

An Embedded Controller for Speaker Dependent Door Security System

V. Ramya

Assistant professor,
Dept of CSE,
Annamalai University,
Annamalainagar, India

C. Gayathri Devi

Department of Computer Science
and Engineering,
Annamalai University,
Annamalainagar, India

M. Kokilasundari

Department of Computer Science
and Engineering,
Annamalai University,
Annamalainagar, India

Abstract-Speech recognition system makes human interaction with computers possible and an automated service or process can be initiated through a speech. In this work, an embedded speaker dependent based door security is implemented. MFCC algorithm is used to extract the speech features of a speaker. Then classification is carried out by using the Euclidean distance with the minimum calculation for matching and recognizing the input speech command with the database. Then the system identifies the speaker, if it matched with the database and it provides authorization notification to the user. The authorized information is sent to the embedded board which consists of microcontroller, door controller and buzzer using XBee pro. XBee pro is used as a wireless data transmission media which supports for up to 90m to 60m from transmitter section to receiver section. The unauthorized notification is identified, when the speech signal is not matched with the database, then the system immediately raises the audio alarm (buzzer). The audio alert will continue until the reset switch is pressed. This developed system take only 0.7 seconds to verify one speech sample.

Key words: *Speech recognition, speaker dependent, MFCC, XBee pro, Microcontroller.*

I. INTRODUCTION

1.1 Role of speech recognition

Speech is essential tool for communication. Speech Recognition is audibly detects information from user speech and parses that speech to produce a word. From that to identify what a person told. Speech is like as sounds, words or phrases. The human speech is transformed through microphone and is digitized by an analog to digital converter, after which that the data is processed by system. These signals are converted into coding patterns. Various sounds are generated by changing the shape of the vocal track. This system is changing time slowly. Changes occur slowly compared to the pitch period, but different sounds are different periodic signals. There are two types of speech recognition system [1, 2],

1. Speaker dependent
2. Speaker independent

A speaker-dependent system is developed to work for a single speaker. These systems are typically simpler to create, less expensive to purchase and more precise. The system is prepared to comprehend one user's pronunciation, expressions and accents, and can run considerably all the more productively and precisely. It requires users to participate in preparing sessions that "instruct" the computer to perceive the user's voice. The computer then makes a voice profile that matches the require training. So it is called "Speaker dependent". Speaker independent system matches the user voice to common voice. These systems that do not use training, are called "Speaker independent". Speaker independent is normally less accurate than speaker dependent.

A. Order of discourse distinguishment framework

The speech recognition system can be ordered into four separate classes, and are, isolated words, connected word, continuous speech, and spontaneous Speech. The developed system supports both secluded and associated words.

1. Isolated words: A isolated word speaker recognition systems acknowledges that the break quickly between words. It perceives single word at a time. This system have "Listen/Not Listen" states, where the speaker to hold up between words.
2. Connected words: A connected word is same to disconnected articulation yet permits isolated words to be run together with a minimal pause between them.
3. Continuous discourse: Speaker recognizers permit users to talk naturally, while the computer checks the content.
4. Spontaneous Speech: It is a speech that is common sounding and not practiced. This work is to be developed by taking into consideration the isolated words.

B. System Applications

1. Automotives: In manual control data, for instance a finger control on the guiding wheel, empowers speech

recognition system and signaled to drive by a sound. Latest auto models offer natural language speech recognition in place of a fixed set of commands, permitting driver to utilize full sentences and phrases.

2. Military:

Test and evaluation of speech recognition are utilized as a part of contender flying machine. Speech recognizers have been working effectively in flight aircraft, with application including: setting radio frequencies, commanding a system, weapons discharge parameter, controlling flight display.

3. Education:

For language learning, speech recognition can be helpful for a second language. It can teach proper pronunciation, in addition helping a person develop familiarity with their talking abilities. Students who are visually impaired can benefit from utilizing the innovation to convey words and after that hear the computer recite them.

II EXISTING SYSTEM

S.R. Suralka et al. "Speech Recognized Automation System Using Speaker Identification through Wireless Communication" Mel Frequency Cepstral Coefficient (MFCC) algorithm is used to recognize the speech of the speaker and to extract features of speech. If security is not a big issue, then Speech processor is used to control the appliances without speaker identification.

