IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

Automated Voice Recognized Wheelchair

Sidharth S[1], Saptha Surendran[2], Nimmu Ann George[3], Jose J. Edathala[4]

Department of ECE, Amal Jyothi College of Engineering, Kottayam, Kerala, India[1, 2,3,4]

Abstract: This paper presents an automated voice recognized wheelchair (AVRW) based on speech recognition. It's common that physically challenged people find it hard to move the wheelchair without external aid so this AVRW can be used by them to move around without any difficulty. The voice of the person is detected by voice capture module which will be compared by the voice recognition module with predefined voices loaded into the system. Vector Quantization technique is used to recognize the command. A microphone is used for taking code words from the patient to the PC. The signals from the PC are processed into code words. These code words cannot be directly sent to the PIC Microcontroller, so a ZigBee module is used for transmission and reception. According to the codes, the motors are rotated. These codes are compared with the MATLAB codes, if they match with each other, then the movement of motor occurs. There is also a motor driver for the control of the motor. It is also equipped with obstacle avoidance technique to help when the person may not be able to provide proper voices at the right time.

Keywords: Speech recognition, MATLAB, Vector Quantization, ZigBee, PIC Microcontroller.

I. INTRODUCTION

Speech Recognition is the process of automatically recognizing a certain word spoken by a particular speaker based on individual information included in speech waves. A speech recognition system operates in two modes to recognize a particular person and to verify the person's claimed identity,

(i) Text dependent
(ii) Text independent

Here a text dependent speech recognition system is used to enhance an ordinary powered wheelchair to interpret commands. Signal processing front end for extracting the feature set is an important stage in any speech recognition system. The optimum feature set is still not yet decided though the vast efforts of researchers. There are many types of features, which are derived differently and have good impact on the recognition rate.At the highest level, all speaker recognition systems contain two main modules feature extraction and feature matching. Feature extraction is the process that extracts a small amount of data from the voice signal that can later be used to represent each speaker. Feature matching involves the actual procedure to identify the unknown speaker by comparing extracted features from his/her voice input with the ones from a set of known speakers.

Intelligent wheelchair will play an important role in the future welfare society. The population of people with disabilities has risen markedly during the past century. These people, suffering from motor deficits, disorientation, amnesia, or cognitive deficits, are dependent upon others to push them, so often feel powerless and out of control.

The automated voice recognized wheelchair has the potential to provide these people with effective ways to alleviate the impact of their limitations, by compensating for their specific impairments.

Literature Survey

The key is to convert the speech waveform to some type of parametric representation (at a considerably lower information rate) for further analysis and processing. This is often referred as the signal-processing front end. There have been various approaches proposed for parametrically representing the speech signal for the speaker recognition task, such as Linear Prediction Coding (LPC), Mel-Frequency Cepstrum Coefficients (MFCC), and others. MFCC[4] is perhaps the best known and most popular, and is the technique used in this paper. MFCCs are based on the known variation of the human ears critical bandwidths with frequency filters spaced linearly at low frequencies and logarithmically at high frequencies have been used to capture the phonetically important characteristics of speech necessary in the feature extraction module. In the feature matching module neural logic or vector quantization technique can be used to confirm the command spoken by matching with the previously extracted features of the command.

The rest of this paper is organized as follows. Section II describes the system design and implementation. The proposed system architecture is presented in the Section III. Finally, the conclusion is drawn in Section IV.

II. SYSTEM DESIGN AND IMPLEMENTATION

The wheeled robot system[1]was used in the making of the prototype. Generally, an electronic wheelchair has two motors and an extra set of wheels to navigate the different directions. On the similar lines, the mobile robot platform had the same functionality as that of the wheelchair. Two 12V-60rpm average torque motors were mounted on either side of the chassis and a castor wheel which is an omnidirectional steel ball was mounted at the front side. The motors are fitted with the nut-bolts and metal brackets. L293D was the motor driver IC used. Supply

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

voltages for the motor, microcontroller and Zigbee module are 12, 5 and 3.4 respectively.

The proposed embedded module can be controlled through voice command and sensors are used for obstacle avoidance purpose. The analog voice command given by the user is obtained through a microphone. Once voice signal is received by the microphone, voice acquisition and recognition is achieved by the MATLAB code. At that time when the controller gets initiated, IR sensor also gets initialized. After sensor is initialized, obstacle detection process will be started and will frequently intimate the controller about the distance of obstacle[6].

