your mouse, typing on your keyboard, or pressing a key on the phone keypad provides input to an application, speech recognition allows you to provide input by talking. This robot/Wheelchair is an automatic manipulator, which is programmable and controlled by basic voice command. A human can easily recognize a familiar voice however; getting a computer to distinguish a particular voice among others is a more difficult task.

Large numbers of vocabularies or many similar sounding words make recognition difficult for the system. Uncertainty is a major problem for navigation systems in mobile robots. Interaction with humans in a natural way, using English rather than a programming language, would be a means of overcoming difficulties with localization.

II. LITERATURE REVIEW

A number of issues define the technology of speech recognition systems. These include [2], [3]:

The manner:
There are generally three modes of speaking, including:

- Isolated word (or phrase) mode: In which the user speaks individual words (or phrases) drawn from a specified vocabulary.
- Connected word mode: In which the user speaks fluent speech consisting entirely of words from a specified vocabulary (e.g., telephone numbers);
- Continuous speech mode: In which the user can speak fluently from a large (often unlimited) vocabulary.

The size of the recognition vocabulary:
- Small vocabulary systems which provide recognition capability for up to 100 words;
- Medium vocabulary systems which provide recognition capability for from 100 to 1000 words;
- Large vocabulary systems which provide recognition capability for over 1000 words.

The knowledge of the user's speech patterns, including:
- Speaker dependent systems which have been custom tailored to each Individual talker.
- Speaker independent systems which work on broad populations of talkers, most of which the system has never encountered or adapted to.
- Speaker adaptive systems which customize their knowledge to each individual user over time while the system is in use.

The degree of dialogue between the human and the machine, including:
- One-way communication: In which each user spoken input is acted upon;
- System-driven dialog systems: In which the system is the sole initiator of a dialog, requesting information from the user via verbal input;
- Natural dialogue systems: In which the machine conducts a conversation with the speaker, solicits inputs, acts in response to user inputs, or even tries to clarify ambiguity in the conversation

III. BLOCK SCHEMATIC

This paper Of Speech Recognition for Robotic Control shown in "Figure 1", basically works on the principle of voice control through matlab. Robot/Wheelchair will be controlled by the voice of the user. On the initial stages the voice samples of the user will be analyzed and depending on the analysis the commands will be given to the microcontroller, which it turn will control the physical moments of the wheelchair. On chair Infrared sensors will be installed. Total 4 sensors will be installed for detection of wall in the forward, backward, left & right direction. These sensors will help the user to avoid collision from the obstacles as shown in "Figure 2". The purpose of this project is to control a robot and directs it to the desired location by sending some commands using speech instead of

other control tools. So, an algorithm is tried to be developed with MATLAB by using some Speech Recognition techniques.
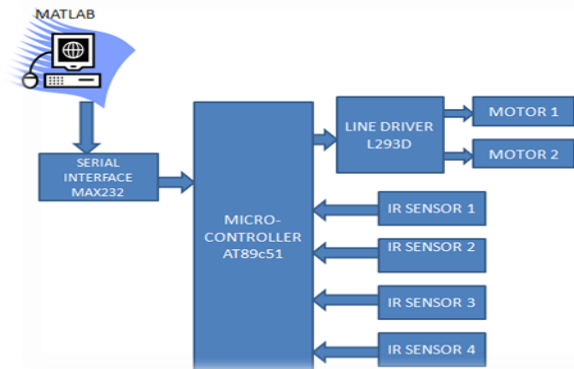

Figure 1: Block Diagram of speech recognition system.

To be able to send some codes to the robot a communication channel should be used and the robot in this project uses an RF serial communication channel to receive data. Communication is one directed. Only computer sends some commands related with directions. Sending commands to the robot can be done by using different ways. One of them may be to enter a character from a keyboard and then sending the value handled from the pressed key to the serial port. But, in this project control of the robot is wanted to be realized by using speech recognition. Speech recognition process will be done by using MATLAB. After recognition, the desired command will be sent to the serial port of the computer, and finally to the RF communication link of the robot. As you guess, there is a command table loaded to the microcontroller of the robot and the robot is directed to the desired location by sending appropriate data.
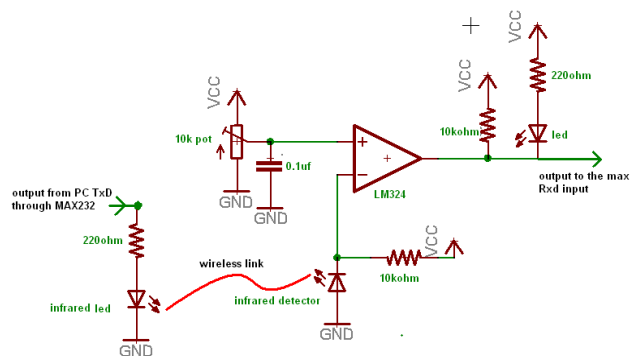

Figure 2: Infrared obstacle detector circuit.

