

# Implementation of Robotic System Using Speech Recognition Technique based on Neuro-Fuzzy Controller

Zeinab M. Omar<sup>1</sup>, Mohammed Tahseen Alagha<sup>2</sup>, Tawfig Abas<sup>3</sup>, Atif Osman Al Tahir<sup>4</sup>

<sup>1</sup>Department of Electrical and Control Engineering, Faculty Of Engineering, University Of Bahri, Khartoum, Sudan  
(E- Mail: Zoba\_Alberedgar@Hotmail.Com)

<sup>2</sup>Department of Electrical and Control Engineering, Academy of Engineering & Medicine Sciences  
(E- Mail : mohd\_ghal@hotmail.com)

<sup>3</sup>Department of Electrical and Control Engineering, Academy of Engineering & Medicine Sciences

<sup>4</sup>Department of Electrical and Control Engineering, Faculty Of Engineering, University Of Bahri, Khartoum,

**Abstract**— Recently, voice becomes one of the methods commonly used to control the electronic appliances, because of easily being reproduced by human compared to other efforts needed to operate to control some other appliances. There are many places which are hard or dangerous to approach by human and there are many people with disabilities do not have the dexterity necessary to control a keypad or a joystick on electrical devices. The aim of this study is to build a mobile robotic system, which can be controlled by analyzing the human voice commands. The robot will identify the voice commands and take action based on received signal. In general, the robotic system consists of the voice recognition module (AD-VR3) which serves as the ear that will listen and interpret the voice command, while the Arduino serve as the brain of the system that will process and coordinate the correct output of the input command to control the robot motors to perform the action.

**Keywords**— Arduino, robot; voice recognition, brain, disabilities, joystick.

## I. INTRODUCTION

Voice recognition robotic System would be an advanced control system that uses human voice/audio speech to identify the speech command. It has many applications and advantages such as providing support to disabled people, Alerts/warning signals during emergencies in airplane, train and/or buses, Develop of educational games and smart toys, Automatic payment and customer service support through telephones.

All that with No key required for devices such as personal computer and laptops, automobiles, cell phones, door locks, smart card applications, ATM machines etc [1].

The aim of this paper is to study the develop a voice driven control robot using artificial intelligence and speech recognition method, where the motors are going to be voice driven. Then the action can be taken based on the given commands. Generally, these kinds of systems are known as Speech Controlled Automation Systems (SCAS). Our system will be a prototype of the same [1]. Speech recognition is the process of electronically converting a speech waveform (as the realization of a linguistic expression) into words (as a best-decoded sequence of linguistic units). Converting a speech waveform into a sequence of words involves several essential steps[2]:

- 1) the microphone picks up the signal of the speech to be recognized and converts it into an electrical signal. A modern speech recognition system also requires that the electrical signal be represented digitally by means of an analog-to-digital (A/D) conversion process, so that it can be processed with a digital computer or a microprocessor
- 2) This speech signal is then analyzed
- 3) (in the analysis block) to produce a representation consisting of salient features of the speech. The most prevalent feature of speech is derived from its short-time spectrum, measured successively over short-time windows of length 20–30 milliseconds overlapping at intervals of 10–20 milliseconds. Each short-time spectrum is transformed into a feature vector, and the temporal sequence of such feature vectors thus forms a speech pattern.
- 4) The speech pattern is then compared to a store of phoneme patterns or models through a dynamic programming process in order to generate a hypothesis (or a number of hypotheses) of the phonemic unit sequence. (A phoneme is a basic unit of speech and a

phoneme model is a succinct representation of the signal that corresponds to a phoneme, usually embedded in an utterance.) A speech signal inherently has substantial variations along many dimensions [3].

## II. DESIGN HARDWARE IMPLEMENTATION

The most challenging part of the entire system is designing and interfacing various stages together. Our approach was to get the analog voice signal being digitized in the microphone. 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 received, the car will perform the required task. Since we wanted the car to be wireless, we used TX & RX wireless module. The address was decoded using decoder in microcontroller and then applied to TX module. This together with driver circuit at receivers end made our complete intelligent systems.

