View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Periodicals of Engineering and Natural Sciences (PEN - International University of...

Periodicals of Engineering and Natural Sciences ISSN 2303-4521
Vol. 8, No. 1, February 2020, pp.58-66

Data hiding using integer lifting wavelet transform and DNA computing

Firas A. Abdullatif[1], Alaa A. Abdullatif[1], Namar A. Taha[1]
[1] Department of Computer Science, College of Education for Pure Sciences, University of Baghdad, Iraq

ABSTRACT

DNA computing widely used in encryption or hiding the data. Many researchers have proposed many developments of encryption and hiding algorithms based on DNA sequence to provide new algorithms. In this paper data hiding using integer lifting wavelet transform based on DNA computing is presented. The transform is applied on blue channel of the cover image. The DNA encoding used to encode the two most significant bits of LL sub-band. The produced DNA sequence used for two purpose, firstly, it use to construct the key for encryption the secret data and secondly to select the pixels in HL, LH, HH sub-bands for hiding in them. Many measurement parameters used to evaluate the performance of the proposed method such PSNR, MSE, and SSIM. The experimental results show high performance with respect to different embedding rate.

Keywords: Data hiding, Integer lifting wavelet, Data encryption, DNA coding

Corresponding Author:

Alaa A. Abdullatif,
Department of Computer Science,
College of Education for Pure Sciences,
University of Baghdad, Iraq
E-mail: alaa.a.h@ihcoedu.uobaghdad.edu.iq

1. Introduction

When the information is transmitted through the internet, the most important factor to be considered is the security. Two security features used to prevent unauthorized access to the secret data, these features are known as cryptography and steganography.
The cryptography deals with two main process, encryption and decryption process of the transmitted data through the network. The cryptography needs the key value that is used by the sender to encrypt the data. The key value also used by recipient to decrypt the received encrypted message [1].

Cryptography can be classified into two systems, symmetric systems (private-key) that use a single key for encryption and decryption process, while the second type is asymmetric systems (public-key) that use different key for encryption and decryption process. There are many algorithms for symmetric cryptography for example Data Encryption Standard (DES), which is the one of most common algorithms in use today. Examples of asymmetric cryptography algorithms are Hellman, Digital Signature Algorithm (DSA), and RSA. The symmetric and asymmetric cryptography systems have advantages and disadvantages [2, 3].

DNA cryptography is process of encrypt data based on DNA molecular, which can be done using several DNA encoding techniques. In addition, data hiding approaches based on the DNA sequence attracted much attention [4]. DNA computing is a new method of computing and simulating the bimolecular structure of DNA based on molecular biology. The scientist believe that DNA computing can used to solve many difficult problems, which require large amount of parallel computing. Many scientists have produced a number of theoretical and practical developments in this area. The main target of DNA based cryptography is used the concept of DNA computing to enhance the security of cryptographic algorithms. It may be symmetric or asymmetric system [5].
The steganography is the second scheme of providing the security to the secret information. This achieved by hiding it within the cover mediums like the image. It is possible to use images to hide secret data, either in the

spatial domain or in frequency domain. The secret data was hidden within pixel values in the spatial domain; while in the frequency domain, the image is transferred then the coefficients are used for hiding the secret information [6, 7]. One of the algorithms that was used to transfer the image to frequency domain is lifting wavelet transform [8].

This paper proposed a method to encrypt the data using DNA encoding algorithm then embedding it in integer lifting wavelet transform coefficients. The embedding process depend DNA computation to ensure the data is hiding not in sequence pixels to increase the security.
The rest of paper is organized as follow; Section 2 explained lifting wavelet transform. Section 3 briefly presents the DNA encoding. Section 4 presents the proposed algorithm to encrypt the data and hide it; the performance measurement explained in Section 5, while the experimental results discussed in section 6. Finally, section 7 is the conclusion.

2. Lifting wavelet transform

In the field of image processing and image steganography, there are many researchers used wavelet for its advantages. Multi resolution capability and optimal representation of signal are the main advantages of the wavelet. Normally, the data hiding using wavelet gives better performance compared to other methods. Lifting transform is a technique used to produce the second-generation wavelet. The convolution of the input image needs many computation processes and large memory space when used the conventional wavelet transform while in lifting method the time and memory space are reduced because the it's simple mathematic operation [9].
A floating-point arithmetic basis is the base of conventional wavelet transform operation. The integer intensity values of an image in the spatial domain transferred into decimal coefficients by applying wavelet transform. When the data hiding in the wavelet coefficients there are many modification happened. The wavelet coefficients may modified in truncated or rounded operation and in reconstruction the image process by applying the inverse wavelet transform. This make the discrete wavelet transform not a good choice for reversible data hiding, because the host image must recover without any distortion in secret payload [10].