There are two parts of this project. One of them is taking the speech signal from the microphone, analyzing and recognizing it and then giving the correct output. The other part of the project is about the hardware (robot).

IV. HARDWARE IMPLEMENTAION

Now to implement hardware we required a machine so chosen the easiest i.e. a robot A large part of the project involved choosing the appropriate hardware components The hardware requirement for robot in speech recognition is, AT89C51 is an 8-bit microcontroller and belongs to Atmel's 8051 family. ATMEL 89C51 has 4KB of Flash programmable and erasable read only memory (PEROM) and 128 bytes of RAM. It can be erased and program to a maximum of 1000 times The MAX232 IC is used to convert the TTL/CMOS logic levels to RS232 logic levels during serial communication of microcontrollers with PC. The controller operates at TTL logic level (0-5V) whereas the serial communication in PC works on RS232 standards (-25 V to + 25V). This makes it difficult to establish a direct link between them to communicate with each other. The transmitters take input from controller's serial transmission pin and send the output to RS232's receiver. The receivers, on the other hand, take input from transmission pin of RS232 serial port and give serial output to microcontroller's receiver pin. MAX232 needs four external capacitors whose value ranges from 1μF to 22μF. L293D is a dual H-bridge motor driver integrated circuit (IC). Motor drivers act as current amplifiers since they take a low-current control signal and provide a higher-current signal. This higher current signal is used to drive the motors.ADC MCP 3208 The Microchip Technology Inc. MCP3208 devices are successive approximation 12-bit Analog to-Digital (A/D) Converters with on-board sample and hold circuitry. Communication with the devices is accomplished using a simple serial interface compatible with the SPI protocol. The devices are capable of conversion rates of up to 100 kbps. Hardware implementation was done in Eagle 5.7.0 professional which is having the facility of PCB layout print.

In the first part of the project, microphone act as the input for the speech signal from the user .A PC sound cord in the microphone performs this function of converting analog speech signal to a suitable digital form for the further processing. Then the spectrum of the obtained digital signal is compared with that of the preprogrammed signal. Here we record the voice of a certain set of people for a certain set of words. Now the spectrum of the speech signal is compared with that of the Pre-programmed signal and the corresponding text output is transferred to the second level. This robot has got serial port connector which is used to connect the robot to PC.. Speech recognition process will be done by using MATLAB. After recognition, the desired command will be sent to the serial port of the computer, and finally to the robot. There is a command table loaded to the microcontroller of the robot and the robot is directed to the desired location .The robot should understand the dialogue, which are mentioned in the "Table 1", show the

sentences/dialogues we have chosen to evaluate our system. These sentences/dialogues are arranged in the tables on the basis of grammar complexity and robotic activities.

Table 1: Simple word for robotic activities.

Robotic Activities	Word
Moves Left	Left
Moves Right	Right
Moves Forward	Forward
Moves Reverse	Backward
Stop	Stop

Now as the recognition of a word or command is complete then the motors of the robot is controlled by sending the respective data through the serial port again through the MATLAB And the respective movement is seen in the robot.

V. SOFTWARE IMPLEMENTATION

Software required for the implementation of the Speech Recognition is Matlab & for the movement of robot is Embedded C. The flowchart for speech recognition program is shown in Figure 3 [4], where, The audio transmits through the microphone to the computer interface. The receiver driver of computer system captures the audio. The capture audio then converted to text string by sound to text converter subsystem of the program. This text string compared to the database for exaction command. If a command is found against the text the implementation command is send to robot for do the activity. And if the command matches with the text it will send a signal to activate the robot for desired movement mention by user.
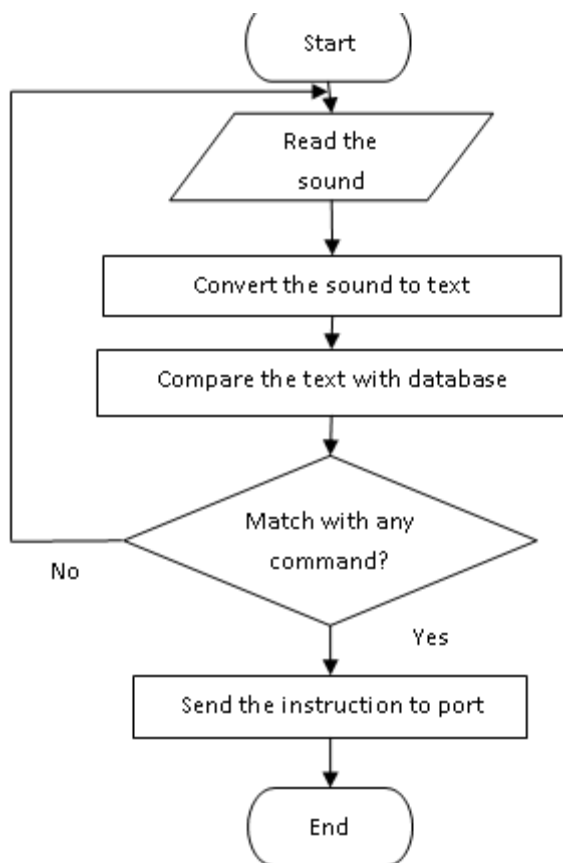
Figure 3: Flow Chart for Speech Recognition.

