Voice recognition Using back propagation algorithm in neural networks

Abdelmajid Hassan Mansour[#1], Gafar Zen Alabdeen Salh[*2], Hozayfa Hayder Zeen Alabdeen [#3]

[#]1 Assistant Professor, Faculty of Computers and Information Technology, King Abdulaziz University, Saudi Arabia
[*]2 Assistant Professor, Faculty of Computers and Information Technology, King Abdulaziz University, Saudi Arabia
[#]3 Lecturer, Faculty of Computer Science and Information Technology, Alneelain University, Sudan

***Abstract*— One of the most terms in the biometric technologies is the voice recognition. It's used to provide an authentication to any system on basis of the voice features instead of images. This research aims to build a system of voice recognition using back propagation algorithm in neural networks, by comparing the voice signal of the speaker with recorded voice signals in the database, and extracting the main features of the voice signal using Mel-frequency cepstral coefficients, which is one of the most important factors in achieving high recognition accuracy.**

***Keywords*— Voice Recognition, Feature Extraction, Feature matching, voice signal. Back propagation, Mel-frequency Cepstral Coefficients "MFCC".**

I. Introduction

Speech and Voice Recognition are the emerging scope of security and authentication for the future. Now-a-days text and image passwords are prone to attacks. In case of the most commonly used text passwords, users are required to handle different passwords for emails, internet banking, etc. Hence they tend to choose passwords such that they are easy to remember. But they are vulnerable in case of hackers. In case of image passwords, they are vulnerable to shoulder surfing and other hacking techniques. Advances in speech technology have created a large interest in the practical application of speech recognition. Therefore this system provides the users with the appropriate and efficient method of authentication system based on voice recognition [1].

Voice is also physiological trait because every person has different pitch, but voice recognition is mainly based on the study of the way a person speaks, commonly classified as behavioural. Speaker verification focuses on the vocal characteristics that produce speech and not on the sound or the pronunciation of speech itself. The vocal characteristics depend on the dimensions of the vocal tract, mouth, nasal cavities and the other speech processing mechanism of the human body [2].

Voice-biometrics systems can be categorized as belonging in two industries: speech processing and biometric security. This dual parentage has strongly influenced how voice-biometrics tools operate in the real world. Speech processing. Like other speech-processing tools, voice biometrics extract information from the stream of speech to accomplish their work. They can be configured to operate on many of the same acoustic parameters as their closest speech-processing relative speech recognition [3].

Voice recognition has two categories text dependent and text independent. Text dependent voice recognition identifies the speaker against the phrase that was given to him at the time of enrollment. Text independent voice recognition identifies the speaker irrespective of what he is saying. This method is very often use in voice recognition as it require very little computations but need more cooperation of speakers. In this case the text in verification phase is different than in training or enrolment phase [4].

II. Artificial Neural Networks

A neural network is a computational structure inspired by the study of biological neural processing. There are many different types of neural networks from relatively simple to very complex [5].

The main characteristics of neural networks are that they have the ability to learn complex nonlinear input-output relationships, use sequential training procedures, and adapt themselves to the data. The most commonly used family of neural networks for pattern classification tasks is the feed-forward network, which includes multilayer perceptron and Radial-Basis Function (RBF) networks. Another popular network is the Self-Organizing Map (SOM), or Kohonen-Network which is mainly used for data clustering and feature mapping. The learning process involves updating network architecture and connection weights so that a network can efficiently perform a specific classification/clustering task. The increasing popularity of neural network models to solve pattern recognition problems has been primarily due to their seemingly low dependence on domain-specific knowledge and due to the availability of efficient learning algorithms for practitioners to use Artificial neural networks (ANNs) provide a new suite of nonlinear algorithms for feature extraction (using hidden layers) and classification (e.g., multilayer perceptron's) [6].

III. MEL FREQUENCY CEPSTRAL COEFFICIENTS "MFCC"

MFCC is used to extract the unique features of human voice. It represents the short term power spectrum of human voice. It is used to calculate the coefficients that represent the frequency Cepstral these coefficients are based on the linear cosine transform of the log power spectrum on the nonlinear Mel scale of frequency. In Mel scale the frequency bands are equally spaced that approximates the human voice more accurate [4]. Equation (1) is used to convert the normal frequency to the Mel scale the formula is used as

$$m=2595 \log10 (1+f/ 700) \text{ -------- } (1)$$

Mel scale and normal frequency scale is referenced by defining the pitch of 1000 Mel to a 1000 Hz tones, 40db above the listener's threshold. Mel frequency are equally spaced on the Mel scale and are applied to linear space filters below 1000 Hz to linearized the Mel scale values and logarithmically spaced filter above 1000 Hz to find the log power of Mel scaled signal. Mel frequency wrapping is the better representation of voice. Voice features are represented in MFCC by dividing the voice signal into frames and windowing them then taking the Fourier transform of a windowing signal. Mel scale frequencies are obtained by applying the Mel filter or triangular band pass filter to the transformed signal [4].

