International Journal of Electrical and Computer Engineering (IJECE)
Vol. 11, No. 2, April 2021, pp. 1311~1318
ISSN: 2088-8708, DOI: 10.11591/ijece.v11i2.pp1311-1318

A flexible method to create wave file features

Ghazi M. J. Qaryouti, Tariq M. Younes
Department of Mechatronics Engineering, Faculty of Engineering Technology, Al-Balqa Applied University, Jordan

Article Info

Article history:

Received Apr 10, 2020
Revised Jul 27, 2020
Accepted Oct 27, 2020

Keywords:

Clustering
Crest factor
Dynamic range
Histogram
K-mean clustering

ABSTRACT

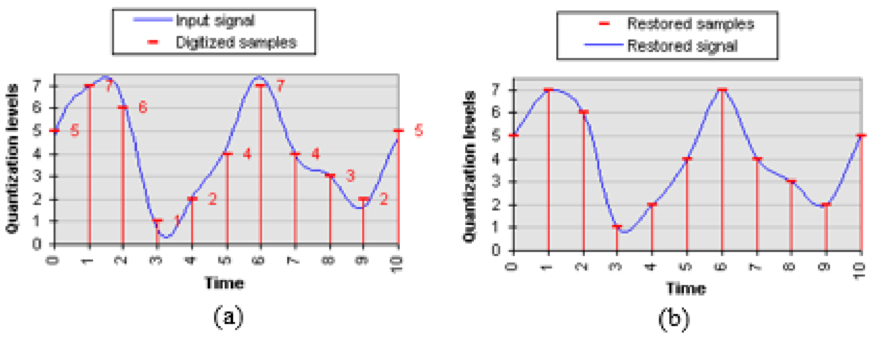
Digital audio signal is one of the most important data type at present, it is used in various vital applications, such as human knowledge, security and banking applications, most applications require signal identification and recognition, and to increase the efficiency of these applications we must seek a method to represent the audio file by a small set of values called a features vector. In this paper research we will introduce an enhanced method of features extraction based on k-mean clustering. The method will be tested and implemented to show how the proposed method can reduce the efforts of voice identification, and can minimize the recognition time a set of voice extracted features must be used instead of using the voice wave file.

Corresponding Author:

Ghazi M. J. Qaryouti
Department of Mechatronics Engineering
Faculty of Engineering Technology
Al-Balqa, Applied University, Jordan
Email: ghazi_qaryouti@bau.edu.jo

1. INTRODUCTION

Digital audio signal is one of the most popular types of digital data utilized in a variety of bioengineering applications. Many of these applications depend on voice tags to distinguish different sounds and words and different people [1, 2]. The analog audio signals is captured using special units, where this signals is passed over the analog to digital converter. As shown in Figure 1, sampling, quantization and encoding analysis are carried out through this unit.



Level	Code word
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111

(c)

Figure 1. Analog to digital converting process: (a) Sampling, (b) quantization, (c) encoding analysis

Digital audio files have large sizes [3, 4], due to the multiplicity of audio samples taken in the specified period of time and due to the possibility of the audio file being mono or stereo [5-7]. Which is organized in one or two columns, which will increase the size of the file. Therefore, comparing audio files in the form of samples will take a long time [8, 9], as shown in Table 1.

Table 1. Matching time for different wave files

Wave file	Type	Channel size	Total size (bytes)	Matching time (Seconds)
1	Stereo	36787	294296	0.006
2	Stereo	39730	317840	0.008
3	Stereo	33844	270752	0.0076
4	Stereo	17658	141264	0.005
5	Stereo	41202	329616	0.0079
6	Stereo	36787	294296	0.006
7	Stereo	63274	506192	0.014
8	Stereo	48049	384392	0.01
9	Stereo	55916	447328	0.013
10	Stereo	89760	718080	0.018
Average		46300	370405	0.0097
Cost of 1 sample			9700/370405=0.0262 microseconds/byte	

Based on the foregoing, the importance of searching for a method to represent the audio file with a set of values that can be used as a depressor for the audio file becomes an urgent task [10-12]. Many researchers have developed a group of methods based on the use of sound suppressors by calculating some of the factors specific to the sound such as the crest factor [13], dynamic range [14-16], and using them as a lowering devices [17-19].

