MINAR International Journal of Applied Sciences and Technology

Article type : **Research Article**

Date Received : **21/09/2021**
Date Accepted : **10/10/2021**
Date published : **01/12/2021**
: **www.minarjournal.com**
http://dx.doi.org/10.47832/2717-8234.4-3.5

A NOVEL METHOD FOR GENERATING DIGITAL IMAGE OF ENCRYPTED SPEECH SIGNAL BASED ON FWHT AND WPLCM

Alyaa M. Abdul MAJEED[1]

Abstract

This paper suggests a novel algorithm for encrypting speech signals in common image formats and retrieve them from these image files. The speech signal is encrypted in three levels. In the first level,the sample positions are permuted based on keys generated using Game of Life matrix and Piecewise linear Chaotic Map (PWLCM) in order to reduce the correlation between adjust samples. In the second level, the resulting samples are then converted to Fast Walsh Hadamard Transform (FWHT) and their transactions are encrypted by using circular transformations in the row and column depending on the generated key. At the third level, the values of encrypted samples are converted to color pixels, which are then arranged in a puzzled manner and put in a 2-D matrix to achieve the secured data transfer across networks, with the image file contains the encrypted speech signal. Several objective measures have been used to evaluate the performance of the suggested method. The experimental results and numerical analyses show that the algorithm gives a high degree of security and robust against brute force attackers, statistical attack, strong diffusion and ambiguity so that the encrypted message has been saved in a different format from the original signal, and finally give the good quality of the reconstructed speech signal from image files.

Keywords: Speech Signal, Image Format, FWHT, Game Life, PWLCM, RGB.

[1] Mosul University, Iraq, alyaahaleem@uomosul.edu.iq

1. Introduction

Today, multimedia applications (such as voice, image and speech) are essential, and its services are the foundation of the telephone industry, video conferencing and broadcast news. Therefore, protecting such systems become very necessary to prevent unauthorized listener attacks [1]. Encryption and Steganographyare not the identical things. While encryption is scrambling a message to prevent it from being understood,it has increased because the current voice communication scheme requiresa large amount of sensitive data to be exchanged over an open orshared network every single minute [2]. On the other hand, Steganography hides messages to prevent itfrom being seen. Undetectability, robustness and hidden data capacitythese are the major characteristics that separate steganography from encryption [3].

The Speech signal has some fundamental features, such as redundancy of speech data and bulk data capacity. therefore, traditional cryptosystems such as the Data Encryption Standard (DES) and Advanced Encryption Standard (AES) are unsuitable for providing fast communication applications [4]. With the consideration of these features, designing a different algorithm to encrypt the voice needs less computational power that decreases the "residual intelligibility"of encrypted signals, sincethe good quality of the signals is retrievedwith a high level of security [5][6].

2. Related Work

In recent years, many speech signal encryption methods have been proposed. In 2021 [2], this study introduces a novel speech encryption technique. It is determined that the quantum chaotic map and k-means clustering are both used in the key generation. Furthermore, two scrambling layers were utilized: the first was based on bits using the suggested method ("binary representation scrambling BiRS"). The second is based on k-means using the proposed method ("block representation scrambling BlRS"). The objective examination evaluated the proposed system using statistical analysis measures. It has been shown that the proposed technique is secure, dependable, and efficient for use in secure speech transmission, in addition to having a high level of clarity in the recovered speech signal. The Ref. [3] in 2017has proposed a novel method for hiding secret image utilizing Discrete Cosine Transform(DCT) components according to a Linear Support Vector Machine .The DCT components are utilized to reduce the amount of redundant information in the image. After that, the secret messages are embedded in the least significant bits. Each bit in the cover image is changed to the extent that it cannot be perceived by human eyes. The SVM is utilized as a classifier to accelerate the hiding process using DCT characteristics. Utilizing three layers of a color image relied on RGB over DCT characteristics adds considerable improvements in performance and accuracy. The Ref. [5] the samples and their values are divided into four levels; L0 =-1 to-0.5, L1 =-0.5 to 0, L2 = 0 to 0.5, and L3 = 0.5 to 1. Each level is altered through four chaotic generators: the logistic map, the tent map, the quadratic map, and the Bernoulli’s map. A chaotic shift keying mechanism assigns a logistic map for L0, a tent map for L1, a quadratic map for L2, and Bernoulli’s map for L3 for shuffling the speech samples at every level. Ref. [6] provided a new way of confusing the plain image by applying the Game of LIFE replacement at the first level. PWLCM is used to diffuse the image at the second level. The test results show that the approach provides effective encryption and that the key size is large enough to withstand an assault .The Ref. [7] describes a new speech encryption method relying on a novel 3D Lorenz-Logistic map introduced by putting the Logistic map into the 3D Lorenz map to create three separate random number sequences. The primary eight and regulatory factors of the 3D Lorenz-Logistic map are used as critical values. Then FFT is implemented on the input voice to produce both real and imaginary values. The input voice's real and imaginary values are altered using two random number sequences from the 3D Lorenz-Logistic map. The last random number sequence is used to permute a reference voice sample, which is then utilized to

replace permuted real values of the voice signal .The Ref.[8] in 2019, proposes an algorithm for audio encryption that is based on permutation of the samples with a discrete modified"Henon map"succeeded by a key stream replacement operation created by the modified Lorenz-hyper-chaotic method. In this paper, the audio file is compressed using the"Fast Walsh Hadamard Transform "(FWHT) to eliminate residual intelligibility in the transform domain .The results show that the suggested algorithm is secure, fast, and resistant to a variety of cryptography challenges.