Mitali patil et al. "The Design and Implementation of Voice Controlled Wireless Intelligent Home Automation System Based on ZigBee" This System uses SAPI (Speech Application Programming Interface) a Microsoft Application to enable voice recognition when a user gives a voice command to the system. This system contains three main components: i) Intelligent Home Server with ZigBee module, ii) Intelligent environment detection sensor modules and iii) Voice command controlling module. The various features of the system include turning any home appliances or devices, playing media applications; downloading RSS feeds, sending mail, etc. The Voice controlled Intelligent Home Automation System we have implemented is called Intelligent Home Server (IHS).

Dhawan.s "Embedded Speech Recognition System "Voice controlled embedded system (VCES) is a Semi - autonomous system whose actions can be controlled by the user by giving specific voice commands. Raw speech is typically sampled at a high frequency, e.g., 16 KHz over a microphone or 8 KHz over a telephone. This yields a sequence of amplitude values over time. In acoustic model templates have major drawbacks. So we are implementing PCA.

Sweatha kn1 et al. "An ARM based Door Phone Embedded System for Voice and Face Identification and Verification by Open CV and Qt GUI Framework" The main aim of this project is to recognize the voice and the

face of the user at the door based on ARM processors. The algorithms were implemented in the system using C++. The system uses the Intel OpenCV library for image processing. However, when we use the OpenCV library to detect a frontal face in an image using its Haar Cascade classifier face Detector, this will increase the human computer interaction by using real time face recognition. The Phonon multimedia framework is used to display the image of the use. For voice recognition we use the pocket sphinx library, face detection uses a Haar Cascade Classifier and for face recognition uses Principal Component analysis. The PCA algorithm uses Eigen faces nothing but an Eigen value and Eigenvector.

III SYSTEM ANALYSIS

For designing efficient speaker identification, the system has to identify the speaker from the prepared database if he/she is a selected speaker or not. Distinguishing a single user among N users in the database requires a much proficient features. After getting the input sound signals, can remove the features by MFCC and speaker recognized by Euclidean separations [2]. The speaker Recognition consists of two phases,

- Training phase
- Testing phase

In training sessions, the speech signals are recorded in multiple modulations and it is stored in the databas. The noise is removed by applying the Mel filter and the feature is extracted here by using MFCC. In the testing phase, the computer checks the speaker speech to explore whether the speaker is authorized or not. By determining the Euclidean distance of the unknown user speech signal with the authorized speech signal, which is stored in the database. If it matches it shows that the user is authorized, this system enable with microcontroller to control the door otherwise it shows that the user's speech signal is unauthorized and start up buzzer beep sound.

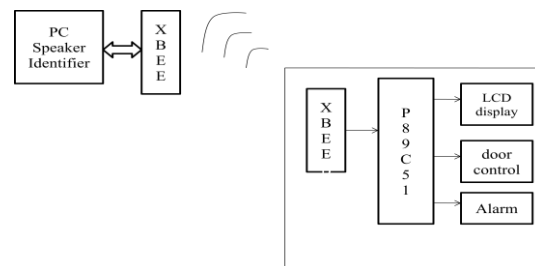


Figure 1: Block Diagram of Proposed system

IV FEATURE EXTRACTION

The input speech signal is blocked and divided into a number of frames (Figure 3). In this work, hamming window is used to get the higher frequency values. The magnitude spectrum is obtained by implementing the FFT to the window coefficients, then Mel spectrum is created

using Mel scale. The inverse FFT is applied to the Mel spectrum to get the MFCC features.

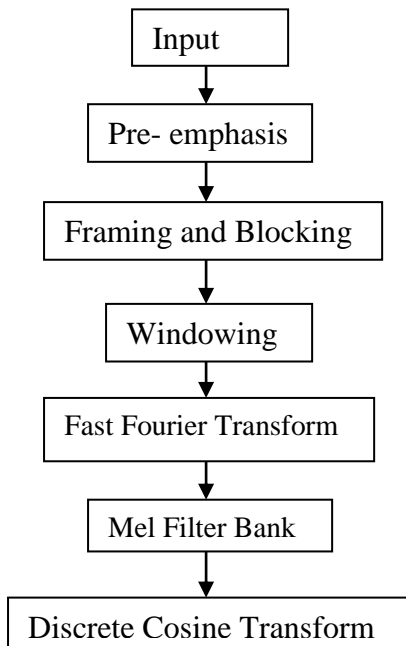


Figure 2: Block diagram of MFCC Algorithm

A. MFCC

The extraction of the best representation of speech signals is a paramount task to create a superior recognition performance. The MFCC algorithm is utilized to extract the features. The MFCC is chosen for the following reason.

- It gives accuracy for clean speech.
- MFCC can be regarded as the standard features in speaker as well as speaker recognition.
- MFCC is the most important features, which are required among various kinds of speech applications.