III. SYSTEM ARCHITECTURE

A. Software Architecture

A text dependent automatic speech recognition system is developed and simulated using MATLAB in the software section. The software layout is shown in Figure 1.

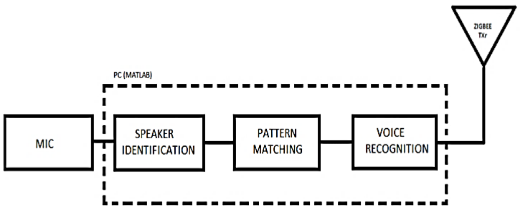

Fig.1 Software layout

In text dependent speech recognition system [6], the phrase is known to the system and can be fixed or prompted orally. All Recognition systems have two phases,

(i) Training phase
(ii) Testing phase

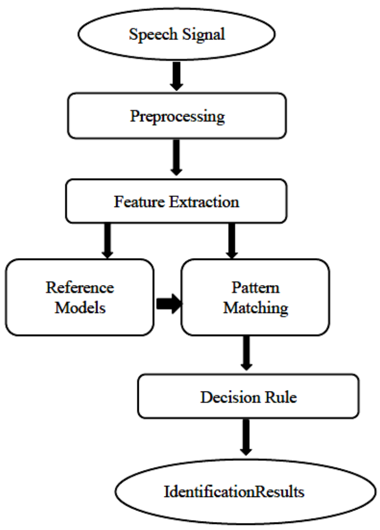

Fig.2 General Diagram of any speaker identification system

Each speaker has to provide a sample speech for each of the commands in the training phase for feature extraction and any specific features of these command words are stored in the vocabulary. In the testing phase, when a command is spoken into the microphone, feature extraction is again done and pattern matching commences with each of the earlier stored features of different commands for the recognition decision to be made.

1) Training phase

a. Pre-processing

Analogue speech signal is sampled and quantized using ADC to be in digital form. The sampling rate and quantization bits are the key parameters to be set optimally to avoid aliasing due to sampling and to obtain precise values of digitalized samples. The next step is to enhance the input speech signal to higher quality and appropriate characteristics by the following processing steps.

- Silence and noise removal: In recorded signals, some portions do not contain any information i.e., they are the silent areas. Hence it is necessary to remove these silence part so as to decrease the unnecessary data. Automatic de-noising method of ID signal using wavelets is used to make it more efficient.

- Frame blocking: The continuous input speech signal is blocked into predefined frames with N samples. Each adjacent frame has N-M overlap samples(N>M) after M samples of current frame, next frame will begin. Need for overlap is that information may be lost at the frame boundaries, so frame boundaries need to be within another frame. This process will continue until all the speech is considered within one or more consecutive frames. N=256 and M=100 and length of frame should be in powers of two so that FFT and DFT is easier.

- Windowing: The next step in the processing is to window each frame so as to minimise signal discontinuities at the beginning and end of the frames. The concept here is to minimise the spectral distortion by using window to taper the signal to zero at the beginning and end of the frame. Mathematically framing is equivalent to multiplying the signal with a series of sliding rectangular windows. The problem with rectangular windows is that the power contained in the side lobes is significantly high and may give rise to spectral leakage. In order to avoid this we use Hamming window.

$$W(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right)$$

where $0 \le n \le (N-1)$

- Fourier Transform: Analysis of signals is easier in frequency transform; therefore the signal is converted from time to frequency domain. The FFT is a fast algorithm to implement DFT which is defined on the set of N samples X(n) as follows

$$X(n) = \sum_{k=0}^{N-1}\left(x(k)e^{\frac{-2\pi jkn}{N}}\right)$$

where n = 0, 1, 2, …, N-1

b. Feature Extraction (Front-end analysis)

After the pre-processing of signal, the features of voice signal are extracted using the technique MFCC (mel frequency cepstrum coefficients) and then stored in the

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

codebook. It aims to extract acoustic features from the speech waveform.

Speech signal is a slowly time varying signal, therefore over long periods of time, the signal's characteristics tend to vary to reflect the different speech sounds being spoken so a short time spectral analysis is the most common way to characterize the speech signal. Thus MFCC is used to give parametric representation of speech signal for the speaker recognition task.

