iJETRM
International Journal of Engineering Technology Research & Management

PROCEDURES FOR SPEECH RECOGNITION USING LPC AND ANN

Dr. Amjad Hindi
Dr. Majed Omar Dwairi
Prof. Ziad Alqadi
Albalqa Applied University, Faculty of engineering technology, Jordan, Amman
amjadhindi@bau.edu.jo
majeddw@gmail.com
natalia_maw@yahoo.com

ABSTRACT
Digital speech recognition computer system can be used in various vital applications, it can be used in security system to identify the human and the spoken word or phrase, it can be used also in a control system to select an action to be executed depending on the spoken speech.
In this research paper we will introduce a procedures needed to build SRCS, this procedure will include speech features extraction based on linear prediction coding, creating and training ANN, Testing and running ANN.
Keywords: Speech, features, FIR, LPC, ANN, CFANN, features database, SRCS.

INTRODUCTION

Digital data such as digital images [1], [2], [3] and digital speeches [8], [9] are very important types of data at present. They are used in several applications [4], [5], such as human knowledge and many security and banking applications [6]. A digital speech signal is usually of a large size where the acoustic signal consists of a set of values distributed in one column (one channel) (mono signal) or distributed in two columns (two channels) (stereo signal), these values usually are the results of encoding of the operations sampling and quantization of the original analogue speech signal [7], [8], [9]]. Since the volume of the speech signal is large, [7] it is difficult to conduct the matching of two speeches using all the values, where the process of matching will require a large amount of time, which in turn leads to delay in the process of digital signal recognition [5-7].
To build an efficient Speech recognition computer system (SRCS), we have to pay attention on representing the signal by a set of small number of values which are called signal features [10].
Digital signal features extraction is an important task perform in SRCS [11], [12], [13] as shown in figure 1

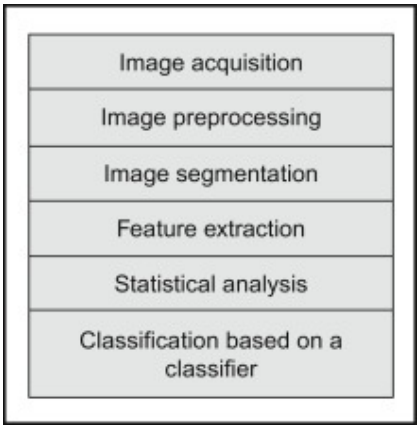

Figure 1: The need of features extraction

The signal extracted feature must require a small size to be stored, simple to process, unique , and must require small extraction time [14[, [15,[16].

iJETRM
International Journal of Engineering Technology Research & Management

Many methods are used to create digital signal features [17], many of them are based on local binary pattern (LBP) [25], [26] and its modifications such as CSLBP [19][26] , RLBP[22], [23], [24], MLBP[25], some of the introduced methods were based on data clustering [18], [19], [20].
Regardless the used method for feature extraction this task must provides the following:

1. Uniqueness of the features, so as we can use them as a signature to identify the signal [18].
2. Efficient by minimizing the extraction time [19].
3. Small size to reduce the storage size and the size of artificial neural network (ANN) used for classification purposes [20], [27], [28].
4. Simplicity to be able to process them.

In this research paper we will introduce a method which provides the above mentioned advantages and abilities, this method will based on linear prediction coding (LPC) [29] by applying a special matlab function lpc, processing this function will retrieve the coefficients of the finite impulse response (FIR) filter which uses a speech signal as an input signal [33], the retrieved coefficients then can be easily used as a signal features [30], [31], [32].

The proposed procedures to build SRCS

Creating and building SRCS can be done by applying the procedures described below:

2-1 Creating speech signal features

This step can be implemented as shown in figure 2 applying the following operations:

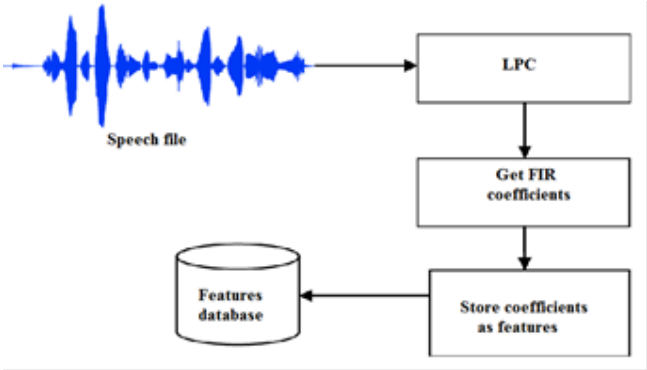

Figure 2: Features extraction phase

1. Get the speech signal.
2. Select FIR filter order.
3. Apply LPC using the selected filter order.
4. Add the retrieved coefficients as a speech features to the features data base.

