

Using color image to encrypt-decrypt wave file

Anwar Abadi¹, Prof. Ziad A. Alqadi²

Computer Engineering Department, Faculty of Engineering Technology, Albalqa Applied University^{1,2}

Abstract: Digital audio files are one of the important types of data to be used in many vital applications, which leads us to protect this type of data from intruders and snoopers, as it may contain confidential information that not allowed to be viewed by any unauthorized third party . In this paper, we present a new method for protecting audio files that depends on the use of the digital color image or part of it as a key to the encryption or decryption process. The digital image used as a key is characterized by its enormous size, in addition to that, the image is not circulated through social media, but rather is agreed upon between the speaker and the listener, which leads to difficulty being known by the interlopers, and thus we provide a safe way to deal with audio files. The proposed method will be implemented and tested to measure the efficiency and quality factors.

Keywords: Speech file, encryption, decryption, WPT, YIQ, MSE, PSNR.

INTRODUCTION

The audio signal is an analog signal [1], [2] that is converted into a digital signal as shown in Figure 1 using an analog-digital converter, which is described by the functions shown in Figure 1. The digital audio signal is consisted of a set of values which represents the signal amplitude in various times, these values are arranged in one column (for the mono voice signal) or two column (for the stereo digital voice signal) as shown in figure 2, getting this matrix of sample values makes it easy to process the voice signal [3], [4], [5].

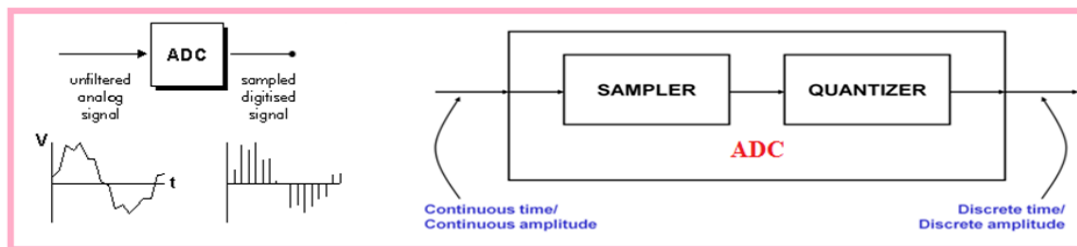


Figure 1: Converting analog voice signal to digital

0.22693878	-0.0062561035
0.22696145	-0.0043945312
0.22698413	-0.0068664551
0.22700680	-0.0115661620
0.22702948	-0.0145568850
0.22705215	-0.0145416260
0.22707483	-0.0121917720
0.22709751	-0.0123901370
0.22712018	-0.0145416260
0.22714286	-0.0144958500
0.22716553	-0.0147705080
0.22718821	-0.0157012940
0.22721088	-0.0129547120
0.22723356	-0.0127105710
0.22725624	-0.0181579590
0.22727891	-0.0191497800
0.22730159	-0.0145721440
0.22732426	-0.0122375490
0.22734694	-0.0124359130
0.22736961	-0.0108184810

Figure 2: Samples from a digital stereo voice signal

The digital audio file is used in many vital applications and may be carrying some important, bitter information that requires protection from snoopers or intruders who try to penetrate it to identify the information that the voice carries [6].

There are now many ways [9], [10] to encrypt the sound and protect it from any third party that has nothing to do with it, and when talking about the encryption method, the following things must be taken into account [7], [8], [9]:

- The method used is easy to apply and does not require high requirements.
- The method is efficient so that the encryption and decryption times are very short.
- That there is no loss of information between the original audio file and the audio file that was decrypted, and this is achieved by making the mean square error (MSE) value between the two files equal to zero or the value of peak signal to noise ratio (PSNR) equal infinite.
- The method used must destroy the encrypted audio file so that no third party can understand it. Therefore, the MSE value between the original and the encrypted files must be very high and the PSNR value is very low [23], [24], [25].

