International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 4, pp: 504-509 www.ijaem.net ISSN: 2395-5252

Control of Electrical appliances through Speech recognition for Disabled

Nikhil Mehta[1], Harsha S Hanamshet[2] and Shivani Kamat[3]
Nitte Meenakshi Institute of Technology, Bangalore-560064, India
Corresponding author; Nikhil Mehta

Date of Submission: 01-08-2020 Date of Acceptance: 15-08-2020

ABSTRACT-Voice has become one of the effective and efficient communication medium among people. This communication method may also be a useful interface for interacting with machines which understand only electrical signals. Thus, the system 's reliance on speech recognition has increased dramatically over the last year. Various methods of speech recognition exist, such as Dynamic Time Warping (DTW), Vector Quantization (VQ),Hidden Markov Model (HMM), etc. This paper offers an detailed analysis over the use of Artificial Neural Networks (ANN)in speech recognition. Here we used a different class of neural networks called Multilayer Perceptron (MLP), which uses the back propagation for error algorithm. Once data has been collected, the voice signal is pre-processed and fed to the MLP for classification.

This paper introduces the prototype of speech recognition-based automation system to control the various electrical devices for the differently abled people suffering from quadriplegia or paraplegia (who can not move their body parts, but can talk and listen). Using the MATLAB (2018a) environment to monitor electrical equipment, the automatic speech recognition (ASR) system is used to accurately and effectively transform a voice signal into a text message.

Keywords: Speech Recognition,ANN,MFCC, Feature Extraction,MLP,MATLAB.

I. INTRODUCTION

Voice is the most effective and natural way to communicate. Human being also wants to possess an effective way of communicating with machines.Thus, they like voice as a media to interact with devices instead of using the other hectic interfaces like mouse and keyboards. Speech recognition is the process of spotting or identifying any phrase or isolated vocabulary from a continuous speech. Basic speech recognition software posses a very short speech recognition technique and it will identify words if only spoken very clearly as human articulators and speech organs aren't under control and they varies biologically. Automatic Speech Recognition (ASR) plays a key role in communications between people and computers. For this purpose various techniques are used, such as LPC, MLP, MFCC and ANN.

Some systems require "training" to read text or short vocabulary into the system by a private speaker in the process of speech . Systems which do not use instruction are referred to as "speaker independent" systems. Training systems are called "speaker dependent".Microphone is used to record the speech signal. Using the Mel Frequency Cepstrum Coefficient (MFCC), this signal is pre-processed at the front-end for feature extraction. Using ANN, the features obtained are compared at the back end for pattern matching. After comparing recorded voice features with database, command control is sent to microcontroller to control electrical devices.

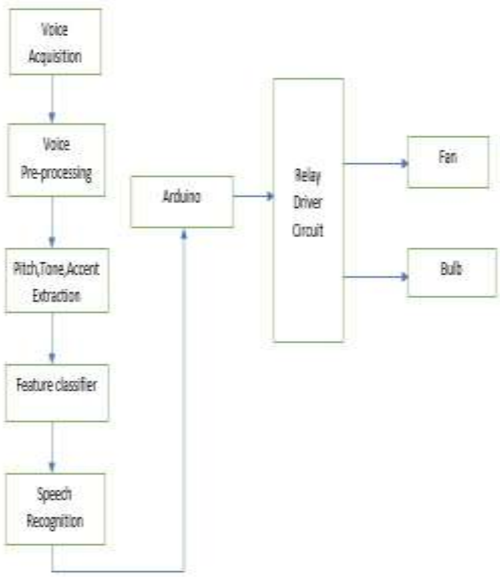

Figure 1: Block diagram of the proposed model

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 4, pp: 504-509 www.ijaem.net ISSN: 2395-5252

II. PROPOSED WORK

Speech recognition process is a complex and hectic task.. For the method of speech recognition we areusing mfcc feature extraction technique and Artificial Neural Networks.

2.1 Speech Recognition Process

A.Voice acquisition: The speaker's voice is obtained in waveform at this step and stored in ".wav" format. There are several applications available which is used to record human speech, for this purpose we used MATLAB. The ambient atmosphere and electronic equipment can have a great influence on the generated discourse. Background noise and room vibrations also present which is of no use.

B.Voice Pre-processing: Pre-processing plays a crucial role in removing the unwanted signal. Ultimately it increases speech recognition quality. Typically, voice pre-processing involves noise filtering, signal smoothing, point-to - point monitoring, signal framing, signal windowing, & echo removal and deletion. It only processes original data for extraction of the features

C.FeatureExtraction: In various people the voice differs. It is because, in utterance, each individual has different characteristics. In theory, some possibility of recognizing speech from the digitized waveform should be there. But there is a need to conduct some extraction of features to minimize those variations due to the wide variability in speech signal. In this project we use extraction of the MFCC feature,which is generally 13 in number.