In this paper, we attained a high level of security which is based on encrypting speech signal files in image formats like PNG, JPEG, and TIF. This can be achieved by shuffling samples of the speech signals at every level. In the last stage, the coefficients in each encrypted block are converted to a colored pixel within an intensity of 0 and 255. The pixels in each block are arranged in an image format of row and column pixels based on the number of frames of the original speech signal.

3. piecewise linear chaotic map (PWLCM)

In the chaotic style, it is well-known that the variation of functions in the logistic mapping is non-uniform, that is, the balance is poor, while the variation of density function in the PWLCM is uniform and has better uniformity.For this reason, the PWLCM algorithm is used to improve the performance of the entire speech encryption stage [9][10]. This map is presented in Eq. (1):

$$X_{i+1} = F(X_i, \alpha) = \begin{cases} X_i/\alpha, & 0 < X_i < \alpha \\ (X_i - \alpha)/(0.5 - \alpha), & \alpha \le X_i < 0.5 \\ F(1 - X_i, \alpha), & 0.5 \le X_i < 1 \end{cases} \quad (1)$$

Where the control parameter$\alpha \in (0,0.5)$, and$X_{i-1} \in (0,1)$. This map can be easily computed and chaotic in the entire scope of the parameter of [12].

4. The Game of Life

The Game of Life is a cellular automaton (CA) that is used in many applications, for example, games, random number generation, and pattern recognition. CA is represented as a symmetrical infinite two-dimensional cell map. Every cell may choose one of two probable states: either live or dead. Each one of the cells is connected to its eight neighbors, where the cells are adjacent vertical, horizontal, or diagonal. [6][11] atevery time step, each cell has a new state based on the transition rules:

1. Any dead cell becomes alive if three neighborhoods of the cell are alive.
2. Any live cell becomes dead if four or more of its neighbors are dying.
3. Any live cell becomes dead if one or none of its neighborhoods are alive.
4. Any live cell will remain alive if two or three of its neighborhoods are alive.

5. The Key Generation Structure

The permutation key is used to eliminate the correlation between samples as much as convenient. In this paper, two keys with a length of N (N: the length of the frame) needed to be generated; the first key was generated by using PWLCM, and the output was then passed to the game of life matrix to produce the second key (key1).

The following steps represent the algorithm for generating the keys.

Input: the Initial value X, as well as the control parameter α
Output: Permute keys (key1, key2).

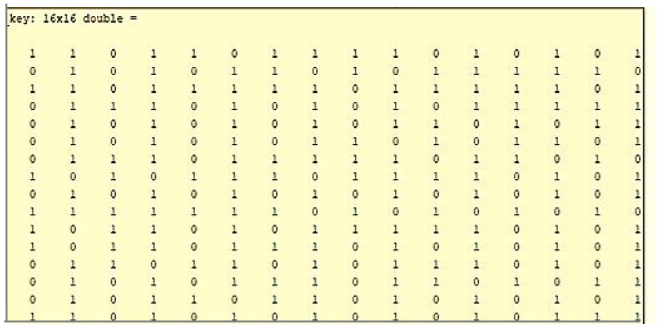
Step1: Generate 1_D array by using PWLCM (illustrated in Eq.1) that contains a sequence of real numbers.
Step2: Convert the output from step1 into 2_D array of size (n*n) and called CA (cellular- array).
Step3: Convert the value in CA to zero or one by applies the first step of the game life matrix:

If CA (I, J) > 0.5 then
CA (I, J) is dead
Else
CA (I, J) is alive
CA represents the initial key.
Step 4:Set the counter to zero, set the Key to zero.
For counter = 1: (N divided by 2)
For r =1: n
For c=1:n
If (CA(r,c) == 0 && key (r,c) == 0)
Key (r,c) = counter

Step 5: Convert the key into a one-dimensional vector called key1.
Step 6:The second key (key 2) is obtained from the initial key by:

1. Taking the first two rows and the last two rows of key.
2. Each row is dividedinto two halves.
3. Reversing and exchanging the two halves called (key2).

