ICSE 2010

PROCEEDINGS OF THE SECOND INTERNATIONAL CONFERENCE ON SCIENCE AND ENGINEERING

Volume - 1

Electronics
Electrical Power
Information Technology
Engg: Physics

Sedona Hotel, Yangon, Myanmar
December 2-3, 2010

PROCEEDINGS OF THE
SECOND INTERNATIONAL CONFERENCE
ON
SCIENCE AND ENGINEERING

Organized by
Ministry of Science and Technology

DECEMBER 2-3, 2010
SEDONA HOTEL, YANGON, MYANMAR

Design and Construction of Voice Recognition Security System

Ei Ei Toe[#1], Maung Maung Latt[*2], U Clement Sandahna[3]

[#]*Technological University (Dawei), Myanmar*

[*]*Technological University (Meiktila), Myanmar*

[1]eieitoe2008@gmail.com

***Abstract*— Since new technologies are advanced in this century, many of the security are developed. Most of the old security systems are lack of human interacting. As advanced technology results, today security system such as voice command control system, voice recognition system, video surveillance system, human face recognition system are still in developed in run way. Among the most of the new technology, human biometric featured security system such as eye signature, face processing and voice processing security systems are most popular because of their advantages. The aim of this paper is to develop a "Voice Recognition Security System". By using this system, a door is controlled by human voice signature. Voice recognition software can process and verify human voice that the voice of the human can pass the door. The system is composed of two main parts, personal computer based voice processing system and parallel (LPT) port interfacing and control circuit. PC based voice processing system is attached with a microphone to PC's soundcard and running with human voice recognition system and processing with engine software.**

***Keywords*— Video Surveillance, Eye signature, Recognition, Biometrics, Parallel (LPT), Engine software**

I. INTRODUCTION

At the end of the century, new technology security systems are being developed according to technology revolution. The more the crime and criminal cases cause, the more population is bombing. Most of the advanced technology features are migrated to existing mechanism for more effective. Many public, private and government associations are concerned on various technologies which related to human identity. One of the advanced security methods is well known as "Biometric Signature" methods [1]. Theory and idea are simple. Everyone has his own natural identity and they are not same to others. There are no possible human features and bio-identity are equal. The most simple and long used method is "Finger Print" methods which are used by worldwide police and national security associations. There are different several identity in human. They are;

1. Finger Print
2. Eye
3. Face
4. Voice/Speech
5. Handwriting/signature (Sign)

The main purpose of this paper is to develop a "Voice Recognition Security System" which is based on human voice identification and recognition. System is based on voice processing engine software to filter, processing and finger out, synthesis and controlling external authority for security. Hardware circuit is needed to develop for demonstration to access and control a door or a gate. A target controlled door can be opened or closed by human voice signature identity. A motor assembly is used to open and close the door. Software based voice identification system controls this motor driven door system.

II. DESCRIPTION OF THE SYSTEM

Fig. 1 shows the schematic of voice based access control system. This system basically consists of three main components namely voice input, user identification/ verification system and door access control. A low–cost microphone is commonly used to record the human's voice. A voice is sent to the identification/verification system which will identify/verify the authenticity of the person based on his/her voice.

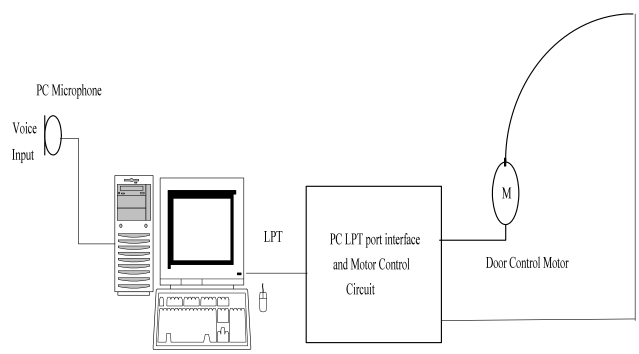

Fig. 1 Block diagram of voice recognition security system

In this paper, a personal computer (PC) 2.4 GHz, 248 MB RAM of Pentium IV processor equipped with sound card is used for user verification implementation. The sound card records the voice data based on the sampling frequency of 8 KHz. In this system, all of the voice data processing and user verification/identification algorithms are implemented in the PC using MATLAB and its toolboxes. As a result of voice-based verification, a decision signal which will access and reject the access will be sent through the parallel port of the PC to the door access control.