MSE and PSNR can be calculated using equations 1 and 2:

$$MSE = \frac{1}{N} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [S(i, j) - R(i, j)]^2, N = m * n \quad (1)$$

S: source signal
R: encrypted/decrypted signal

$$PSNR = 10 * \log_{10} \frac{(MAX_I)^2}{MSE_t} \quad (2)$$

Audio file consists of a set of samples [22], [26], the size of this file is the total sizes of the samples recorded in a time (Time:RT), the number of samples will depend on a sampling frequency (FS: sample rate:SR) and RT, each sample has double type (sample size), thus the audio file size can be calculated using equation 3 for mono audio and 4 for stereo audio:

$$\text{Mono} \quad \text{Sample Rate} \times \text{Sample Size} \times \text{Time} \quad (3)$$

$$\text{Stereo} \quad (\text{Sample Rate} \times \text{Sample Size} \times \text{Time}) \times 2 \quad (4)$$

Table 1 shows the audio files used in our research paper:

Table 1: speech files information
FS=44100, sample size=8 bytes

Speech number	Spoken phrase	RT	Samples	Size in byte
1	Please send me a message	2.8843	127198	1017584
2	Amman is the capital city of Jordan	4.1620	183544	1468352
3	Albalqa applied university is one of the best universities in Jordan	7.5212	331685	2653480
4	The Jordanian city of Petra is one of the most important ancient cities in the world	7.8202	344870	2758960
5	Aqaba is one of the Jordanian cities which is famous for diving	6.1316	270402	2163216
6	Jerash is one of the Jordanian cities with ancient Roman monuments	6.8160	300587	2404696
7	Wadi Rum is one of the most beautiful desert areas	5.0661	223413	1787304
8	The College of Technological Engineering is one of the leading colleges in AlBalqa Applied University	9.2423	407585	3260680
9	Encryption to protect important speech files	4.4628	196811	1574488
10	Using color image to encrypt-decrypt speech signals	6.0488	266753	2134024

Pixels in digital color image [11], [12], [13] (RGB color image) usually have integer values within the range 0_to-255, to adopt these values for double values calculation we have to convert RGB_image to YIQ image, This image will contains the values within the range -1_to_1, and this can be done by using equation 5, figure 3 shows an RGB_image example, while figure 5 shows the equivalent YIQ image [14], [15], [16]:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (5)$$

