

Design of Security System by Means of Sound Print

Ismail Abdul-Jabbar Hasan

Department of Electrical Technology, Technical Institute of Baghdad/MTU, Baghdad, Iraq

Email: ismail_baghdadi@yahoo.com

How to cite this paper: Hasan, I.A.-J. (2018) Design of Security System by Means of Sound Print. *World Journal of Engineering and Technology*, 6, 903-913. <https://doi.org/10.4236/wjet.2018.64060>

Received: September 15, 2018

Accepted: November 25, 2018

Published: November 28, 2018

Copyright © 2018 by author and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Sound Recognition becomes an important tool for intrusion detection or for the monitoring of public premises exposed to personal hostility. It could further identify different sounds. The main idea of the sound recognition process in a security system is to store samples of different sound signals in the memory of the computer as references, and to be analyzed with respect to their frequencies components. In this paper, the sound signal of an unknown source would be analyzed and compared with all the available reference samples, and then recognition is made according to the closest sample. The developed security system consists of two main parts: the spectrum analyzer that converts the sound signal to spectrograms. It is designed based on the real-time analyzes, and the recognizer which compares the spectrograms and gives the decision of the recognition by using a special criterion. Experimental results prove that the accuracy of the proposed system can be 98.33% for the selected sample of signals.

Keywords

Security System, Sound Signal, Spectrum Analyzer, Spectrogram

1. Introduction

The speech is the most prominent and essential mode of communication among that of human being. Technological interest in the mechanism of mechanical recognition of human speech capabilities has attracted a great deal of consideration over the past five decades [1] [2]. At the present time, many residential areas and institutes are consuming different types of security system to make sure that their belongings are secured, for instance using password and User ID/Pin for safeguard. Nevertheless, these security systems are not definitely secured because the pin code can be easily hacked, and the ID card can be pinched

and reproduced, etc. Regarding these reasons, a different technology of security system must be emphasized to increase the intimate of the private citizen about the security system [3]. A remarkable development has been detected in the use of speech recognition in security systems, where better protection of confidential of the civilian is needed [4]. Biometric technology is known to make use of the costumer attributes element as the password. The attribute element of everyone is specific, even if the costumers are twins. Therefore, the voice recognition system is protected for the administrator user (or costumer). In this paper, the matter of voice recognition is considered and a voice analysis and recognition system is constructed for definite spoken words.

Automatic Speech Recognition (ASR) is the technique of depicting human speech in a computer. Another technical definition is provided by Gurafsky as [5], the building of system for mapping acoustic signal to a string of words. The main objective of an ASR system is to transform a speech gesture into a text copy message of the verbal words accurately and efficiently without considering the characteristics of the speaker, environment or the scheme used to record the speech (*i.e.* microphone). This process begins when a speaker decides what to say and actually speaks a sentence. Then, a speech wave form will be produced by the software. This speech wave contains the words of the sentence along with the inessential sounds and pauses in the spoken input. Subsequently, the software attempts to decode the speech into the best estimate of the sentence. First, the speech signal will be changed into a sequence of vectors that are dignified all through the period of the speech gesture. Then, using a syntactic decoder it produces a valid sequence of representations [6].

The voice recognition system comprises two main modules which are feature extraction and feature matching. Feature abstraction is to abstract a slight amount of data from the speech signal to be subsequently utilized for signifying each consumer. On the other hand feature matching includes the step to recognize the unidentified user by matching extracted features from admin voice input with that from a set of recognized user [7].

The speech signal and its characteristics can be represented in two different domains which are time and frequency domain. An utterance is the vocalization of a word or words that represent a single meaning to computer [7] [8].

Meanwhile, the genetic development of spectrum analysis facilitates the development of modern communication systems. Spectrum analyzer is a part of the system that divides a signal according to its frequency component. Patterns recognition represents the spectrum of specified sound signals that are stored as a reference patterns. If an object sound has sufficient similarity to any stored reference pattern then it is specified as that of the reference. The computer is used to perform the correlation operations between the input signal and the reference signal [9] [10].

