SECURE SPEECH CONTROLLED ROBOT USING MATLAB AND ARDUINO

Hardik Chhatbar[1], Janak Trivedi[2], Rahul Chauhan[3], Darshan Bhatt[4]

[1]*Student of GEC, Bhavnagar, Department Of Electronics and Communication Engineering*
[2]*Ass. Prof. of GEC, Bhavnagar, Department Of Electronics and Communication Engineering*
[3]*Student of GEC, Bhavnagar, Department Of Electronics and Communication Engineering*
[4]*Student of GEC, Bhavnagar, Department Of Electronics and Communication Engineering*

Abstract— The major goal of this project is speaker identification with security. We provide security or encryption to open or access the system using correct 6 digit password so to access the robot user must be known the 6 digit password .This speech identification system consists of comparing speech signal from speaker to the object database library. After comparing the system can identify the speaker. Based upon the identification of the speech we handle or control any appliances or object like ROBOT. For example if we speak forward in the microphone then speech identification system identify the word 'forward' and it indicates to the ROBOT to move forward.

Keywords— Speech identification, voice identification, string matching in matlab, voice controlled robot

I.   INTRODUCTION

The speech is primary mode of communication among human being and also the most natural and efficient form of exchanging information among human in speech.[1]The aim of this research is to describe the main aspects influencing control robot using MATLAB and ARDUINO. By using matlab we can identify the speaker voice and also provide security system for robot using string matching. Its advantage is in the easy way of controlling where the templates of the whole commands are used, which makes it very suitable in case of the automatic speech identification systems with small dictionary of voice commands. Voice samples of the users or authorised persons are first recorded and stored in the computer memory using a microphone. When the program detects the same voice next time, it displays the templates (or commands) being detected and other information. You need a 32-bit operating system with MATLAB R2012b software, microphone. The main problem is in the preparation of reliable reference templates for the set of commands should be identify. Fast Fourier transform (FFT) is an important operation in speech identification, in which a great number of floating-point arithmetic operations are performed. Here we implement speech identification using two methods first is FFT and second is discrete cosine transform (DCT) .Using FFT and DCT minimum error comparison between recorded voice and templates. To controlling robot there is requirement of microcontroller so we choose the ARDUINO UNO controller. it is open source software and programming and interfacing with MATLAB was very easy.

II.   OBJECTIVE AND GOALS

The main objective of this research is to develop an Speaker Identification based automation system capable of controlling many devices inside a home and office using speech commands with security of voice command of respective user only with speech commands transmitted and received through Matlab and Arduino.

- To create speech database templates of users in order to compare the speech signal to the given voice command.
- To design Serial communication media using ARDUINO interfacing with MATLAB to controlling the device (Robot)
- To evaluate and select the most appropriate speech processing method by MATLAB programming for identify the audio signal.
- To giving specific robot control instruction after identifying the audio signal we select the ARDUINO UNO developing board for programming of device (robot) controlling application.

