Security System Using Biometric Technology: Design and Implementation of Voice Recognition System (VRS)

Rozeha A. Rashid, Nur Hija Mahalin, Mohd Adib Sarijari, Ahmad Aizuddin Abdul Aziz
Department of Telecommunication and Optics, Faculty of Electrical Engineering
Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, MALAYSIA
rozeha@fke.utm.my, adib_sairi@fke.utm.my

Abstract

Biometric technology is fast gaining popularity as means of security measures to reduce cases of fraud and theft due to its use of physical characteristics and traits for the identification of individuals. The earliest methods of biometric identification included fingerprint and handwriting while more recent ones include iris/eye scan, face scan, voice print, and hand print. Biometric voice recognition and identification technology focuses on training the system to recognize an individual's unique voice characteristics (i.e., their voice print). The technology lends itself well to a variety of uses and applications, including security access control for cell phones (to eliminate cell phone fraud), ATM manufacturers (to eliminate pin # fraud) and automobile manufacturers (to dramatically reduce theft and carjacking). In this paper, we present an implementation of a security system based on voice identification as the access control key. Verification algorithm is developed using MATLAB (SIMULINK) function blocks which is capable of authenticating a person's identity by his or her voice pattern. A voice match will produce logic '1' while a mismatch, logic '0'. A microcontroller circuit controlling access to a door is built to test the reliability of this voice controlled security system. It is found out that the developed voice recognition software has successfully activated the door opening mechanism using a voice command that ONLY works for the authenticated individual. The system is proven to be able to provide medium-security access control and also has an adjustable security level setting to account for the variations in one's voice each time a voice identification occurs.

Keywords*:* Biometric technology, voice recognition and identification process

I. INTRODUCTION

Previously, the most popular methods of keeping information and resources secure are to use password and UserID/PIN protection. These schemes require the users to authenticate themselves by entering a "secret" password that they had previously created or were assigned. These systems are prone to hacking, either from an attempt to crack the password or from passwords which were not unique. A Biometric Identification system is one in which the user's "body" becomes the password/PIN. Biometric characteristics of an individual are unique and therefore can be used to authenticate a user's access to various systems.

The word 'Biometric' is taken from the Greek word, of which 'Bio' means life and 'Metric', means measure. By combining these two words, 'Biometric' can be defined as the measure (study) of life, which includes humans, animals, and plants [1]. Collectively, biometric technologies are defined as 'automated methods of verifying or recognizing the identity of a living person based on a physiological or behavioral characteristic' [1]. In analyzing the definition of biometrics, several distinct terms must be elaborated upon to completely understand the framework of biometric technology. The phrase "automated methods" refers to three basic methods connected with biometric devices: (1) a mechanism to scan and capture a digital or analog image of a living personal characteristic; (2) compression, processing and comparison of the image to a database of stored images; and (3) interface with applications systems.

These methods can be configured in a number of different topographies depending upon the biometric device and application. For example, a common issue is whether the stored images (reference templates) reside on a card, in the device or at a host or database. When referring to a biometric technology, it is important to distinguish between physiological and behavioral human characteristics. A physiological characteristic is a relatively stable human physical characteristic, such as a fingerprint, hand silhouette, iris pattern, or voice print. This type of measurement is unchanging and unalterable without significant duress [2].

The application of biometric technology is limitless. Four to five years ago biometric technology

978-1-4244-1692-9/08/$25.00 ©2008 IEEE

was still considered too "fictional" for many. However, with the advancement of microprocessors and signal processing hardware and software, the usage becomes increasingly widespread as security feature in the area of access control, law enforcement and confidential transactions.

The basis for voice or speech identification technology was pioneered by Texas Instruments in the 1960's [2]. Since that time, voice identification has undergone aggressive research and development to bring it into mainstream society.