IV. RELATED WORKS

Voice Recognition is an emerging scope of security and authentication for the future, there are numerous studies and researches on this area. N. Revathy and T.Guhan were used back propagation neural network for implementation. It is an information processing system that has been developed as a generalization of the mathematical model of human recognition [5]. N.AYSHWARYA, G.LOGESHWARI and G.S.ANANDHA MALA were discussed A Feed Forward Back Propagation Neural Network (FFBPNN) was used to classify the voices of various speakers in the learning or training phase. The network was tested with samples from the various speakers [7]. Siddhant C. Joshi and Dr. A.N.Cheeran were presented MATLAB based feature recognition using backpropagation neural network for ASR. The objective of this research is to explore how neural networks can be employed to recognize isolated-word speech as an alternative to the traditional methodologies [8]. Md. Ali Hossain, Md. Mijanur Rahman, Uzzal Kumar Prodhan and Md. Farukuzzaman Khan were concerned with the development of Back-propagation Neural Network for Bangla Speech Recognition. In this paper, ten bangla digits were recorded from ten speakers and have been recognized [9]. Presenting the viability of MFCC to extract features and DTW to compare the test patterns were proposed by Lindasalwa Muda, Mumtaj Begam and I. Elamvazuthi [10]. Md. Akkas Ali, Manwar Hossain and Mohammad Nuruzzaman, were presented some technique for recognizing spoken words in Bangla. In this work we use MFCC, LPC, GMM and DTW [11]. MarutiLimkar, RamaRao & VidyaSagvekar were proposed an approach to recognize spoken English words corresponding to digits zero to nine in an isolated way by different male and female speakers [12]. The secure system that deploys the voice recognition for a natural language (Tamil) by combining the digital and mathematical knowledge using MFCC and DTW to extract and match the features to improve the accuracy for better performance were proposed by Dr. Kavitha. R, Nachammai. N, Ranjani. R, Shifali. J [13]. Ms. Savitha and S Upadhya were presented the two template matching techniques namely the Single and Average template matching techniques that were developed to recognize the English digits spoken in isolation. The algorithm was implemented both for speaker dependent and speaker independent type of isolated digit recognition and a comparison of recognition accuracies for the two template matching techniques was made [14]. An overview of major technological perspective and appreciation of the fundamental progress of speech recognition and gives overview technique developed in each stage of speech recognition and also summarize and compare different speech recognition systems and identify research topics and applications which are at the forefront of this exciting and challenging field were presented by Om Prakash Prabhakar, Navneet Kumar Sahu [15].

V. PROPOSED SCHEME

This paper build a system of voice recognition using back propagation algorithm in neural networks. By comparing the voice signal of the speaker with recorded voice signals in the database for the purpose of verifying, and extracting the main features of the voice signal using Mel-frequency cepstral coefficients, which is one of the most important factors in achieving high recognition accuracy. The fig. 1 show the general description of the system.

The proposed scheme depending on two basic phases that each recognition system composed of it's a high level. The first phase is the features extraction, which is called the process of analysis. It create the sources for each digital voice from a set of vocabulary for the forming the voice data base, which is acoustic signal for the pronounced voice called the source signal which is the name of each person. So each signal is divided into templates of equal length samples from the beginning to the end, then each template converted to attributes vector that extract the signal features in that template. Include those vectors in groups called Features Vectors. This processing repeated to each digital voice in the vocabulary group. The second phase is feature matching, also called the process of recognition, where converted the coming signal to be recognized, that called the test signal to a series of vectors using the conversion (beginning - end) itself in the feature extraction processing and these features compared with all the existent probability in the database by using patterns matching method. Given the recognition decision from the matching by a Euclidean distance function between two series of features vector. One represents the vector of source signal feature and the other is the vector of test signal feature.

In this scheme has been created a database for a set of voices from 40 person, by taking 5 samples of each individual person, each of them recording a sample by pronouncing his name five times, in the system training phase.