B. Pre-emphasis

This step, process the passing signal through a filter which highlight higher frequencies. This process will increase the energy of the signal at higher frequency. The wide range of the input speech signal doesn't follow the linear scale. So Mel filter bank is used.

C. Framing and blocking

The methodology of decomposing the recognition sample is acquired from analog to computerized transformation (digital) (ADC) into a small frame within the length of 20 to 40 msec. The continuous 1D signal is obstructed into the small frame of N samples, with neighbor frame differentiated by M examples (M<N) with this the neighbor frames are overlapped by N-M sample. The given 1D signal is split into a small frame to get sufficient samples for recovering the enough information.

D. Windowing

Windowing is defeated minimizing the interruptions at the beginning and at the end the edge, the casing and window capacity is generally duplicated. If the of the window being defined as,

$$W_n(m), 0 \leq m \leq Nm - 1$$

Where Nm = quantity of samples within frame, then

$$Y(m) = X(m) * W(m), 0 \leq m \leq Nm - 1$$

Y(m) = output signal X(m) = input signal

E. Hamming Window

The Hamming window is additionally one period of a raised cosine. Be that as it may, the cosine is raised so high, that its negative peaks are over zero, and the window has a discontinuity in amplitude leaving the window (stepping discontinuously from 0.08 to 0). This makes the side-lobe roll off rate moderate.

F. Fast Fourier Transform (FFT)

FFT is utilized for making change from the spatial domain to the frequency domain. Each frame having N samples and are changed over into the frequency domain. The Fourier transform is a quick calculation algorithm to apply Discrete Fourier Transform (DFT). FFT and DFT are same, however the yield for this transformation contrasts just in computational complexity. For this situation DFT, each one frame N-M samples are specifically utilized as an issue for Fourier transforms. In FFT this frame will be separated into small DFT's and computation will be done with this isolated Dft's. This computation is quick and easy. By calculating DFT we can acquire the magnitude spectrum.

V DESIGN AND IMPLEMANTATION

5.1 Embedded Main Board Module

The security system consists of two modules, embedded main board module and the wireless module. This is important module for the proposed system and provides security for home and is shown in the Figure 5.1. The embedded main board module designed with the help of P89C51 Micro controller and the controller is programmed with the Embedded C language. A device is design of Single-Chip 8-Bit Microcontroller manufactured in advanced CMOS process and is a derivative of earlier 80C51 microcontroller family. The instruction set is 100% well-matched with the 80C51 instruction set. P89C51 device contains a non-volatile Flash program memory that is both parallel programmable and serial In-System and In-Application Programmable. In-System Programming (ISP) allows the user to download new code while the microcontroller sits in the application. In-Application Programming (IAP) means that the microcontroller fetches new program code and reprograms itself while in the

system. It allow for remote programming through a modem link. A default serial loader (boot loader) program in ROM allows serial In-System programming of the Flash memory via the UART without the need for a loader in the Flash code.

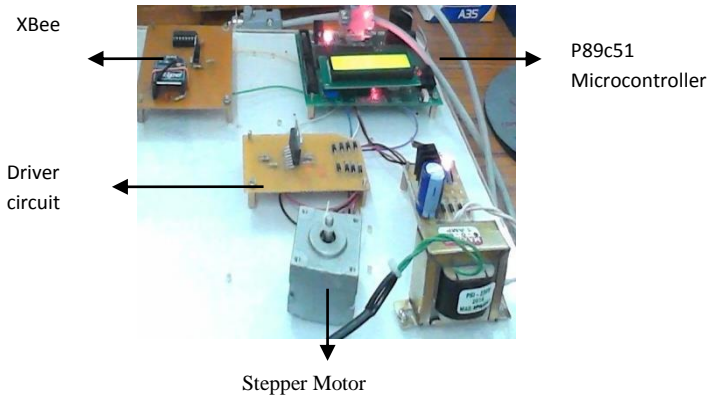


Figure 3: Embedded Main Board Module

5.2 Wireless Module

The Wireless modules consist of LCD display and XBee Pro (Figure 5.3 and Figure 5.4). The main purpose of XBee pro is transmitting the data between transmitter and receiver through wireless.it is designed for 3.3V system. When the data are transmit or receive that time LCD display blink of lights. XBee-PRO module provides cost-effective wireless connectivity to devices in mesh network. Its support for indoor up to 90m to 60m and outdoor 1600m to 750m. Technically XBee PRO is better in communication range and in the data rate.