An engineer looking at a speech signal might characterize it as follows, [5]

- The bandwidth of the signal is 4 KHz
- The signal is periodic with a fundamental frequency between 80Hz and 350 Hz.
- There are peaks in the spectral distribution of the energy at $(2n-1)*500$ Hz; n=1, 2, 3…
- The envelope of the power spectrum of the signal shows a decrease with increasing frequency (-6dB per octave).

VI . RESULTS
As we run the MATLAB program, the GUI of speech recognition system appears on the screen that is Figure 4.
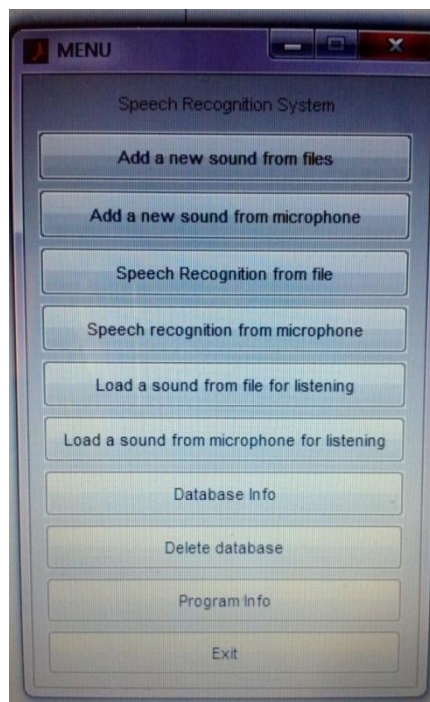

Figure 4: GUI of Speech Recognition System

This system includes
- Add a new sound from files
  This will add new sound (word) to our current directory of database, here files indicate the reference sound stored (left, right, forward & backward). From which we will select a word & add to the current directory ,that means if we select a word 'backward', label it as 1,which means we converted sound to text, as the sound is added we get a reply on screen which is shown in Figure 5.
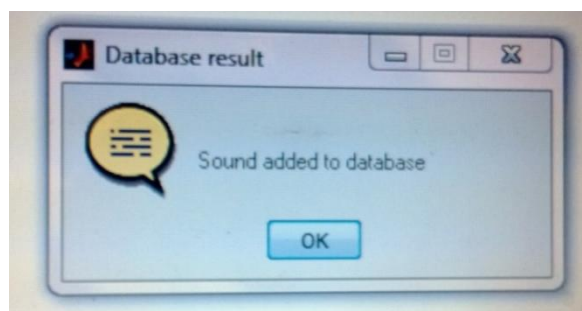

Figure 5: New sound added

- **Add a new sound from microphone**
  This is a sound which we will add directly from microphone with respect to the database which is previously stored. The system will request for the time required for the recording, & then Class ID will be given by user for future recognition. This is shown as below. & finally the sound will be added to database.The following parameters will be used during recording:

Sampling frequency22050
Bits per sample8
Insert the duration of the recording (in seconds):3
Now, speak into microphone...
Recording...
Recording...
Recording...
Recording...
Recording...
Recording...
Recording stopped.
Insert a class number (sound ID) that will be used for recognition: 3
  Sound added to database

- **Speech recognition from file**
  After adding the files to the database Speech Recognition will be done, that means the word given as a input from the files will be compared with the database. The text used by user will be compared with the database and the matching results will be displayed on screen which is shown in Figure 6. This in ID 1 indicates that the sound for text 1 is backward which is added in new sound, is recognized by the system.

Sound selected for recognition:
File:b.wav
Location
D:\component\Shradha_speech\speechcode1
Database scanning...
Warning: Calling MEX-file
'D:\component\Shradha_speech\speechcode1\dpcore.dll'.
MEX-files with .dll extensions will not execute in a future version of
MATLAB.
Sound #1
Sound #2
Sound #3
Sound #4
Sound #5
Matching sound:
File:b.wav
Location:
D:\component\Shradha_speech\speechcode1
Recognized speech ID:1

Figure 6: Recognized sound

- **Speech recognition from microphone**
  After adding the files to the database Speech Recognition will be done, that means the word given as a input from the microphone will be compared with the database which is stored in the file.