This article attempts to provide a design of sound recognition system for the Frequency range up to 20 KHz. **Figure 1** shows the sound spectrum (representation of sound in terms of the amount of vibration of each individual frequency).

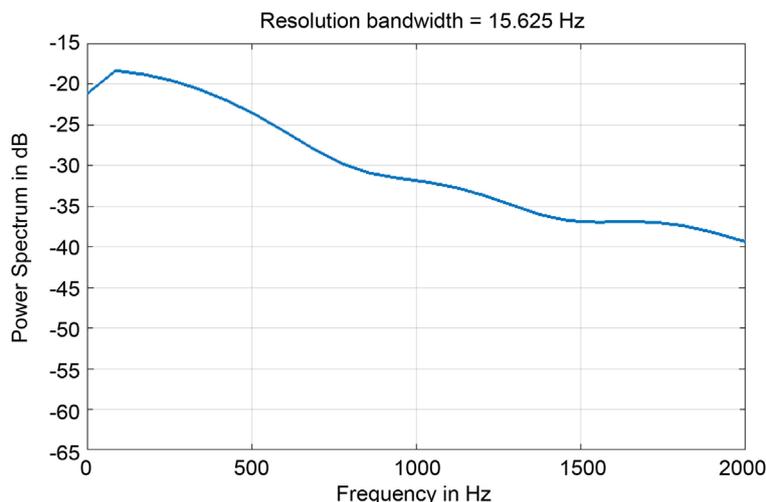


Figure 1. Sound spectrum.

By comparing the spectrum of different signals, we can recognize between different sounds.

Object identification is a pattern-recognition and it is worthwhile to consider some of the general properties of a pattern-recognition system. It operates by measuring a set of features from a representation of an object and then applying a classification to make a classification decision. In ARS, the comparison is done by the computer for any other recognizer. The general block diagram of such a system is shown in **Figure 2**.

2. Real Time Analyzer

Sound signal is a type of continuous signal and such signal will normally be analyzed in short section (short time). The most suitable analyzer for the sound signal is the real-time analyzer. The real-time analyzer obtains the whole spectrum in parallel from the same section of signal, and this is not only able to flow variable signals but can also obtain the spectrum very much faster because it searches for all frequencies within its frequency range at the same time. The most direct way of performing such system is simply to apply the signal to a parallel bank of filter/amplifier/detector channels as shown in **Figure 3**.

The electrical signal coming from microphone is pre-amplified to the BPF by means of a signal operational amplifier. Very high “Q-factor” centers at selective frequencies, then to the detector which detect the RMS value of the signal. The detection accuracy is governed by how narrow a band of frequencies of the filters is, and by how slowly the analyzer sweeps over the desired range of frequencies. In this system, each channel measures the power of its band, increasing the system channel will increase the analysis’s precision, *i.e.* the accuracy will increase. The sharpness of the filter will increase the accuracy.

2.1. Digital Analyzer

A digital spectrum analyzer takes advantage of the ability to convert a sampled

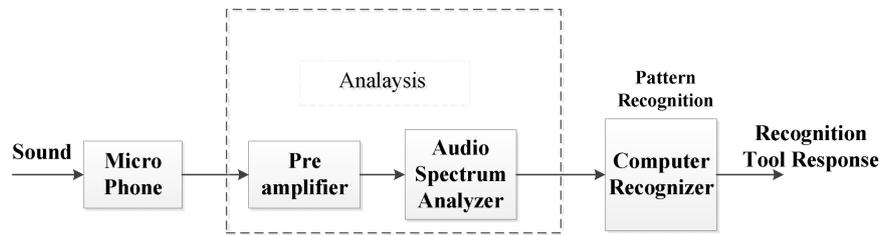


Figure 2. General sound recognition system.