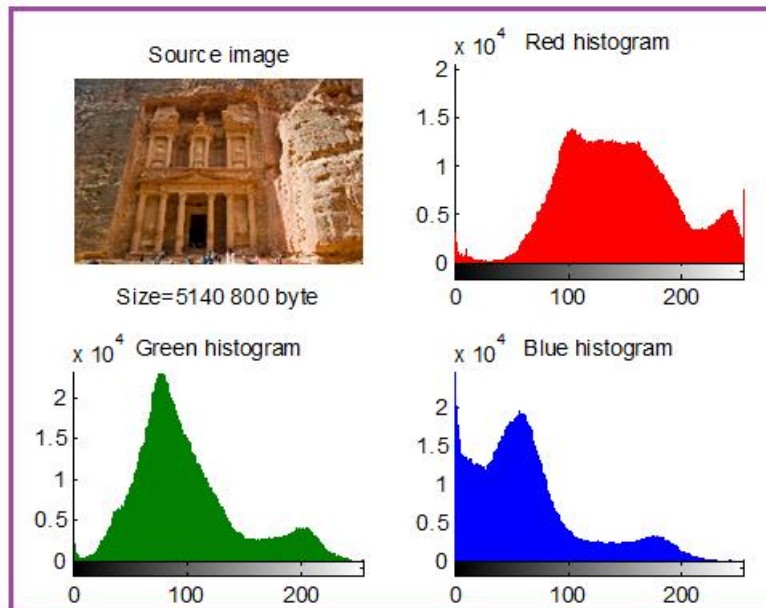


Figure 3: RGB_image example

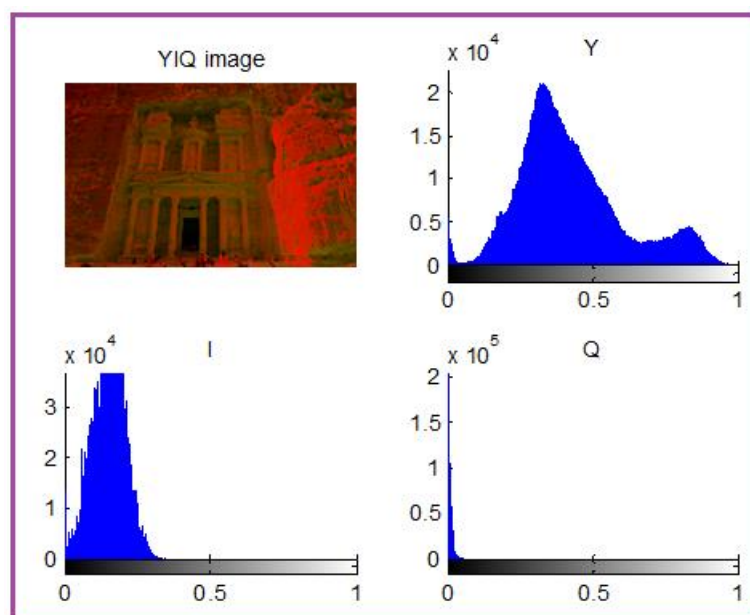


Figure 4: YIQ version

The image example shown in figure (Petra city image) has a huge size which is equal to 5 mbytes, this size can be used for encryption_decryption, or we can use a selective part by decomposing the image (using a selected number of levels) into approximations and details using the wavelet packet tree decomposition as shown in in figure 5 [17], [18].

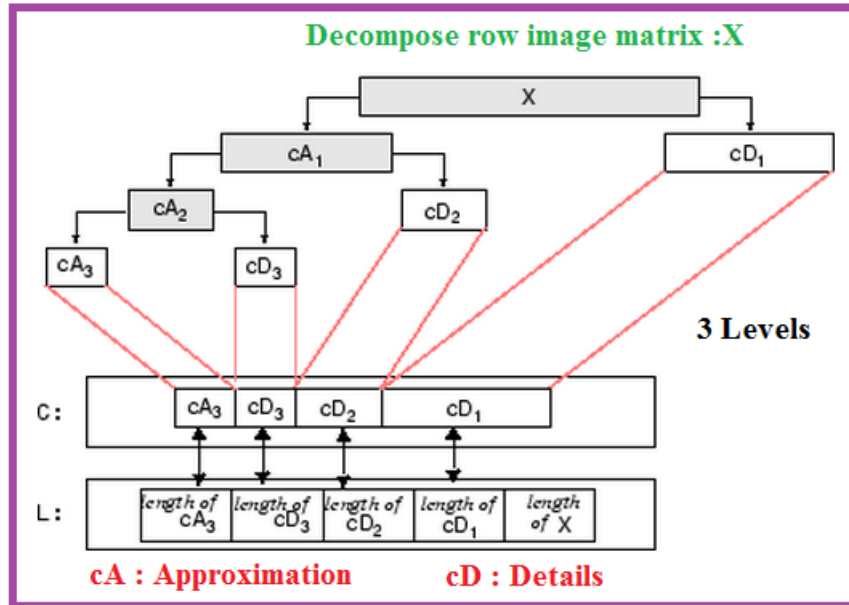


Figure 5: Image decomposition

The approximation or the detail [21] at any level can be used as an image part and they are easily calculated [19], [20].

II. THE PROPOSED METHOD

The proposed method can be implemented applying the following steps:

Encryption phase: (see figure 6)

This phase can be implemented apply the following steps:

- 1) The speaker (sender) and the listener (receiver) must agree on the digital image to be used as a private key.
- 2) Get the image and the image size.
- 3) Reshape the image into one column matrix.
- 4) Get the wave file and retrieve its size.
- 5) If the wave file represents a stereo file reshape it to one column matrix.
- 6) Apply image decomposition using the desired number of levels (here the sender and receiver must agree on this number).
- 7) Select the agreed details with size greater or equal wave file size.
- 8) Get the necessary detail and add it to the wave file to get the encrypted wave file.

Decryption phase: (see figure 7)

This phase can be implemented apply the following steps:

- 1) The speaker (sender) and the listener (receiver) must agree on the digital image to be used as a private key.
- 2) Get the image and the image size.
- 3) Reshape the image into one column matrix.
- 4) Get the encrypted wave file and retrieve its size.
- 5) If the wave file represents a stereo file reshape it to one column matrix.
- 6) Apply image decomposition using the desired number of levels (here the sender and receiver must agree on this number).
- 7) Select the agreed details with size greater or equal wave file size.
- 8) Get the necessary detail and subtract it from the encrypted wave file to get the decrypted wave file.