Insert the duration of the recording (in seconds):3
Now, speak into microphone...
Recording...
Recording...
Recording...
Recording...
Recording...
Recording...
Recording stopped.
Database scanning...
Sound #1
Sound #2
Sound #3
Sound #4
Sound #5
Matching sound:
File:Microphone
Location: Microphone
Recognized speech ID: 1

- Load a sound from file for listening
  This is to listen the sound/word uttered form the file.
- Load a sound from microphone for listening
  This is to listen the sound/word uttered form the microphone
- Database Info
  This shows the current directory of Database which we have made for current recognition system.

Database has #4words:
File:b.wav
Location:
D:\component\Shradha_speech\speechcode1
Sound ID: 1
File:f.wav
Location:
D:\component\Shradha_speech\speechcode1
Sound ID: 2
File:Microphone
Location: Microphone
Sound ID: 3
File:Microphone
Location: Microphone
Sound ID: 4

- Delete Database
  Clear the database for new words or Recognition system.
- Program Info
  It is about the Matlab codes.
- Exit
  To exit from the Speech Recognition System.

The algorithm was implemented using Matlab 7.9.0(R2009b) 32-bit (win 32). The project

proposed, to design a speech recognition system was successfully designed and working of it was checked with successful output. This work included the conversion of this text data into a mechanical action by using a microcontroller.

Experimental Results of speech recognition system using 05 words with 05 samples is 47.4%, & for 20 samples are 66.2% are shown in Table 2. In the Table 1, 20 samples of each word have been taken. Consider i.e. Forward, for which the recognition rate is 55%, those samples were taken from five different persons, there was movement in robot for two samples perfectly and for remaining two sample it can't recognize the word, and the last one sample was sometimes recognized. In this way, all the words were recognized as per recognition rate given in the Table 2.

Table 2: Results of Speech Recognition System using 5, 10, 15 & 20 samples of word.

Word	Recognition Rate%			
	05 samples	10 samples	15 samples	20 samples
Forward	33	47	53	55
Left	41	58	63	60
Right	59	68	70	76
Stop	46	53	62	62
Backward	58	60	72	78
Avg, Rec. rate	47.4%	57%	64%	66.2%

VII. CONCLUSION

The speech recognition for robotic control has been achieved in this paper. The highest recognition rate that can be achieved is 66.2%. This result is achieved by the system using 20 samples and 05 samples per word for data training. The overall performance of the system can be greatly improved if we explore the options more completely. It has been found that this robot can readily be used by handicapped people .It is found that for obtaining higher levels of performance we need to use microcontroller of higher memory levels. And for the application it is needed to include the basic functions such as turn around, move 180 degree etc. so as to enable the user to use it for day to day .

Systems designed for use with particular set of words, but system users may not know exactly which words are in system vocabulary, leads to certain percentage of out-of vocabulary words in natural conditions. Systems must have some method of detecting such out-of-vocabulary words. Speech has difficulties to be recognized by an application. Because speech is different for every speaker, May be fast, slow, or varying in speed. May have high pitch, low pitch, or be whispered. Have widely-varying types of environmental noise. Changes depending on the semantics of the utterance, has an unlimited number of words, has phonemes that can be modified, inserted, or deleted.

REFERENCES

[1] P. Nauth, *"Speech and Image Recognition for Intelligent Robot Control with Self Generating Will", IWSSIP 2010 - 17th International Conference on Systems, Signals and Image Processing*

[2] R. Jain and S. K. Saxena, "Voice Automated Mobile Robot*", International Journal of Computer Applications* (0975 – 8887) vol. 16, no.2, February 2011

[3] H. Huang, "Manual Wheelchair Automator: Design of a Speech Recognition System with Hidden MarkovModels and Joystick Steering Control" McMaster University Digital Commons@McMaster.

[4] Md. A. Al Ahasan, Md. A. Awal *and* S. S Mostafa *"*Implementation of Speech Recognition Based Robotic System". COPYRIGHT © 2011 IJCIT, ISSN 2078-5828 (PRINT), ISSN 2218-5224 (ONLINE), vol. 01, no. 02, Manuscript Code: 110108

[5] M. Balaganesh, E. Logashanmugam, and R. Manikandan, Sathyabama University, Chennai, India and C.S. Aadhitya Loughbrough University UK "Robotic Arm Showing Writing Skills by Speech Recognition", 978-1-4244-9005-9/10/$26.00©2010,IEEE