III. SOFTWARE DEVELOPMENT

There are several main important in this software development. These several functions parts are kept to perform a complete solution for processing and analysis. The important functions are:

1. User Voice Acquisition
2. Pre-processing
3. Training Voice Database
4. Saving Voice Data Codes (Signatures)
5. Analyzing User Voice
6. Hardware Interfacing with Real World

The paper's objective is to design a close-set system that would recognize several designated speakers by using some of the DSP techniques by using MATLAB. To do this, MATLAB has very important and good performance tools for both DSP and voice acquisition. MATLAB, the language of technical computing provides us with a plethora of useful functions such as fft, dct, autocorrelation, etc.

Moreover handling audio files in MATLAB is a tool with built in functionality to read, record and manipulate PCM encoded wave files. The graphics support is excellent resulting is instant analysis of the procedures involved, and fine tuning, thus speeding up the job.

A. Choosing Methodology

The most commonly employed parameters in speech/voice recognition are pitch frequency, vocal tract area functions, format frequencies, and spectra. Combined together they generate varying properties of nasals, vowels and consonants in a speech waveform for each individual. The representativeness of each feature is rather objective and requires intensive research within a large group of people to generate a decent comparison. The complexity and speed for calculation for each of their analysis, on the other hand, vary with the algorithms chosen for implementation. Simple comparisons of several methods are as followed.

TABLE I
COMPARISON OF SEVERAL METHODS

Method	Complexity	Storage	Computation
LPC	Medium	Small	Fast
Cepstrum Analysis	Medium	Medium	Medium
Time-domain Analysis	Simple	Large	Slow

The listed efficiency for each algorithm is rather generalized. In practice, some voice features can only be implemented with certain algorithm, and some algorithms may not seem beneficial from the above table but their performance in accuracy and stability may compensate for the extra storage space/computational time.

For this paper MFCC (Mel Frequecy Cepstrum Coefficience) is used for this system. MFCC is perhaps the best known and most popular will be utilized in order to realize a system with satisfactory performance in different areas.

B. Choosing Main Program Flow Chart

Program is composed of two main routine and three sub-routines. The main routines are "User Identification "and "New User Registration". If a user want to try access to security, identification loop is tasked an identify "Who is user".

If a new user is try to register in voice database, registration routine is performed. The five new users can be used in this system. User voice is acquired first. Then, voice file is analyzed and generate code signatures. Then, these signatures are saved in voice database. Fig. 2 shows the main program flow chart of the voice access security control system.

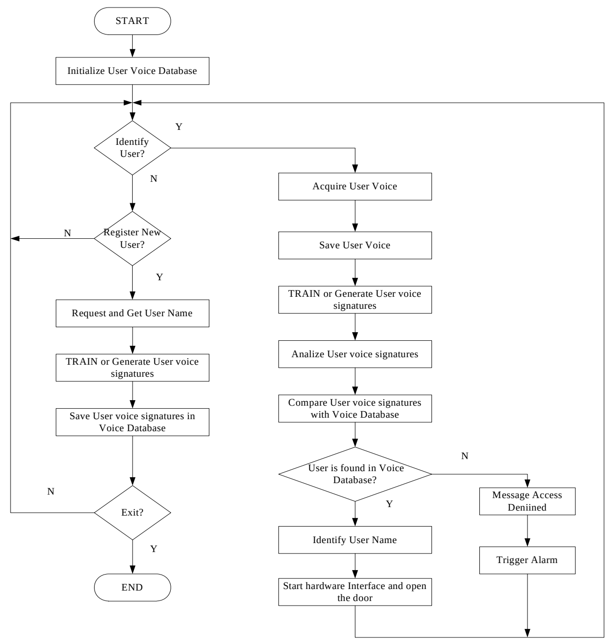

Fig. 2 Main program flow chart

Subroutine act initiating voice database, acquiring user voice and hardware interfacing with LPT port. If user is accessed or denied, LPT port send required signal to open the door or alarm triggering.