Example: Extract key2



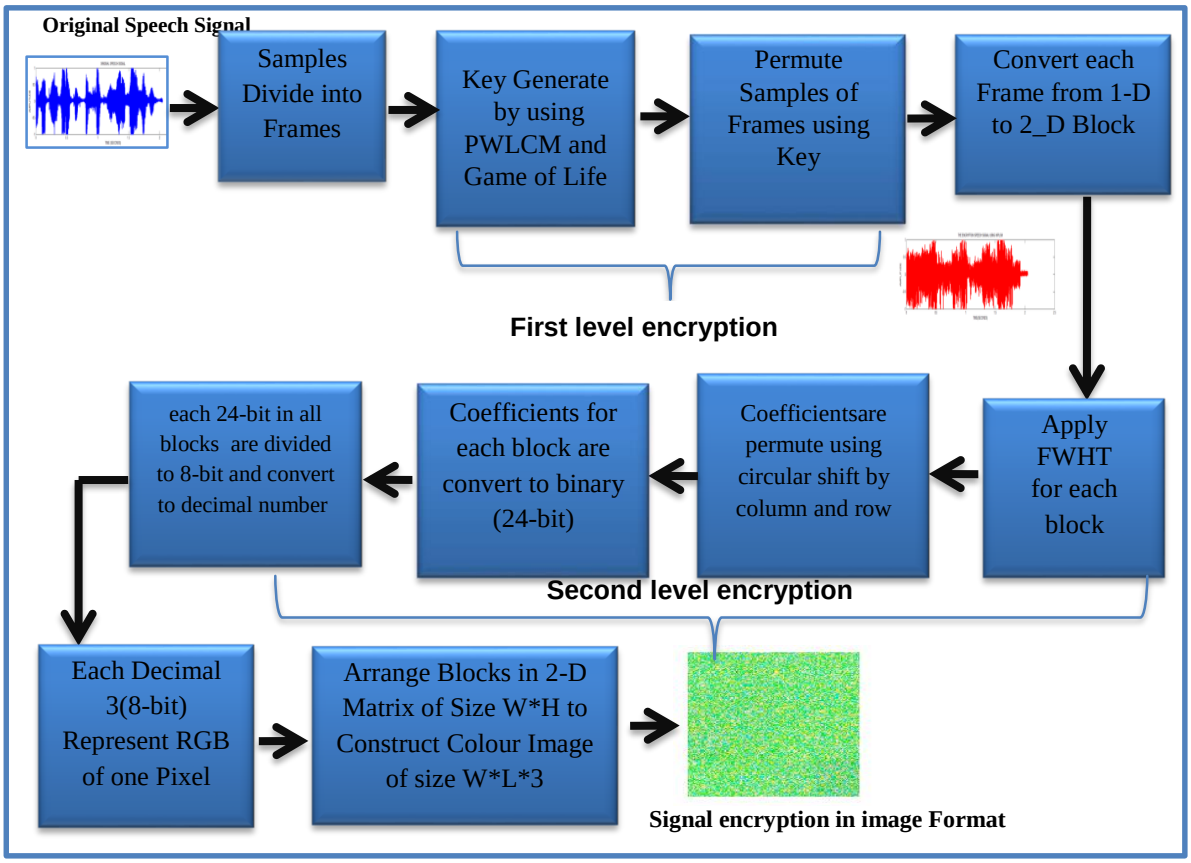
Row1: 1101101111010101
Row2: 0101011010111110
Row15: 0101101101010101
Row16: 1101010101010111
Key2:0010101000100100010000011010100110101010101001001010100000101010

6. **Speech Encryption in Image Format**

In this section the paperexplains the basic concept of the speech encryption process. Initially, the original speech signal is broken into a number of non-overlapping frames of a fixed size, each frame containing N samples. The location of samples in each frame is permuted on the first encryption layer that is based on the key generated by using PWLCM and the game of life. It takes the advantage of the use of one-dimensional piecewise linear chaotic map (PWLCM) to diffuse the samples values. The encrypted speech signal from the previous layer is encrypted in image format in the second layer by performing the following steps: each encrypted frame is converted from one

dimensional vector to 2-D array, each block is compressed using discrete Walsh Hadamard Transform (FWHT), and the coefficients are encrypted using circular transformations in the row and column based on the generated encryption key (key2). The coefficients in each encrypted block are converted to the binary representation (24-bit) in the last layers, then these 24-bits are separated into 3-bytes and each byte is converted to decimal data type in order to represent RGB as each sample in the block denotes a different color pixel within an intensity of 0 and 255. The pixels in all blocks are arranged in the image format of row by column pixels depending on the number of frames in the original speech signal, Figure (1) illustrates the proposed speech encryption algorithm.

Figure (1) the block diagram of speech encryption



7.the Proposed algorithm

The general structure of the algorithm needed for encrypting and decrypting the speech message is summarized as follows:

The steps of encryption speech signal are:

Input: Speech signal file.

Step1: Divided speech samples into frames of size 1024 samples to produce array (M*N), Here M and N denotes the number of frame and the total number of samples in every frame respectively.

Step2: Generate of permutation key1 by using PWLCM and game of life.

Step3: Permute samples of each frame with permutation key1 (first level encryption).

Step4: Reshape each encrypted frame into 2_D square array of size n*n called ENC_MATM.

Step5: Convert ENC_MATM from time domain to FWHT called ENC_FWHT_MATM.

Step6: Permute the coefficients of each ENC_FWHT_MATM using key2 and circular shift by column and row, as shown in the following figure (2).

Step7: Convert the each coefficients in the produced matrix from the previous step from real number to Binary number of size 24-bits then divide it into 8-bits that represent Red, Green and Blue in order to construct the pixel (RGB) then produce the sub-image matrix of size r*c*3.