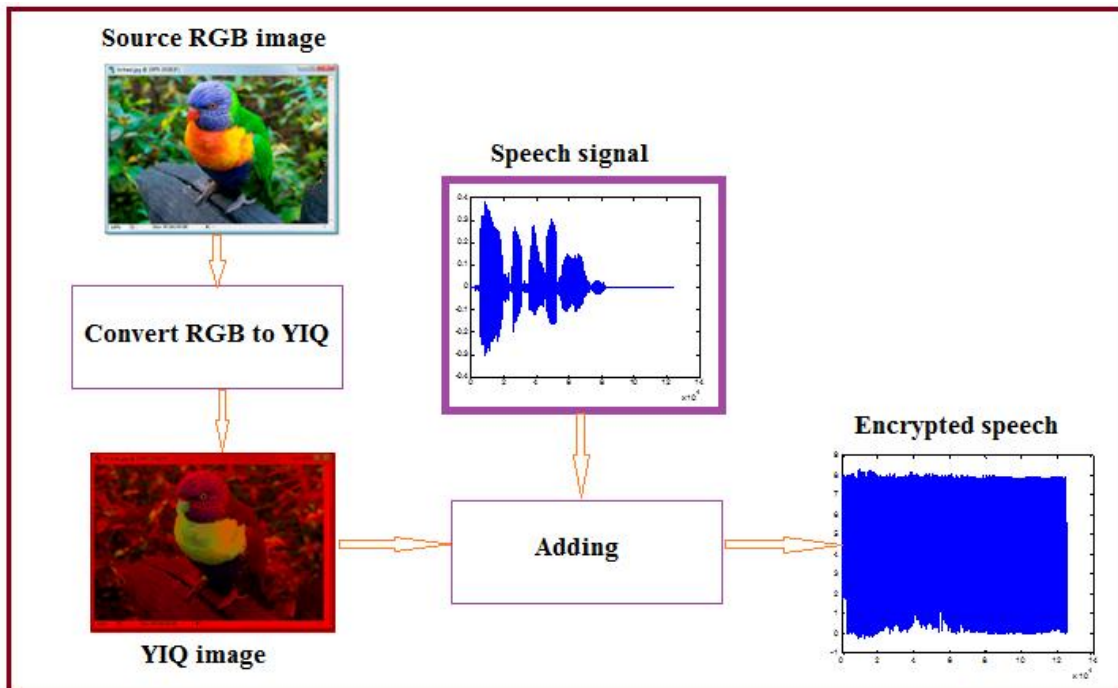


Figure 6: Encryption phase

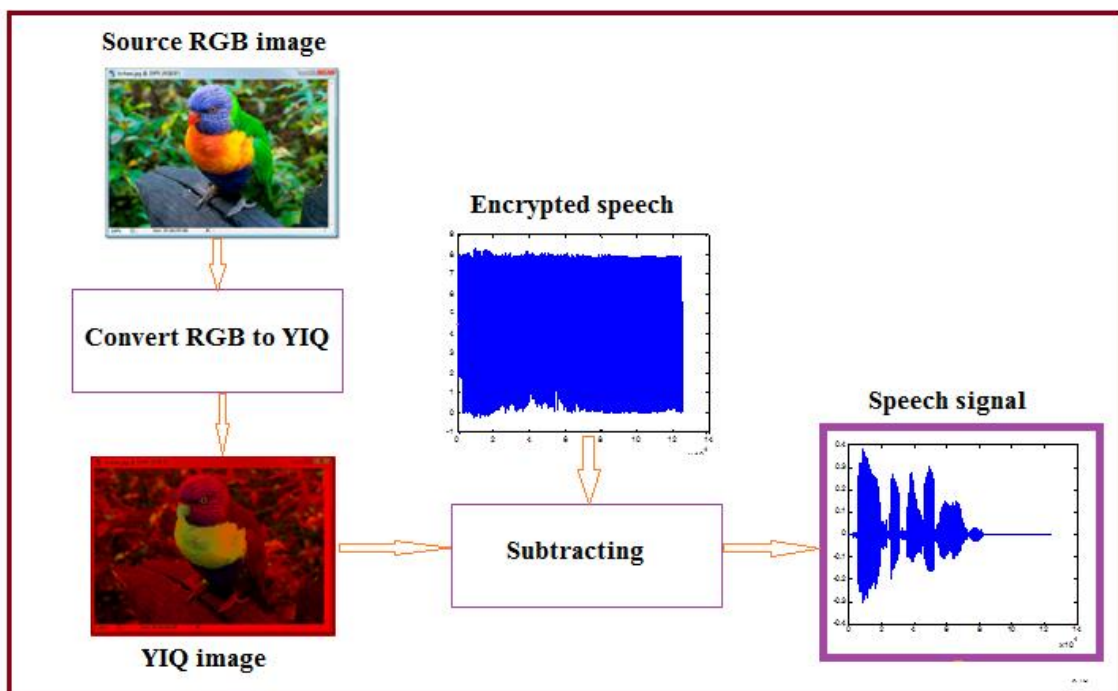


Figure 7: Decryption phase

III.IMPLEMNTATION AND EXPERIMENTAL RESULTS

Ten spoken phrase(which are shown in table 1) were recorded using sampling rate equal 44100 samples per second, the obtained wavw files were encrypted –decrypted using the proposed method, table 2 shows the obtained experimental results:

Table 2: Obtained results using the recorded phrases

Speech number	MSE	PSNR	Encryption time(seconds)	Decryption time(seconds)
1	8.3325	7.9530	1.371000	1.371000
2	8.2582	8.2472	1.373000	1.373000
3	8.3248	7.9682	1.391000	1.391000
4	8.2962	7.9934	1.398000	1.398000
5	8.2856	8.0462	1.381000	1.381000
6	8.3398	7.9559	1.383000	1.383000
7	8.2809	8.0079	1.379000	1.379000
8	8.4205	8.3285	1.392000	1.392000
9	8.2616	8.0871	1.384000	1.384000
10	8.2789	8.0403	1.425000	1.425000

From table 2 we can see the following facts:

- ✚ The proposed method is very efficient by providing small time for implementing encryption and decryption processes.
- ✚ Obtained MSE values are very high, and PSNR values are very small and the encrypted file was always a destroyed version of the original audio file, and no body can understand it.
- ✚ MSE between the original wave file and the decrypted one was always zero, while PSNR was always infinite, which means that there no any loss of information after applying the decryption process, and the decrypted file is the same as the original one.

Five speeches were load from the internet, and the wave files were treated using the proposed method, table 3 shows the obtained experimental results

Table 3: Obtained results using Five voices from the internet

Audio file	FS	Size(byte)	MSE	PSNR	Encryption time(seconds)	Decryption time(seconds)
cat	11025	139776	8.4393	7.1499	1.226000	1.226000
dog	44100	227784	8.3683	8.0860	1.184000	1.184000
gorilla	11025	514536	8.6836	7.4046	1.203000	1.203000
horse	11025	375104	8.5115	7.7866	1.159000	1.159000
tiger	11025	605808	8.6991	7.6236	1.145000	1.145000
Average		372600				1.1834
Throughput(byte per second)		314860				

The results that we obtained, which are shown in the table 3, confirm the facts that we mentioned, and they show that the proposed method has a high throughput with average equal 314860 byte per second.

IV. CONCLUSION

A new method for encrypting and decrypting audio files was proposed, and this method was tested and the obtained results proved that it is impossible to penetrate this method, as it uses a huge and anonymous digital image as a private key.

Also, the obtained practical results proved the efficiency of this method through the small values needed to implement the encrypting and decrypting processes in addition to achieving the required values for each of the MSE and PSNR.