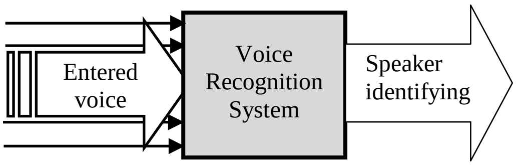

Fig. 1 general description of the system

The proposed system move through two phase:

A. *Training phase*

On this phase the system will be trained by creating training groups composed of different voices samples. Through which the system can create its own voice database by selecting the more accuracy samples and purity, as shown in Fig. 2.

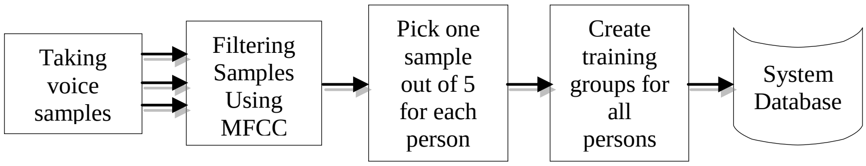

Fig. 2 Steps of Training Phase

B. *Testing Phase*

On this phase the system will be trained to recognize the voice of the speaker after performing the process of matching with the samples taken on the training phase, then the is capable to take a decision on the existence of the voice in the database or not, as shown in Fig. 3.

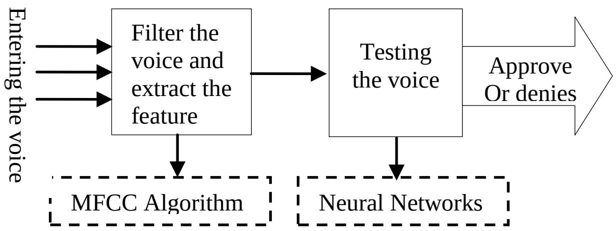

Fig. 3 Steps of Testing Phase

The voice recognition move through phases as shown in Fig. 4.

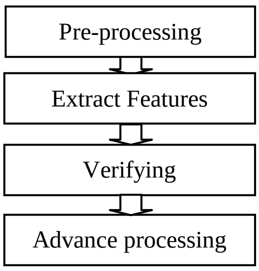

Fig. 4: Voice recognition phases

A. *Pre-processing phase*

At this stage, they will removing the noisiness that surrounding the digital voice signals to extract the distinctive features of the voice clearly. Also consolidating the sounds frequency in many ways.

B. *Extraction of acoustic signal Features phase*

This phase will perform the processing acoustic signal until we get a clear features for voice, through it we can differentiate the voice. This process has three stages: spectral Shaping, Spectral Analysis and Filtering.

C. *Verifying phase*

On this phase done the recognition of voices within classifier training, then come testing phase, which they testing the capability of the classifier for recognizing of voices that has not been trained on it before. In order to know the accuracy of the system.

D. *Advanced processing phase*

The processing occurs after the voice was recognized. that user can be able to know the result of the system after using it. This stage is that the pronunciation was correct way or not.

The system was trained on 20 recorded sample on the database from 20 different person and stored in 16 bit by sample rate equal 11025Hz in a .wave voice file format, the Fig. 5 show the back propagation network training, and the Fig. 6 show learning rate on the Back propagation network.

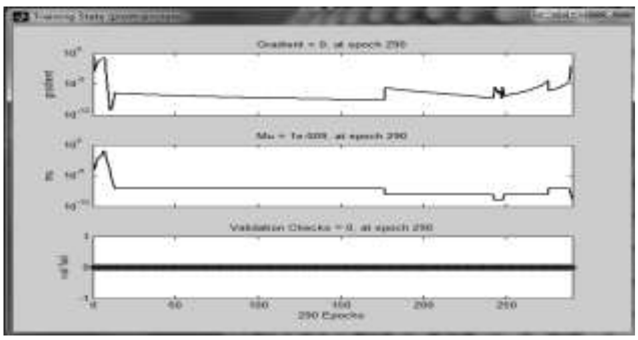

Fig. 5: Back propagation network training

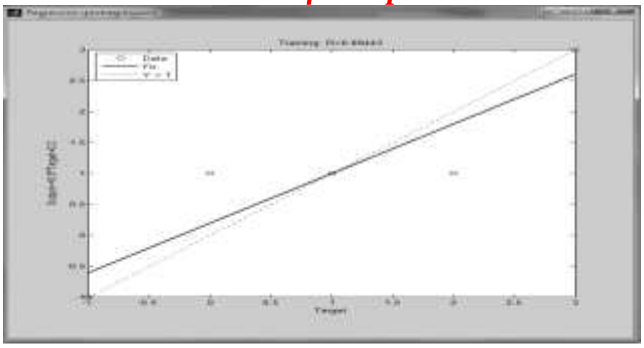

Fig. 6: Learning rate in Back propagation network

The following figures showing the models of recording the word (Hozayfa) 5 times and selecting the best of them to store it in the database, this process repeated to all required sample to be stored.