C. Pre-Processing

The first stage of voice processing is to collect user voice in well known from such as wave file format. For speech signature data acquisition, a database of speech files is maintained for various users. In this system four speech files are used, each containing the spoken equivalent of the some words for three seconds.

A user should store samples at different times of the day rather than all at one go to get good variability. Moreover the speaker should speak clearly into the microphone without any specific emphasis on any particular word. The recording should be done in a situation when there is very less ambient noise.

Software functions of MATLAB can sample the original time signal, which is real and continuous, and use instead a vector of vales discretized both in time and amplitude. User voice is stored the file in wave file format (PCM), with a sampling rate of 8 KHz sampling rate and 16- bit resolution. All recorded files are stored in “test “folder for future analyzing.

MATLAB data acquisition toolbox is used for user voice acquisition. Function “wavewrite” is used to save acquired data as wave file. Function “waveread” is used to read saved file to memory.

D. Pre-Processing MFCC Function or User Voice Signature Generation

To extract user voice signature, there are two steps of processing. The first step is FFT function over recorded signal and creates filter banks with mel scales. Last function is to generate MFCC components or power spectral of all filter bank passing frequencies. It is shown in Fig. 3.

In this step the Fast Fourier of each resultant frame after windowing is performed, thus obtaining the frequency domain representation (spectrum). The spectrum is now filtered using Mel spaced filter bank. 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, f, measured in Hz, a subjective pitch is measured on a scale called the Mel scale. The mel-frequency scale is linear frequency spacing below 1000 Hz and a logarithmic spacing above 1000 Hz, much similar to the perception model of our ears. As a reference point, the pitch of a 1 kHz tone, 40 dB above the perceptual hearing threshold, is defined as 1000 Mels. Therefore the following approximate formula can be used to compute the Mels for a given frequency f in Hz:

$$\mathrm{mel}(f) = 2595 * \log_{10}(1 + \frac{f}{700})$$

Each bank is filtered as FFT function to get time domain to power spectral values. Each filter bank passing frequency is retransformed by DCT functions to get power spectral of frequency components. Using DCT the log Mel spectrum of the signal can be decorrelated to generate the Mel Frequency Cepstral Coefficients for each frame, thus generating the most important factor for speaker recognition viz. the feature vectors, which indicates the various discontinuities, repetitions and fundamental frequencies in the speech. In this final step, the log mel frequency can be converted back to time. The result is called the mel frequency cepstrum coefficients (MFCC).

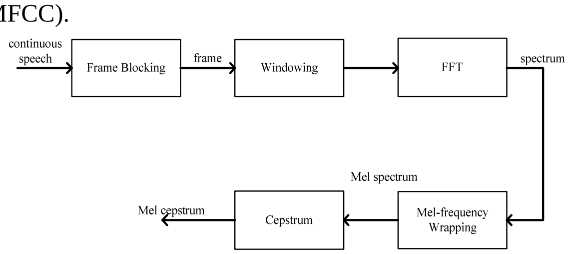

Fig. 3 Block diagram of MFCC processing

E. Analyzing and Identifying User Voice

Last task of voice recognition is to identify unknown user voice with voice database. After user is training or registering his/her voice, his pattern of signatures is stored in voice database. In MATLAB, these data is stored “mat” file format .Every time program is running, these mat file is reloaded as code in program memory for fast accessing.

In this paper, the VQ approach will be used, due to ease of implementation and high accuracy. VQ is a process of mapping vectors from a larger vector space to a finite number of regions in that space. Each region is called a cluster and can be represented by its center called codeword. The collection of all code words is called a codebook [4].

In the training phase, a speaker-specific VQ codebook is generated for each known speaker by clustering his/her training acoustic vectors. The distance from a vector to the closed codeword of a codebook is called a VQ-distortion. In the recognition phase, an input utterance of an unknown voice is “vector-quantized” using each trained codebook and the total VQ distortion is computed [5]. The speaker corresponding to the VQ codebook with smallest total distortion is identified.