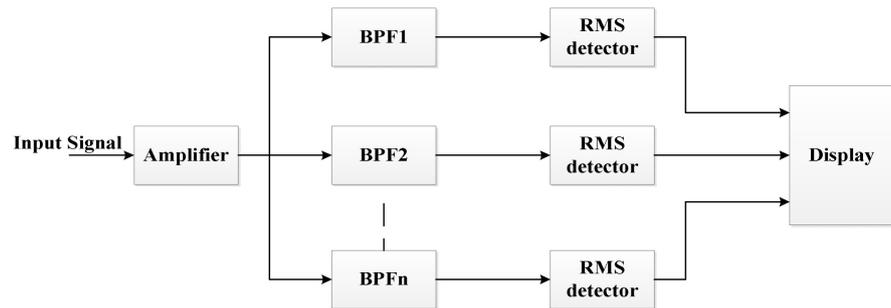


Figure 3. Real time analyzer.

time series into a frequency spectrum directly, two principles are used: digital filter and Fast Fourier Transform (FFT) technique both of which can be realized either as a fast software on computer or hard ware using (DSP) processor. FFT is the efficient method to compute Discrete Fourier Transform (DFT), where it performs as a digital processing tool [11].

The analyzer samples incoming signal at a high rates, saving the data into a bloke. The unit then performs a digital FFT operation on the data block to generate the power spectrum, and the resolution depend on the number of sample points in the block.

2.2. The Implemented System

Figure 4 shows the block diagram of the system. The microphone and pre-amplifier output signal is analyzed into octave band by means of a filter bank channels. Each filter in the bank allows a narrow frequency band to pass when the filter have equal pass band, they form a uniform filler bank; otherwise they form a non-uniform filtering.

The pass band filter helps to define the band of energy to be detected according to the standard specifications using non-uniform bandwidth filtering, sound band (125 - 1600) Hz divided into eight octave band:

$$f_c(i+1) = 2f_c(i) \tag{1}$$

$$F_u(i) = \sqrt{2f_c(i)} \tag{2}$$

$$f_L(i) = \frac{f_c(i)}{\sqrt{2}} \tag{3}$$

where $f_c(i)$ is the centre frequency, $f_u(i)$ is the upper frequency, and $f_L(i)$ is the