II. VOICE RECOGNITION SYSTEM DESIGN

Speech processing and language technology contains lots of special concepts and terminology. To understand how different speech synthesis and analysis methods work; we must have some knowledge of speech production, articulatory phonetics, and some other related terminology. Speech signals of the three vowels (/a/ /i/ /u/) are presented in time- and frequency domain in Figure 1. The fundamental frequency is about 100 Hz in all cases and the formant frequencies f_1, f_2, and f_3 with vowel /a/ are approximately 600 Hz, 1000 Hz, and 2500 Hz respectively. With vowel /i/, the first three formants are 200 Hz, 2300 Hz, and 3000 Hz, and with /u/, 300 Hz, 600 Hz, and 2300 Hz. The harmonic structure of the excitation is also easy to perceive from frequency domain presentation.

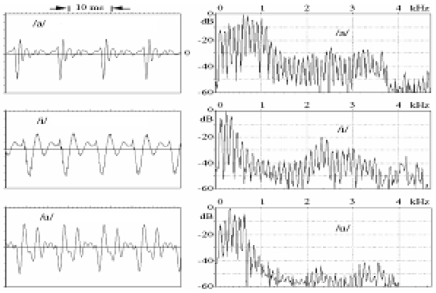

Figure 1. The time-and frequency-domain presentation of vowels /a/, /i/, and /u/.[3]

As can be observed from Figure 1, the rate and pitch of each pronounced vowel is not the same. This poses major concern for voice identification systems, that is how to account for the variations in one's voice each time voice identification occurs. Furthermore, they tend to have a high false reject rate because of background noise and other variables. A simple yet reliable voice recognition system (VRS) (software) has been built in this project. The system was created using Simulink block sets from MATLAB. Basically, a 'voice reference template' is constructed so that it can be compared against subsequent voice identifications. To construct the "reference template", an individual must speak his/her name and this is recorded in the form of *.wav* file.

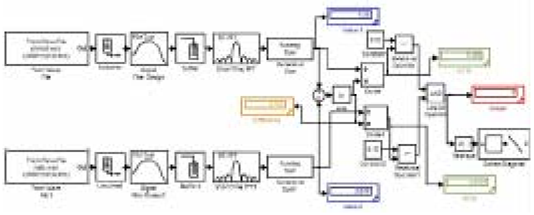

Figure 2. Voice Recognition System

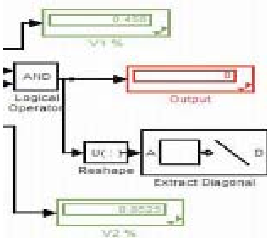

Figure 3. Function Blocks for Access Control

VRS incorporates several variables or parameters in the recognition of one's voice/speech pattern including pitch, dynamics, and waveform. The verification algorithm involving these parameters is executed using the function blocks available in Simulink as shown in Figure 2. There are several steps of speech signal processing involved. First, the measurement of energy levels of short duration of the signal compared to energy level of silence. Second is the removal of noise or any unwanted signal by passing it through *Digital Filter Design* block, which serves as digital FIR band-pass filter, and also produces the frequency components of the signal. The next step is the feature extractions. These include determination of pitch contour by computing autocorrelation on short time basis, determination of format frequencies 1st, 2nd and 3rd and determination of average energy spectral density using autocorrelation and FFT. The comparison procedure is done by equating the sizes of reference and input patterns. Different statistical parameters like standard deviations and covariance were calculated to verify the final result.

The main strength of the developed VRS is that the system provides adjustable level of security for user's convenience. For example, 15% deviation (85% matches) is deemed sufficient to allow normal variation of human voice when they are taken ill. This much deviation might also be allowed in the case of time and attendance systems where only medium-security level is required. Otherwise, the security setting can be set to a higher match value when it is required so; i.e. access to confidential information. Figure 3 shows the function blocks for access control. The user's voice input will be compared with the voice reference template. If the deviation is less than the set security level, logic '1' is produced which means access is granted. Otherwise, logic'0' is produced, meaning that access is denied. The resultant logic output will be transmitted through the parallel port to the microcontroller circuit to perform the next task (opening the door or vice versa) accordingly.

III. Hardware Implementation

The main purpose of the hardware is to implement the control process of locking and unlocking the door. The hardware consists of two parts which are the communication part and the control circuit part.