The lifting wavelet transform decomposes the image into four frequency sub-bands, which contain approximation (LL) and detail coefficients (HL, LH, and HH). The Lifting Scheme can be converted easily into a transform that maps integers to integers with good reconstruction property. In data hiding schema, the advantage of using integer lifting wavelet transform is the guarantee of reverse data hiding. For that, the embedding data in integer lifting wavelet domain satisfies many good properties like security, robustness, and imperceptibility [11-13].
The three main steps of lifting wavelet transform is explained as follows:
(1) Split: One method for signal split is dividing signals into two halves (right and left halves). There is low relativity between the two halves signals for that the result will be unsatisfying. Divide the data $x\,[n]$ into two sets (even and odd sets) is another effective method. The even set $xe\,[n] = x\,[2n]$ and odd set $xo\,[n] = x\,[2n + 1]$, where n is the number of data.

(2) Prediction: to eliminate the low frequency components of signals and preserve the high frequency part, the prediction step is necessary. The predicted of odd set $xo\,[n]$ from the even set $xe\,[n]$ done using the prediction operator $\mathbf{P} = [p_1, p_2, \ldots, p_N]$. The prediction value is $\mathbf{P}\,(xe\,[n])$. The $d\,[n]$ represent the difference between the two values of practical and the prediction as follows:

$$d\,[n] = xo\,[n] - \mathbf{P}\,(xe\,[n]) \qquad (1)$$

Where the detailed signal ($d\,[n]$) which reflects the high frequency component of the signals while the smoothness of the interpolation function is determined by dual vanishing moment (N).

(3) Update: in this step the detail signal $d\,[n]$ and update operator $\mathbf{U} = [u_1, u_2, \ldots, u_N']$ are used to update the set $xe\,[n]$ for reducing the frequency aliasing effect. The approximation signal $c\,[n]$ is the result of this step as shown in the following equation:

$$c\,[n] = xe\,[n] + \mathbf{U}\,(d\,[n])\,\mathbf{P}\,(xe\,[n]) \qquad (2)$$

The iterative operation of $c\ [n]$ is used to decompose the signals by lifting wavelet transform. The lagrange interpolation formula can calculate the prediction coefficients. The value of update coefficient becomes half of the prediction coefficient in case the length of the update operator is equal with the prediction operator. The transformed signal will be half of its length in the previous layer for that the conventional lifting wavelet transform used the down sampling algorithm to frequency aliasing. The frequency aliasing cannot be eliminated but can reduce it in update algorithm [14].

3. DNA Encoding

A DNA sequence consists of four different basic nucleotides, called Adenine, Cytosine, Thymine, and Guanine (i.e., A, C, T, and G), but the pairing is allowed just between A and T and C and G only. The DNA bases (A, C, G, and T) are used to encode the digit pairs 00, 01, 10, and 11 that are also complementary. The twenty-four types can be obtained of this coding scheme, but the pairing rules must satisfied in the program, which A bases must be paired with T and C must be paired with G, for that only eight types of coding are effective. These are shown in Table 1 [15-17].

Table 1. The rules of DNA encoding

	Rule No.							
	1	2	3	4	5	6	7	8
A	00	00	01	01	10	10	11	11
C	01	10	00	11	00	11	01	10
G	10	01	11	00	11	00	10	01
T	11	11	10	10	01	01	00	00

The color image consist of three channels, each pixel in one channel can be represented in binary number with eight bits. These bits are encoded into four bases by using one of DNA encoding rule. After the process of DNA encoding, the methods of confusion and diffusion DNA will be used in encryption steps. DNA method consider as an excellent method in text and image encryption.
Traditionally, the same encoding rule is applied to all pixels in one image which make the DNA encoding scheme is fixed. For that, some improvements have been made to use different rules in encryption; this called dynamic DNA encoding [18].

Many researchers have proposed many development of DNA computing like DNA sequences algorithm depend on arithmetic operations (addition, subtraction and XOR operations). These operations corresponds to eight different types of DNA encoding schemes [19, 20]. The XOR operation shown in Table 2 [19].