III. FLOW CHART
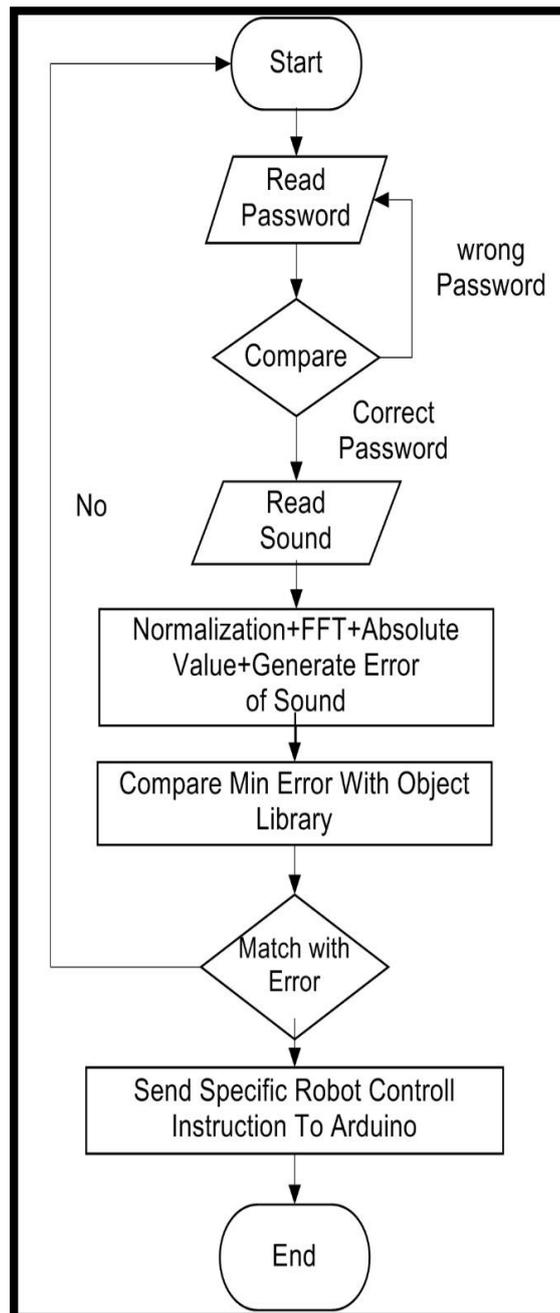
[Figure 1. The flow chart or the process flow diagram of the speech identification process and controlling the ROBOT.]

Description:

Step 1: Start

The flow chart begins with Start, which mentions that the system is ready to work or it is ready to accept commands from user.

Step 2: Read Password

In our project we build GUI module of Speech identification process. But for accessing it we provide a password protection to the system using MATLAB commands. So, system forced a user to enter 6 digit password to access or to open the system for identification. So the next step in flow chart is Read Password.

Step 3: Compare

The system match the user entered password with original correct one which is already defined in the system. If the password match, than system opens the speech identification module. By comparing if the password is wrong than system does not provide access to user for speech identification.

Step 4: Read Sound

Now if password is correct and user enter in the speech identification module. Then user feed the sound to the system and system tries to identify that voice using some computational techniques.

Step 5: Normalization + FFT + Absolute Value + Generate error of sound

Speech Identification is done by using Fast Fourier transform algorithm and normalization with respect to zeros method in MATLAB. FFT is used for the analysis of sound waves and used to get the peak power distribution of the sound waves. Here Normalization is used to get the average and peak amplitude to a target level (norm level).

After applying FFT algorithm to the normalized sound waves MATLAB computes Absolute Value of that complex sound waves. These absolute value is nothing but the magnitude of any complex number or we can say that it is the square root of the squares of real and imaginary parts of the complex number. These processes generate errors of sound and then after it will compare with Min error with Object Library.

Step 6: Compare Min Error with Object Library

The result of comparing errors in Object library is match with error than the specific command is send to the ROBOT using ARDUINO. But if the generated errors are not match with object library than no result is generated and the process flow will again start from the first block.

Step 7: Send Specific ROBOT control Instruction to Arduino

As in the flow chart, the second last block is Send Specific ROBOT control Instruction to Arduino. That means now the specific commands are send to the ARDUINO to control and to instruct the ROBOT. At the last ROBOT will work according to the instructions and commands coming from ARDUINO.

IV.   DISCRIPTION OF METHODS

1.   Fast Fourier Transform (FFT)

The Fast Fourier Transform is a mathematical method for transforming a function of time into a function of frequency. The FFT is the discrete Fourier transform that is used to reduce the no. of calculation of Fourier transform. In this we are using the FFT technique for the input signal for speech identification.[3]

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi k \frac{n}{N}} \qquad k = 0, \ldots, N-1.$$

…Equation - 1

2.   Normalization

Audio normalization is the application of a constant amount of gain to an audio recording to bring the average or peak amplitude to a target level (the norm). Because the same amount of gain is applied across the given range, the signal-to-noise ratio and relative dynamics are generally unchanged. Normalization differs from dynamic range compression, which applies varying levels of gain over a recording to fit the level within a minimum and maximum range. Normalization is one of the functions commonly provided by a digital audio workstation. [2].