### A. Circuit Construction

Most ANFIS models are designed using software. The ease of manipulating data and changing the architecture make software a popular choice. An often unlooked at side of ANFIS is when they are created using hardware. The first goal of this study was to create a circuit implementing ANFIS technology that utilized stand-alone hardware to perform the functions instead of the more commonly used software.

The AD-VR3 and all the other components comprising this circuit were assembled and wired on a car model of FL-330 Breadboard. Table 1 below shows the parts list used in creating this circuit including components of transmitter and receiver circuit.

Table.1: Components Used In Circuit's Construction

Components	Type
Arduino uno	Microcontroller
AD_VR3	Voice recognition module v 3.1
Microphone	AVR's microphone
HT12E	Encoder
HT12D	Decoder
CDT-88	Transmitter module
CDR-C05A	Receiver module
L293D	Driver module
FL-330	Car module
SPL-003010	Multipurpose circuit board
1 M $\Omega$	Resistor
56 K $\Omega$	Resistor
LED	Green led
5V battery	Power bank

The circuit below was wired according to the schematic shown in Fig. 1. and 2 below, and a picture of the completed circuits is shown in Fig. 7. and 8.

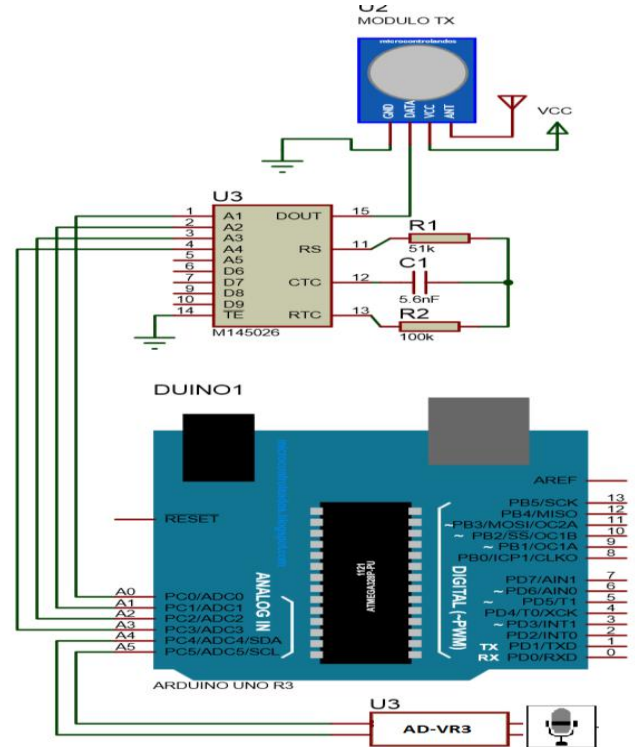


Fig.1: Transmitter Circuit

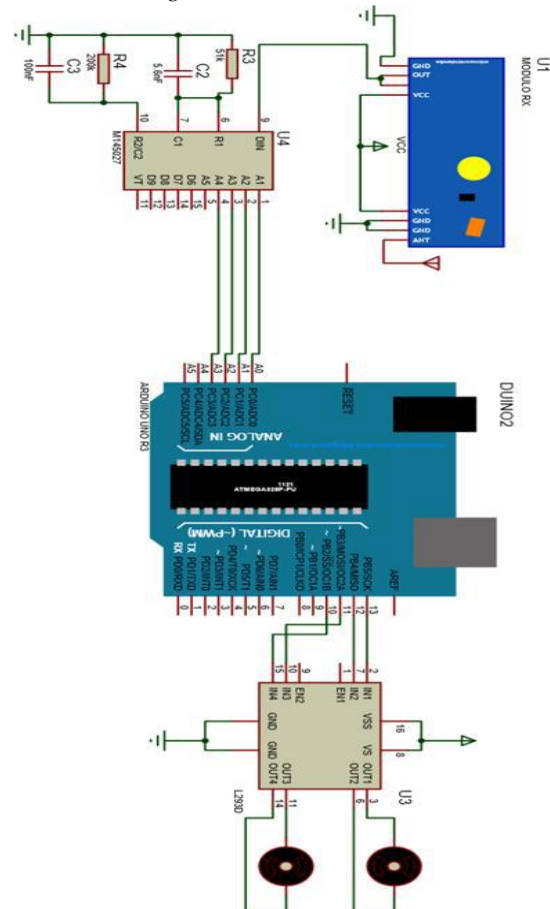


Fig.2: Receiver Circuit

## B. Schematic Descriptions

### 1) Microphone

The user will speak voice commands (“Right”, “Left”, “Forward”, etc.) through the microphone which will pick up and use as an input. The microphone will be an electrets condenser, so it will need a voltage input of about 4 to 5V. The microphone itself contains a built-in field effect transistor amplifier stage, so the sound is amplified before it is sent to the voice IC for further amplification.