Mel Frequency Cepstrum Coefficients: It one of the most common and powerful toolwhich is used for feature extraction during speech recognition. This is based on the human ear scale-dependent frequency domain that is expected on Mel scale. MFCCs, being features of a frequency domain, is more reliable than features of time domains.This process gives the real valued cepstral of the short time framed signal from FFT.

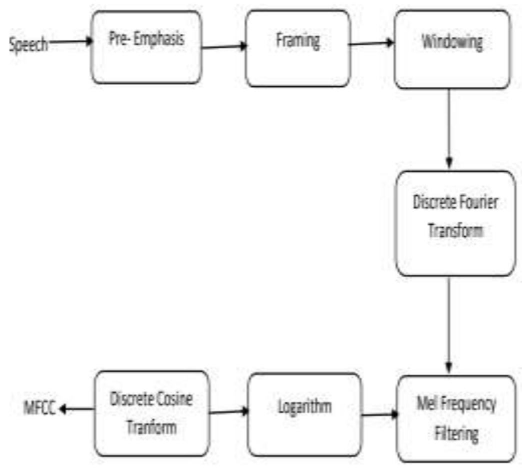

Figure 2: Block diagram of MFCC steps

- **Pre-emphasis**:Noise impacts more on high frequency signals compared to lower frequencies. Pre-emphasis increases higher frequency signals energy. For this purpose, HPF(High pass filter) is used.
- **Framing**-There is a continuous shift in an audio signal and for simplificationwe take assumption that the audio signal does not shift much on short time scales. And we frame the signal into frames of 20 to 40 ms.
- **Windowing**-To remove the unnecessary discontinuities in the segment of speech and the spectrum distortion caused by framing ,operation of windowing is carried out.Mostly Hamming window is used when processing speech.
- **Discrete Fourier Transform (DFT**)- FFT is a very efficient DFT implementing algorithm. This method transforms signal to frequency domain from time domain. As frequency domain includes more signalling information
- **Mel Frequency Filtering**-This process is required to convert frequency of speech signal from linear scale to Mel scale. As human ear works in decibel so this process take the logarithm of speech signal.
- **Discrete Cosine Transform (DCT**)-Conversion of Mel coefficients back to time domain takes place in this step.It is basically a compression technique,with DCT we keep only first few coefficients. As a result of this step we get Mel frequency Cepstral Coefficients.First 13 coefficients,which contains most of the information of speech, are used for training neural networks.

D. Feature Classification-Common techniques used for feature classification are

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 4, pp: 504-509 www.ijaem.net ISSN: 2395-5252

HMM,DTW,VQ.We are using artificial neural network for feature classification. Various complex mathematical function involved in these systems helps in taking out information which is hidden in the processed input signal.

2.2 Artificial Neural Networkfor ThePerspective of Speech Recognition

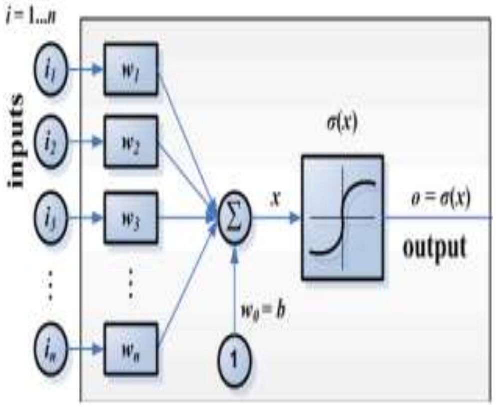

Figure 3: Basic artificial neuron

In The above figure, the mathematical character indicates various inputs, i.e. i(n). Multiply each input by the respected connecting weights w(n). Generally, to get the desired results these products are simply added and given to the transfer function.

C.TYPES OF ANN

Various different structure of Artificial Neural Network have been found by researchers across the world.like Convolutional neural network, Deep neural network, Recurrent neural network. New class of feedforward layered network known as multilayer perceptron has been used.