IV. SOFTWARE HARDWARE AND SOFTWARE IMPLEMENTATION FOR DOOR CONTROL

The goal of this system is to access the door by user voice signature. To demonstrate the program aim and features, a prototype of small door, door control motor and door control software are required to develop.

MATLAB programming can also be used to hardware interfacing and supports all IO port access over PC such as serial, LPT and GPIO cards. LPT port is used in this system design to control door motor assembly. By using LPT, hardware requirement and code complexity can be avoided.

A. Software Development for LPT Control

There are four types of signal is required to apply the door controller hardware. These are:

1. Motor Left Direction Signal
2. Motor Right Direction Signal
3. Motor Stop Signal
4. Alarm Triggering Signal

If user has access to the door or user voice is found in voice database, door open (left) signal is applied first. After some time delay (4 seconds), door close signal (right) need to apply to close the door.

After these processes motor stop signal is required to stop motor. LPT pin-2 and 3(Output pin 1, 2) are used to control door motor in this design. Alarm triggering signal is required to apply when user is invalid database record or no access the door. Output pin 3 or pin 4 of LPT is used for this purpose.

When door is open, digital signal “01” is applied to motor control. So, output line-1 or DIO-1 is required to set “1” and DIO-2 is required to set “0”. Associated LPT port‘s pin-2 is high and pin-3 is low in this condition. Then, motor will be driven to left direction and door will be opened.

For reverse direction or right direction of motor, output register will be put with binary “10” or decimal “2”. This will drive motor to right and the door will be closed.

If user has no sufficient permission to open the door, the alarm will trigger from LPT port pin-4 or output register no-3.

This can be done by setting port by binary number “100” or decimal “3”.

B. Hardware Implementation for Motor Driver

A small prototype door is controlled by a small 6V DC motor. By driving motor to forward direction, the door can be opened. By driving reverse direction, the door can be closed.

750 rpm Motor loading is important for devices selection. To drive 500 mA load, maximum current rating of switching device should be at least 1 A (twice of load) to avoid overloading and starting torque. H1061 is a NPN type mid

Range driver transistor and it has collector emitter voltage is 80 V and I_{max}= 8 A [6].

Another problem is that LPT pins can sink 5-25 mA of output current. This current is not enough to directly drive motor base bias. So, a buffer is needed to interface between LPT port and H-Bridge. A 4050 HEX NOT buffer is used in this design to interface for current gain and simplicity. This IC has six numbers of NOT gates and each gate can supply source current and input sink current are ± 25 mA. This specification is identically to system design H-Bridge requirement. Since IC has NOT gate, two NOT gates are to invert signal to get normal inputs. Base resistors are used to limit base current overloading if motor is overloaded.

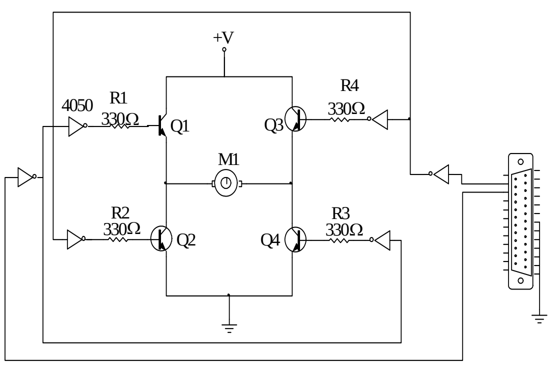

Fig. 4 Door motor driver circuit

C. Alarm Circuit

A piezo buzzer is used to trigger alarm when user who is not allowed to the door. The ratings of buzzer are 5 V and 100 mA. Also there is not enough current to drive directly buzzer from motor pin (output-3).

A C1384 transistor is used to drive buzzer as a switch. NPN type general purpose transistor and rating of 10 V_{ce} and I_{cmax} 500 mA. This is enough for driving buzzer load. Alarm circuit is as shown in Fig. 5.