If you have a group of audio files at different volumes you may want to make them all as close as possible to the same volume. It may be individual snare hits or even full mixes. [2]

Normalization can be done automatically without changing the sound as compression does. While this is a huge advantage, it can't replace compression as it can't affect the peaks in relation to the bulk of the sound. This means you have far less control. Often Normalizing audio just won't work for matching volume levels, mastering engineers need not lose any sleep.

3.   Discrete Cosine Transform (DCT)

A **discrete cosine transform** (**DCT**) expresses a finite sequence of data points in terms of a sum of cosine functions oscillating at different frequencies. DCTs are important to numerous applications in science and engineering, from lossy compression of audio (e.g. MP3) and images (e.g. JPEG) (where small high-frequency components can be discarded), to spectral methods for the numerical solution of partial differential equations. The use of cosine rather than sine functions is critical for compression, since it turns out (as described below) that fewer cosine functions are needed to approximate a typical signal, whereas for differential equations the cosines express a particular choice of boundary conditions.[5]

The DCT, and in particular the DCT-II, is often used in signal and image processing, especially for lossy compression, because it has a strong "energy compaction" property. in typical applications, most of the signal information tends to be concentrated in a few low-frequency components of the DCT. [5]

A related transform, the *modified* discrete cosine transform, or MDCT (based on the DCT-IV), is used in AAC, Vorbis, WMA, and MP3 audio compression.

DCT-I

$$X_k = \frac{1}{2}(x_0 + (-1)^k x_{N-1}) + \sum_{n=1}^{N-2} x_n \cos\left[\frac{\pi}{N-1} nk\right] \qquad k = 0, \ldots, N-1.$$

…Equation - 2

DCT-II

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right] \qquad k = 0, \ldots, N-1.$$

…Equation - 3

V. DETAIL ABOUT HARDWARE

In this project we need microcontroller and motor driver ic module. Here we choose ARDUINO UNO board as controller and L293D Motor Driver IC to amplify the current from the controller because the current from the controller is very low to start up the dc motor. So it need to be amplify.

- ***ARDUINO UNO:***

The ARDUINO Uno is a microcontroller board based on the ATmega328. It has 14 digital input/output pins (of which 6 can be used as PWM outputs), 6 analog inputs, a 16 MHz ceramic resonator, a USB connection, a power jack, an ICSP header, and a reset button. The ARDUINO Uno can be powered via the USB connection or with an external power supply. The power source is selected automatically.[6]

[Figure 2(a)- Arduino UNO Board]

Specifications:

Microcontroller	ATmega328
Operating Voltage	5V
Input Voltage (recommended)	7-12V
Input Voltage (limits)	6-20V

Digital I/O Pins	14 (of which 6 provide PWM output)
Analog Input Pins	6
DC Current per I/O Pin	40 mA
DC Current for 3.3V Pin	50 mA

- ***MOTOR DRIVER module IC L293D***

[Figure 2(b)- Motor driver IC Module]

Motor driver is basically a current amplifier which takes a low-current signal from the microcontroller and gives out a proportionally higher current signal which can control and drive a motor. In most cases, a transistor can act as a switch and perform this task which drives the motor in a single direction.

VI. DETAIL ABOUT MATLAB GUI

The user must be known about the 6 digit login password. After entering correct password the user can control the robot.


[Figure 3- User login Panel]

Here we added one more feture for security purpose if unknown user enter wrong password more than five times then the login panel will be locked. Figure show as below if user enter wrong password more than five times.
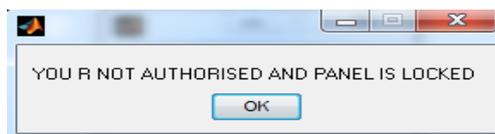
[Figure 4- Entering wrong password more than 5times]