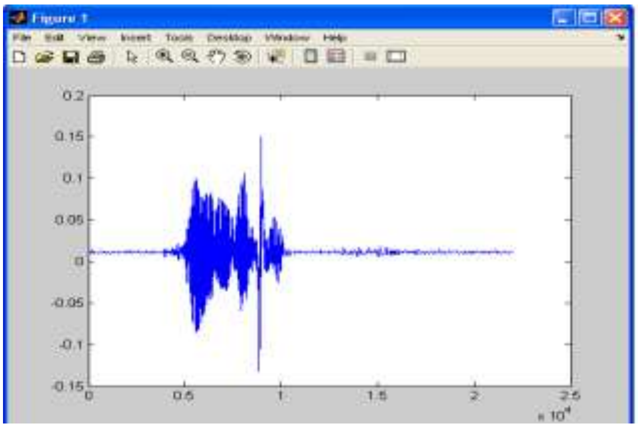

Fig. 7: the 1st model of the word “Hozayfa”

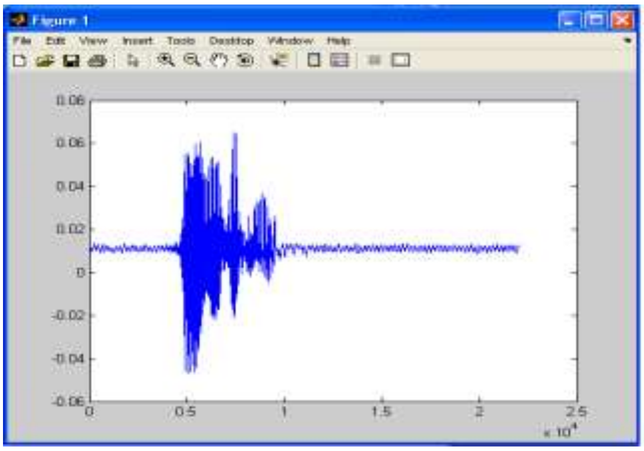

Fig. 8: the 2nd model of the word “Hozayfa”

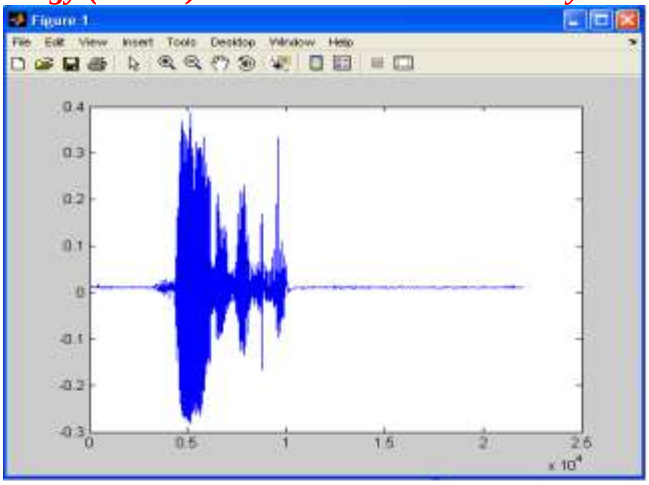

Fig. 9: the 3rd model of the word “Hozayfa”

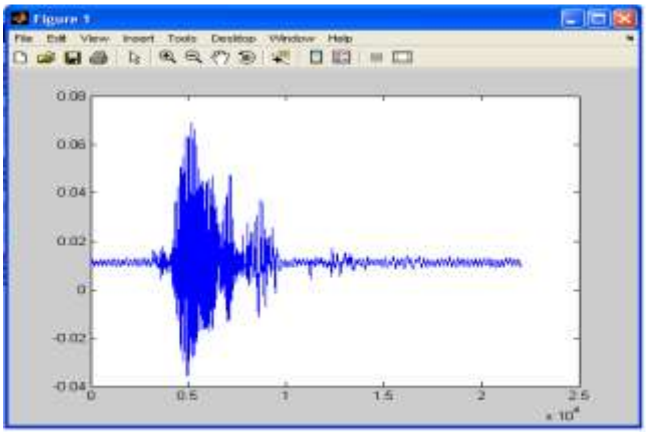

Fig. 10: the 4th model of the word “Hozayfa”