2. WAVE FILE HISTOGRAM

Data histogram [20-23] is an array of elements, each of which points to the repetition of one value in the data set [24-27]. Calculating the wave file histogram is an initial task of the proposed later in this paper method of features extraction. The wave file histogram can be calculated using the following MATLAB function. First we have to set the size of the histogram, here we use (1):

$$nBins = round(datasetlength/15) \tag{1}$$

Then we start arranging the wave file values, by calculation the repetition of each value, saving this repetition in the corresponding index of the histogram. Figure 2 shows the calculated histogram of a wave file example:

```
function [Hn,Xn] = WAV_Hist(Data)
nBins = round(length(Data) / 15)
[H,X] = hist(Data, nBins);[Max,IMax] = max(H);sumH = sum(H);
curSum = 0.0;T = 0.990;i1 = IMax;i2 = IMax;I1 = i1;I2 = i2;
stop1 = 0;stop2 = 0;
while ((i1>=1) || (i2<=length(X)))
    if (i1~=i2)
        if (stop1~=1) curSum = curSum + H(i1); end
        if (stop2~=1) curSum = curSum + H(i2); end
    else          curSum = curSum + H(i1);
    end
    if (curSum <= sumH * T)          I1 = i1; I2 = i2;
        if (i1>1)     i1 = i1 - 1;
        else          stop1 = 1;
        end
        if (i2<length(X))  i2 = i2 + 1;
        else               stop2 = 1;     end
    else          break;      end
end
Xn = X(I1:I2);Hn = H(I1:I2);
```

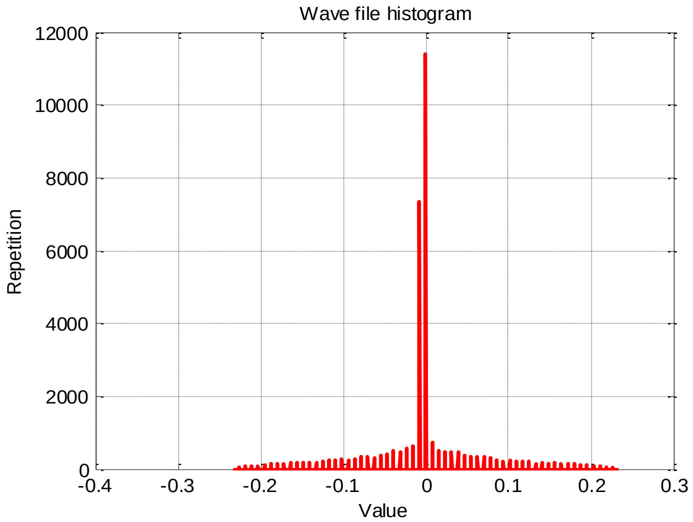


Figure 2. Wave file histogram (example)

3. K-MEAN CLUSTERING

Clustering means arranging data set values in groups (clusters), then the sums of values in each cluster, or the number of points in each cluster can be used as features for the data set [22]. K-mean clustering is implemented by applying a set of procedures which can be explained by the following example:

- Initialization:
  Here we have to select the data set, number of clusters, and the centroid of each cluster:
  Data set = 15, 15, 16, 19, 19, 20, 20, 21, 22, 28, 35, 40, 41, 42, 43, 44, 60, 61, 65
  Clusters=2;
  C1=16, C2=22
- Perform the following tasks while centroid changing:
  - Find distances to each cluster by taking the absolute value of the deference between the data item and the cluster centroid.
  - Select the cluster to which the data item belongs by selecting the nearest cluster depending on the distance.
  - Calculate the new centroid by averaging the data items belong to the cluster.

Table 2 and Table 3 shows the results of calculations:

Table 2. Calculation results of passes 1 and 2

	Pass 1(Centroid 16 and 22)			Pass 2		
x_i	Distance 1 $\mid x_i - c_1 \mid$	Distance 2 $\mid x_i - c_2 \mid$	Nearest Cluster and new Centroid 15.33 and 36.25 15.33 =(15+15++16)/3	Distance 1 $\mid x_i - c_1 \mid$	Distance 2 $\mid x_i - c_2 \mid$	Nearest Cluster and new Centroid 18.56and 45.9;
15	1	7	1	0.33	21.25	1
15	1	7	1	0.33	21.25	1
16	0	6	1	0.67	20.25	1
19	3	3	2	3.67	17.25	1
19	3	3	2	3.67	17.25	1
20	4	2	2	4.67	16.25	1
20	4	2	2	4.67	16.25	1
21	5	1	2	5.67	15.25	1
22	6	0	2	6.67	14.25	1
28	12	6	2	12.67	8.25	2
35	19	13	2	19.67	1.25	2
40	24	18	2	24.67	3.75	2
41	25	19	2	25.67	4.75	2
42	26	20	2	26.67	5.75	2
43	27	21	2	27.67	6.75	2
44	28	22	2	28.67	7.75	2
60	44	38	2	44.67	23.75	2
61	45	39	2	45.67	24.75	2
65	49	43	2	49.67	28.75	2

Table 3. Calculation results of passes 3 and 4

x_i	Pass **3** Distance 1 $\lvert x_i - c_1 \rvert$	Distance 2 $\lvert x_i - c_2 \rvert$	Nearest Cluster and new Centroid 19.50and 47.89	Pass **4** Distance 1 $\lvert x_i - c_1 \rvert$	Distance 2 $\lvert x_i - c_2 \rvert$	Nearest Cluster and new Centroid 19.50and 47.89
15	3.56	30.9	1	4.50	32.89	1
15	3.56	30.9	1	4.50	32.89	1
16	2.56	29.9	1	3.50	31.89	1
19	0.44	26.9	1	0.50	28.89	1
19	0.44	26.9	1	0.50	28.89	1
20	1.44	25.9	1	0.50	27.89	1
20	1.44	25.9	1	0.50	27.89	1
21	2.44	24.9	1	1.50	26.89	1
22	3.44	23.9	1	2.50	25.89	1
28	9.44	17.9	1	8.50	19.89	1
35	16.44	10.9	2	15.50	12.89	2
40	21.44	5.9	2	20.50	7.89	2
41	22.44	4.9	2	21.50	6.89	2
42	23.44	3.9	2	22.50	5.89	2
43	24.44	2.9	2	23.50	4.89	2
44	25.44	1.9	2	24.50	3.89	2
60	41.44	14.1	2	40.50	12.11	2
61	42.44	15.1	2	41.50	13.11	2
65	46.44	19.1	2	45.50	17.11	2

4. THE PROPOSED METHOD

The proposed method of wave file features extraction is based on k-mean clustering and it can be implemented applying the following steps:

- Get the wave file.
- Calculate the wave file histogram to be used as an input data set for clustering.
- Initialization by selecting the number of clusters and a centroid for each cluster.
- Apply k-mean clustering.
- Save the clusters as a feature for the wave file.

4.1. Implementation and experimental results

A necessary MATLAB codes were written to create a features for a wav files using statistical method and k-mean method, below we will discuss the obtained experimental results.

4.1.1. Statistical method

a. Experiment 1

We took a sinusoidal signal and for deferent parameter values (amplitude, frequency and phase shifting) we calculate some statistical parameters, Table 4 shows the results of this experiment. From Table 4 we can see:

- Changing the signal parameters leads to changing the features set.
- Changing the features set means that the modified signal will be considered as a new signal thus will increase the memory space required to store the signals, and increase the required time for signal identification.

Table 4. Experiment 1 results

Signal	Features set (x=-4pi:0.001:4pi) Dynamic range (db)	Crest factor (db)	Mean	RMS value
1) Y1=sin(10x+5)	80.7663	3.0098	-9.8742e-006	0.70712
2) Y2=5sin(10x+5)	68.0299	0.38655	-2.3672e-005	0.95644
3) Y3=5sin(20x+5)	76.3293	0.38656	-2.3711e-005	0.95644
4) Y4=5sin(20x+15)	76.3293	0.38655	-3.038e-006	0.95644

b. Experiment 2

Here we took the first version of the digital signal, and used it to create wave file with deferent sampling frequencies, Table 5 shows the results of this experiment. From the results shown in Table 5 we can see that the features set remain the same for the same wave file recorded with deferent sampling frequencies, which mean that all the wave file versions can be considered as one file with a stable set of features.