The human perception of the frequency contents of sounds for speech signals does not follow a linear scale. Thus for each tone with an actual frequency of 'f' Hz, a subjective pitch is measured on a scale called 'mel' scale.

$$f(\text{mel}) = 2595 \log\left(1 + \frac{f}{700}\right)$$

MFCC is based on human perceptions which cannot detect frequencies above 1 kHz. It is computed by using a bank of triangularly shaped filters, with centre frequency of the filter spaced linearly for frequencies below 1000Hz and logarithmically for above 1000Hz

c. Codebook Formation
The codebook contains the features of voice signals from all the speakers. The features are stored in the form of clusters. The similar featured clusters reside nearby in the codebook.

2) Testing phase
For testing the same procedure is followed till the formation of codebook and then feature matching or pattern recognition is done. The goal of pattern recognition is to classify objects of interest into one of a number of categories or classes. Here the objects of interest are patterns which are the sequences of acoustic vectors that are extracted from an input speech using the techniques described before and classes are the individual speakers. Using Vector Quantization technique the detection of fed input is done.

a. Vector Quantization
Vector Quantization[7] is a process of mapping vectors from a large vector space to a finite number of regions in that space. Each region is called a cluster and can be represented by its centre called centroid. The collection of all code words is called codebook.

In the training phase a speaker specific Vector Quantization codebook is generated for each known speaker by clustering his training acoustic vectors.

The distance from a vector to the closest code word of a codebook is called a VQ distortion or Euclidean distance. In the recognition phase, an input utterance of an unknown voice is vector quantized using each trained codebook and total Euclidean distance is computed. The command word corresponding to the VQ codebook with the smallest Euclidean distance is identified.

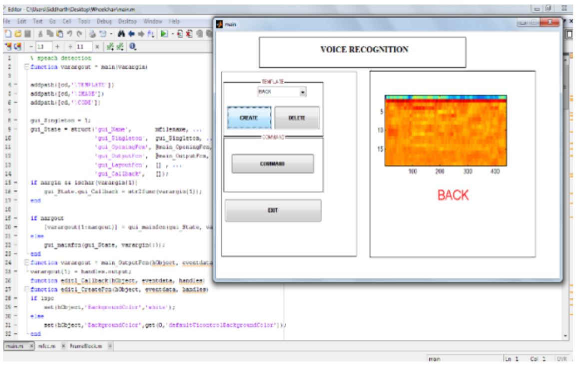

Fig.3 GUI

B. Hardware Architecture
The hardware layout of the embedded system is shown in Figure 3.

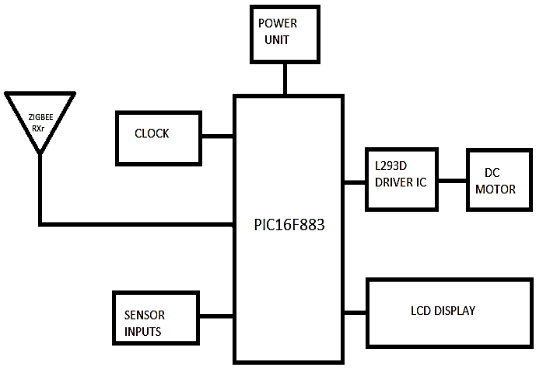

Fig.3 Hardware layout

1) PIC16F883 Microcontroller
This powerful yet easy-to-program (only 35 single word instructions) CMOS FLASH-based 8-bit microcontroller packs Microchip's powerful PIC architecture into a 28 pin package. The PIC16F883 features.

- 256 bytes of EEPROM data memory,
- self-programming, an ICD,
- 2 Comparators,
- 11 channels of 10-bit Analog-to-Digital (A/D) converter,
- 1capture/compare /PWM and 1 Enhanced capture/compare /PWM functions
- Synchronous serial port that can be configured as either 3-wire Serial Peripheral Interface (SPI) or the 2-wire Inter-Integrated Circuit (I²C) bus and an Enhanced Universal Asynchronous Receiver Transmitter (EUSART).

2) Stepper Motor Driver IC (L293D)
Motor driver IC L293D act as an interface between microprocessors and the motors in the robot. Since motor needs more driving supply than microcontroller, the voltage given will not be quite enough.