### 2) Voice recognition module

The AD-VR3 will get the voice input from the microphone and perform the speech processing (training & testing) as required, the AD-VR3 Speech Recognition chip is the basis of the voice recognizing. For every command that has been trained into the chip, the corresponding output will be set high. Using the output, the correct action (“Right”, “Left”, “Forward”, etc.) can be implemented.



Fig.3: Microphone and voice recognition module

### 3) DC Power Source

The main power supply to all parts of the robot is a chargeable 9 V battery that is in the form of power bank for the receiver part and a direct DC supply of 5V for the transmitter part

### 4) Microcontroller (Arduino uno)

Arduino is an open-source electronics prototyping platform based on flexible, easy-to-use hardware and software. [7] The Arduino microcontroller is essential to the design of the robotic car as it provides communication between the voice recognition components and the motors.

Also the microcontroller will be essential for the integration of the rest of the system blocks. It will receive its inputs from the AD-VR3 voice module and interpret the signals to direct the motors to follow its next instructions. It will be integral for the microcontroller to decide when to send the input to the motor driver, which signal to send, and which motor to move at what time in order to get the robot to move in the correct direction and in a fluid motion.

### 5) DC Motor

The motion part contains two motors (right and left) with their drivers, the motors are DC types allowing the

wheelchair to move forward, backward, turn right, and turn left. The motors will be synchronized so that the movement of each side will be one fluid motion.

DC motors are very simple to use and control, which make them a short design-in item, generally two different styles of high torque DC motors: Brush Commutated and Gear Motor where the last one has high torque at load affects.

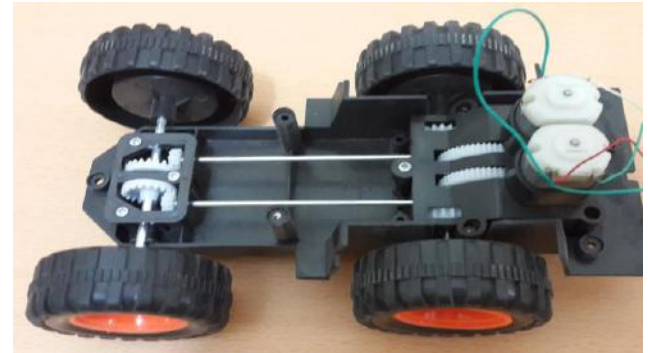


Fig.4: DC motor and main chassis of the robotic car

### 6) Motor Driver Controller (L293D)

The L293D is a quadruple high-current half-H driver. It is designed to provide bidirectional drive currents of up to 1 A at voltages from 4.5 V to 36 V. The L293D is also designed to provide bidirectional drive currents of up to 600-mA at voltages from 4.5 V to 36 V. Both devices are designed to drive inductive loads such as relays, solenoids, dc and bipolar stepping motors, as well as other high-current/high-voltage loads in positive-supply applications.

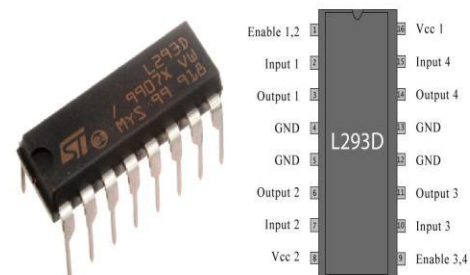


Fig.5: L293 driver

### 7) Transmitter and receiver modules (TX & RX)

Wireless Transmitter Modules allow the arduino to wirelessly communicate with other arduino, or with radio frequency (RF) controlled devices that operate in the same frequency (433Mhz in this case).