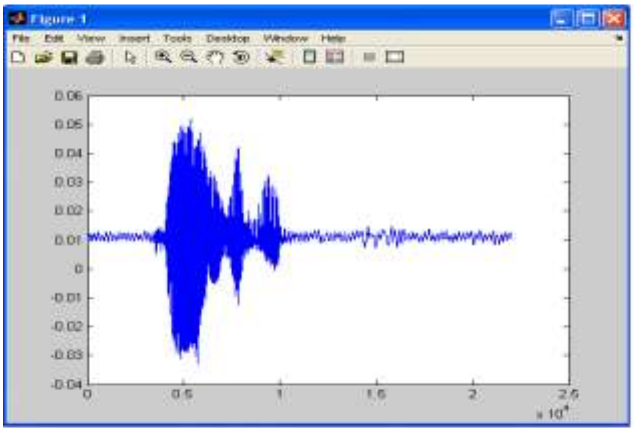

Fig. 11: the 5th model of the word “Hozayfa”

The Fig. 12 show the models of the voice recording for the word (Hozayfa) at the process of it’s feature Extracting using MFCC.

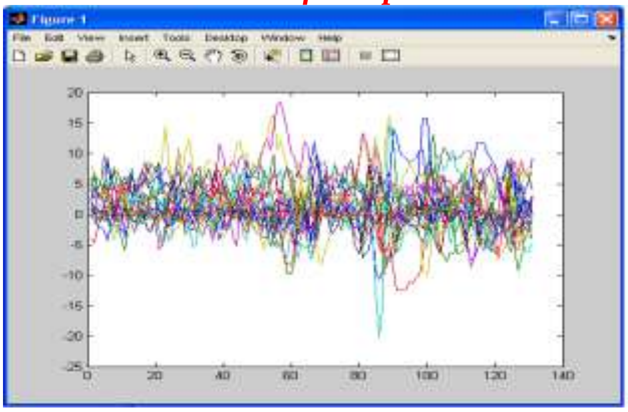

Fig. 12: model of the word “Hozayfa” after feature extraction

For training purpose was taken 5 voice samples from every person, to be compared each sample with the samples of trainee person for taking the sample that has highest match and more voice purity and save it in the voice database for the system. They has been taken model to 3 person from the training groups. The training process as in the following tables.

TABLE 1. SYSTEM TRAINING ON THE WORD “HOZAYFA”

Training sample	Matching Rate					Rate Average
	1	2	3	4	5	
1	100	94	91	95	93	94.6
2	94	100	96	93	95	95.6
3	91	96	100	97	94	95.6
4	95	95	97	100	93	96
5	93	95	94	93	100	95

From the table 1 above the sample number 4 achieved the highest rate of matching with 96%. For that has been selected within the system database to represent the word “Hozayfa”.

TABLE 2. SYSTEM TRAINING ON THE WORD “ARWA”

Training sample	Matching Rate					Rate Average
	1	2	3	4	5	
1	100	94	97	95	95	96.2
2	94	100	96	93	95	95.6
3	91	96	100	98	94	95.8
4	95	95	97	100	93	96
5	93	95	94	93	100	95

From the table 2 above the sample number 1 achieved the highest rate of matching with 96.2%. For that has been selected within the system database to represent the word “Arwa”.

TABLE 3. SYSTEM TRAINING ON THE WORD “LEMA”

Training sample	Matching Rate					Rate Average
	1	2	3	4	5	
1	100	94	91	95	93	94.6
2	94	100	96	93	92	95
3	91	93	100	97	94	95
4	95	95	97	100	93	96
5	96	95	98	97	100	97.2

From the table 3 above the sample number 5 achieved the highest rate of matching with 97.2%. For that has been selected within the system database to represent the word “Lema”.

VI. TEST AND RESULTS

This system has been tested for about 20 person by taking a voice sample for each of them and make matching process to verify the identity of the persons. The basic features that affect the system efficiency, and the accuracy identifying and verifying rate of the person identity is the number of MFCC coefficients and the recognition rate according to the number of MFCC coefficients. As shown in Table .4 and Fig. 13.

TABLE 4. NUMBER OF MFCC COEFFICIENTS VS IDENTIFICATION RATE

No. of MFC coefficients	ID rate (%)
5	76%
12	91%
20	91%

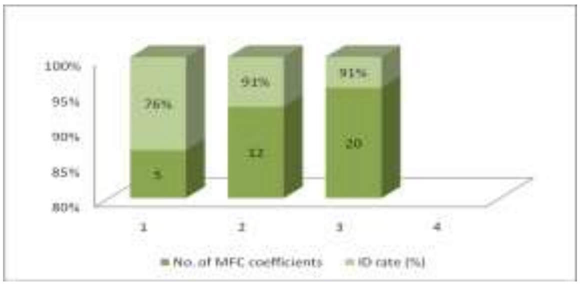

Fig. 13: Number of MFCC coefficients VS Identification rate

From the table and figure above, we notice increasing the number of frequencies in the MFCC coefficients affect positively in improving the rate of identification, but at the expense of overall time for the process of computation, usually MFCC coefficients in the range (12-15) to give the best results.