The voice can be recorded through the graphical user interface (GUI) program given here. The recorded voice acts as a reference database for the speech identification model. When you run the program and speak in front of the microphone, the program extracts the unique feature from runtime-recorded audio signal using normalization of absolute value of Fourier transform. During this voice pattern matching process, the program compares with the voice in the database and creates an array of errors or difference outputs. Errors that satisfy the decision rule will be accepted, else rejected. The minimum error object detected will be considered as recognised information.
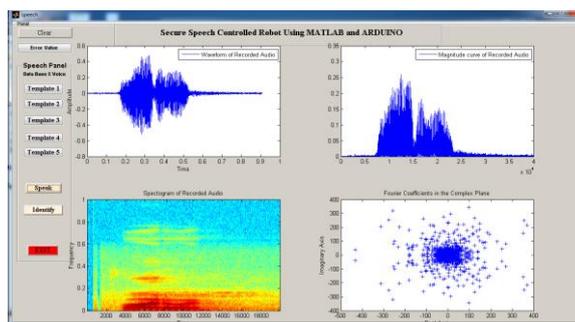

*[Figure 5 – GUI of Speech controlled robot]*
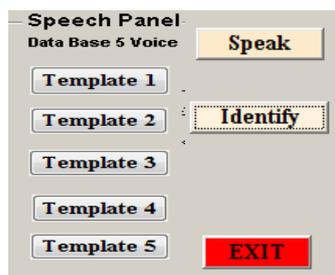
[Figure 6- Template and Speech identification Panel of GUI]

Now, we can check the recorded voice samples by clicking Speak button. After successful voice identification, the program displays the detected template and its sound properties. It generates the following plots of speech signal:

1. Waveform of recorded signal: Time verses Amplitude (Fig 8)
2. Magnitude Curve of recorded audio: Magnitude plot (Fig 9)
3. Spectrogram of recorded signal: Time verses Frequency (Fig 7)
4. Fourier coefficient in complex plane Using FFT: Real verses imaginary axis (Fig 10)
5. Fourier coefficient in complex plane Using DCT: Real verses imaginary axis (Fig 11)
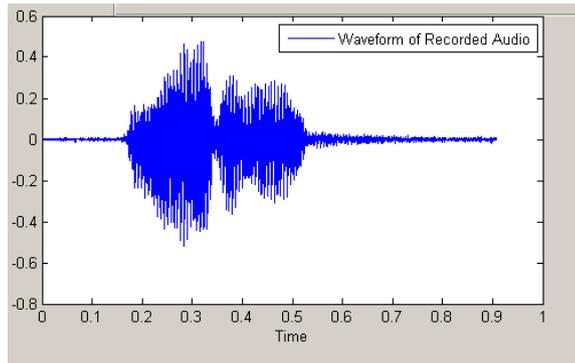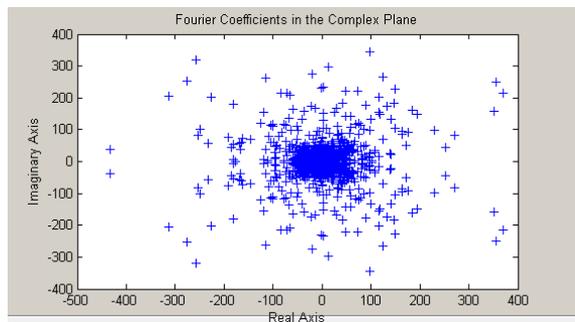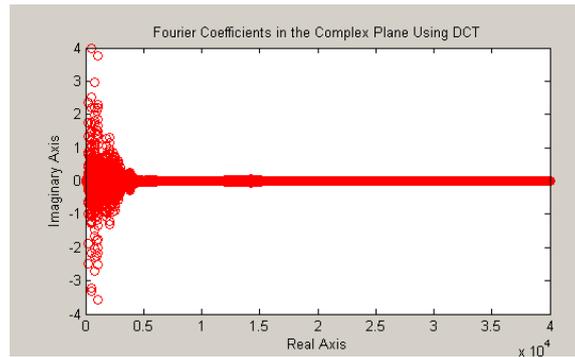
[Figure 8- Waveform of recorded signal]
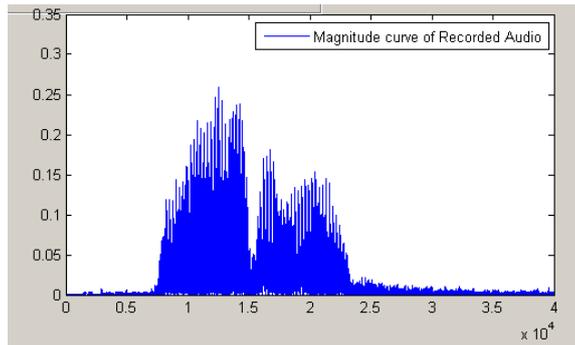
[Figure 9- Magnitude curve of recorded signal]
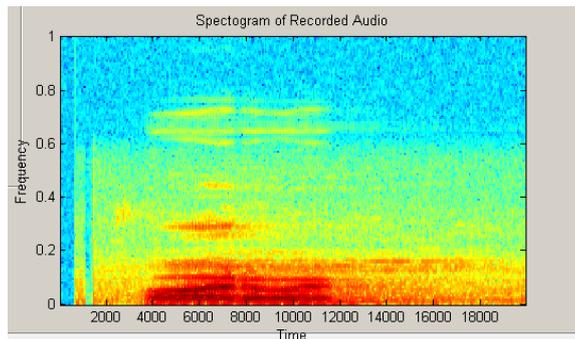
*[Figure 7- Spectrogram of recorded signal]*