Step8:the sub-image blocks are arranged into image of size r*c based on the number of frame.

Step9:Save encryption matrix in the image format.

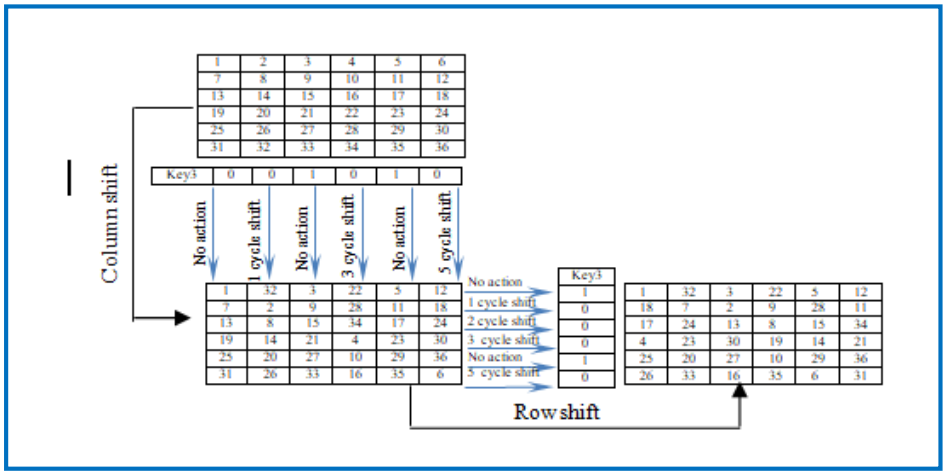


Figure (2) Column and Row Permutation step

The steps of Decryption speech signal are:

Input: Load an image.

Step1: Generate permutation key1, key2.

Step2: Divide image into blocks of size n*n depend on the length of the frame that agreed between the sender and receiver.

Step3: Convert RGB to binary number, and then convert to decimal.

Step4: Invers Permutation the coefficients of each block using key2 and circular shift column and row.

Step5: Apply inverse FWHT for each block.

Step6: Reshape each block from 2-D into 1-D to construct the frame.

Step7: Apply Inverse Permutation samples of the frame by key1.

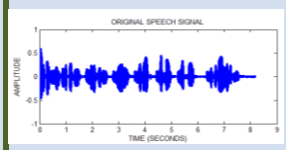
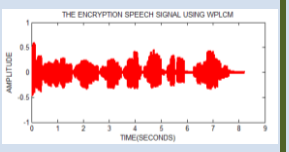
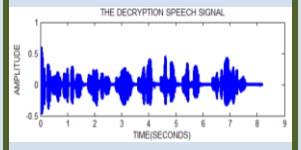
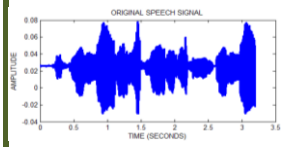
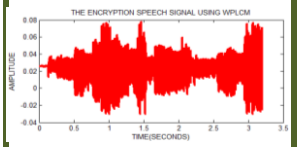
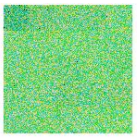
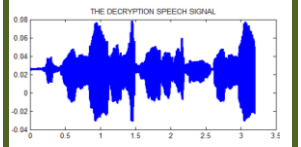
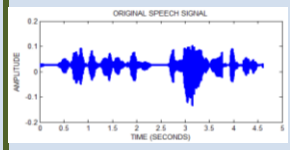
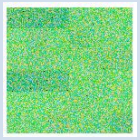
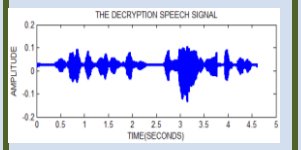
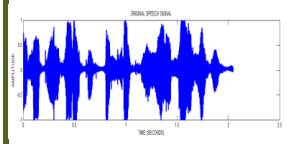
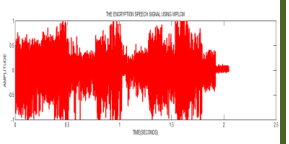
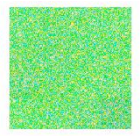
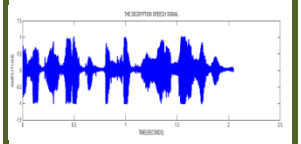
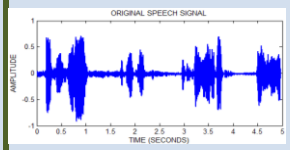
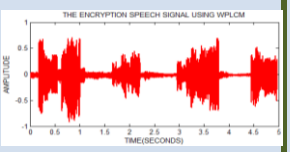
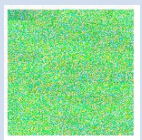
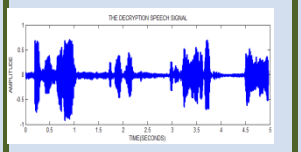
Step8: Construct the decrypted frame from the set of the frames by repeating the 5th step for all frames.

Step9: Get the decrypted message.