The data base size will depend on the number of speech file and the number of selected FIR coefficients, for 1000 speech signal and using 4 orders FIR the database size will equal 1000*32=32 Kbytes.
This phase was implemented using 3 random signals and various orders FIR, the results of implementation are shown in table 1. From table 1 we can see the following:

5. Each signal has unique features.
6. FIR coefficients remain stable for different running.
7. Increasing the order of FIR filter will increase the number of features.
8. For various FIR orders the features remain unique for each signal.
9. Features extraction time is significantly small, and the extraction process throughput is very high.

iJETRM
International Journal of Engineering Technology Research & Management

Signal	FIR order	FIR coefficients (Features)							
X1	2	-0.4240				-0.4370			
X2	2	-0.4427				-0.4284			
X3	2	-0.4261				-0.4291			
X1	4	-0.2656		-0.2592		-0.1646		-0.2291	
X2	4	-0.2884		-0.2578		-0.1869		-0.1885	
X3	4	-0.2080		-0.2210		-0.2598		-0.2368	
X1	6	-0.1946		-0.1973	-0.0981	-0.1412	-0.1311	-0.1825	
X2	6	-0.2179		-0.1994	-0.1132	-0.0953	-0.1658	-0.1557	
X3	6	-0.1207		-0.1324	-0.1761	-0.1640	-0.1870	-0.1715	
X1	8	-0.1607	-0.1624	-0.0686	-0.1138	-0.0990	-0.1378	-0.0868	-0.1281
X2	8	-0.1819	-0.1632	-0.0852	-0.0712	-0.1278	-0.1069	-0.1153	-0.1079
X3	8	-0.0909	-0.0964	-0.1413	-0.1305	-0.1570	-0.1473	-0.1014	-0.0963
Average extraction time(seconds)	0.0014								
Throughput(Samples per second)	714290								

Artificial neural network training

Artificial neural network is a set of fully connected neurons [33], [34], and it is a powerful tool which can be used for classification. Several types of ANN are now used in various applications [35[, [36] as shown in figure 3, but in our research paper we will use cascade forward ANN, because it has a better performance [37], [38], [39].

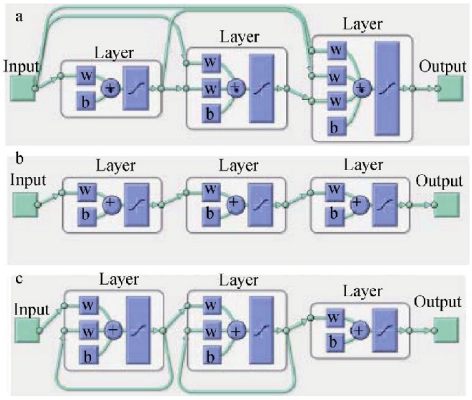

Figure 3 Topology (Architecture) of neural networks: a cascade-forward backpropagation network, b feed-forward back-propagation network, and c layer recurrent network

To train CFANN we have to follow the following:

1. Select the speech signal features data base, which is a 2D matrix, one column for each speech signal and 4 rows for the features as shown in figure 4:

iJETRM
International Journal of Engineering Technology Research & Management

```
Columns 1 through 8

 -0.9468   -0.9915   -0.9796   -0.9847   -0.8794   -0.9920   -0.9488   -0.9756
 -0.1671   -0.0417   -0.0867   -0.0664   -0.2851    0.0310   -0.1660   -0.1031
 -0.2640   -0.0899   -0.1662   -0.1370   -0.3999    0.0593   -0.2380   -0.1761
  0.3905    0.1305    0.2398    0.1969    0.5932   -0.0883    0.3663    0.2610

Columns 9 through 12

 -0.9774   -0.9885   -0.9463   -0.9690
 -0.0881   -0.0508   -0.1771   -0.1106
 -0.1803   -0.1175   -0.2695   -0.2021
  0.2552    0.1654    0.3993    0.2935
```

Features database

```
     1     2     3     4     5     6     7     8     9    10    11    12
```

Targets

Figure 4: Features database and targets

2. Select the targets, an index for each speech signal.
3. Create ANN by defining the number of layers, the number of neurons in each layer, and the activation function for each layer.
4. Initialize all the weights to zero.
5. Define the goal (error) which must be closed to zero.
6. Train ANN using the input database and the output targets.
7. If the error not acceptable retrain ANN or modify ANN architecture and repeat the training process as shown in figure 5.
8. Save ANN.