The effect of filters in the Filter-bank on the identity identification rate (increasing Filter-bank size). As shown in Table. 5 and Fig. 14.

TABLE 5. EFFECT OF FILTERS NUMBER ON IDENTIFICATION RATE EFFICIENCY

Number of Filter-banks	Identification Rate
12	18.75%
16	63.75%
20	90.625%
29	98.75%
36	99.375%
40	100%

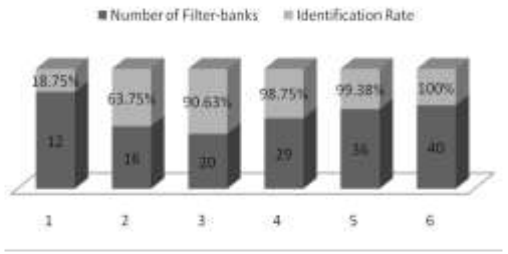

Fig. 14: Effect of filters number on improving identification accuracy

It is clear that increasing the number of filters in the filter-bank plays a key role in improving the recognition accuracy. From the table we can see that is possible to get 100% filtering rate by using 40 filters in the filter- bank, but that it needs to very high speed and largest time to complete the process.

TABLE 6. EFFECT OF CODE-BOOK SIZE VS IDENTIFICATION RATE

Code-book size	Identification rate (%)
2	84.375
8	95
16	98.75
64	98.75
256	98.75

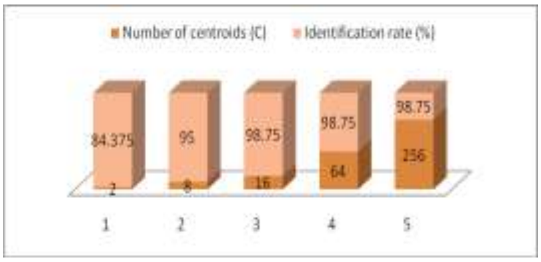

Fig. 15: Effect of code-book size on identification accuracy rate

From the table 6 and Fig. 15 above, It is clear that the increasing of the data increases the efficiency of identification rate but at the expense of the total time for the decision-making process.

VII. IDENTIFYING THE THRESHOLD

To identify the threshold value, has been tested 6 values for threshold by 20 sample, and according to the results it has been determined the threshold value used in this paper, as shown in the Table 7.

TABLE 7. IDENTIFYING THRESHOLD

Threshold Value	True Acceptance and rejection	False acceptance	False rejection	False acceptance rate	False rejection rate
90	16	4	0	%20	%0
91	16	3	1	%15	%5
92	17	1	2	%5	%10
93	19	0	1	%0	%5
94	18	0	2	%0	%10
95	15	0	5	%0	%25

According to the result in the above table have been selected the optimal value for the threshold 0.93. The system has tested 5 times on 20 samples, which represent the system database. The verification result in the range of 0 to 1.

A. *Result of the first experiment*

Was selected the voice signal "Omer" and compared with pre-recorded samples in the system as shown in the Table 8.

TABLE 8. FIRST EXPERIMENT OF THE WORD "OMER"

Speaker	Rate
Hozayfa	0.62
Ali	0.70
Omer	0.94
Ali	0.66
Sara	0.70
Sahar	0.49
Ithar	0.55
Abdulrahim	0.72
Mohammed	0.66
Salah	0.71
Abdelmanie	0.64
Lama	0.76
Arwa	0.82
Soaad	0.52
Madina	0.46
Mahmoud	0.71
Haytham	0.64
Mohanad	0.77
Osman	0.61
Amna	0.61

B. *Result of the second experiment*

Was selected the voice signal "Ahmed" and compared with pre-recorded samples in the system as shown in the Table 9.

TABLE 9. SECOND EXPERIMENT OF THE WORD "AHMED"

Speaker	Rate
Hozayfa	0.48
Omer	0.53
Salah	0.57

Ali	0.36
Sara	0.80
Sahar	0.83
Ithar	0.37
Abdulrahim	0.62
Mohammed	0.43
Ahmed	0.93
Abdelmanie	0.63
Lama	0.39
Arwa	0.40
Soaad	0.89
Madina	0.31
Mahmoud	0.64
Haytham	0.50
Mohanad	0.44
Osman	0.76
Amna	0.30

C. *Result of the third experiment*

Was selected the voice signal "Mohammed" and compared with pre-recorded samples in the system as shown in Table 10.