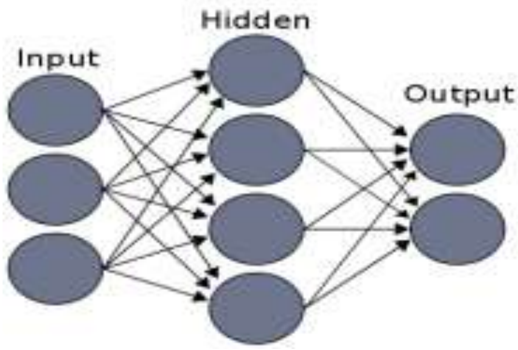

Figure 4: A fully connected Feedforward

III. METHODOLOGY

We have implemented this using MATLAB and commands are given to Arduino through serial communication.

Artificial Neural Networks (ANN) are the electronic models which is based on brain structure. Ultimately the human brain learns from the interactions.In the same way a neural network has various interconnected neurons which can perform any complex task with ease using mathematical calculations.

3.1 Software Implementation

We have written code for modeling the system in MATLAB. This includes manipulation of matrix, visualization of functions and data, implementation of various algorithms to the user interface, etc. This integrates with multi-language written programs, including C++ , Java, and FORTRAN.

3.2 Hardware Implementation

The hardware implementation of proposed system is explained below.

A. ARDUINO UNO- The Arduino Uno is an available to all microcontroller board created by Arduino.cc and based on the user friendly and compact controller Microchip ATmega328P. The board is equipped with both the optical and analog pins

B. MICROPHONE- A microphone, also known as mic, is a transducer that converts sound energy into an electrical energy. Microphones are used in many fields, such as telephones, development of motion pictures, voice recording, two-way radios, megaphones, speech recognition etc.

C.RELAY-The switching operation is done using relays.

D. LED-.It produces light using one or more light-emitting diodes (LEDs) in light fixtures. LED lamps have longer span than comparable incandescent lamps, and are much more powerful than most of the ongoing used bulbs.

E. Arduino Uno Microcontroller: Using various types of sensors and actuators, the Arduino microcontroller offers a simple, inexpensive, platform for college kids and professionals to create devices that communicate with valued environment.

The ATmega328 is one type of single-chip microcontroller built within the megaAVR family together with Atmel. This Arduino Uno architecture may be a customized Harvard architecture with a core 8 bit RISC processor.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 4, pp: 504-509 www.ijaem.net ISSN: 2395-5252

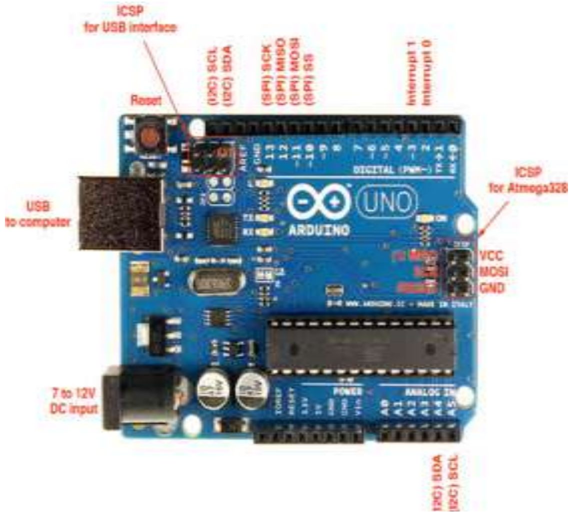

Figure 5: Pin Diagram of Arduino

4.2.2 Main Characteristics of Arduino:
- Voltage used for working :5V
- Voltage provided at input :7-12V
- Digital I/O Pins : 14
- Analog I/P Pins :6
- DC Current per I/O Pin :40 mA
- DC Current of 3.3V Pin :50 mA
- Flash Memory : 32 KB
- SRAM :2 KB
- EEPROM :1 KB
- CLK Speed :16 MHz

3.3 Serial Communication using MATLAB Command Window

It is the direct method for setting up serial communication between Arduino and MATLAB. Here we will simply send the MATLAB info serially to the Arduino using the command window then Arduino will read the incoming serial data. This transmitted serial data can then be used to power everything connected to the Arduino. In this project we have connected an LED and A FAN (dc motor) with Arduino, which will be switched on and off according to Arduino's serially received data.

IV. TESTING RESULTS

SPEECH SPECTROGRAM:Visual representation of spectrum of frequencies of signal as it varies over time is called speech spectrogram.Lighter region shows less frequency while darker region has highest frequency.