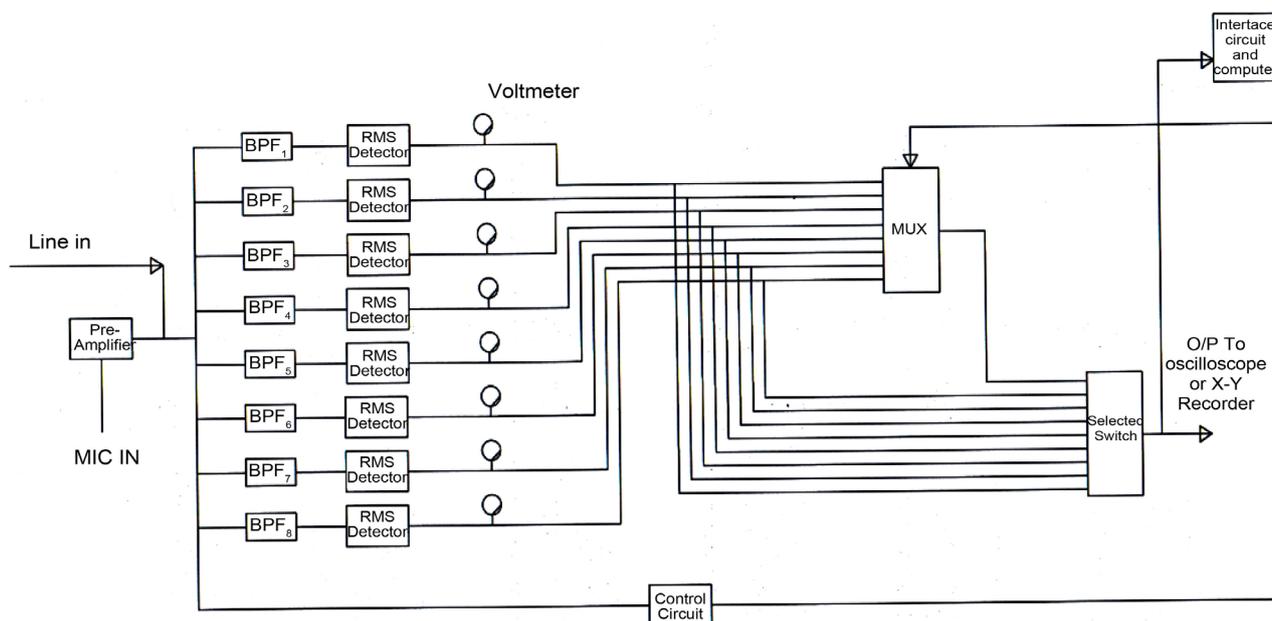


Figure 4. Block diagram of the implemented circuit.

lower frequency.

2.2.1. Automatic Control System

The objective of the control circuit is to detect the existence of the sound signal, and then it gives an order to the other circuits in the system to operate.

The detection is done by using a (710) comparator and (7404) inverter, the output RMS detector circuit is connected to the comparator to detect the sound signal, thus every conservative threshold voltage is chosen, so that the sound level always exceeds it. At the same time it is greater than the noise level in the air. The value of this threshold level is (100 mv) which is exacted from (+15V) supply by means of a voltage divider. The output of the comparator is at logic "0" when there is no signals, but when detect sound signal the output at logic "1".

Figure 5 shows the timing diagram of the control circuit, signal (A) represent the RMS value of the input sound signal, while signal (B) is used to turn the multiplexer on, signal (C) is used to operate the counter type (74293) connected to the multiplexer.

2.2.2. The Interface Circuit

The interface block diagram is showing in **Figure 6**. The heart of the sound system is a special purpose processor called the Digital Signal Processor (DSP) which has a special architecture with an instruction set designed to process analogue information which has been converted to digital.

The RAM chip is usually used for storing temporary sound data. The ROM chip contains all the instructions necessary to operate the DSP. The sound card adapter can process and store audio signal directly from microphone or from the audio spectrum analyzer.