8. Experimental Results

The suggested algorithm is applied to a set of speech signals which get a good result. Here, the result of five test files is applied. The efficiency of the encrypted signal and the quality of the recovered signal of the proposed algorithm are evaluated using a number of objective measures. Table (1) shows the waveform that is plotting of the original signal, the encrypted signal from the first level, the encrypted signal in the image format from the second level with its size of the image and the recovered signal.

Table (1) the results of various speech signals as original signal, encryptedsignal (first level), encrypted signal in image format (second level), and decryption signal

File name	Original signal	Encrypted signal First level	Encrypted signal as image	Decrypted Signal	Image Size
Man.WAV					**256*256**
A3.WAV					**160*160**
Aya.WAV					**192*192**
Dish.WAV					**160*160**
Test3.WAV					**384*384**

The value of MSE (Mean Square Error), PSNR (Peak Signal-to-Noise-Ratio), and the correlation between the original and encrypted speech signal (for first level) are calculated. MSE, PSNR, and the correlation between the original and recovered speech signals are also calculated. Table (2) shows these results. The residual intelligibility is the measurement that is used to measure the residual intelligibility of the encryption and the quality of the retrieved speech signals. The peak signal to noise ratio (PSNR) measure is used. The lower values of the PSNR indicate the low residual intelligibility of encrypted signals, while the higher values of PSNR indicate the very strong quality of the retrieved speech signals.

A linear relationship between the two variables is represented by the correlation coefficient. It compares similar samples of the encrypted signal with the original speech signal to evaluate the performance of the encryption algorithm. The correlation between the two adjacent samples is very high. As a result, the encryption algorithm breaks this correlation. f the correlation coefficient is equal to 1, this indicates that the two variables have a very strong linear relationship. If the correlation coefficient is equal to 0, this means there is no connection. When the correlation coefficient is equal to -1, this indicates that one of the variables is the negative of the other.As

shown in Table (2), we can see the results of the correlation between the original and encrypted signal. The result indicates that the correlation value is very small, indicating that the algorithm is able to stand a statistic attack. While if the result indicates that the correlation value is perfect and positive, this means that there is a high correlation between the original and the decrypted signals.

Table (2) Show the result of PSNR, MSE, CORRELATION, for many speech signals

File Name	PSNR First Level Encryption	MSE First Level Encryption	Correlation First Level Encryption	PSNR Decryption	MSER Decryption	Correlation Decryption
Man.wav	10.5792dB	1.4192	-0.0070	81.6589dB	0.00039633	1.0000
A3.wav	-7.4090dB	1.4277	-0.0184	56.5386dB	0.00090627	1.0000
Aya.wav	-1.2302dB	1.4109	0.0047	65.0537dB	0.00068438	1.0000
Dish.wav	9.6040dB	1.4116	0.0037	80.3058dB	0.00041174	1.0000
Test3	15.4176dB	1.4189	-0.0067	86.0919dB	0.00041519	1.0000

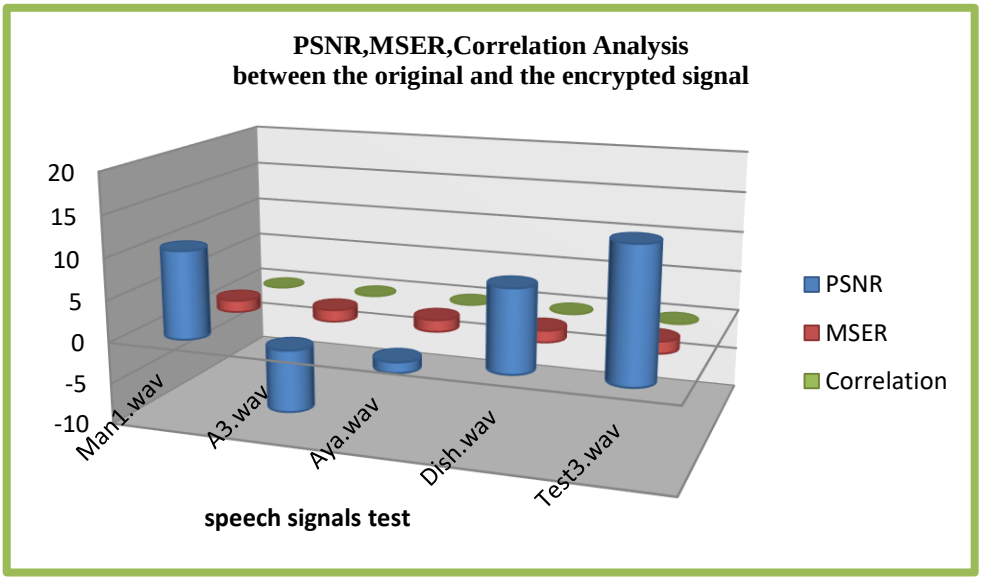


Figure (3): PSNR, MSER, and correlation analysis chart between the original and the encrypted signal (at first level).

In Figure (3), we note that the encryption algorithm at the first level gives a large negative value of PSNR, which is a strong indicator of low residual intelligibility of the encrypted speech signal, making it hard to detect. While the value of the correlation among the original signal and the encrypted signal is close to zero for all the encrypted signals. This means that the original and the encrypted speech signals are uncorrelated.