Table 2. The XOR operation for DNA sequences based on Rule 1

	A	C	G	T
A	A	C	G	T
C	C	A	T	G
G	G	T	A	C
T	T	G	C	A

4. Proposed method

The proposed method used frequency domain by applying integer lifting wavelet transform on cover image. The transform was applied on blue channel of the color cover image to obtain four sub-band (LL, LH, HL, and HH). The DNA encoding applied on two most significant bits in LL sub-band to get one base in each pixel. The produced DNA sequence used for two purposes, firstly to extract key for encryption the secret data by applying XOR operation between the key and DNA encoding of secret data. While the second one used to select the pixels in (LH, HL, and HH) to hide the data within it. The selection process is depend on the base with max redundancy and used its pixels locations to hide secret bits in (LH, HL, and HH) sub-bands pixels.

This way to ensure secrete bits is hiding not in sequence pixels, which increase the security. Final step in the proposed method is applied inverse integer wavelet transform to get the stego image. The details of embedding process shown in Figure 1.

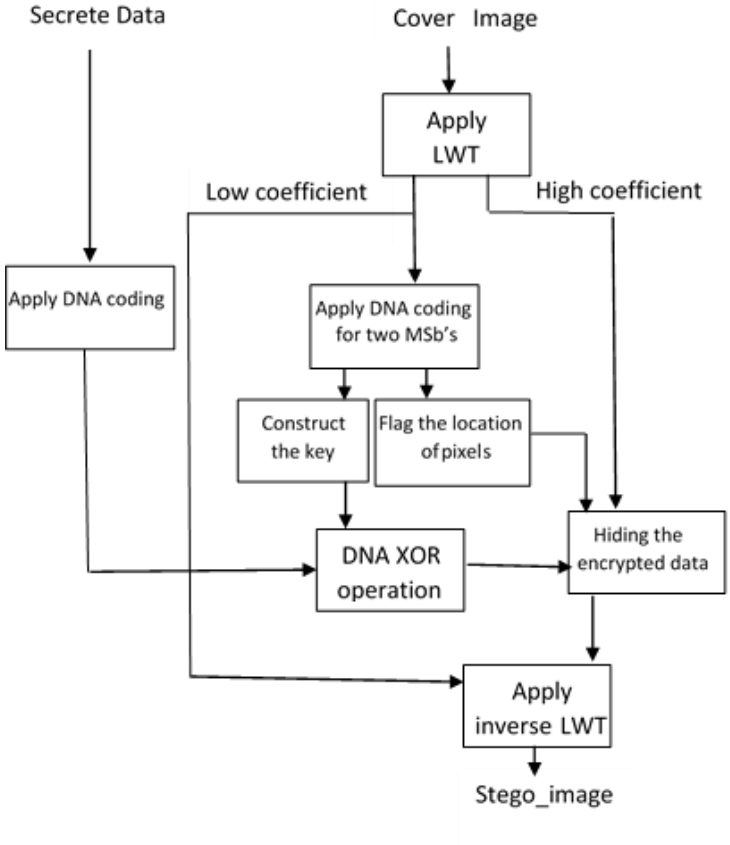


Figure 1. The block diagram of proposed method

4.1 Embedding process

The text encryption and embedding in an image steps are as follows:

Input: cover image *I* of size [M×N], secret data of size [K] bits.
Output: stego_imge *I'*.

Step 1: apply integer lifting wavelet transform to blue channel of cover image *I*, which results in four sub-band (LL, LH, HL, and HH) of size M/2 × N/2 each.
Step 2: apply DNA encoding for two most significant bits in each pixel in LL sub- band with consequence rules using table (1).
Step 3: construct the encryption key called enc_key of size K/2 from the DNA sequence, which obtained from step 2.
Step 4: flag the locations pixels with high redundancy of base to use it locations for hiding data in high coefficient sub-band.
Step 5: encoding the secret data with DNA encoding consequence rules using table (1) to produce DNA sequence with size K/2.
Step 6: XOR the DNA sequence of secret data with the enc_key to get enc_data using table (2).
Step 7: convert the enc-data to binary using table (1) then hiding it in (HL, LH, and HH) sub-bands depend on flagged location of pixels.
Step 8: Apply inverse integer lifting transform on LL coefficients and the modified LH, HL, and HH coefficients.
Step 9: Combine all the three color channels to get stego_image (*I'*).
End.