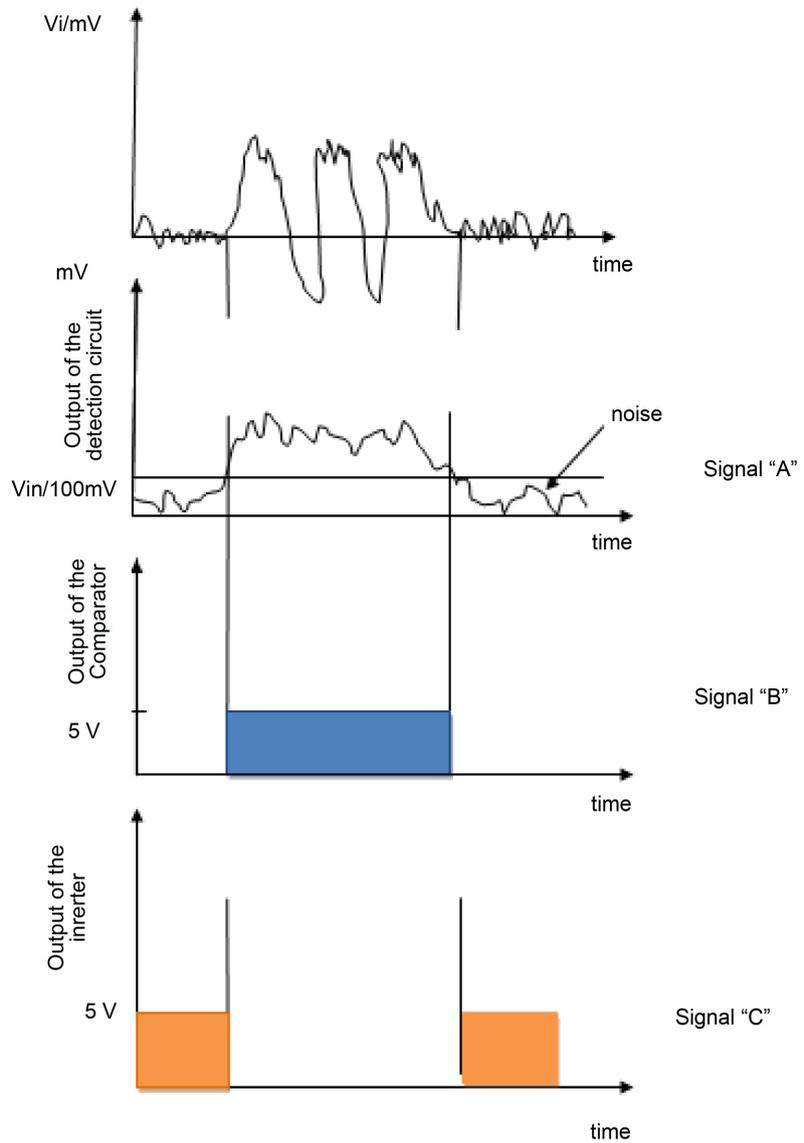


Figure 5. Timing diagram of the control circuit.

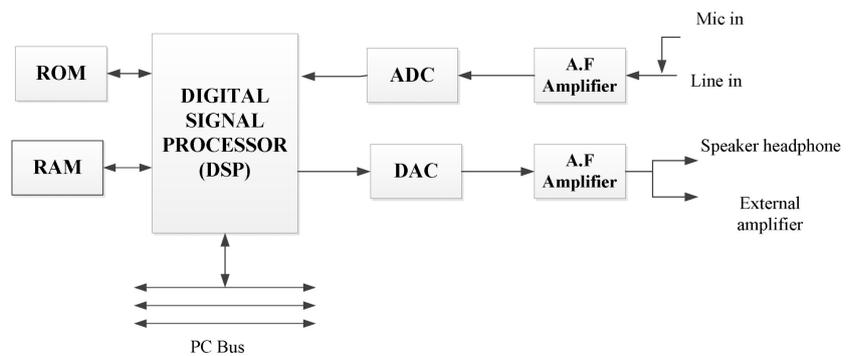


Figure 6. The interface circuit block diagram.

The use of the sound card as an interface circuit is very useful and easy, because it is available in all type of the modern computer, and it's software is very

flexible to be used, providing some necessary options like: a number of samples/sec, a number of bit/sample, etc.

3. Experimental Results

In this Section, the obtained experimental data will be demonstrated. Then experimental results will be discussed. A simple MATLAB program has been built to evaluate the Relative Average Voltage (RAV) of each measured signal.

3.1. Preparation of Experimental Data

The recognition system in this work is designed to recognize the objects by their sound signals. The data of the source sound signals are used in testing the system. Data are collected according to the following conditions:

- The environment recording condition of the sound samples must be the same, otherwise considerable difference in spectrogram will appear.
- The recording of the sound should be made in a quite free field in order to minimize the undesired noise.
- The recording of the one sound source is repeated many times in order to minimize the dissimilarities of the spectrograms.
- Using a high quality microphone in addition to the use of a wind-screen over the microphone to attenuate wind noise.

The spectrogram of the sound signal that obtained by the spectrum analyzer can be described as a three dimensional time-frequency-amplitude patterns, if the patterns of the spectrum vary with time then the signal will be time variant, as in **Figure 7(a)** otherwise it will be time invariant which means it has the same spectrum patterns at any time, as in **Figure 7(b)**.