Table 5. Experiment 2 results

Sampling frequency	Features Y1=sin (10x+5)			
	Dynamic range(db)	Crest factor(db)	Mean	RMS value
1000	80.7663	3.0098	-9.8742e-006	0.70712
1500	80.7663	3.0098	-9.8742e-006	0.70712
2000	80.7663	3.0098	-9.8742e-006	0.70712
2500	80.7663	3.0098	-9.8742e-006	0.70712
3000	80.7663	3.0098	-9.8742e-006	0.70712

c. Experiment 3

Statistical method of wave file features extraction was implemented using various wave files, Table 6 shows the results of this experiment. From the results shown in Table 6 we can see that statistical method is good for wave file features extraction, each wave file has a unique features set, which can be used as a signature or a key to identify or recognize the wave file.

Table 6. Experiment 3 results

Wav file	Features set			
	Dynamic range(db)	Crest factor(db)	Mean	RMS value
bird	32.0412	12.7755	-0.0033595	0.071792
bear_growl_y	42.0761	11.737	-0.039608	0.25689
bird_caw1	42.0761	15.8556	-0.0091737	0.15989
bird_caw2	42.0761	14.6131	-0.0093481	0.18448
bird_chirp	42.0761	13.5213	-0.0092506	0.20918
bird_chirping2	42.0761	14.0603	-0.0036374	0.1966
bison	42.0761	11.1476	-0.006455	0.27493
cat_big_x	42.0761	14.0567	0.0010605	0.19668
cat_fight	42.0761	10.935	-0.043354	0.28174
chicken	41.1381	18.2011	-0.0089245	0.10956
cow1	42.0761	8.8379	-0.0044406	0.35867
dog_x	42.0761	11.3235	-0.0031504	0.26941

4.1.2. Proposed k-mean of features extraction

a. Experiment 4

We took a sinusoidal signal and for deferent parameter values (amplitude, frequency and phase shifting), then we implemented k-mean method. Table 7 shows the results of this experiment.

Table 7. Experiment 4 results

Signal	Features(x=-4pi:0.001:4pi) Clusters=4			
Y1=sin(10x+5)	213.1222	530.0000	795.0000	47.3315
Y2=5sin(10x+5)	213.1222	530.0000	795.0000	47.3315
Y3=5sin(20x+5)	213.1222	530.0000	795.0000	47.3315
Y4=5sin(20x+15)	213.1222	530.0000	795.0000	47.3315

From Table 7 we can see:

- Changing the signal parameters does not lead to changing the features set.
- Changing the features set means that the modified signal will be considered as the new same signal thus this will not affect the memory space and the recognition time.

b. Experiment 5

Here we took the first version of the digital signal, and used it to create wave file with deferent sampling frequencies, Table 8 and Table 9 shows the results of this experiment. From the results shown in Table 8 and Table 9 we can see that the features set remain the same for the same wave file recorded with deferent sampling frequencies, which mean that all the wave file versions can be considered as one file with a stable set of features.

Table 8. Experiment 5-1 results

Sampling frequency	Features set			
1000	1374	3348	5022	316.9
1500	1374	3348	5022	316.9
2000	1374	3348	5022	316.9
2500	1374	3348	5022	316.9
3000	1374	3348	5022	316.9

Table 9. Experiment 5-2 results

Sampling frequency	Features set(*1.0e+003)			
	Bird wave file			
11025	1.2416	3.3080	4.9620	0.7537
12000	1.2416	3.3080	4.9620	0.7537
10000	1.2416	3.3080	4.9620	0.7537
8500	1.2416	3.3080	4.9620	0.7537
6000	1.2416	3.3080	4.9620	0.7537
	Bison wave file			
15000	1.5593	4.3880	6.5820	0.9327
12000	1.5593	4.3880	6.5820	0.9327
11025	1.5593	4.3880	6.5820	0.9327
10000	1.5593	4.3880	6.5820	0.9327
6000	1.5593	4.3880	6.5820	0.9327
3000	1.5593	4.3880	6.5820	0.9327
1000	1.5593	4.3880	6.5820	0.9327

c. Experiment 6

K-mean method of wave file features extraction was implemented using various wave files, Table 10 shows the results of this experiment. From the results shown in Table 10 we can see that k-mean method is good for wave file features extraction, each wave file has a unique features set, which can be used as a signature or a key to identify or recognize the wave file.