REFERENCES

- [1] Ziad A AlQadi Amjad Y Hindi, O Dwairi Majed, PROCEDURES FOR SPEECH RECOGNITION USING LPC AND ANN, International Journal of Engineering Technology Research & Management, vol. 4, issue 2, pp. 48-55, 2020.
- [2] Holwa Fayeeg Taha Prof. Ziad Alqadi, SPEECH SIGNAL CRYPTOGRAPHY USING WPT, International Journal of Engineering Technologies and Management Research, vol. 7, issue 7, pp. 71-79, 2020.
- [3] Prof. Ziad Alqadi Prof. Yousif Eltous, Dr. Jihad Nader, Dr. Mohammad S. Khrisat, Saleh A. Khawatreh, Creating Sound Signal Histogram to be Used in Signal Decomposition, IJARCCE, vol. 9, issue 2, pp. 118-124, 2020.
- [4] Dr. Majed Omar Dwairi Prof. Ziad Alqadi, Dr. Mohammad S. Khrisat, Dr. Amjad Hindi, Analysis of speech signal MLBP features, International Journal of Latest Research in Engineering and Technology (IJLRET), vol. 6, issue 4, pp. 22-27, 2020.
- [5] Prof. Ziad Alqadi, Dr. Mohammad S. Khrisat, Dr. Amjad Hindi, Dr. Majed Omar Dwairi, USING SPEECH SIGNAL HISTOGRAM TO CREATE SIGNAL FEATURES, International Journal of Engineering Technology Research & Management, vol. 4, issue 3, pp. 144-153, 2020.
- [6] K Matrouk, A Al-Hasanat, H Alasha'ary, Z Al-Qadi, H Al-Shalabi, Speech fingerprint to identify isolated word person, World Applied Sciences Journal, vol. 31, issue 10, pp. 1767-1771, 2014.
- [7] Amjad Y Hindi, Majed O Dwairi, Ziad A AlQadi, A Novel Technique for Data Steganography, Engineering, Technology & Applied Science Research, vol. 9, issue 6, pp. 4942-4945, 2019.
- [8] Rashad J Rasras, Mohammed Abuzalata, Ziad Alqadi, Jamil Al-Azzeh, Qazem Jaber, Comparative Analysis of Color Image Encryption- Decryption Methods Based on Matrix Manipulation, International Journal of Computer Science and Mobile Computing, vol. 8, issue 3, pp. 14-26, 2019.
- [9] Jihad Nadir, Ashraf Abu Ein, Ziad Alqadi. A Technique to Encrypt-decrypt Stereo Wave File, International Journal of Computer and Information Technology, vol. 5, issue 5, pp. 465-470, 2016.
- [10] Ziad A AlQadi Amjad Y Hindi, O Dwairi Majed, PROCEDURES FOR SPEECH RECOGNITION USING LPC AND ANN, International Journal of Engineering Technology Research & Management, vol. 4, issue 2, pp. 48-55, 2020.
- [11] 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.
- [12] 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.
- [13] Qazem Jaber Ziad Alqadi, Jamil azza, Statistical analysis of methods used to enhance color image histogram, XX International scientific and technical conference, 2017.
- [14] Bassam Subaih Ziad Alqadi, Hamdan Mazen, A Methodology to Analyze Objects in Digital Image using Matlab, International Journal of Computer Science & Mobile Computing, vol. 5, issue 11, pp. 21-28, 2016.
- [15] Mazen A. Hamdan Bassam M. Subaih, Prof. Ziad A. Alqadi, Extracting Isolated Words from an Image of Text, International Journal of Computer Science & Mobile Computing, vol. 5, issue 11, pp. 29-36, 2016.
- [16] Dr. Amjad Hindi, Dr. Majed Omar Dwairi, Prof. Ziad Alqadi, Analysis of Procedures used to build an Optimal Fingerprint Recognition System, International Journal of Computer Science and Mobile Computing, vol. 9, issue 2, pp. 21 – 37, 2020.
- [17] Mohammad S Khrisat, Ziad Alqadi, Saleh A Khawatreh, Improving WPT color image decomposition, International Journal of Computer Science and Information Security (IJCSIS), vol. 18, issue 7, 2020.
- [18] Prof. Ziad Alqadi, Dr. Mohammad S. Khrisat, Two ways to improve WPT decomposition used for image features extraction, European journal for scientific research, vol. 157, issue 2, pp. 195-205, 2020.
- [19] Ismail Shayeb, Naseem Asad, Ziad Alqadi, Qazem Jaber, Evaluation of speech signal features extraction methods, Journal of Applied Science, Engineering, Technology, and Education, vol. 2, issue 1, pp. 69-78, 2020.
- [20] Amjad Y Hindi, Majed O Dwairi, Ziad A AlQadi, Creating Human Speech Identifier using WPT, International Journal of Computer Science and Mobile Computing, vol. 9, issue 2, pp. 117-123, 2020.
- [21] Ziad Alqadi, Dr. Mohammad S. Khrisat, Dr. Amjad Hindi, Dr. Majed Omar Dwairi, VALUABLE WAVELET PACKET INFORMATION TO ANALYZE COLOR IMAGES FEATURES, International Journal of Current Advanced Research, vol. 9, issue 2, pp. 2319, 2020.
- [22] 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.
- [23] Ziad AlQadi, M Elsayyed Hussein, Window Averaging Method to Create a Feature Vector for RGB Color Image, International Journal of Computer Science and Mobile Computing, vol. 6, issue 2, pp. 60-66, 2017.
- [24] Bilal Zahran Belal Ayyoub, Jihad Nader, Ziad Al-Qadi, Suggested Method to Create Color Image Features Vector, Journal of Engineering and Applied Sciences, vol. 14, issue 1, pp. 2203-2207, 2019.
- [25] 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.
- [26] 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.