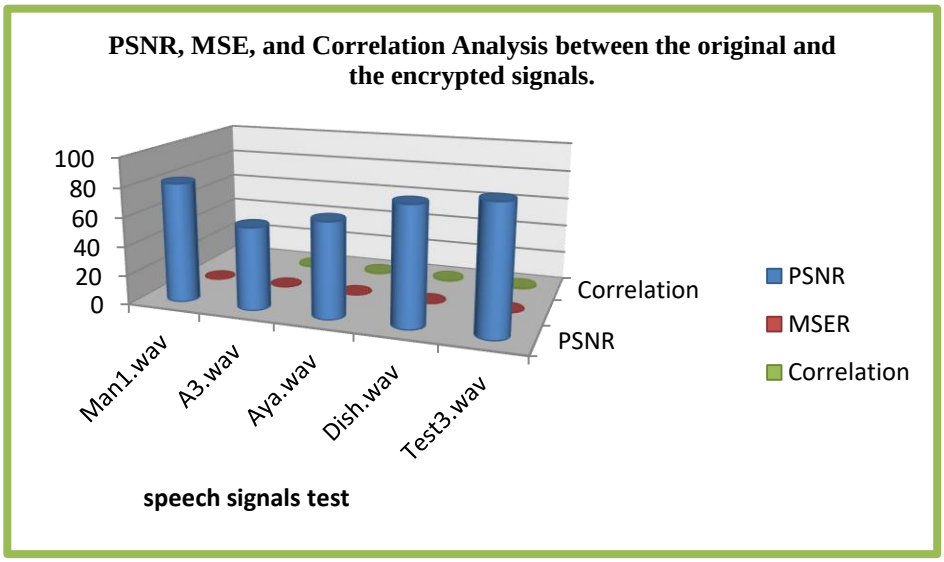
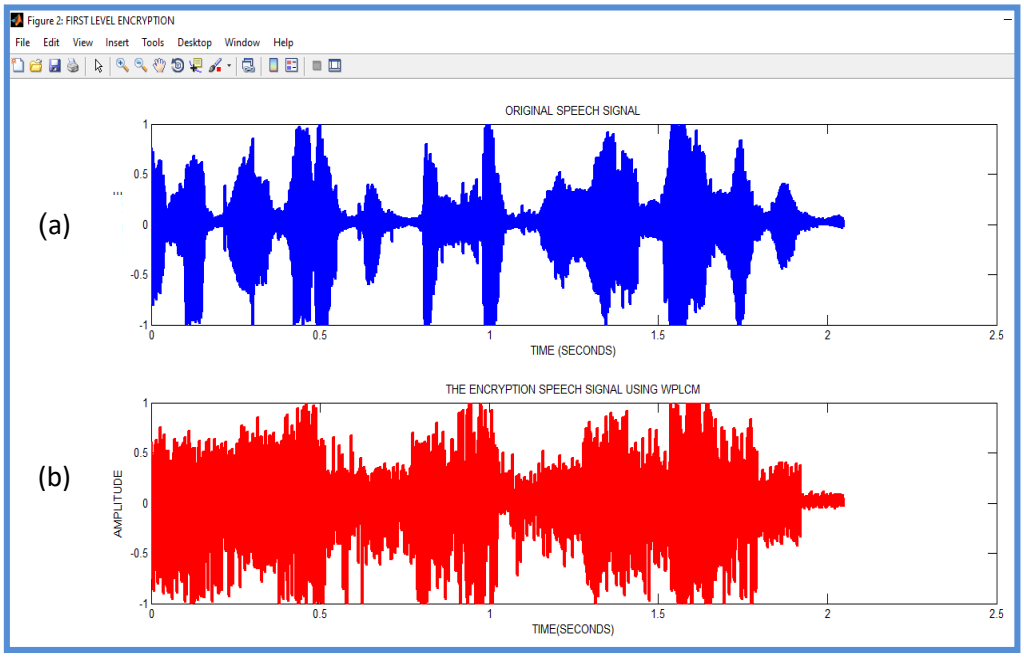


Figure (4): PSNR, MSER, and the correlation analysis chart among the original and the recovered signal.

When examining the results in Figure (4), we note that the PSNR scale value for all retrieved speech signals is high, which gives a good quality of the retrieved speech signal. We also note that the value of the correlation among the original and the retrieved signal is equal to one for all the recovered signals. This indicates a very strong linear relationship between the two signals. As for MSER, the value of all signals is close to zero, which gives a clear impression of the quality of the algorithm.

Figure (5) shows the samples of the speech signal before encryption, then after encryption, while Figure (6) shows the entire image file that contains the encrypted speech signal. Figure (7) shows the retrieved speech signal and as we observe, it is very close to the waveform of the original signal.