In communication part, we use parallel port as the communication interface between the PC and the microcontroller hardware as it is easier to program and faster compared to the serial port. However, in parallel transmission, all the 8 bits of a byte will be sent to the port simultaneously while an indicator signal will be sent through another line. Basically, data transmission through a parallel port will involve several data lines, together with a few control and handshaking lines. In the design of the control circuit, we used electronics components such as transistor, capacitor, resistor, oscillator, Light Emitting Diode (LED), switch and a microcontroller.

In this project we have decided to use PIC16F84A microcontroller which comprises of a program memory of 1K words, which translates to 1024 instructions, since each 14-bit program memory word is the same width as each device instruction. The data memory (RAM) contains 68 bytes while 64 bytes for EEPROM. There are also 13 I/O pins that are user-configured on a pin-to-pin basis. Some pins are multiplexed with other device functions [4]. These functions include external interrupt, changes on PORTB interrupts and Timer0 clock input.

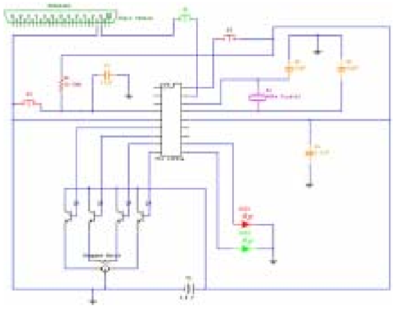

Figure 4. Control circuit design

Figure 4 shows the control circuit design. The designed control circuit is attached to a miniature model of a door, to demonstrate the access control mechanism.

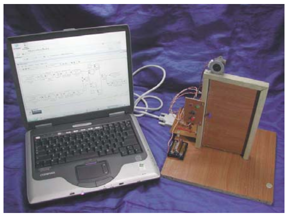

Figure 5. The Developed Hardware

While Figure 5 shows the overall design of the PC based voice controlled security system with door access control application. The following section will discuss the system output consistency which validates the reliability of the verification process.

IV. Results and discussion

When a man speaks a simple word, such as his own name, his voice will produce a waveform. This waveform is known as a voice pattern. Just like fingerprints, voice patterns of human beings are different from one another. Therefore, voice patterns can be used to identify a person identity.

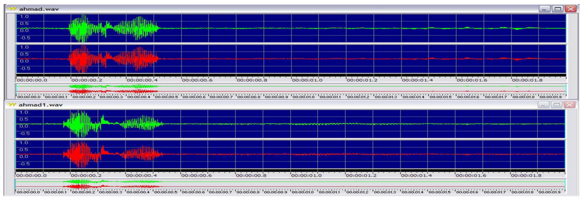

Figure 6. Same Person Voice Patterns*

Figure 6 shows two voice patterns of a male saying the word "Ahmad". One is used as the voice reference template. Both voice samples were recorded at a different time. Through VRS, the differences from both samples can be calculated. *Reference Voice* and *Voice ID* are nonparametric estimates of voice power spectrums. Value of the *Reference Voice* is 5.28, while the *Voice ID* produces a value of 5.959. Therefore, the difference (∂) between these two samples is 0.6792. The standard deviations, V1 and V2, compared to *Reference Voice* are 12.79% and 11.33% to *Voice ID* respectively. As both of the differences are less than 15%, the voices are recognized to be of the same voice. Figure 6 shows the output waveforms. If both inputs V1 and V2 produce a difference below 15% (the set

TABLE I. VOICE RECOGNITION RESULTS

Reference Template	**Second Input**	**Reference Voice**	**Voice ID**	***Diff (∂) = | Reference Voice – Voice ID |***	***V1%= (Diff (∂) / Reference Voice) x 100%***	***V2% = (Diff (∂) / Voice ID) x 100%***	**Output**
*AHMAD	SAME PERSON	5.28	5.959	0.6792	12.79	11.33	1
**AHMAD	DIFFERENT PERSON SAME GENDER	5.28	2.846	2.434	45.8	85.25	0
***AZA	DIFFERENT PERSON SAME GENDER	0.1292	0.0995	0.0297	22.95	29.79	0
****AHMAD	DIFFERENT PERSON DIFFERENT GENDER	5.959	0.1394	5.82	97.17	4175	0

security level for this project), the logic output will be '1' which means access is granted. Figure 7 and 8 show voiceprints of different males and females. Both V1% and V2% exceed 15%, producing logic '0' and therefore, access will be denied. Figure 9 shows the waveform for different gender. Table 1 shows results taken from several voice samples to test the reliability and efficiency of VRS.