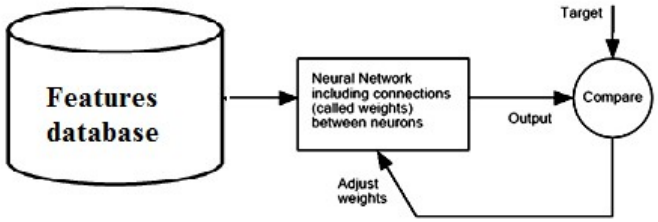

Figure 5: ANN training
ANN running

To use ANN as a SRCS we have to follow the following (See figure 6):

1. Load trained ANN.
2. Get the speech signal.
3. Apply LPC to find speech features.
4. Run ANN using the calculated features as an input.
5. According to the calculated classifier select execute the wanted action

iJETRM
International Journal of Engineering Technology Research & Management

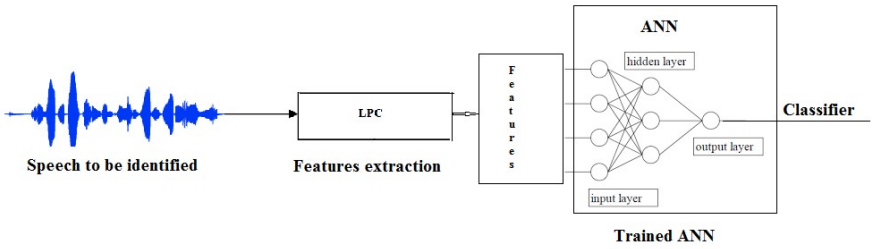

Figure 6: ANN running

Figure 7 summarises the stages required to build SRCS.

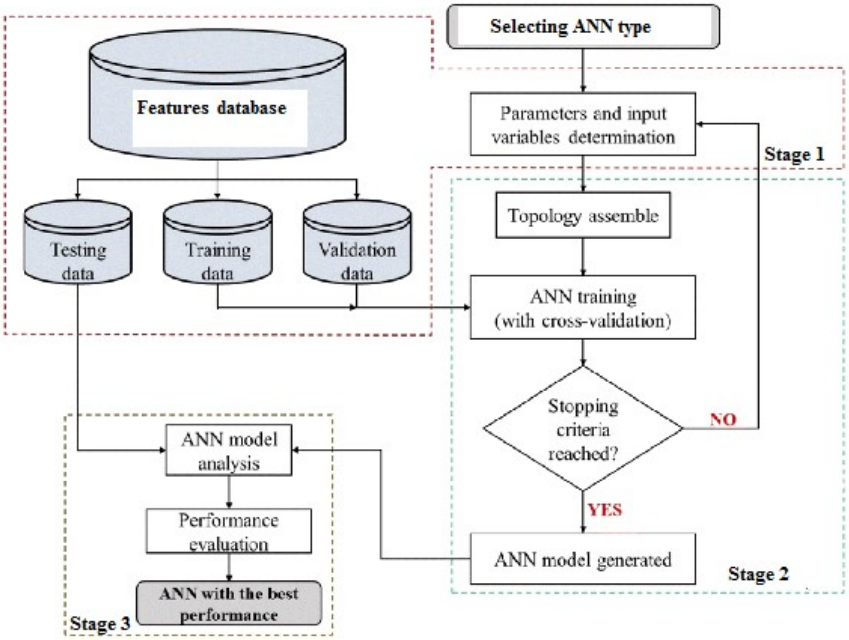

Figure 7: Stages to build SRCS.

Experimental results

iJETRM
International Journal of Engineering Technology Research & Management

Twelve speech signals were selected as shown in table 2, the proposed procedures were applied, and a special ANN was trained and saved.

