# CONTROLLING OF DEVICE THROUGH VOICE RECOGNITION USING MATLAB

1. Sonam Kumari, GBTU; 2.Kavita Arya, GBTU;3. Komal Saxena, GBTU


## Abstract

Speech Recognition is the process of automatically recognizing a certain word spoken by a particular speaker based on individual information included in speech waves. This technique makes it possible to use the speaker's voice to verify his/her identity and provide controlled access to services like voice based biometrics, database access services, voice based dialing, voice mail and remote access to computers. In this project, the algorithms for the speech recognition has been developed and implemented on MATLAB .These algorithms can be used for any security system in which the person authentication is required.


## Introduction

Speech is the most natural way to communicate for humans. While this has been true since the dawn of civilization, the invention and widespread use of the telephone, audio-phonic storage media, radio, and television has given even further importance to speech communication and speech processing. The advances in digital signal processing technology has led the use of speech processing in many different application areas like speech compression, enhancement, synthesis, and recognition. [1]

The concept of a machine than can recognize the human voice has long been an accepted feature in Science Fiction. From „Star Trek to George Orwell's „1984 - "Actually he was not used to writing by hand. Apart from very short notes, it was usual to dictate everything into the speak writer." - It has been commonly assumed that one day it will be possible to converse naturally with an advanced computer-based system.

The voice recognition technology enables the severely deaf or hearing impaired people, who cannot recognize by the aids. This is expected to be a remarkable innovation for the life quality of the hearing-impaired. Lowering the gate length from the current size of voice merely by amplifying the sound, to see the words recognized 40nm to 10nm in the semiconductor process technology within the next 20 years will bring about a reduction in the size of the hearing aids. The theory was

So simple that a voice was generated through the trachea and the speech was decoded in the brain. Even the voice spectrogram was not considered.

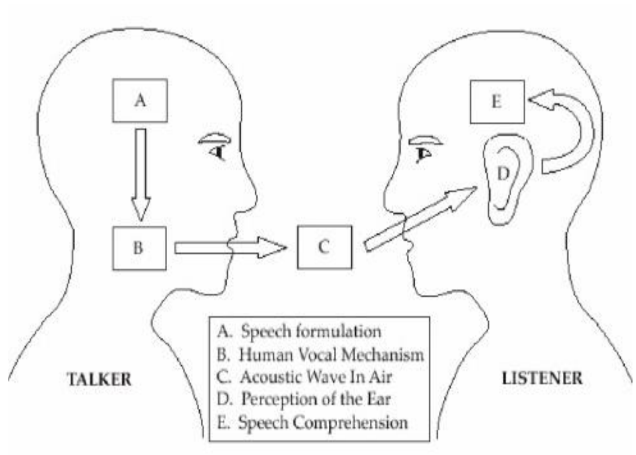


Figure 1.Diagram of Voice Production/PerceptionProcess

## Description

### Hardware

In hardware the components are serial port, MAX232 voltage level converter controller to take input and generate output. To drive the relays we have used ULN 2803 IC which has arrays of 8 Darlington pair of transistors. Darlington pair of transistors are capable of provide larger amount of current to drive the relays. 8 LEDs are used as indicator each corresponding to 8 data lines of the data port.

### Circuits

- RS-232 Level converters circuit for UART communication.
- ULN based relay driver circuit
- Programmer Circuit for Microcontroller

# Flow Diagram

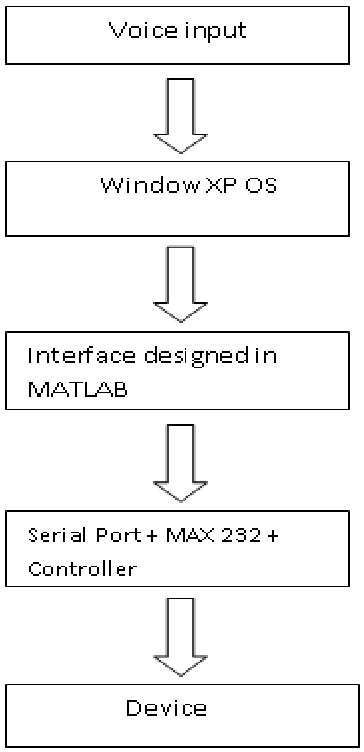


Figure 2. block diagram of hardware

## Software

The software section is completely based on MATLAB. In our interface we have used MATLAB for voice reorganization. It can be used into three different modes viz. 'text to speech', 'speech to text' and as 'voice command recognizer'. We have used it in third mode. In this mode of operation we can add predefined commands. It listens command and matches it from the given list. If matching occurs it generates an event corresponding to the matching. This event is used to control the device by giving the controller input to control the output and thus control the system.