Figure(5) the speech signal "Dish3.wav"; (a)Original signal, (b) encrypted signal (at first level of Encrypt)

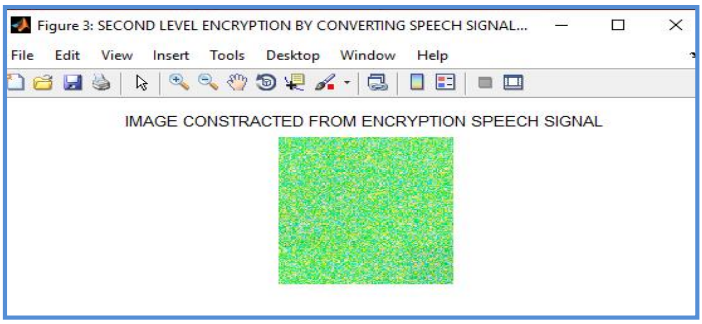


Figure (6) the image constructed from the encryption signal (second level encryption)

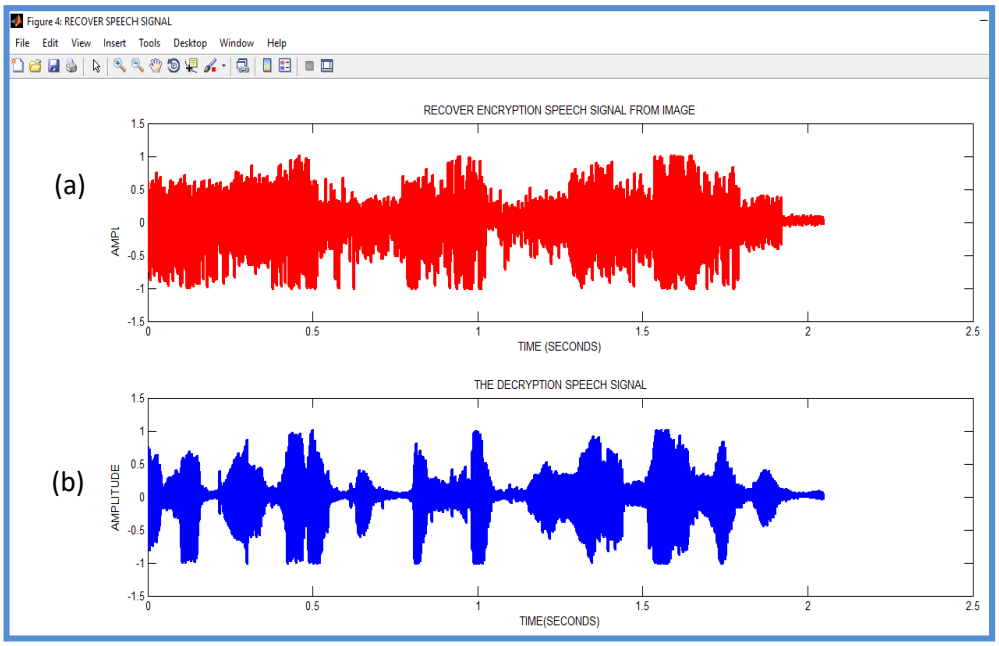


Figure (7) the recover speech signal "Dish3.wav"; (a) recover signal, (b) decrypt signal

In order to calculate the correlation coefficients of the encrypted image pixels, adjacent pixels in the horizontal, vertical, and diagonal directions are chosen. The correlation coefficients of the encrypted image are listed in the table (3) and Figure (8) shows the distribution of adjacent pixels for the encrypted image.

Table (3) the results of CORRELATION COFFICIANTS for many encrypted speech signals in image format

File Name	Correlation Coefficients of Encryption Signal in image		
	H	V	D
Man.wav	0.3876	0.4066	0.3584
A3.wav	0.3126	0.3554	0.3510
Aya.wav	0.3774	0.3480	0.3125
Dish.wav	0.4203	0.4650	0.3974
Test3.wav	0.3816	0.3561	0.3299

H: Horizontal Correlation Coefficients V: Vertical Correlation Coefficients
D: Diagonal Correlation Coefficients