Figure 4: XBee Pro

5.3 Components Description

- RS232

It is typically for a serial communication transmission of data. It is describe signal connection between data terminal equipment and computer terminal equipment. This identifies electronic characteristic, timing of signals, physical size and Pinout connectors.

- LCD display

Liquid Crystal Display is defined as a flat panel display, electronic display. It is light adjust properties of

liquid crystal. It is used for display images, fixed images such as digit, word, 7segment in a digital clock.

- Buzzer

Buzzer is audio is design for an audio signaling device such as door bell. It may be a mechanical, electromechanical.

Stepper motor

A stepper motor is considered as electromechanical device which converts electrical pulses into discrete mechanical movements. The shaft or spindle of a stepper motor rotates in discrete step increments when electrical command pulses are applied to it in the proper sequence. The motors rotation has several direct relationships to these applied input pulses. The sequence of the applied pulses is directly related to the direction of motor shafts rotation. . The speed of the motor shafts rotation is directly related to the frequency of the input pulses and the length of rotation is directly related to the number of input pulses applied. The rotor will require 24 pulses of electricity to move the 24 steps to make one complete revolution. Motor can be interface with driver circuit.

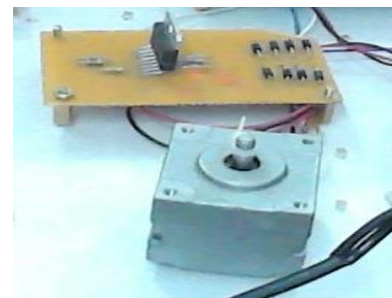


Figure 6: Stepper Motor

6. RESULT AND DISCUSSION

6.1 Creating a database

During the database, the input speech is trained by user and stored in database. The database of first person p1 for 5 words stored and for the second person p2 for 5 words stored.

A. Training phase

Step 1

This is the first stage of the proposed system. In this stage, the system will be getting the input speech signals from the authorized user. There are five samples or modulation of the speaker speech will be taken and analyzed. Finally, the speech signals are stored in the database. The user has to press the recording button to record the speech signals lively Figure 9.2.

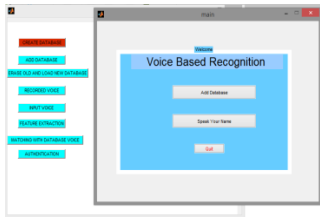


Figure 7: Training phase

Step 2

In this step, already entered data erased using erase old and add new database button then enter number which one have to delete that can be erased.

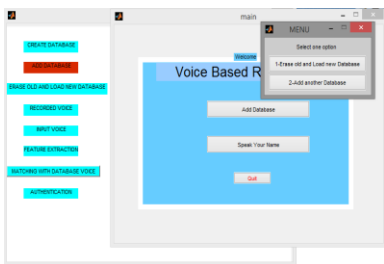


Figure 8: Add Database

Step 3:

In this step, the new user's speech signals is given as input to the system and it is preprocessed and checks the new speech signals into the database.

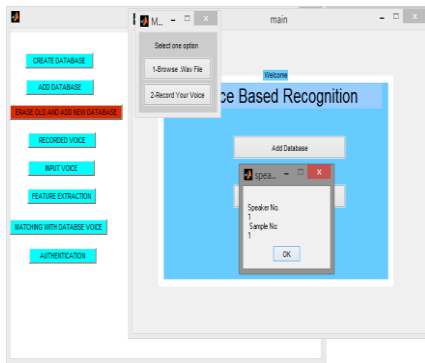


Figure 9: Erase old and add new database

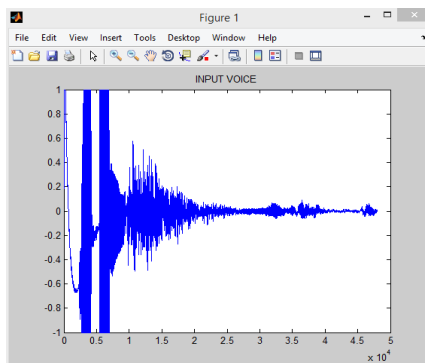


Figure 10: Recorded Voice

6.2 Matching the speaker recognition

Input speech is matched with the database, if it is matched, then it identifies the authorized person else it takes as unauthorized.