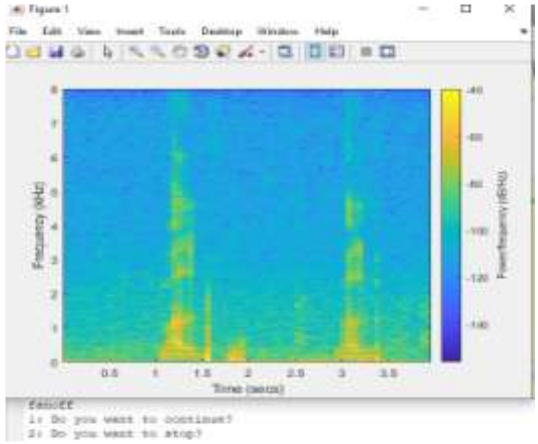

Figure 6: Spectrogram showing speech signal "fanoff"

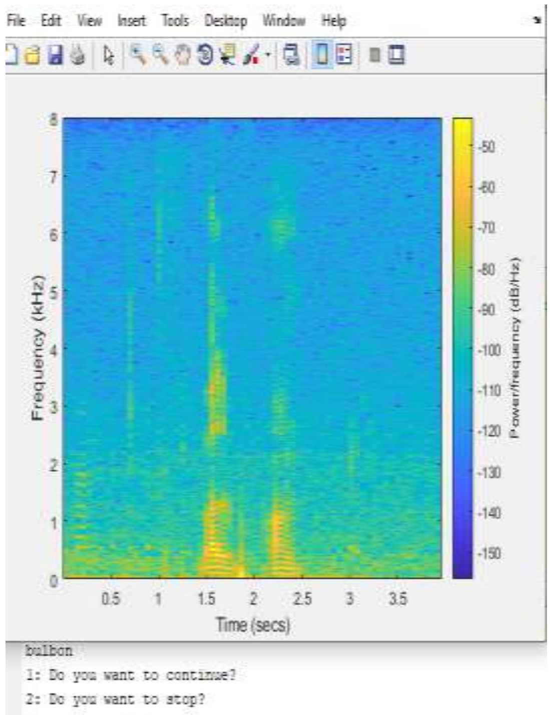

Figure 7:Spectrogram showing speech signal "bulbon"

MAPE: MAPE stands for mean absolute percentage error, It is a numerical method of measuring accuracy in statistics of a prediction system.It is also used used in machine learning regression for loss function.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 4, pp: 504-509 www.ijaem.net ISSN: 2395-5252

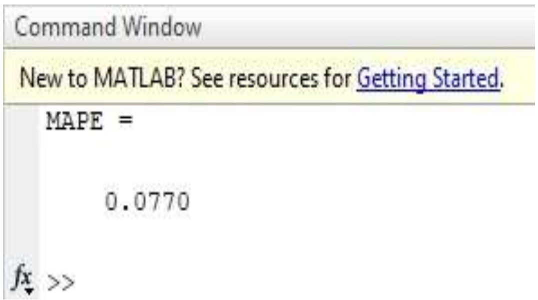

Figure 8: Large MAPE value showing less accuracy of trained model.

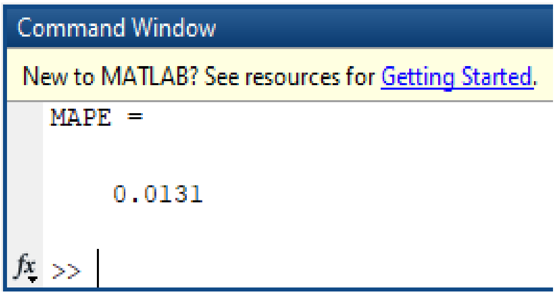

Figure 9: Less MAPE value showing better accuracy of trained model.

CONFUSION MATRIX: A confusion matrix often used to explain the output of a classification model on a collection of test data in the form of a table for which the true values are known. Visualization of an algorithm 's performance may be done using the matrix.

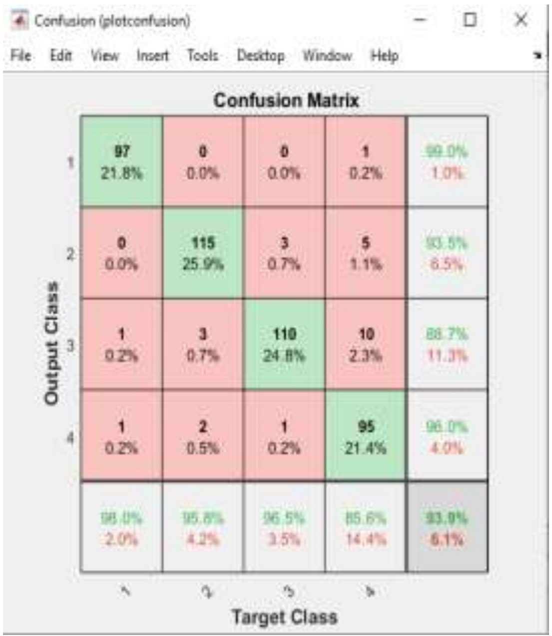

Figure 10: Confusion matrix showing the system is 93.9% accurate when trained and tested in proper surrounding.