Table 10. Experiment 6 results

Wav file	Features set(*1.0e+003)			
bird	1.2416	3.3080	4.9620	0.7537
bear_growl_y	1.7797	5.3800	8.0700	1.1491
bird_caw1	1.1913	3.3080	4.9620	0.7665
bird_caw2	1.5103	4.2960	6.4440	1.0081
bird_chirp	0.8042	2.2280	3.3420	0.5109
bird_chirping2	0.8007	2.2312	3.3468	0.5261
bison	1.5593	4.3880	6.5820	0.9327
cat_big_x	1.5197	4.4880	6.7320	1.0143
cat_fight	1.1934	3.7468	5.6202	0.7035
chicken	0.3662	1.0000	1.5000	0.2420
cow1	0.6767	1.8120	2.7180	0.3821
dog_x	0.7145	1.8352	2.7528	0.4367

d. Experiment 7

Here we took the bird.wav wave file, and then we applied k-mean method of features extraction using the original file, amplified version of the file, amplified with addition version of the file, here the features remain the same without any changes as shown in Table 11.

Table 11. Experiment 7 results

Wave file	Features set(*1.0e+003)			
	Bird wave file			
X=bird	1.2416	3.3080	4.9620	0.7537
X=4X	1.2416	3.3080	4.9620	0.7537
X=X+5	1.2416	3.3080	4.9620	0.7537
X=X-10	1.2416	3.3080	4.9620	0.7537
X=2X+7	1.2416	3.3080	4.9620	0.7537

As a conclusion of these experiments we can summarize the advantages of k-mean method of features extraction comparing with statistical method as shown in Table 12, and from this table we can see that k-mean method is more flexible especially when dealing with deferent versions of the original wave file.

Table 12. Comparisons between k-mean and statistical methods

Wave file	K-mean method	Statistical method
Original	Good	Good
Modified parameters	Good	Bad
Various sampling frequencies	Good	Good
Amplified	Good	Bad
Attenuated	Good	Bad

5. CONCLUSION

Experimental investigations of statistical and k-mean methods of wave file features extraction were proposed. Experimental results showed that k-mean method is more flexible by maintaining a stable set of features for the original wave file and other modified versions, which leads to minimizing the memory space and the required processing time needed for voice identification or recognition.

REFERENCES

[1] Brijesh N. S. and Jaymin K. B., "Comparative analysis and implementation of structured edge active contour," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 2, pp. 1842-1848. 2020.

[2] Barker J., et al., "The third 'CHiME' speech separation and recognition challenge: Dataset, task and baselines," *IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*, 2015, pp. 504-511.

[3] S. K. Dastoor, "Comparative Analysis of Steganographic Algorithms impacting the information in the Speech Signal for enhancing the Message Security in next Generation Mobile devices," *World Congress on Information and Communication Technologies*, pp. 279-284, 2011.

[4] Bin, et al., "Dynamic range estimation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 25, no. 9, pp. 1618-1636, 2006.

[5] B. Wu, et al., "An analytical approach for dynamic range estimation," *Proceedings of the 41st annual Design Automation Conference*, 2004, pp. 472-477.

[6] A. Marathe, P. Jain, and V. Vyas, "Iterative improved learning algorithm for petrographic image classification accuracy enhancement," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 1, pp. 289-296, 2019.

[7] F. Gouyon, et al., "Classifying percussive sounds: a matter of zero-crossing rate in Proceedings of the COST G-6," *Conference on Digital Audio Effects (DAFX-00)*, 2000.

[8] D. D. Jaslene, "Feature Selection and Extraction of Audio," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 5, no. 3, pp. 3148-3155, 2016.

[9] Sanjivani S., et al., "An Overview of Technical Progress in Speech Recognition," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 3, pp. 488-497, 2013.

[10] M. I. Ouloul, et al., "An Efficient Face Recognition Using SIFT Descriptor in RGB-D Images," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 5, no. 6, pp. 1227-1233, 2015.

[11] S. Narang and D. Gupta, "Speech Feature Extraction Techniques: A Review," *International Journal of Computer Science and Mobile Computing*, vol. 4, no. 3, pp. 107-114, 2015.

[12] M. J. Fadhil, R. A. Fayadh, and M. K. Wali, "Design and implementation a prototype system for fusion image by using SWT-PCA algorithm with FPGA technique," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 1, pp. 757-766, 2020.

[13] Rybach, et al., "The RWTH Aachen University Open Source Speech Recognition system," *Tenth Annual Conference of the International Speech Communication Association*, 2009.