They work in pairs, meaning you need both a receiver and a transmitter to communicate with each other.

The receiver has 4 pins, but we actually use 3 of them: GND (Ground), VCC (5V) and one DATA pin.

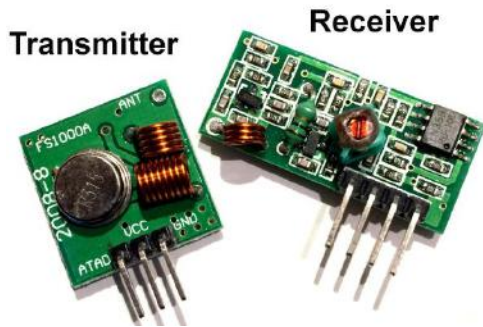


Fig. 6: Tx & Rx wireless modules

#### 8) Encoder and decoder

The HT12E Encoder IC is used at the end of the sending arduino for Remote Control of the system. They are capable of Encoding 12 bit of information which consists of N address bits and 12-N data bits. Each address/data input is externally binary programmable if bonded out. The HT12D decoder IC is placed before the receiver arduino to decode encoded the signals. For proper operation a pair of encoder/decoder with the same number of address and data format should be selected. The Decoder receive the serial address and data from its corresponding decoder, transmitted by a carrier using an RF transmission medium and gives output to the output pins after processing the data.

#### 9) Voice Recognition Circuit

The most important elements used in this research consist of a microphone which acts as a voice sensor connected to AD-VR3 for audio processing. The AD-VR3 is then connected to arduino microcontroller for wireless command communication to the robot which is received at the receiver end using another arduino circuit. The main transmitter board is shown in Fig 3.8 below. Voice Recognition Module is a compact and easy-control speaking recognition board. It is a speaker-dependent voice recognition module. It supports up to 80 voice commands in all. Max 7 voice commands could work at the same time. Any sound could be trained as command. Users need to train the module first before let it recognizing any voice command.

This board has 2 controlling ways: Serial Port (full function), General Input Pins (part of function). General Output Pins on the board could generate several kinds of waves while corresponding voice command was recognized. The other arduino circuit will act as a receiver, then transmit the received data to a microcontroller that drives the robot motors using L293D motor driver IC.

The circuit can be powered from a 5 volt battery or directly from a laptop usb slot that is connected to the Arduino Positive Voltage Regulator to limit and stabilize the board voltage to +5.0 volts. All ICs are powered from this regulated +5.0 volts. The microphone consists of the

only user interfaces with the circuit. The microphone is a standard microphone which acts as the transducer converting the pressure waves to an electrical signal. The microphone is coupled to the AD-VR3 module which is attempting to classify each word into the different trained categories.

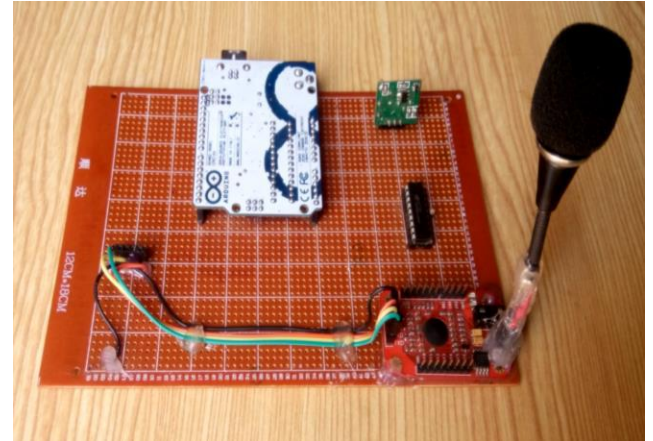


Fig. 7: Transmitter Circuit with Microphone and AD-VR3 module

#### C. Microcontroller Based Circuit

The microcontroller is the brain of the system and nothing can be done if it isn't fully functioning. It has the ability to send different signals to the DC Motors to reach the appropriate tasks [5].