*[Figure 10- Fourier coefficient in complex plane Using FFT]*

[Figure 11- Fourier coefficient in complex plane Using DCT]

After clicking on the identify button the program will be execute and it match the recorded speech signal with the templates data base which has been stored in the computer by using the minimum error with the templates matching. The identify window shown as below.
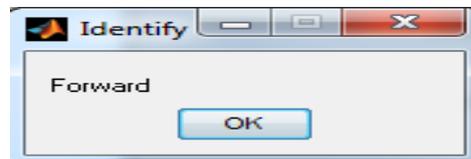
[Figure 12- Identified voice command for Robot]

As per identified command the robot moving forward direction. Particular templates detection has unique coding given to the arduino for movement of motor of the robot.
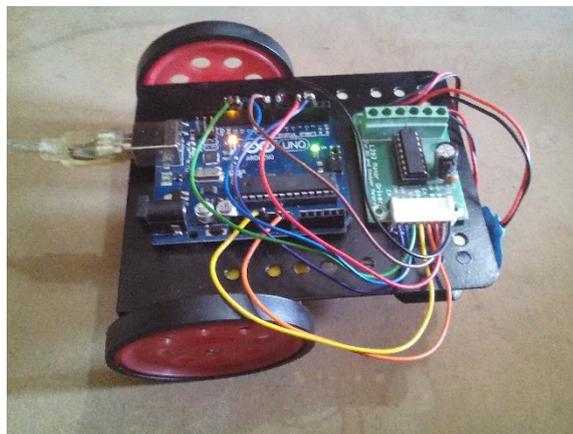
[Figure 13- Voice controlled robot]

Here on the above figure show the voice controlled robot. The robot interface with matlab via USB serial communication with Arduino. First identified command fed to the controller and then the current from the controller pass through the motor driver ic then the motor driver ic amplify the current from the controller and amplified current will be given to the motor.

VII.   RESULT ON DIFFERENT CONDITION OF SOUND AND DIFFERENT METHOD

On this research we tries different condition of Sound by using different methods. The result table of shown as below.

Using FFT Method:

[Table 1: Result Using FFT Methods on Different Conditions]

Number of Person(s)	Number of Attempt	Correct Identification	False Identification
One Person	10	9	1
Two Persons	10	8	2
More than Two persons	10	6	4

Using DCT Method:

[Table 2: Result Using DCT Methods on Different Conditions]

Number of Person(s)	Number of Attempt	Correct Identification	False Identification
One Person	10	7	3
Two Persons	10	6	4
More than Two persons	10	4	6

From The above results we observe the final average result by mathematical calculations. The final Result table shown as below.

[Table 3: Final Average Result]

Sr. No	Methods	Successfully Identification (%)	False Identification (%)
1	FFT	76.66 %	23.34 %
2	DCT	56.66%	43.34%

VIII.   FUTURE WORK:

The program can be modify to interface with embedded systems through parallel or serial port of the PC for monitoring on LCD/LED displays. If a Audio Signal match is found, you can enable a lock or other external devices.