TABLE 10. THIRD EXPERIMENT OF THE WORD "MOHAMMED"

Speaker	Rate
Hozayfa	0.30
Omer	0.41
Ahmed	0.81
Ali	0.41
Sara	0.36
Sahar	0.29
Ithar	0.45
Abdulrahim	0.52
Mohammed	0.96
Salah	0.33
Abdelmanie	0.29
Lama	0.29
Arwa	0.36
Soaad	0.54
Madina	0.31
Mahmoud	0.90
Haytham	0.50
Mohanad	0.88
Osman	0.42
Amna	0.31

D. *Result of the fourth experiment*

Was selected the voice signal "Osman" and compared with pre-recorded samples in the system as shown in the Table 11.

TABLE 11. FOURTH EXPERIMENT OF THE WORD "OSMAN"

Speaker	Rate
Hozayfa	0.26
Omer	0.89
Ahmed	0.33
Ali	0.57
Sara	0.47
Sahar	0.44
Ithar	0.53
Abdulrahim	0.37
Mohammed	0.61
Salah	410.
Abdelmanie	220.
Lama	0.35
Arwa	0.58
Soaad	0.39
Madina	0.51
Mahmoud	0.61
Haytham	0.49
Mohanad	0.43
Osman	0.91
Amna	0.66

E. *Result of the fifth experiment*

Was selected the voice signal "Amna" and compared with pre-recorded samples in the system as shown in the Table 12.

TABLE 12. FIFTH EXPERIMENT OF THE WORD "AMNA"

Speaker	Rate
Hozayfa	0.65
Omer	0.42
Ahmed	0.39
Ali	0.41
Sara	0.25
Sahar	0.29
Ithar	0.39
Abdulrahim	0.76
Mohammed	0.71
Salah	440.
Abdelmanie	350.
Lama	0.74
Arwa	0.66
Soaad	0.23
Madina	0.79
Mahmoud	0.32
Amna	0.98
Mohanad	0.56
Osman	0.78
Haytham	0.59

VIII. EVALUATION

In reference to the previous results, we found that the system was identified to the speaker in 4 attempts of total 5 attempts. In addition, we find that in the case of did not identifying the speaker for not accessing acceptance threshold. Because it can DTW compute the similarities to the registered speakers and non-registered and accept the result according to the amount of the threshold used, which in turn determines the specific range for acceptance. Moreover, the threshold used in this system are 0.93 this indicate the range of acceptance group be limited from 0 to 1.

IX. CONCLUSION

The DTW is flexible mathematical method and give high accuracy results and represent it is possible to improve performance by selecting sources the carefully, since has a significant role in influencing in the identification accuracy. The use of comprehensive restrictions reduce the search area and decrease the necessary calculations to find a match, which lead to not occupancy a very big space during the execution.

The signal analysis by using MFCC provide us by spectrum coefficients. Which represents the accurate vocal system for the stored word, which any MFCC provide us with a high level of perception for the human voices, it works to remove all information that is not important, then give the best representation for the signal, which leads to a higher accuracy in the performance of recognition using this processing.

The recognition accuracy increases extrusive increase with duration of sample recording and number of the Filter bank. From that we can be conclude very important fact, that the spectrum coefficients has a high specification show their importance depending on the speaker himself and the method of producing the voice and pronunciation and style, which can be used in many applications such as security systems, where the voices of people differ as fingerprints differ. So the back propagation algorithm has been tested for the voices recognition, and has achieved a success rate about 95%.

REFERENCES

[1] Dr. Kavitha. R, Nachammai. N, Ranjani. R, Shifali. J," Speech Based Voice Recognition System for Natural Language Processing", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (4), 2014, ISSN: 0975-9646, pp. 5301-5305

[2] Debnath Bhattacharyya, Rahul Ranjan, Farkhod Alisherov A. and Minkyu Choi, "Biometric Authentication: A Review", International Journal of u- and e- Service, Science and Technology Vol. 2, No. 3, September, 2009.

[3] Judith A. Markowitz, "Voice Biometrics", September 2000/Vol. 43, No. 9, Communications of the ACM.

[4] Shumaila Iqbal, Tahira Mahboob and Malik Sikandar Hayat Khiyal, "Voice Recognition using HMM with MFCC for Secure ATM", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 6, No 3, November 2011, ISSN (Online): 1694-0814, pp. 297-303.