Table 2: Speech signals features

Speech	Features			
Display image 1	-0.9468	-0.1671	-0.2640	0.3905
Display image 2	-0.9915	-0.0417	-0.0899	0.1305
Display image 3	-0.9796	-0.0867	-0.1662	0.2398
Display image 4	-0.9847	-0.0664	-0.1370	0.1969
Display image 5	-0.8794	-0.2851	-0.3999	0.5932
Display image 6	-0.9920	0.0310	0.0593	-0.0883
Display image 7	-0.9488	-0.1660	-0.2380	0.3663
Display image 8	-0.9756	-0.1031	-0.1761	0.2610
Display image 9	-0.9774	-0.0881	-0.1803	0.2552
Display image 10	-0.9885	-0.0508	-0.1175	0.1654
Display image 11	-0.9463	-0.1771	-0.2695	0.3993
Display image 12	-0.9690	-0.1106	-0.2021	0.2935

ANN then was used using each of the speech signal, after each run the output was correct, figure 8 shows one output when the selected speech was : Display image 12

Figure 8: Output of running ANN using the speech: Display image 12

Here we have to mention the following:

- ANN size = 28661 bytes.
- Features database size= 384 bytes.
- ANN type: CFANN.
- ANN architecture: 2 layers with 4neurons and 1.
- ANN running time in average =1.016000 seconds: this time includes the feature extraction time, ANN running time, and image display time.

Conclusion

A simple and efficient procedures to build SRCS was proposed, various stages of building and running SRCS were tested, the obtained results shows the accuracy of the system, and it provides a 100% recognition ratio.

The proposed SRCS can be expanded easily for any number of speech signals, it can be used to select an action depending on the spoken speech by identifying the person and the spoken word or phrase.

References

iJETRM
International Journal of Engineering Technology Research & Management

[1]Qazem Jaber Ziad Alqadi, Jamil azza, Statistical analysis of methods used to enhance color image histogram, XX International scientific and technical conference, 2017.

[2]Ziad Al-Qadi, Musbah Aqel, Performance analysis of parallel matrix multiplication algorithms used in image processing, World Applied Sciences Journal, vol. 6, issue 1, pp. 45-52, 2009.

[3] Majed O Al-Dwairi, Ziad A Alqadi, Amjad A Abujazar, Rushdi Abu Zneit, Optimized true-color image processing, World Applied Sciences Journal, vol. 8, issue 10, pp. 1175-1182, 2010.

[4]Jamil Al Azzeh, Hussein Alhatamleh, Ziad A Alqadi, Mohammad Khalil Abuzalata, Creating a Color Map to be used to Convert a Gray Image to Color Image, International Journal of Computer Applications, vol. 153. Issue 2, pp. 31-34, 2016.

[5]J Al-Azzeh M Abuzalata, Ziad Alqadi, Modified Inverse LSB Method for Highly Secure Message Hiding, International Journal of Computer Science and Mobile Computing, vol. 8, issue 2, pp. 93-103, 2019.

[6]Akram A Moustafa, Ziad A Alqadi, A Practical Approach of Selecting the Edge Detector Parameters to Achieve a Good Edge Map of the Gray Image, Journal of Computer Science 5 (5): 355-362, 2009.

[7]Jamil Al-Azzeh, Bilal Zahran, Ziad Alqadi, Belal Ayyoub, Muhammed Mesleh, A Novel Based On Image Blocking Method To Encrypt-Decrypt Color, JOIV: International Journal on Informatics Visualization, vol. 3, issue 1, pp. 86-93, 2019.

[8] Jihad Nader Ahmad Sharadqh , Ziad Al-Qadi, Bilal Zahran, Experimental Investigation of Wave File Compression-Decompression, International Journal of Computer Science and Information Security, vol. 14, issue 10, pp. 774-780, 2016.

[9]Ashraf Abu-Ein, Ziad AA Alqadi, Jihad Nader, A TECHNIQUE OF HIDING SECRETE TEXT IN WAVE FILE, International Journal of Computer Applications, pp. 1-4, 2016.

[9]K Matrouk, A Al-Hasanat, H Alasha'ary, Ziad Al-Qadi, H Al-Shalabi, Speech fingerprint to identify isolated word person, World Applied Sciences Journal, vol. 31, issue 10, pp. 1767-1771, 2014.