4.2 Data extraction process

The extraction process is blind that mean the existence of cover image not required. The extracting process is similar to that of embedding process in the reversed order. The following steps describe the details of this process.

Input: stego_ image *I'* of size [M×N].
Output: secret data of size [K] bits.

Step 1: apply integer lifting wavelet transform to blue channel of stego_image *I'* which results in four sub-band (LL', LH', HL', and HH') of size M/2 × N/2 each.
Step 2: apply DNA encoding for two most significant bits in each pixel in LL' sub-band with consequence rules using table (1).
Step 3: construct the encryption key called enc_key of size K/2 with the DNA sequence, which obtained from step 2.
Step 4: flag the locations pixels with high redundancy of base to use it for extract data in high coefficient sub-band.
Step 5: extract the least significant bit from the pixels of high coefficient depend on flagged locations
Step 6: encoding the extracted data with DNA encoding consequence rules using table (1).
Step 7: XOR the DNA sequence of extracted data with the enc_key using table (2).
Step 8: decode the result of step 7 using table (1) to convert it to binary.
End.

5. Performance measurement

The metrics used to measure the performance of proposed method are as follows:

1) Peak Signal to Noise Ratio (PSNR)

This term is defined via the mean square error (MSE) as in the following equation [21, 22]:

$$PSNR = 10\ log_{10}\frac{R^2}{MSE} \quad (3)$$

$$\boldsymbol{MSE = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}[I(i,j) - S(i,j)]^2} \quad (4)$$

Where m and n are represent the image size, R is the maximum value of the image's pixels. The MSE represents the pixel differences between two images (cover image (*I*), and stego-image (*S*)).

2) Structural Similarity Index Measure (SSIM)

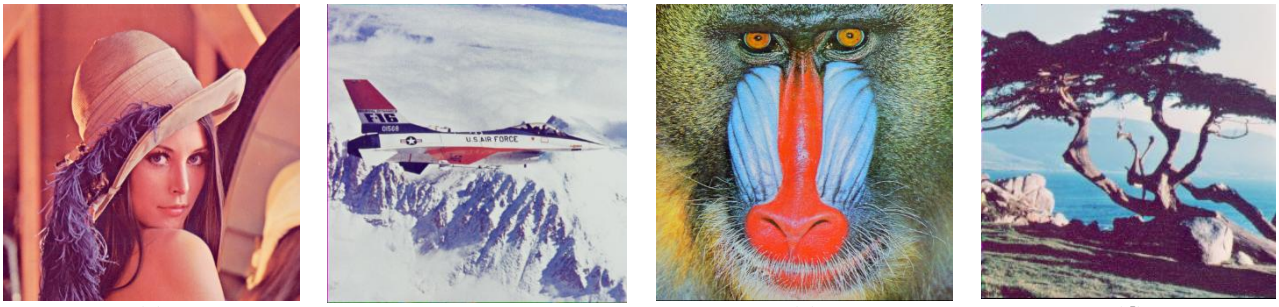
The SSIM is one of the most powerful methods of performance evaluation. It is a full reference matric; that mean it is based on initial or a reference image (distortion free image). The higher SSIM means high closeness between the two images, and it's calculated as the follows [23]:

$$\text{SSIM}\ (x, y) = \frac{(2\ \mu_x\ \mu_y + C_1)\ (2\ \sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (5)$$

Where μ_x , μ_y is the average x and y respectively; σ_x , σ_y is the variance of x and y respectively; σ_{xy} the correlation coefficient of x and y; $c_1 = (k_1 L)^2$; $c_2 = (k_2 L)^2$; $L = (2^{\text{no.of bits per pixel}}) - 1$; $[k_1, k_2] = [0.01, 0.03]$ by default.