[5] N. Revathy, T.Guhan, "FACE RECOGNITION SYSTEM USING BACK PROPAGATION ARTIFICIAL NEURAL NETWORKS", International Journal of Advanced Engineering Technology E-ISSN 0976-3945, IJAET/Vol.III/ Issue I/January-March, 2012, pp. 321-324.

[6] Jayanta Kumar Basu, Debnath Bhattacharyya, Tai-hoon Kim2," Use of Artificial Neural Network in Pattern Recognition", International Journal of Software Engineering and Its Applications, Vol. 4, No. 2, April 2010, pp. 23-34.

[7] N.AYSHWARYA, G.LOGESHWARI, G.S.ANANDHA MALA, "FEED FORWARD BACK PROPAGATION NEURAL NETWORK FOR SPEAKER INDEPENDENT SPEECH RECOGNITION", International Journal of Industrial Electronics and Electrical Engineering, ISSN: 2347-6982 Volume-2, Issue-8, Aug.-2014, pp. 36-39.

[8] Siddhant C. Joshi, Dr. A.N.Cheeran, "MATLAB Based Back-Propagation Neural Network for Automatic Speech Recognition", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, ISSN (Online): 2278 – 8875, Vol. 3, Issue 7, July 2014, pp. 10498- 10504.

[9] Md. Ali Hossain, Md. Mijanur Rahman, Uzzal Kumar Prodhan, Md. Farukuzzaman Khan, "Implementation Of Back-Propagation Neural Network For Isolated Bangla Speech Recognition", International Journal of Information Sciences and Techniques (IJIST) Vol.3, No.4, July 2013, pp. 1-9.

[10] Lindasalwa Muda, Mumtaj Begam and I. Elamvazuthi," Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques", JOURNAL OF COMPUTING, VOLUME 2, ISSUE 3, MARCH 2010, ISSN 2151-9617, pp. 138 -143.

[11] Md. Akkas Ali, Manwar Hossain and Mohammad Nuruzzaman Bhuiyan," Automatic Speech Recognition Technique for Bangla Words", International Journal of Advanced Science and Technology Vol. 50, January, 2013, pp. 51-60.

[12] MarutiLimkar, RamaRao & VidyaSagvekar, "Isolated Digit Recognition Using MFCC AND DTW", International Journal on Advanced Electrical and Electronics Engineering, (IJAEEE), ISSN (Print): 2278-8948, Volume-1, Issue-1, 2012, pp. 59-64.

[13] M.A.Anusuya, S.K.Katti, Classification Techniques used in Speech Recognition Applications: A Review -– Int. J. Comp. Tech. Appl., July-August 2011 -Vol 2 (4), 910-954.

[14] Ms. Savitha and S Upadhya," Digit Recognizer Using Single and Average Template Matching Techniques", International Journal of Emerging Technologies in Computational and Applied Sciences, 3(3), Dec.12-Feb.13, ISSN (Online): 2279-0055, pp. 357-362.

[15] Om Prakash Prabhakar, Navneet Kumar Sahu," A Survey On: Voice Command Recognition Technique", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 5, May 2013 ISSN: 2277 128X, pp. 576-585.

[16] Sahil Verma, Tarun Gulati, Rohit Lamba," RECOGNIZING VOICE FOR NUMERICS USING MFCC AND DTW" International Journal of Application or Innovation in Engineering & Management (IJAIEM), Volume 2, Issue 5, May 2013 ISSN 2319 – 4847, pp. 127 -130.

[17] Nidhi Desai 1, Prof.Kinnal Dhameliya2, Prof. Vijayendra Desai, "Recognizing voice commands for robot using MFCC and DTW", International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 5, May 2014, ISSN (Online), 2278-1021, pp. 6456 6459.

AUTHOR'S PROFILE

Dr. Abdelmajid Hassan Mansour Emam, Assistant Professor, Department of Computers and Information Technology, University of Jeddah, Faculty of Computers and Information Technology, Khulais, Jeddah, Saudi Arabia.

Permanent Address: Department of Information Technology, Faculty of computer Science and Information Technology, Alneelain University, Khartoum, Sudan.

Dr. Gafar Zen Alabdeen Salh Hassan, Assistant Professor, Department of Computers and Information Technology, University of Jeddah, Faculty of Computers and Information Technology, Khulais, Jeddah, Saudi Arabia..

Permanent Address: Department of Information Technology, Faculty of computer Science and Information Technology, Alneelain University, Khartoum, Sudan.

Hozayfa Hayder Zeen Alabdeen Mohammed, Lecturer, Department of Information Technology, Faculty of computer Science and Information Technology, Alneelain University, Sudan.