The next would be to recognize continuous speech rather than isolated words. The next step after controlling an object the future scope of this application is also in the field of security systems, home automation, and healthcare, and telephony, field of Robotics and for hands free computing.

IX.   CONCLUSION

From the above work we can conclude that by using FFT and normalization we can successfully identify the sound using minimum error comparison. By performing this project with different algorithms and computational techniques we get different results. In the DCT method we observe that the system could not identify the voice accurately. And also too much error components are present in the characteristics of the method. So as compared to DCT, FFT algorithm was much better and gave an accurate result.

We also observe that in FFT & DCT algorithm there are such limitations like in FFT algorithm, FFT technique is fairly complex for signal compression and filter design. Now a days by developing more advanced techniques for speech identification like Mel frequency Cepstrum Coefficient (MFCC) and Hidden Markov Model (HMM), FFT technique is not more accurate and efficient.

X.   ACKNOWLEDGMENT

REFERENCES

[1] "Speaker Recognition A Tutorial" Joseph P. Campbell, Jr. Senior Member, *IEEE Proceedings of the IEEE, VOL. 85, NO. 9*

[2] "Normalization of the Speech Modulation Spectra for Robust Speech Recognition" Xiong Xiao, Siong Chng, Haizhou Li. *IEEE Transactions on Audio, Speech And Language Processing, Vol. X, No.* X.

[3] "Real-Time Musical Applications using FFT-based Resynthesis" Zack Settel & Cort Lippe. *IRCAM, 1 place Stravinsky, Paris, 75004, France*

[4] "Speech Recognition Based Learning System" Lavin Jalan, Rahul Masrani, Roshan Jadhav,Tejaswini Palav. *Vidyavardhini's College Of Engineering And Technology-Mumbai University Vidyavardhini's Campus. International Journal of Engineering Trends and Technology- Volume4Issue2- 2013*

[5] "Audio Compression in MPEG Technology" Manas Arora, Neha Maurya. *International Journal of Scientific and Research Publications, Volume 3, Issue 12, December 2013*

[6] "Getting Started with Arduino" By Massimo Banzi. Publisher: Maker Media. *September 2011: Second Edition ISBN: 978-1-449-309879*

[7] "A Review on Speech Recognition Technique" by Santosh K.Gaikwad, Bharti W.Gawali, Pravin Yannawar. *International Journal of Computer Applications (0975 – 8887) Volume 10– No.3, November 2010*

[8] "Speaker Identification Based Automation System through Speech Recognition" by Prabhakar V. Mhadse, Amol C.Wani. *International Journal of Emerging Technology and Advanced Engineering Volume 3, Issue 1, January 2013*

[9] "Audio signal identification via pattern capture and template matching" Martin Kermit, Age J. Eide. *Pattern Recognition Letters 21 (2000) 269-275*

[10] "Robotic Automation through Speech Recognition" Ms. Neerja S. Dharmale, Mr. Rupesh S. Mahamune. *International Journal of Scientific and Research Publications, Volume 3, Issue 6, June 2013. ISSN 2250-3153*

[11] "Robust Speech Recognition by Normalization of the Acoustic Space" Alejandro Acero and Richard M. Stern. *Carnegie Mellon University Pittsburgh, Pennsylvania 15213*

[12] "Voice and Touch Screen Based Direction and Speed Control of Wheel Chair for Physically Challenged Using Arduino" by M.Prathyusha, K. S. Roy, Mahaboob Ali Shaik. *International Journal of Engineering Trends and Technology (IJETT) -Volume4Issue4- April 2013*

[13] "Controlling Of Device Through Voice Recognition Using Matlab" by Kavita Arya, Sonam Kumari, Komal Saxena. *International Journal of Advanced Technology & Engineering Research (IJATER) Volume 2, Issue 2, March 2012*

[14] "A Gaussian Mixture Model Based Speech Recognition System Using Matlab" Manan Vyas. *Signal & Image Processing: An International Journal (SIPIJ) Vol.4, No.4, August 2013*

[15] "Design of an Automatic Speech Recognition System Using Matlab" by Jamel Price. *University of Maryland Eastern Shore Princess Anne, MD 21853 August 2005*