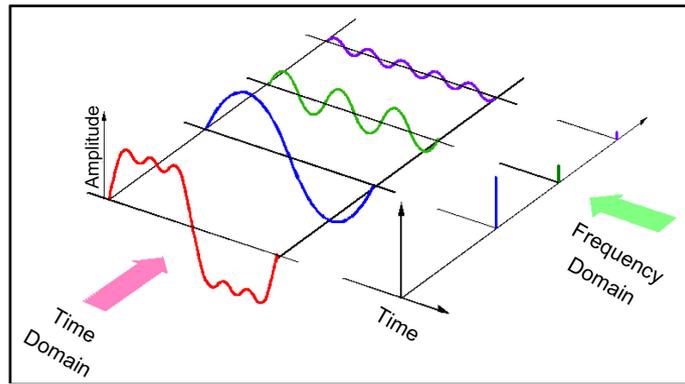
For speech signal the short time spectrum analysis must be accomplished where the recognition process is done with a part or short time of the sound signal, as shown in **Figure 8**.

The system of this work uses the RMS spectrum representation, the sound of the source are converted into digital data (samples) and delivered to the computer then they can be written as matrix, whereas the rows represent the eight frequencies of the spectrum analysis, while the columns represent the number of spectrum patterns and each element in matrix represents the RMS amplitude of the signal in the individual frequency, using "RAV" approach in which the spectrum patterns are converted into a single patterns vector

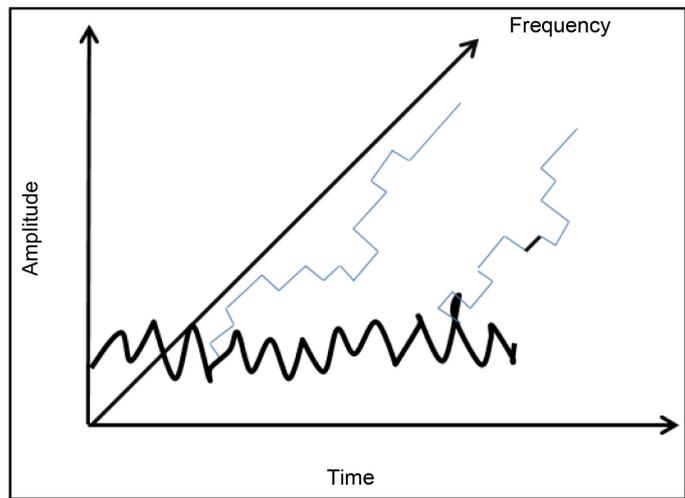
$$\text{RAV} = \frac{1}{N} \sum_{i=1}^N A_k(i) \quad (4)$$

where $A_k(i)$ is Amplitude of the pattern(i) and N is Total number of samples per channel.

According to Nyquist theory (which states that the analogue signal must be sampled at a rate of twice the highest frequency) [12], therefore the sampling rate of the A/D convertor must be 32 KHz (2×16) since the highest frequency in the system is 16 KHz.



(a)



(b)

Figure 7. (a) Spectrum pattern. (b) Spectrum pattern.