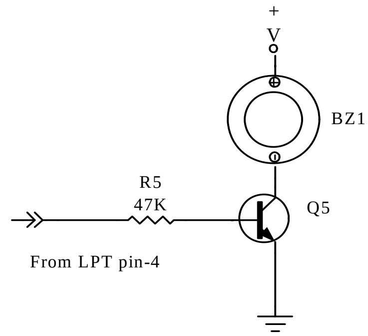

Fig. 5 Alarm circuit

V. TEST AND RESULT

The system is created in the following GUI design. When the system starts and runs the M file, And then, the GUI box is appeared. The pushbuttons, list box and edit box are used for the purpose result. When the system started, the following message is appeared in the Message box as shown in Fig. 6.

Fig. 6 Door motor driver circuit

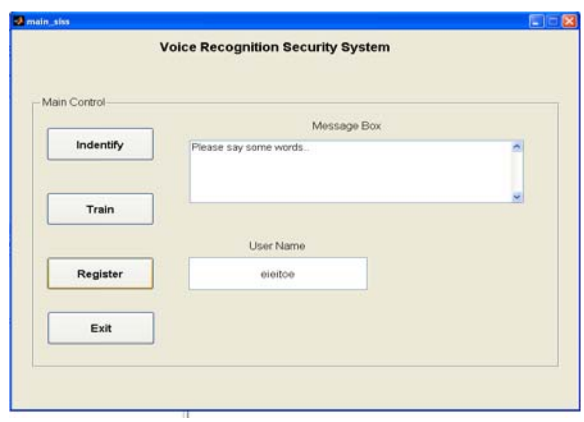

Fig. 7 Register pushbutton is activated

Register push button is used to add the user's name and then message is appeared in the message box. No sooner User's Name is added to the User Name edit box than this user will say the words within three seconds as shown Fig. 7. After the register is added, Train button is used to train the user' name and his/her voice is added to the train file as shown in Fig. 8.

To check the user has in the voice train file, In-identify pushbutton is used. If this user has in the train file, the system access the door and then the door close as shown Fig. 9. If the user has not in the train file, the message is appeared in the Message box as shown Fig. 10.

Fig. 8 Train pushbutton is activated

Fig. 9 In-identify pushbutton is activated

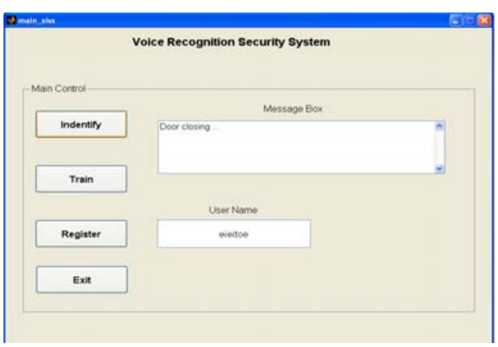

Fig. 10 When the user's voice has not in the data base

VI. Conclusion

The system is very robust, but with the use of good quality microphone and some additional features such as pitch, power spectral density, etc the system can be improved the robustness where discrepancies arise. This paper has presented a simple and efficient MFCC method for voice signature generation. The proposed method has achieved a significant improvement in recognition accuracy. LPC and Time Domain Analysis method can be also used for voice recognition system. But these systems haven't been analyzed in this proposed system. Moreover, four user registers are tested by saying same words in this system. In the future work, it is necessary to develop the system that can be used more users by saying same words or different words.

References

[1] "Which is the Best Biometric Technology?"

[2] http://voicerecognition.com

[3] Campbell, J.P. "Speaker Recognition: a Tutorial", Proceedings of the IEEE, Vol. 85, no. 9, 1997, p. 1437-1462

[4] A. Gersho, R. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Boston, 1992.

[5] F. Soong, E. Rosenberg, B. Juang, and L. Rabiner. "A Vector Quantization Approach to Speaker Recognition". AT&T Technical Journal, vol. 66, March/April 1987, pp. 14-26

[6] http:// www.barello.net/Paper/H-bridge.pdf