MEAN SQUARED ERROR: Mean squared error is just one way of estimating the error, there are also other ways of measuring error. It is important to square the error, because the error can be positive or negative, and if we add up the error of all the nodes, the positive and negative ones can be cancelled. So, we square the errors to prevent the cancelling effect.

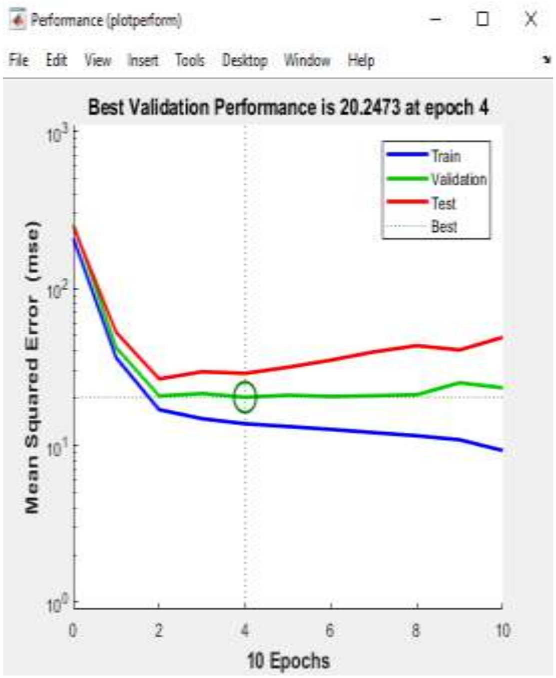

Figure 11: Mean square error plot

V. CONCLUSION

All automation systems are currently acting in accordance with a series of common commands or procedures to communicate with their home appliances.Complicated procedures brings distance between user and technology. Our project proposes voice commands to communicate with home electrical appliances using Arduino to solve all the current problems and discuss the solution

Using MATLAB, we implemented a speech recognition system with success. The commands that are recognized are shown in text format. As MATLAB recognizes voice command, the specific data which is mainly binary is sent to microcontroller. Thus the operation of controlling devices are carried out according to the data received by the microcontroller. The system is geared towards the elderly, the blind and the quadriplegic.

Training and testing of the speech recognition model was done in noise and noise free surroundings with various speakers involved.

REFERENCES

[1]. Yogita H. Ghadage, Sushama D. Shelke, "Speech to Text Conversion for Multilingual Languages", International Conference on Communication and Signal Processing, April 6-8, 2016, India
[2]. Bhushan C. Kamble, "Speech Recognition Using Artificial Neural Network – A Review", Int'l Journal of Computing, Communications & Instrumentation Engg. (IJCCIE) Vol. 3, Issue 1 (2016) ISSN 2349-1469 EISSN 2349-1477.
[3]. Chee Peng Lim, Siew Chan Woo, Aun Sim Loh, and Rohaizan Osman, "Speech Recognition Using Artificial Neural Networks".
[4]. Mrs. Paul Jasmin Rani, Jason Bakthakumar, Praveen Kumaar.B, Praveen Kumaar.U and Santhosh Kumar, "VOICE CONTROLLED HOME AUTOMATION SYSTEM USING NATURAL LANGUAGE PROCESSING (NLP) AND INTERNET OF THINGS (IoT)", 2017Third International Conference on Science Technology Engineering & Management (ICONSTEM).
[5]. Md. Saiful Islam, Md. Shajid-Ul-Mahmud, Md. Akhtaruzzaman, "Voice Command Based Matlab GUI for Microcontroller", 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC) 21 - 23 Dec 2017, Dhaka, Bangladesh
[6]. Nathan David, AbaforChima, Aronu Ugochukwu, "Design of a Home Automation System Using Arduino", International Journal of Scientific & Engineering Research, Volume 6, Issue 6, June-2015.
[7]. Mukesh Kumar, Shimi S.L," Voice Recognition Based Home Automation System for Paralyzed People", International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 4, Issue 10, October 2015

International Journal of Advances in Engineering and Management
ISSN: 2395-5252

IJAEM

Volume: 02 Issue: 01 DOI: 10.35629/5252

www.ijaem.net
Email id: ijaem.paper@gmail.com