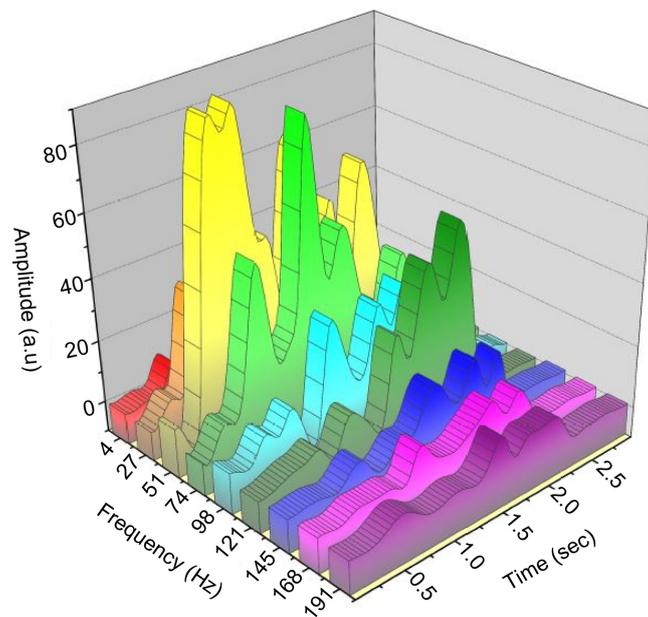


Figure 8. Spectrum the short time spectrum analysis for speech signal.

3.2. Recognition of Signals

All the patterns vectors of the reference sound are stored when a new analyzed signal of the unknown is received then it will be converted into RAV vector. This vector is then compared with all reference vector by calculating the absolute differences between its element and the element of the other vectors. The decision rule is that of the “nearest neighbor”.

Sound of eight words were analyzed and converted into RAV vectors and then stored as a reference patterns, these vectors are shown in **Table 1**.

Sound of six signals for people will be repeated again to obtain vectors that will be considered as test patterns this is shown in **Table 2**.

Table 3 repeated the final results by calculating the average absolute difference between the test vector and each of the reference vectors. The highlighted in red numbers indicated the lowest value in the row, where the best accuracy can be obtained.

The accuracy of the recognition depends on the similarity in the acoustic properties of the sound signal. Greater accuracy will be obtained whenever there are larger differences between these sounds.

Using a large amount of the sounds in the test might further improves the investigation of the system accuracy. It has been observed that greater accuracy

Table 1. Reference pattern.

Frequency band	A	B	C	D	E	F	G	H	I	J
1	31.00	33.00	39.10	44.00	30.20	36.80	36.00	35.90	16.00	15.00
2	64.90	50.00	35.10	50.00	41.00	40.10	33.00	31.00	35.00	33.30
3	65.40	59.00	52.00	53.00	58.00	55.5	41.00	33.00	45.50	44.00
4	37.00	35.00	53.80	42.00	53.10	55.10	44.00	32.00	51.00	54.00
5	15.80	20.00	28.00	10.50	8.30	22.00	42.00	29.00	17.00	25.00
6	5.40	8.00	12.10	5.20	5.20	16.90	20.90	24.10	10.10	9.10
7	0.0	0.0	4.70	2.30	2.50	2.00	5.00	8.00	3.00	4.00
8	0.0	0.0	6.70	2.00	8.40	3.20	1.50	6.70	2.50	3.50

Table 2. Test patterns.

Frequency band	B	C	D	E	F	G
1	28.00	40.00	41.50	30.10	26.80	19.90
2	44.80	51.60	44.50	40.00	39.10	34.40
3	57.80	50.00	63.00	44.00	39.20	30.00
4	30.00	46.00	25.20	39.00	39.30	27.00
5	14.00	28.00	13.00	22.00	20.60	16.50
6	50.00	8.00	5.60	7.70	18.40	15.40
7	0.00	4.30	4.60	4.70	6.80	2.00
8	0.00	2.00	0.90	5.50	3.00	7.00

Table 3. Average differences between references and test patterns.

R/T	A	B	C	D	E	F	G	H	I	J
B	4.78	3.175	10.36	5.72	3.51	6.77	11.90	12.43	10.125	10.60
C	7.40	5.36	4.525	4.58	8.33	5.66	12.12	12.625	8.10	7.90
D	6.46	5.33	5.22	4.93	5.025	5.45	10.57	9.62	11.25	12.925
E	7.61	5.55	6.23	6.77	5.23	6.45	7.05	9.91	6.185	5.437
F	13.40	10.70	11.31	12.35	10.33	11.175	6.90	6.88	6.137	6.76
G	13.73	8.06	7.41	9.22	6.99	6.35	5.90	10.65	7.03	8.411

would be obtained if each sound source has series of the patterns produced by several repetition of the recording process in different environmental conditions.