With the allowable deviation set at 15%, the system gives full accuracy of authentication. As mentioned, the allowable deviation can be of smaller value for the implementation of higher security access control.

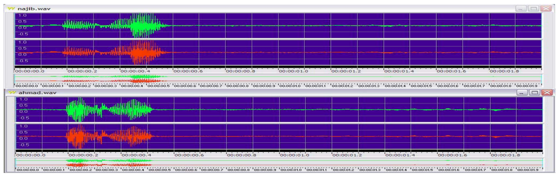

Figure 7. Different Male Voice Patterns**

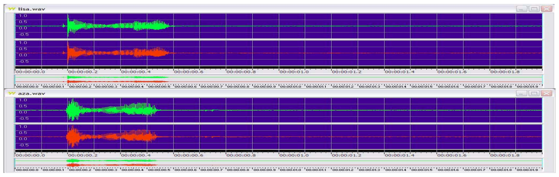

Figure 8. Different Female Voice Patterns***

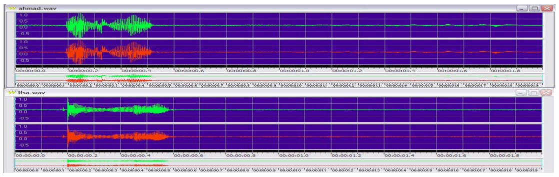

Figure 9. Voice Patterns from a Male and Female****

V. FUTURE RECOMMENDATIONS AND CONCLUSION

Nowadays, biometric technology is becoming increasingly popular due to the use of unique physical traits, such as fingerprints, iris scans, voiceprints, faces, signatures or the geometry of the hand as identification and verification mechanisms. The technology helps business and governments to fight identity theft and fraud, secure transactions, protect confidential information, reduce costs and enhance levels of service. Biometric voice identification technology is still slow to take off in many markets. One reason is it is not as accurate as other biometric technologies due to the tendency to have a high false reject rate because of background noise and other variables. However, with the advancement of microprocessor and signal processing technologies, better vocal measurements using sophisticated algorithms can be taken and converted into a voice print – a unique digital representation of an individual's voice.

A voice recognition system (VRS) has been developed where it has functioned and performed as security feature for access control mechanism with full accuracy at an allowable deviation set at 15%. Although this project can be claimed as successful, there are still opportunities for improvement such as utilizing the voice command control into a stand-alone, self-contained, microprocessor based access control device. Such improvement is necessary to enhance the market value of this project.

REFERENCES

[1] Thomas F. Quatieri, "*Speech Signal Processing – principles and practice*", New York, Prentice Hall PTR, 2002.

[2] Lajos Hanzo, F. Clare A. Somerville and Jason Woodard, "*Voice Compression and Communications*", New Jersey, John Wiley and Sons, 2000.

[3] Martin S. Roden, "*Analog and Digital Communication Systems – forth edition*", New York, Prentice Hall, 1996.

[4] Chi Tsong Chen, "*Digital Signal Processing*", New York, Oxford University Press, 2003.

[5] Richard E. Haskel, "*Design of Embedded Systems Using 68HC12/11 Microcontrollers*", New Jersey, Prentice Hall, 2000.

[6] Stephen J. Chapman, "*MATLAB Programming for Engineers*", Canada, Brooks/Cole, 2002.

[7] Alberto Cavallo, Roberto Setola and Francesco Vasca, "*Using MATLAB, SIMULINK and Control System Toolbox*", New York, Prentice Hall, 1996.

[8] http://www.machinegrid.com/content/view/51/108

[9] http://www.cs.indiana.edu/~zmcmahon/biometrics-history.htm

[10] http://www.voice-security.com/