[14] K. Matrouk, Ziad Alqadi, et al., "Speech Fingerprint to Identify Isolated Word-Person," *World Applied Sciences Journal*, vol. 31, no. 10, pp. 1767-1771, 2014.

[15] S. Khawatreh, Ziad Alqadi et al., "A Novel Methodology to Extract Voice Signal Features," *International Journal of Computer Applications*, vol. 179, no. 9, pp. 40-43, 2018.

[16] E. Benmalek, et al., "Voice Assessments for Detecting Patients with Parkinson's Diseases in Different Stages," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 6, pp. 4265-4271, 2018.

[17] V. Jaiswal, V. Sharma, and S. Varma, "MMFO: modified moth flame optimization algorithm for region based RGB color image segmentation," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 1, pp, 196-201, 2020.

[18] F. P. George, et al., "Recognition of emotional states using EEG signals based on time-frequency analysis and SVM classifier," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 2, pp. 1012-1020, 2019.

[19] Majed O. Al-Dwairi, et al, "A new method for voice signal features creation," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 5, pp. 4092-4098, 2019.

[20] Ayman Al-Rawashdeh, Ziad Al-Qadi, "Using wave equation to extract digital signal features," *Engineering, Technology & Applied Science Research,* vol. 8, no. 4, pp. 1356-1359, 2018.
[21] J Al Azzeh, Z Alqadi, and Qazem M Jabber, "Statistical Analysis of Methods Used to Enhanced color Image Histogram," *XX International Scientific and Technical Conference,* 2017, pp. 8-15.
[22] Tareq Khan, "Smart Microwave Oven with Image Classification and Temperature Recommendation Algorithm," *International Journal of Electrical and Computer Engineering (IJECE),* vol. 8, no. 6, pp. 4239-4252, 2018.
[23] Jihad Nadir, Ashraf Abu Ein, and Ziad Alqadi, "A Technique to Encrypt-decrypt Stereo Wave File," *International Journal of Computer and Information Technology,* vol. 5, no. 5, pp. 465-470, 2016.
[24] Ashraf Abu-Ein, Ziad AA Alqadi, Jihad Nader, "A technique of hiding secrete text in wave file," *International Journal of Computer Applications,* pp. 0975-8887, 2016.
[25] A. S. Jamil Al-Azzeh, et al., "Adaptation of matlab K-means clustering function to create color image features," *International Journal of Research in Advanced Engineering and Technology,* vol. 5, no. 2, pp. 10-18, 2019.
[26] M. Q. Shatnawi, M. Alrousan, S. Amareen, "A new approach for content-based image retrieval for medical applications using low-level image descriptors," *International Journal of Electrical and Computer Engineering (IJECE),* vol. 10, no. 4, pp. 4363-4371, 2020.
[27] Vijaya S. M. and Suresh K, "An efficient design approach of ROI based DWT using vedic and wallace tree multiplier on FPGA platform," *International Journal of Electrical and Computer Engineering (IJECE),* vol. 9, no. 4, pp. 2433-2442, 2019.

BIOGRAPHY OF AUTHOR

Ghazi M. J. Qaryouti received his B.Sc. in electrical Engineering from Damascus university in 1982, and M.S in Mechatronics Engineering from Al-Balqa Applied University Jordan in 2006, and Ph.D. degrees in Mechatronics Engineering from De Montfort University in Leicester, UK in 2015 His research interests are in Power Electronics and Electric Machine Drives, CNC Machine, 3D Printer, Energy Efficiency and Management, Renewable System, Image processing, and signals conditioning. Email : ghazi_qaryouti@bau.edu.jo

Tariq M. Younes was born in Irbid, Jordan, in 1974. He received the bachelor and Master Degrees in Biotechnical and medical apparatus and systems Engineering from Tver State Technical University (Russian Federation) in 1998, and Ph. D degree in Instrumentation and measurement from Moscow State University of Environmental Engineering in 2004. In 2004, he joined the Department of Computer Engineering, Faculty of Engineering Technology Al Balqa Applied University, as a part time Lecturer, and in 2005 he became a full-time lecturer at the same department. Since September 2007, he has been with the Department of Mechatronics Engineering at the same university, where he was an Assistant Professor, and he became an Associate Professor in 2012