The design of the mobile robot is simple yet convenient for the system. The main board and the arduino module along with the motor driver are placed on the upper outside of the vehicle as shown in Fig. 4. below

The mobile robot consists of a chassis mounted on four wheels out of which two are dummy wheels and the other two are attached to 12V gear motors. The complete circuit for the robot operation is placed on the chassis. The gear motors are driven by motor controller driver IC L293D for forward, backward, left and right movements. The chassis also holds a power bank as battery for power supply.



Fig.8: Robotic Car and Receiver Part Circuit

#### D. Training and Recognition

The important step in this stage is to select an appropriate type of microcontroller language for the programming; here it will use C language for Arduino and the compiler provided by Arduino Company.

Since the microcontroller is the Arduino Uno microcontroller, the number of pin is limited, thus, the number of appliances that can be controlled are just a few but still sufficient.

To record or train a command, the AD-VR3 chip stores the analog signal pattern and amplitude and saves it. In recognition mode, the chip compares the user- inputted analog signal from the microphone with those stored already and if it recognizes a command, an output of the command identifier will be sent to the microprocessor through the AD-VR3 ports of the chip.

Steps for Training Words to be recognized

- Open vr\_sample\_train (File -> Examples -> VoiceRecognitionV3 ->vr\_sample\_train)
- Choose right Arduino board (Tool -> Board, UNO recommended), Choose right serial port.
- Click Upload button, wait until Arduino is uploaded.
- Open Serial Monitor. Set baud rate 115200, set send with Newline or Both NL & CR.
- Send command settings (case insensitive) to check Voice Recognition Module settings. Input settings, and hit Enter to send.
- Train Voice Recognition Module. Send sigtrain 0 on command to train record 0 with signature "left" for example. When Serial Monitor prints

"Speak now", you need to speak your voice (can be any word, meaningful word recommended, may 'left' here), and when Serial Monitor prints "Speak again", you need to repeat your voice again. If these two voices are matched, Serial Monitor prints "Success" and "record 0" is trained, or if are not matched, repeat speaking until success.

What is a signature? Signature is a piece of text description for the voice command. ForExample, if our 5 voice command are "0, 1, 2, 3, 4", we could train in the following way:

- Sigtrain 0 left
- Sigtrain 1 right
- Sigtrain 2 forward
- Sigtrain 3 backward
- Sigtrain 4 stop

The signature could be displayed if its command was called. When training the two led on the Voice Recognition Module can indicate the training process.

After sending the training command, the SYS\_LED (yellow) is blinking fast which remind you to get ready. Speak your voice command as soon as the STATUS\_LED (red) light lights on. The recording process ends once when the STATUS\_LED (red) lights off. Then the SYS\_LED is blinking again, get ready for next recording process. When the training process ends successful, SYS\_LED and STATUS\_LED blink together. If the training fails, SYS\_LED and STATUS\_LED blink together, but quickly. Fig. 9 below illustrate the above process.

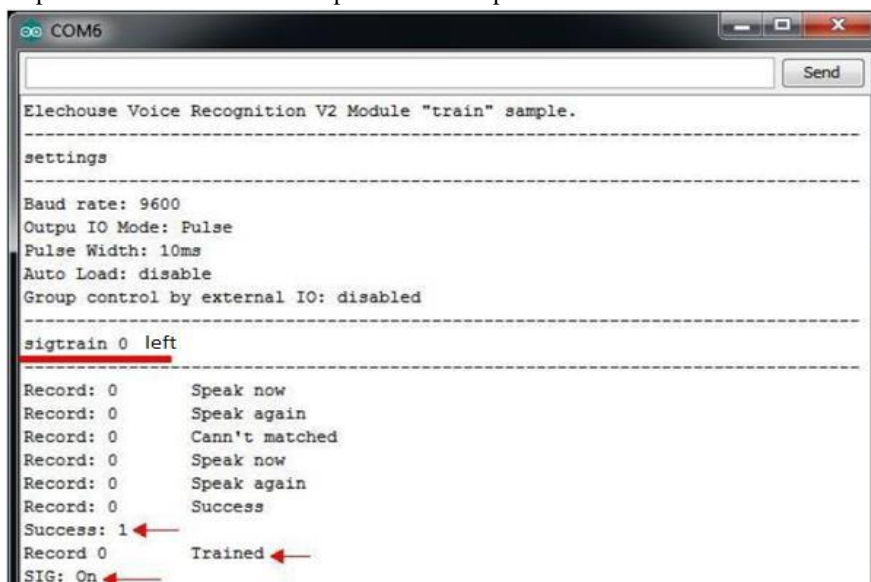


Fig.9: Training of word "left" in the arduino software

- Train another record. Send sigtrain 1 right. Command to train record 1 with signature "right". Choose the required words to train (it can be any word, meaningful word recommended, may be 'right' here)
- Send load 0 1 command to load voice. And say your word to see if the Voice Recognition Module can recognize your words.