6. Results and discussions

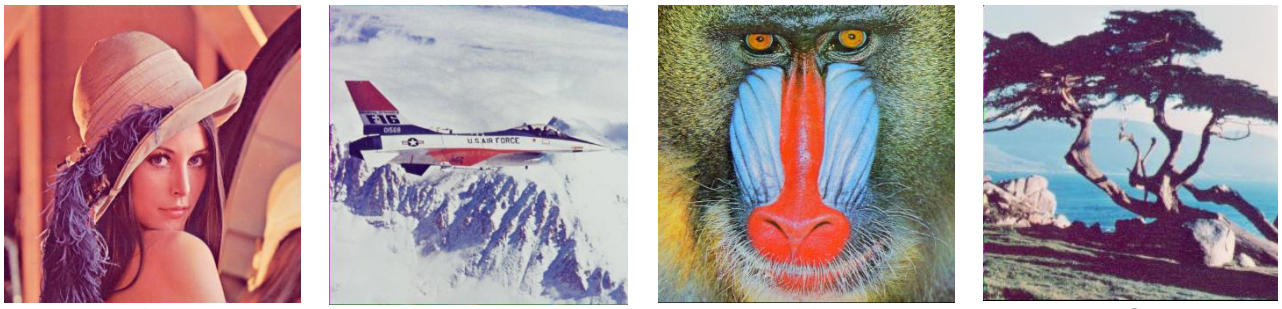
In this paper, many color images from USC-SIPI database are considered for the experiment. These images with the size 256×256 such as (a) Lena, (b) Airplane, (c) Baboon, (d) Tree as shown in Fig. 2



(a) (b) (c) (d)

Figure 2. The cover image (a) Lena, (b) Airplane, (c) Baboon, (d) Tree

The embedding process was applied in blue channel of cover images. After applying integer wavelet transform the embedding process was done on (LH, HL, and HH) sub-ands. The pixels in detail coefficients are selected to embedded depending on the DNA computation on LL sub-band. The stego images with secrete payload 30,000 bits is shown in Fig.3



(a) (b) (c) (d)

Figure 3. The cover image (a) Lena, (b) Airplane, (c) Baboon, (d) Tree

Many evaluation parameters are used to evaluate the proposed method performance such as PSNR, MSE, and SSIM with different embedding rate (ER). The embedding rate (ER) is used to represent the percentage of the embedded secret bits in the cover image pixels as shown in the follows equation:

$$ER = K / (M \times N) \text{ bpp} \quad (6)$$

Where K is the total number of the embedded secret bits in the cover image, and M × N is the size of the cover image.

The steganography method has good performance if it deals with a large value of ER. On the contrary, it has worse performance if it deals with small value of ER. The quality metrics of the proposed method depend on different ER are shown in Table 3.

Table 3. Quality metrics (PSNR, MSE, and SSIM) of the proposed method

Image	Secret payload (bits)	Embedding Rate (bpp)	PSNR (dB)	MSE	SSIM
Lena	15,000	0.228	57.755	0.109	0.9992
	30,000	0.457	52.492	0.366	0.9974
	45,000	0.686	49.999	0.650	0.9955
Airplane	15,000	0.228	58.414	0.093	0.9991
	30,000	0.457	52.861	0.336	0.9971
	45,000	0.686	50.175	0.624	0.9949
Baboon	15,000	0.228	57.097	0.126	0.9997
	30,000	0.457	52.063	0.404	0.9992
	45,000	0.686	49.528	0.724	0.9986
Tree	15,000	0.288	57.707	0.110	0.9994
	30,000	0.457	52.427	0.371	0.9982
	45,000	0.686	49.805	0.680	0.9969

The histogram for the blue channel of the cover images and the stego images present in Fig. 4. It can be seen from Fig. 4 the high correlation of histograms of cover and stego images.