This means that increased accuracy depends on increasing the information available to the recognition process.

The sound of the sources is analyzed and converted into RAV vector and stored as eight reference patterns, while six of sound is repeated and stored as a test patterns as in **Table 1** and **Table 2** respectively. By using the average absolute difference between these tables, result reported in **Table 3**. For example, the value of 4.78 is the average absolute difference between the parameters of the coloumn (B) from **Table 1** in test patterns and each parameter in coloumn in vector (A) in reference patterns from **Table 2**. Each parameter in the row of **Table 3** is the average difference between test vector and reference vector.

4. Discussion

By using real time analysis scheme with eight channels, a very good performance is verified in the evaluation of the spectrograms of the sound signals. Regarding the results in **Table 3** which represent the average absolute difference between the test vector and each of the reference vectors, it is mentioned that the percentage recognition for (B, C, D, E, G) was 100% and for (F) was 90%, this means that the probability of error is 2%. Although there are substantial similarities in the sound of the words, and to improve the investigation of the system accuracy a large amount of the sounds in the test must be used.

Increasing the octave bands at centered frequencies may increase the accuracy of the sound recognition, also using a large number of filters such as a bank of (16) or (24) filters instead of the (8) filters in the spectra analysis may lead to a better accuracy. The system will very easily fail to recognize the sound source if the spectrum analyzer is very precise.

It has been observed that greater accuracy would be obtained if each sound source has series of patterns produced by several repetitions of the recoding process in different environmental conditions.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Gaikwad, S.K., Gawali, B.W. and Yannawar, P. (2010) A Review on Speech Recognition Technique. *International Journal of Computer Applications*, **10**, 16-24.
- [2] Juang, B.H. and Rabiner, L.R. (2005) Automatic Speech Recognition—A Brief History of the Technology Development. 2nd Edition, Elsevier, Amsterdam.
- [3] Martonow, M. (2009) Design of Fusion Classifiers for Voice-Based Access Control System of Building Security. *WRI World Congress on Computer Science and Information Engineering*, Los Angeles, CA, 31 March-2 April 2009, 80-84.
- [4] Dixit, S. and Mulge, M.Y. (2014) Security System in Speech Recognition. *International Journal of Computer Science and Mobile Computing (IJCSMC)*, **3**, 275-284.
- [5] Jurafsky, D. and Martin, J.H. (2000) Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice Hall, Upper Saddle River, NJ.
- [6] Rabiner, L.R. and Juang B.H. (2004) Statistical Method for the Recognition and Understanding of Speech. Rutgers University and the University of California, Santa Barbara; Georgia Institute of Technology, Atlanta.
- [7] Shah, H.N.M., Ab Rashid, M.Z., Abdollah, M.F., Kamarudin, M.N., Lin, C.K. and Kamis, Z. (2014) Biometric Voice Recognition in Security System. *Indian journal of Science and Technology*, **7**,104-112.
- [8] Poornima, S. (2016) Basic Characteristics of Speech Signal Analysis. *International Journal of Innovative Research & Development*, **5**, 169-173.
http://www.internationaljournalcorner.com/index.php/ijird_ojs
- [9] (1997) Hints for Making Better Spectrum Analyzer Measurements. Application Note 1286-1, Hewlett-Packard.
- [10] Keysight Technologies Spectrum Analysis Basics. Application Note 150. HYPERLINK.
<http://materias.fi.uba.ar/6644/info/anespec/basico/AN%20150%202016.pdf>
- [11] Achankunju, S. and Mondikathi, Ch. (2015) Voice & Speech Based Security System Using MATLAB. *International Journal of Emerging Trends in Electrical and Electronics*, **11**.
- [12] Nyquist-Shannon Sampling Theorem, Wikipedia, the Free Encyclopedia.