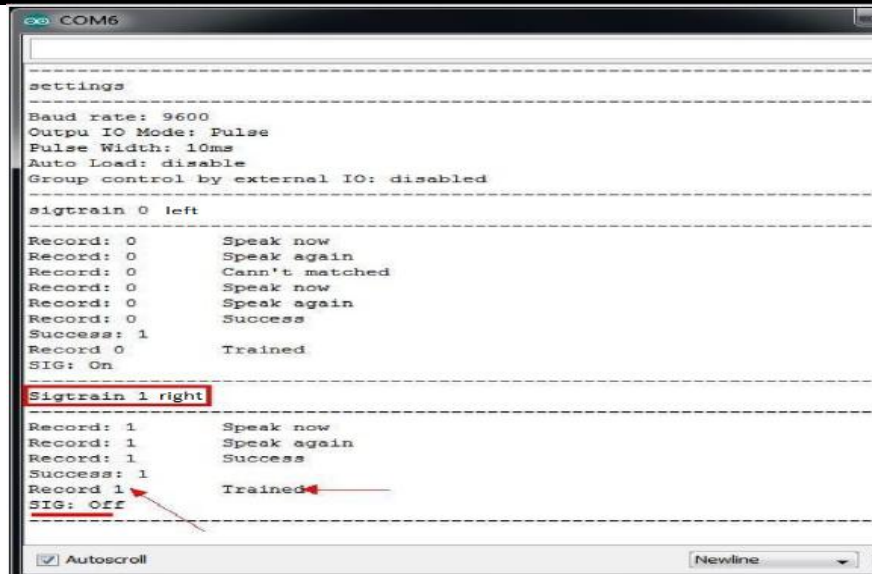


Fig. 10: training of the word "right"

### E. Testing Recognition and Fuzzy Rules Logic

Repeat a trained word into the microphone. the word and its number should be displayed on the monitor display.

For instance, if the word "right" was trained as word number1, saying the word "right" into the microphone will cause the number 1 and the word itself to be displayed.

Based on the recognized command from VR3 four signals are to be sent to the drive from transmitter arduino microcontroller. And the overall membership of the neuro-fuzzy inference system will consist of five rules that control the movement of the motors as follows:

If volt in bin1 is > 100 and volt in bin2 is > 100 and volt in bin3 is > 100 and volt in bin4 is < 100, then a digital signal will be sent to turn on the left motor.

If volt in bin1 is > 100 and volt in bin2 is > 100 and volt in bin3 is < 100 and volt in bin4 is > 100, then a digital signal will be sent to turn on the right motor.

If volt in bin1 is > 100 and volt in bin2 is < 100 and volt in bin3 is > 100 and volt in bin4 is > 100, then a digital signal will be sent to turn on both motors backwardly.

If volt in bin1 is < 100 and volt in bin2 is > 100 and volt in bin3 is > 100 and volt in bin4 is > 100, then a digital signal will be sent to turn on both motors forwardly.

If volt in bin1 is < 100 and volt in bin2 is < 100 and volt in bin3 is < 100 and volt in bin4 is < 100, then a digital signal will be sent to turn off the right and left motor.