# Key Components

## Hardware

- Computer system
- Microphone
- DB 9 connector
- MAX232
- AT89S8253 controller
- Relay driver ( ULN 2803)

Resistors, LED. , PCB, RELAYS etc. [3]

## Software

- Window OS
- MATLAB [3]

# Operation

To recognize the voice commands efficiently different parameters of speech like pitch, amplitude pattern or power/energy can be used. Here to recognize the voice commands power of the speech signal is used. First the voice commands are taken with the help of a microphone that is directly connected to PC. After it the analog voice signals are sampled using MATLAB. As speech signals generally lie in the range of 300Hz-4000 Hz, so according to Nyquist Sampling Theorem, minimum sampling rate required should be greater or equal to 8000 samples/second.

$$\mathbf{F_{S}=2*F_m} \qquad \mathbf{(1)}$$

Where FS is sampling frequency and Fm is frequency of the modulated wave signal.[4] After sampling, the discrete data obtained is passed through a band pass filter having pass band frequency in the range of 300 - 4000 Hz. The basic purpose of using band pass filter is to eliminate the noise that lies at low frequencies (below 300 Hz) and generally above 4000 Hz there is no speech signal. This algorithm for voice recognition comprises of speech templates. The templates basically consist of the power of discrete signals. To create the templates here the power of each sample is calculated and then the accumulated power of 250 subsequent samples is represented by one value. [2]

For recognition of commands first a dictionary is created that consists of templates of all the commands that the device has to follow (like ON and OFF).For creating the dictionary the same command is taken several times and template is created each time. For creating the final template the average of all these templates is taken and stored.

After creating the dictionary of templates, the command to be followed is taken with the help of the microphone and the template of the input command signal is created. Now the template of command received is compared with the templates of dictionary using Euclidian distance. It is the accumulation of the square of each difference between the value of dictionary template and that of command template at each sample points. The formula can be given as

$$\textit{EuclidianDistance} = \sum_{i=1} (\text{dic[i]-com[i]})^2 \qquad (2)$$

Where i denotes the number of sample points, which is 32 in the proposed algorithm. After calculating Euclidian distance for each dictionary template, these distances are sorted in the ascending order to find out the smallest distance among them. This distance corresponds to a particular dictionary template which is the template belonging to a particular dictionary command. Then the device detects that particular command given by the operator and performs the task accordingly. If the command given by the operator does not match with any of the dictionary command then the device should not follow that command.[2]

# Advantages

- Very easy to control the appliances
- Even an ordinary person can control it
- Very much cost effective
- Easy to design

# Limitations

- It can understand pure English only.
- Voice interference can lead to undesired operation.
- Even noise can activate it.

# Application

The maximum application is in our home. The different appliances can be controlled very easily by voice. It can be used in military application where soldiers are given command so it can also be used there. It can be further implemented into robots which operate on voice. This technology is also used in security.

# Conclusion

A system for reliable recognition of voice has been designed and developed. This system can be made highly efficient and effective if stringent environmental conditions are maintained. The setup for maintaining these environmental conditions will be a onetime investment for any real life application.

# Acknowledgement

It gives us an immense pleasure to present the technical paper on topic CONTROLLING OF DEVICE THROUGH VOICE RECOGNITION USING MATLAB undertaken during Bachelor of Technology final year. We would like to express our deep and sincere gratitude to IJATER Journal for the support to develop this document our project guide for providing us excellent guidance, encouragement and inspiration throughout the project work. We would like to extend our special thanks to Head of Department & Project In charge for their immense support in all kind of work.

# References

[1] DARSHAN MANDALIA, PRAVIN GARETA – "Speaker Recognition Using MFCC and Vector Quantization Model".

[2] ADITYA AGARWAL, SUBHAYAN BANERJEE , NILESH GOEL "Design & Implementation of a Person Authenticating & Commands Following Robot."

[3]39609152-voice recognition.pdf/www . troubleshoo t 4 free . com / fyp /

[4] SINGH AND SAPRE –"PRINCIPLE OF COMMUNICATION".

# Biographies

**Sonam Kumari** is pursuing the B. TECH degree in Electronics & Communication Engineering from the Gautam Buddh Technical University, Lucknow, Uttar Pradesh & will complete the degree in the year 2012. Currently she is in B.TECH final year at Kanpur Institute of Technology, Kanpur. Sonam Kumari may be reached at sonam099ec@gmail.com.

**Kavita Arya** is pursuing the B. TECH degree in Electronics & Communication Engineering from the Gautam Buddh Technical University,Lucknow,Uttar Pradesh & will complete the degree in the year 2012. Currently she is in B.TECH final year Kanpur institute of technology,Kanpur. Kavita Arya may be available at aryakavita35@gmail.com.

**Komal Saxena** is pursuing the B. TECH degree in Electronics & Communication Engineering from the Gautam Buddh Technical University,Lucknow,Uttar Pradesh & will complete the degree in the year 2012. Currently she is in B.TECH final year Kanpur institute of technology,Kanpur. Komal Saxena may be available at komalsaxena11@gmail.com.