[10] Prof. Mohammed Abu Zalata, Dr. Ghazi. M. Qaryouti, Dr.Saleh Khawatreh, Prof. Ziad A.A. Alqadi, Optimal Color Image Recognition System (OCIRS), International Journal of Advanced Computer Science and Technology, vol. 7, issue 1, pp. 91-99, 2017.

[11] Yousif El-tous Dr. Ghazi M. Qaryouti, Mohammad Abuzalata, Ziad Al Qadi, A METHOD TO CREATE VOICE SIGNAL HISTOGRAM, IJRSR, vol. 10, issue 12, pp. 36650-36654, 2020.

[12] Majed O. Al-Dwairi, Amjad Y. Hendi, Mohamed S. Soliman, Ziad A.A. Alqadi, A new method for voice signal features creation, International Journal of Electrical and Computer Engineering (IJECE), vol. 9, issue 5, pp. 4092-4098, 2019.

[13]Bilal Zahran Belal Ayyoub, Jihad Nader, Ziad Al-Qadi, Suggested Method to Create Color Image Features Victor, Journal of Engineering and Applied Sciences, vol. 14, issue 1, pp. 2203-2207, 2019.

[14]Aws Al-Qaisi, Saleh A Khawatreh, Ahmad A Sharadqah, Ziad A Alqadi, Wave File Features Extraction Using Reduced LBP, International Journal of Electrical and Computer Engineering, vol. 8, issue 5, pp. 2780, 2018.

[15]Ziad Al-Qadi Ayman Al-Rawashdeh, Creating Color Image Features Using Local Contrast Method, Bulletin of Electrical Engineering and Informatics, vol. 7, issue 3, pp. 367-376, 2018.

[16]BELAL AYYOUB AND MAZEN ABU-ZAHER JAMIL AL-AZZEH, BILAL ZAHRAN, ZIAD ALQADI, A NOVEL ZERO-ERROR METHOD TO CREATE A SECRET TAG FOR AN IMAGE, Journal of Theoretical and Applied Information Technology, vol. 96, issue 13, pp. 4081-4091, 2018.

[17] Prof. Ziad Alqadi Dr. Ghazi M. Qaryouti, Prof. Mohammad Abuzalata, Prof. Yousf Eltous, **Comparative Study of Voice Signal Features Extraction Methods**, IOSR Journal of Computer Engineering (IOSR-JCE), vol. 22, issue 1, pp. 58-66, 2020.

[18] Prof. Yousf Eltous Prof. Ziad Alqadi, Dr. Ghazi M. Qaryouti, Prof. Mohammad Abuzalata, Enhancing Color Image Clustering using K-Means Method, International Journal of Advanced Research in Computer and Communication Engineering, vol. 9, issue 1, pp. 78-84, 2020.

[19] Ziad Alqad, Prof. Yousf Eltous Dr. Ghazi M. Qaryouti, Prof. Mohammad Abuzalata, Analysis of Digital Signal Features Extraction Based on LBP Operator, International Journal of Advanced Research in Computer and Communication Engineering, vol. 9, issue 1, pp. 1-7, 2020.

[20] Yousf Eltous Ziad A. AlQadi, Ghazi M. Qaryouti, Mohammad Abuzalata, ANALYSIS OF DIGITAL SIGNAL FEATURES EXTRACTION BASED ON KMEANS CLUSTERING, International Journal of Engineering Technology Research & Management, vol. 4, issue 1, pp. 66-75, 2020.

iJETRM
International Journal of Engineering Technology Research & Management

[21] Prof. Yousif Eltous, Dr. Ghazi M. Qaryouti, Prof. Mohammad Abuzalata, Prof. Ziad Alqadi, Evaluation of Fuzzy and C_mean Clustering Methods used to Generate Voiceprint, IJCSMC, vol. 9, issue 1, pp. 75 -83, 2020.

[22]Jihad Nader Ismail Shayeb, Ziad Alqadi, Analysis of digital voice features extraction methods, International Journal of Educational Research and Development, vol. 1, issue 4, pp. 49-55, 2019.

[23] Ahmad Sharadqh Naseem Asad, Ismail Shayeb, Qazem Jaber, Belal Ayyoub, Ziad Alqadi, Creating a Stable and Fixed Features Array for Digital Color Image, IJCSMC, vol.8, issue 8, pp. 50-56, 2019.