In order to amplify it and give the supply necessary to motor, driver IC L293D is required. In order to drive stepper motor we have, there is a need of H-bridge circuit, which is provided by L293D. Features of L293D are:

- 600ma output current capability per channel
- 1.2a peak output current (non repetitive) per channel
- Enable facility over
- Temperature protection
- Logical "0" input voltage up to 1.5 v (high noise immunity)
- Internal clamp diodes.

3) Zigbee Module

This is an FSK Transceiver module, which is designed using the Chipcon IC(CC2500). It is a true single-chip transceiver, based on 3 wire digital serial interfaces and an entire Phase-Locked Loop (PLL) for precise local oscillator generation so that frequency could be set. It can be used in UART/ NRZ / Manchester encoding decoding. It is a high performance and low cost module .It gives 30 meters range with onboard antenna.

In a typical system, this trans-receiver will be used together with a microcontroller. It provides extensive hardware support for packet handling, data buffering, burst transmissions, clear channel assessment, link quality indication and wake on radio. It is used to provide wireless connectivity. Operating Range is 30 meters without requiring any external antenna.

Features

- Low power consumption.
- Integrated bit synchronizer.
- Integrated IF and data filters.
- High sensitivity (type -104dBm)
- Programmable output power -20dBm~1dBm
- Operation temperature range : -40~+85 deg C
- Operation voltage: 1.8~3.6 Volts.
- Available frequency at : 2.4~2.483 GHz
- Digital RSSI

The keyword generated out of each voice command is transmitted from a Zigbee transmitter module attached to the PC .The Zigbee receiver at the Wheelchair receives the keywords wirelessly and generates movements accordingly.

4) DC motor

The use of DC motor is to reduce power usage and increase the performance of the wheelchair. 12V 60rpm motors are used. It derives power from the driver IC L293D. According to the voice commands received the motor rotates in the specified direction. [3]

5) LEDS

Light Emitting Diodes are used to indicate the status of the voice recognition. Based on the voice command given, different keywords are generated and transmitted by zigbee module.

These keywords, on reception, light up the corresponding LEDs on the wheelchair. Red, Green, Yellow, White and Blue LEDs correspond to the Stop, Front, Right, Left and Back Commands respectively.

6) Infrared Sensors

An infrared sensor is an electronic device that emits in order to sense some aspects of the surroundings. An IR sensor can measure the heat of an object and detect motion. These types of sensors measure only the infrared radiation, rather than emitting it, so it's called a passive IR sensor.

IV. CONCLUSION

A basic prototype model of a voice recognition module has been successfully designed, tested, trained and implemented using MATLAB and interfaced with a wheeled mobilerobot[5] platform. The mobile robot responds efficiently to the voice commands.

This basic prototype can be a platform for further advanced implementations effectively making it a smart wheelchair. With the inclusion of GSM module, the wheelchair can be modified as a home navigated wheelchair. Also, by including IOT technology and implementing a webcam, it will enable the user to interact with the doctors online and also can provide medical details for the patient

REFERENCES

[1] DilipPhal, D.; Phal, K.D.; Jacob, S., "Design, implementation and reliability estimation of speech-controlled mobile robot," in
[2] Emerging Technology Trends in Electronics, Communication and Networking (ET2ECN), 2014 2nd International Conference on , vol., no., pp.1-6, 26-27 Dec. 2014
[3] Jain S, Jha P, Suresh R, "Design and implementation of an Automatic Speaker recognition system using neural and fuzzy logic in MATLAB," in Signal Processing and Communication (ICSC), 2013 International Conference on vol., no., pp.319-324, 12-14 Dec.2013
[4] Kumaran, M. B., "Implementation of voice based wheelchair for
[5] Differently abled", ICCCNT 2013.
[6] S. Kibria, "Speech Recognition for Robotic Control", Master's Thesis, UMEA University, Sweden, December 2005.
[7] Richard C. Simpson, "Smart Wheelchairs: A Literature Review" J.
[8] Rehabilitation Res.& Dev. 42 (4), pp. 423-438, 2005.
[9] A Reynolds, "An Overview of Automatic Speaker Recognition
[10] Technology", Proc. IEEE, pp. 4072-4075, 2002.
[11] Darshan Mandalia, PravinGareta, Rachna Sharma, "Speaker Recognition Using MFCC, Vector Quantization Model", India, 2011.

DOI 10.17148/IJARCCE.2016.5454