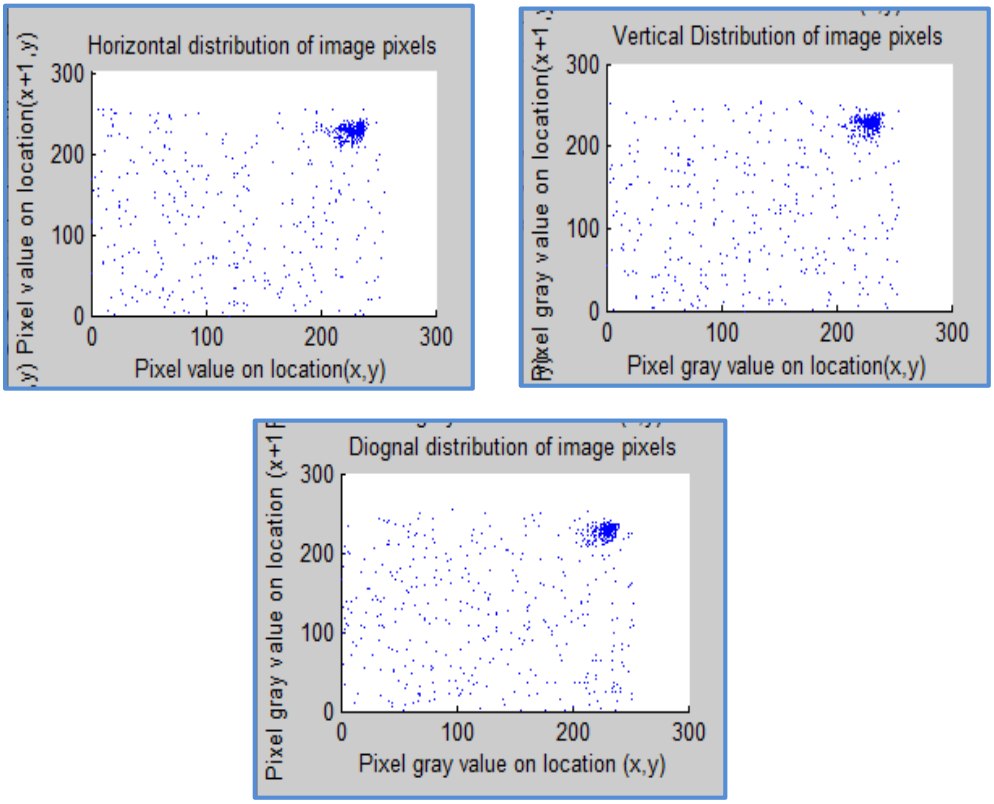


Figure (8) shows the horizontal, vertical, and diagonal pixel distributions of the encryption signal in image format.

Conclusion

This paper has provided an efficient way to transfer the encrypted speech signal .Securely over the networks, the entire color image file is constructed from the encrypted speech samples that are organized randomly and the retrieval end that the algorithm could be provided to extract the speech signal. This algorithm is not only a stenographic tool, but also a data compression method. The execution of the program is very critical to the primary condition and the control parameters, which makes the key very sensitive. For this reason, a tiny change in the key between the encryption and decryption sides will make the encrypted message unable to be decrypted correctly. The experiments indicate that the proposed algorithm gives closer encrypted data, distributions and small values of the correlation between the samples so that the encryption algorithm is able to withstand the statistical assaults.

References

Mohammed, R. S., & Sadkhan, S. B. (2016). Speech scrambler based on proposed random chaotic maps. IEEE International Conference on Multidisciplinary in IT and Communication Science and Applications, Baghdad, 2016,1-6

AmalHameedKhaleel, Iman Qays Abduljaleel , A novel technique for speech encryption based on k-means clustering and quantum chaotic map, Bulletin of Electrical Engineering and Informatics, Vol. 10, No. 1, February 2021, pp. 160~170, ISSN: 2302-9285, DOI: 10.11591/eei.v10i1.2405.

AkramAbdelQader and FadelAlTamimi, A Novel Image Steganography Approach Using multi-layers DCT Features Based on Support vector Machine Classifier. The International Journal of Multimedia & Its Applications (IJMA) Vol.9, No.1, February 2017 DOI : 10.5121/ijma.2017.9101

Mahmoud Farouk , Osama Faragallah, Osama Elshakankiry, Ahmed Elmhalaway Comparison of Audio Speech Cryptosystem Using 2-D Chaotic Map Algorithms , Mathematics and Computer Science 2016 ; 1(4):66-81,doi: 10.11648/j.mcs.20160104.11. http://www.sciencepublishinggroup.com/j/mcs

.P.Sathiymurthi and S.Ramakrishnan, Speech Encryption Chaotic Shift Keying for Secured Speech Communication, EURASIP Journal on Audio,Speech and Music Processing (2017) :20,DOI 10.1186/s136-017-0118-0.

.Xingyuan Wang*,Canqi Jin, Image Encryption Using Game of Life Permutation and PWLCM Chaotic System , Optics Communications. 285(2012) 412-417.

P.Sathiyamurthi and S.Ramakrishnan,Speech encryption algorithm using FFT and 3D-Lorenz–logistic chaotic map, Multimedia Tools and Applications volume 79, pages17817–17835 (2020).

keystreams F.J. Farsana , V.R. Devi , K. Gopakumar , An audio encryption scheme based on Fast Walsh Hadamard Transform and mixed chaotic, , Applied Computing and Informatics, https://doi.org/10.1016/j.aci.2019.10.001.

Yucheng Chen, ChunmingTang,andZongxiang Yi, A Novel Image Encryption Scheme Based on PWLCM and Standard Map , Hindawi Complexity Volume 2020,Article ID 3026972,23 pages https://doi.org/10.1155/2020/3026972.

Xingyuan Wang, Hui-li Zhang, A Color Image Encryption with Heterogeneous Bit-Permutation and Correlated Chaos, Optics Communications 342 (2015) 51-60.

MajidVafaeiJahan, FaezehKhosrojerdi, Text Encryption Based on Glider in the Game of Life, International Journal of Information Science,p-ISSN:2163-1921 e-ISSN:2163-193X2016; 6(1): 20-27, DOI:10.5923/j.ijis.20160601.02.

Hao Li , Lianbing Deng , and ZhaoquanGu , A Robust Image Encryption Algorithm based on a 32-bit Chaotic System, Citation information: DOI 10.1109/ACCESS.2020.2972296, IEEE Access