B. Testing phase

Step 1

In this step, the system displays the outputs the new user's speech signals. And the modulated speech signal is displayed

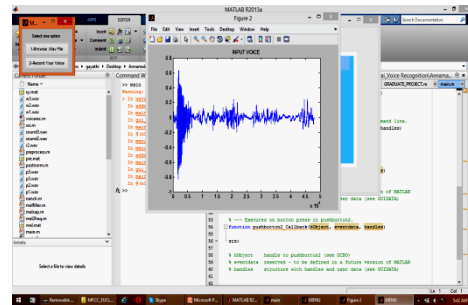


Figure 11: Input voice

Step 2

The most important feature of the proposed work is to display the information about the speech signal, whether it is the authorized person or not. In this step, the system is preprocessed and MFCC algorithm is used to extract features of the speech of a speaker. Then classification is carried out by using Euclidean distance with the minimum calculation, which identifies the authorized person. And provides the notification about the speaker's speech is authorized. This system enable with microcontroller to control the door.

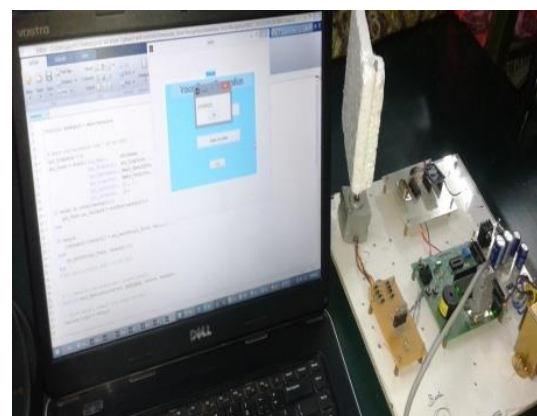


Figure 12: speaker is authorized

Step 3

This step refers, that the speaker's speech signals is not authorized. This may infer the denial of upcoming process. Buzzer starts up with beep sound, when unauthorized person try to access the door.

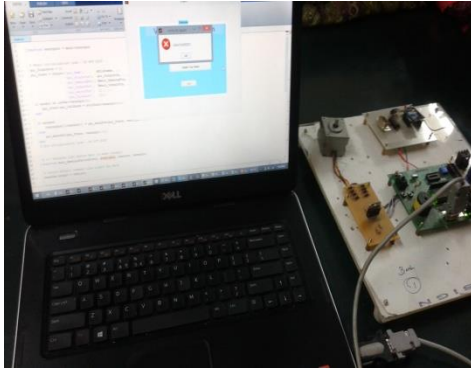


Figure 13: speaker is Unauthorized

V CONCLUSION

The speaker dependent system is introduced and implemented in this work. This system is suitable for highly secured environments like automotive, military, Education. The well known technique MFCC is implemented in this work to recognizing the speaker's speech signal characteristics, furthermore, it is with respect to the speaker discriminative properties of the vocal tract. The objective of this work was to make a speaker dependent system, and apply it to a speech of an unknown speaker. By researching the extracted features of the unknown speech signal and after that compare them with the stored extracted features for each one approved speaker. In this work enable with hardware based an embedded system for the door security system.

REFERENCES

1. Abhishek Thakur, Neeru singla, "Design of MATLAB based Automation Speaker Recognition and control system", International Journal of Advanced Engineering sciences And Technologies, vol no. 8, issue no.1
2. Design of an automation speaker recognition system using MFCC, vector quantization and LBG Algorithm, Ch. Srinivasa Kumar et al. / international journal on computer science and engineering.
3. Lindasalwa Muda, Mumtaj Begam and I.Elamvazuthi, "voice recognition algorithms using Mel frequency cepstral coefficient (MFCC) and Dynamic time Wrapping(DTW) techniques" , journals of computer science, volume 2, issue 3,march 2010, Issn 2151-9617.
4. Jorge MARTINEZ*, Hector Perez, Enrique ESCAMILLA; Masahisa Mabo SUZUKI; "Speaker recognition using Mel Frequency cepstral coefficient (MFCC) and Vector Quantization (VQ) techniques" IEEE 2012.
5. Mohd. Rihan, M. salim Beg, Narayan Gehlot; "developments in home automation and networking", proc. National conferences on "computing for nationaldevelopment", Bharathi vidyapeeth institute of computer and management, new Delhi 23-24 Feb 2007, pp. 61-64.
6. Vovos A. kladis B. and Fakotakis N.D. "speech operated smart-home control system for users with special needs", in proc INTERSPEECH, PP.193-196, 2005.
7. Fezari, M, Khati, A-E "New speech processor and ultrasonic sensors based embedded system to improved the control of a motorized wheel chair, "design and test workshop,2008. IDT 2008.
8. Minho Jin, Frank K. Soong, and Chang D. Yoo, "A Syllable Lattice Approach to Speaker Verification," IEEE Trans. Audio, Speech, and Language Processing, Vol. 15, No. 8, pp. 2476-2484, 2007.