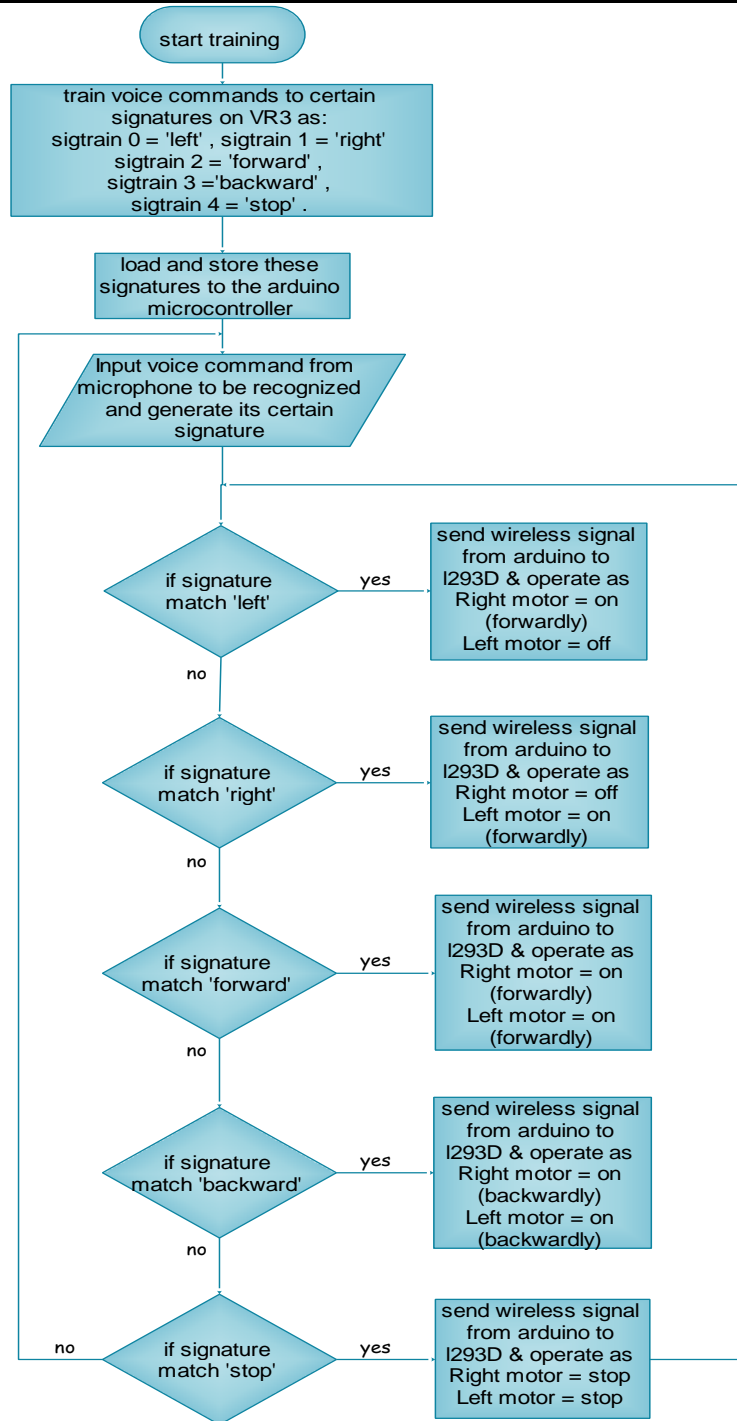


Fig.11: flow chart of the system

Table 2 below shows the used command in our robotic car training and the action that should result from each one

TABLE II. Robotic Car Controlling Commands

Location No	Train Word	Description
0	Left	Right motor on. Left motor off
1	Right	Right motor off. Left motor on

2	Forward	Both motors on forwardly
3	backward	Both motors on/in back direction
4	stop	Both motors off

### III. DISCUSSION THE NATURE OF THE PROBLEMS:

#### 1. Analyzing The Problem:

Speech recognition is the process of finding an interpretation of a spoken utterance; typically, this means finding the sequence of words that were spoken. This involves preprocessing the acoustic signals to parameterize it in a more usable and useful form. The input signal must be matched against a stored pattern and then makes a decision of accepting or rejecting a match. No two utterances of the same word or sentence are likely to give rise to the same digital signal. This obvious point not only underlies the difficulty in speech recognition but also means that we be able to extract more than just a sequence of words from the signal.