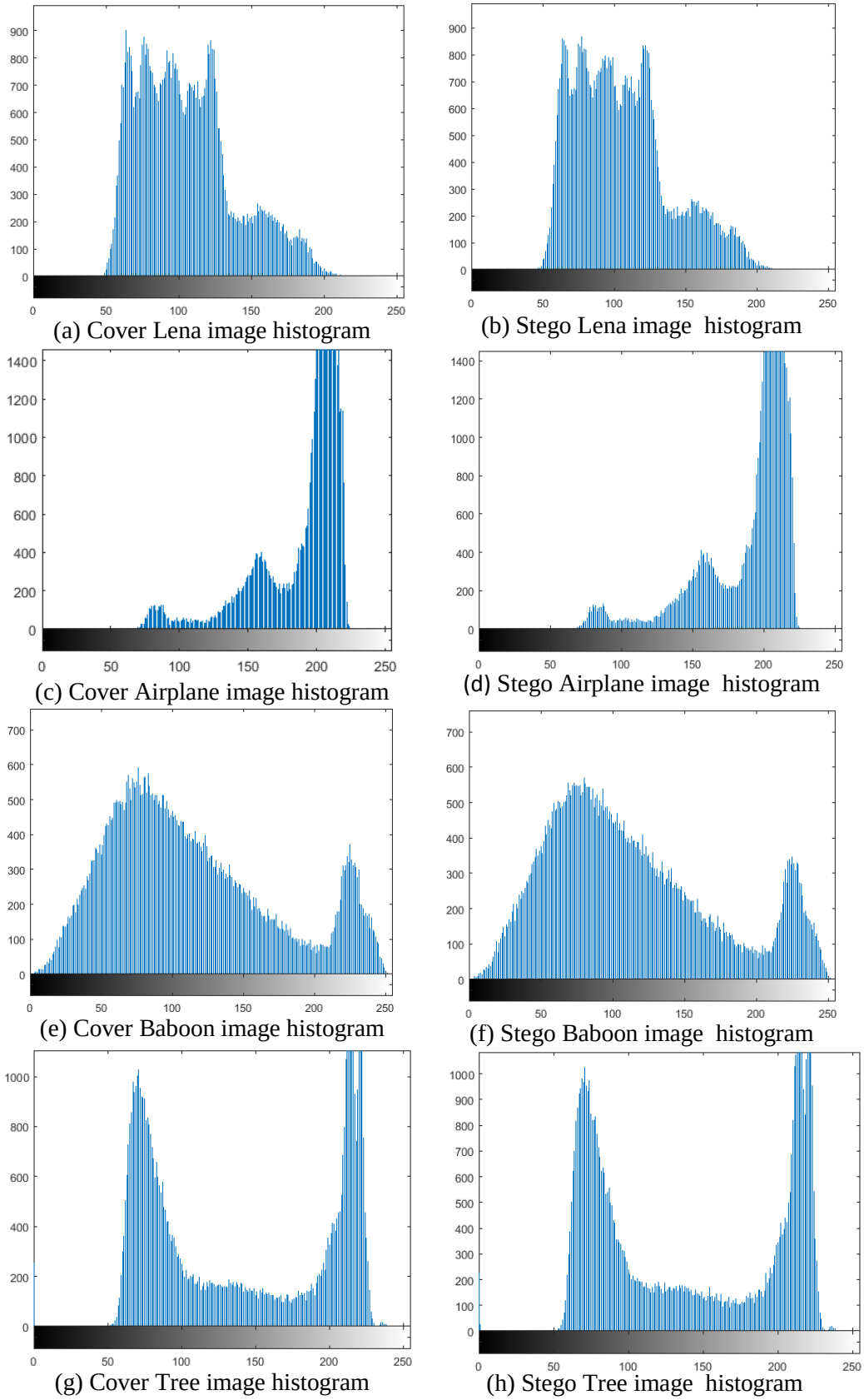


Figure 4. The histogram of cover and stego images

7. Conclusions

In steganography, there are three important factors, quality of stego image, and security of secret data finally the hiding capacity of cover image. The proposed algorithm encrypted the secret data by DNA coding using the key extracted from the cover image. The data is hiding in integer lifting wavelet coefficients. The purpose of using the integer lifting wavelet transform is to guarantees complete reversibility of data. This work provides new way to select the pixels to hide the data in it by using DNA computing to ensure more security than the hiding process in sequence pixels. Experimental results show good results depending on the high value of PSNR and SSIM.

References

[1] P. Malathi, M. Manoaj, R. Manoj, V. Raghavan, and R. E. Vinodhini, "Highly Improved DNA Based Steganography," *Procedia Comput. Sci.*, vol. 115, pp. 651–659, 2017.

[2] A. Khalifa, "LSBase: A key encapsulation scheme to improve hybrid crypto-systems using DNA steganography," *Proc. - 2013 8th Int. Conf. Comput. Eng. Syst. ICCES 2013*, no. September, pp. 105–110, 2013.

[3] F. A. Abdullatif, A. A. Abdullatif, and A. Al-Saffar, "Hiding Techniques for Dynamic Encryption Text based on Corner Point Hiding Techniques for Dynamic Encryption Text based on Corner Point," *IOP Conf. Ser. J. Phys.*, vol. 1003, no. 012027, p. 10, 2018.

[4] E. I. Abd El-Latif and M. I. Moussa, "Information hiding using artificial DNA sequences based on Elliptic Curve," *J. Inf. Optim. Sci.*, vol. 40, no. 6, pp. 1181–1194, 2019.

[5] F. E.Ibrahim, M. I. Moussa, and H. M. Abdalkader, "A Symmetric Encryption Algorithm based on DNA Computing," *Int. J. Comput. Appl.*, vol. 97, no. 16, pp. 41–45, 2014.