[24] Ahmad Sharadqh Jamil Al-Azzeh , Rashad Rasras , Ziad Alqadi , Belal Ayyoub, Adaptation of matlab K-means clustering function to create Color Image Features, International Journal of Research in Advanced Engineering and Technology, vol. 5, issue 2, pp. 10-18, 2019.

[25]Ayman Al-Rawashdeh, Ziad Al-Qadi, Using wave equation to extract digital signal features, Engineering, Technology & Applied Science Research, vol. 8, issue 4, pp. 1356-1359, 2018.

[26]ZIAD ALQADI, A MODIFIED LBP METHOD TO EXTRACT FEATURES FROM COLOR IMAGES, Journal of Theoretical and Applied Information Technology, vol 96, issue 10, pp. 3014-3024, 2018.

[27] Eng. Sameh S. Salma Prof. Ziad A. AlQadi, Eng. Ahmad S. AlOthma , Eng. Mahmoud O. Alleddaw ,Eng. Mohammad Ali Al-Hiar , Eng. Osama T. Ghaza, Investigation of ANN Used to Solve MLR Problem, IJCSMC, vol. 8, issue 5, pp. 38 – 50, 2019.

[28]Eng. Sameh S. Salman Prof. Ziad A. AlQadi, Eng. Ahmad S. AlOthman, Eng. Mahmoud O. Alleddawi, Eng. Mohammad AlHiary, Eng. Osama T. Ghazal, Building Accurate and Efficient Color Image Recognizer, IJCSMC, vol. 8, issue 4, pp. 127 – 135, 2019.

[29] A. K. Paul, D. Das, M. M. Kamal, "Bangla Speech Recognition System Using LPC and ANN", Seventh International Conference on Advances in Pattern Recognition (ICAPR), Kolkata, India, pp. 171– 174, February 4-9, 2009.

[30] Ziad Alqadi, Bilal Zahran, Jihad Nader, Estimation and Tuning of FIR Low pass Digital Filter Parameters, international Journal of Advanced Research in Computer Science and Software Engineering, v. 7, issue 2, pp. 18-23, 2017.

[31] Qazem Jaber Ziad Alqadi, Jamil azza, Statistical analysis of methods used to enhance color image histogram, XX International scientific and technical conference, 2017.

[32] Jihad Nader, Ziad A. A. Alqadi, Bilal Zahran, Analysis of Color Image Filtering Methods, International Journal of Computer Applications, v. 174, issue 8, pp. 12-17, 2017.

[33]Jamil Al-azzeh Belal Ayyoub, Ahmad Sharadqh, Ziad Alqadi, Simulink based RNN models to solve LPM, International Journal of Research in Advanced Engineering and Technology, vol. 5, issue 1, pp. 49-55, 2019.

[34]Jamil Al-Azzeh, Ziad Alqadi, Mohammed Abuzalata, Performance Analysis of Artificial Neural Networks used for Color Image Recognition and Retrieving, International Journal of Computer Science and Mobile Computing, vol. 8, issue 2, 2019.

[35] Prof. Mohammed K. Abu Zalata Dr. Ghazi, M. Qaryouti, Prof. Ziad A.A. Alqadi, A Novel Method for Color Image Recognition, International Journal of Computer Science and Mobile Computing, vol. 5, issue 11, pp. 57 – 64, 2016.

[36] Khaled M Matrouk, Haitham A Alasha'ary, Abdullah I Al-Hasanat, Ziad A Al-Qadi, Hasan M Al-Shalabi, Investigation and Analysis of ANN Parameters, European Journal of Scientific Research, vol. 121, issue 2, pp. 217-225, 2014.

[37] Abdullah Al-Hasanat, Haitham Alasha'ary, Khaled Matrouk, Ziad Al-Qadi, Hasan Al-Shalabi, Experimental Investigations of Training Algorithms used in Backpropagation Artificial Neural Networks to Apply Curve Fitting, European Journal of Scientific Research, vol. 121, issue 4, pp. 328-335, 2014.

[38]Akram A Moustafa, Ziad A Alqadi, Eyad A Shahroury, Performance evaluation of artificial neural networks for spatial data analysis, WSEAS Transactions on Computers, vol. 10, issue 4, pp. 115-124. 2011.

[39]Ziad AlQadi, Yehya Abded Allatif, Musbah J Aqel, A Proposed methodology for image objects recognition using artificial neural networks, IJCSS, Vol.3, No.1, pp. 49-56, 2011.