The different types of problems we faced in our system have been enumerated below:

- 1) Differences in the voices of different people the voice of a man differs from the voice of a woman that again differs from the voice of a child. Different speakers have different vocal tracts and source physiology. Electrically speaking, the difference is in frequency. Women and children tend to speak at higher frequencies from that of men.
- 2) Differences in the loudness of spoken words. No two persons speak with the same loudness. One person will constantly go on speaking in a loud manner while another person will speak in a light tone. Even if the same person speaks the same word on two different instants, there is no guarantee that he will speak the word with the same loudness at the different instants. The problem of loudness also depends on the distance the microphone is held from the user's mouth. Electrically speaking, the problem of difference is reflected in the amplitude of the generated digital signal.
- 3) Differences in the time Even if the same person speaks the same word at two different instants of time, there is no guarantee that he will speak exactly similarly on both the occasions. Electrically speaking there is a problem of difference in time i.e. indirectly frequency.
- 4) Problem due to noise: The robot will have to face many problems, when trying to imitate the ability of humans hearing. The audio range of frequencies varies from 20 Hz to 20 kHz. Some external noises have frequencies that may be within this audio range. These noises pose a problem since they cannot be filtered out.
- 5) Power supply Another important problem which needed to be solved was to provide sufficient current and stable voltage to entire assembly fourth affiliation).

#### 2. Solutions of The Problems

After analyzing the problems, we come out with the solutions which are listed below.

#### A. Amplitude Variation

Amplitude variation of the electrical signal output of microphone may occur mainly due to:

a) Variation of distance between sound source and the transducer.

b) Variation of strength of sound generated by source.

To recognize a spoken word, it does not matter whether it has been spoken loudly or less loudly. This is because characteristic features of a word spoken lies in its frequency & not in its loudness (amplitude). Thus, at a certain stage this amplitude information is suitably normalized.

#### B. Recognition of a word

If same word is spoken two times at different time instants, they sound similar to us; question arises what is the similarity in-between them? It is important to note that it does not matter whether one of spoken word was of different loudness than the other. The difference lies in frequency. Hence, any large frequency variation would cause the system not to recognize the word, so its better if the speaker try to imitate the same frequency of that used in training process. In speaker independent type of system, some logic can be implemented to take care of frequency variation. A small frequency variation i.e. features variation within tolerable limits is considered to be acceptable [2].

#### C. Noise

Along with the sound source of the speech the other stray sounds also are picked up by the microphone, thus degrading the information contained in the signal by using the system in appropriate quit environment.

#### D. Power supply

As mentioned early one of the important problems which needed to be solved was to provide sufficient current and voltage to entire assembly when interfered together specially the receiver part of the system. Since the current drawn from supply was so much that a 9V battery could not last for a longer period, we used a rechargeable power bank on the receiver circuit and used direct power supply on the sender circuit.

#### E. Applications

We believe such a system would find wide variety of applications. Because menu driven systems such as e-mail readers, household appliances like washing machines, microwave ovens, and pagers and mobiles etc. will become voice controlled in future

- 1) The robot is useful in places where humans find it difficult to reach but human voice reaches. E.g. in a small pipeline, in a fire-situations, in highly toxic areas.
- 2) It can be used to bring and place small objects.
- 3) Speech and voice recognition security systems.
- 4) The same system components can be widened and installed in a wheelchair of disabled people which



make their movement much easier when controlled by voice command.

#### IV. CONCLUSION

The goals of the study are accomplished successfully. A circuit was constructed around the AD-VR3 module and interfaced with arduino microcontroller to creating a stand-alone model of speech recognition based on neuro-fuzzy system.

The system performance was measured through various experiments and determined to perform with a excellent recognition accuracy for a clear noiseless commands set.

The recognition can further be improved by increasing the recording and training environment.

#### REFERENCES

- [1] V. D. Patel, "Voice recognition system in noisy environment," 2011.
- [2] V. Kapila and S.-H. Lee, "Science and mechatronics-aided research for teachers," *IEEE control systems*, vol. 24, pp. 24-30, 2004.
- [3] R. Murphy, *Introduction to AI robotics*: MIT press, 2000.
- [4] M. Negnevitsky, *Artificial intelligence: a guide to intelligent systems*: Pearson Education, 2005.
- [5] A. D'Ausilio, "Arduino: A low-cost multipurpose lab equipment," *Behavior research methods*, vol. 44, pp. 305-313, 2011.