[6] A. A. A. Latef, "Color Image Steganography Based on Discrete Wavelet and Discrete Cosine Transforms," vol. 24, no. 3, 2011.

[7] R. J. Essa, N. A. Z. Abdullah, and R. D. Al-dabbagh, "Steganography Technique using Genetic Algorithm," *Iraqi J. Sci.*, vol. 59, no. 3A, pp. 1312–1325, 2018.

[8] A. K. Gulve and M. S. Joshi, "An image steganography method hiding secret data into coefficients of integer wavelet transform using pixel value differencing approach," *Math. Probl. Eng.*, vol. 2015, 2015.

[9] A. Amsaveni and P. T. Vanathi, "Reversible data hiding based on radon and integer lifting wavelet transform," *Inf. MIDEM*, vol. 47, no. 2, pp. 91–99, 2017.

[10] S. Lee, C. D. Yoo and T. Kalker, "Reversible Image Watermarking Based on Integer-to-Integer Wavelet Transform," in IEEE Transactions on Information Forensics and Security, vol. 2, no. 3, pp. 321-330, Sept. 2007.

[11] A. Shaik and V. Thanikaiselvan, "Comparative analysis of integer wavelet transforms in reversible data hiding using threshold based histogram modification," *J. King Saud Univ. - Comput. Inf. Sci.*, 2018.

[12] D. Chen, Y. Li, H. Zhang, and W. Gao, "Invertible update-then-predict integer lifting wavelet for lossless image compression," *EURASIP J. Adv. Signal Process.*, vol. 2017, no. 1, pp. 1–9, 2017.

[13] V. Dhandapani and S. Ramachandran, "Power-optimized log-based image processing system," *Eurasip J. Image Video Process.*, vol. 2014, no. 1, pp. 1–15, 2014.

[14] L. Duan, Y. Wang, J. Wang, L. Zhang, and J. Chen, "Undecimated lifting wavelet packet transform with boundary treatment for machinery incipient fault diagnosis," *Shock Vib.*, vol. 2016, 2016.

[15] J. Zhang, D. Hou, and H. Ren, "Image Encryption Algorithm Based on Dynamic DNA Coding and Chen ' s Hyperchaotic System," *Math. Probl. Eng.*, vol. 2016, 2016.

[16] T. T. Zhang, S. J. Yan, C. Yan Gu, R. Ren, and K. X. Liao, "Research on Image Encryption Based on DNA Sequence and Chaos Theory," *J. Phys. Conf. Ser.*, vol. 1004, no. 1, 2018.

[17] B. Wang, Y. Xie, S. Zhou, X. Zheng, and C. Zhou, "Correcting Errors in Image Encryption Based on DNA Coding," *Molecules*, vol. 23, no. 8, pp. 9–17, 2018.

[18] Z. Liu, C. Wu, J. Wang, and Y. Hu, "A Color Image Encryption Using Dynamic DNA and 4-D Memristive Hyper-Chaos," *IEEE Access*, vol. 7, no. June, pp. 78367–78378, 2019.

[19] X. Zhang, F. Han, and Y. Niu, "Chaotic Image Encryption Algorithm Based on Bit Permutation and Dynamic DNA Encoding," *Comput. Intell. Neurosci.*, vol. 2017, p. 11, 2017.

[20] X. Li, C. Zhou, and N. Xu, "A Secure and Efficient Image Encryption Algorithm Based on DNA Coding and Spatiotemporal Chaos," *Int. J. Netw. Secur.*, vol. 20, no. 1, pp. 110–120, 2018.

[21] A. A. Abdullatif, F. A. Abdullatif, and S. A. Naji, "An enhanced hybrid image encryption algorithm using Rubik's cube and dynamic DNA encoding techniques," *Period. Eng. Nat. Sci.*, vol. 7, no. 4, pp. 1607–1617, 2019.

[22] S. I. M. Ali, M. G. Ali, L. Abd, Z. Qudr, and S. I. M. Ali, "PDA : A private domains approach for improved msb steganography image," *Period. Eng. Nat. Sci.*, vol. 7, no. 3, pp. 1405–1411, 2019.

[23] A. Bovik, Z. Wang, and H. Sheikh, *Structural Similarity Based Image Quality Assessment. Digital Video Image Quality and Perceptual Coding, Ser. Series in Signal Processing and Communications.*, no. November. 2005.