An Image Steganography Algorithm for Hiding Data Based on HDWT, LZW and OPAP

A. Y. Al-Ashwal[(1)], A. H. M. Al-Mawgani[(1)], Waled Hussein Al-Arashi[(1,*)]

Abstract

Image steganography is the art of information hiding, which embeds a secret data into a cover image. However, high capacity of secret data and high quality of stego image are key issues in image steganography. In this paper, an image steganography technique based on Haar Discrete Wavelet Transform (HDWT), Lempel Ziv Welch (LZW) algorithm and Optimal Pixel Adjustment Process (OPAP) is proposed. The HDWT is used to increase the robustness of stego image against attacks. To increase the hidden capacity, LZW algorithm is performed on the secret message. The OPAP is then applied to reduce the embedding error between the cover image and stego image. The experimental results are evaluated by four standard images as covers, and with two types of secret messages. The results demonstrate high visual quality of stego image with large Hidden Capacity (HC) of secret data compared with recent techniques.

Keywords: Steganography, HDWT, LZW, OPAP, Embedding algorithm, Extraction algorithm.

1. Introduction

With the development of communication technologies and the growth of computers and the Internet, a large amount of important data is transmitted in different parts of the world. However, the privacy and security remain an open issue and become a great challenge. To meet this challenge, there are different methods that have been developed, such as cryptography and steganography. Cryptography focuses on the preservation of the contents of a secret message, while steganography focuses on the preservation of the existence of a secret message [1, 2].

Steganography is a technique that hides a secret message within another message [3]. The secret message could be hidden in digital images, audios,

[1] Department of Electronic Engineering, Faculty of Engineering, University of Science and Technology, Sana'a, Yemen

[*] Correspondence Author: w.alarashi@ust.edu

DOI: 10.20428/JST.20.1.2

videos, and other digital files. These digital files are referred as cover files. The result from the combination between cover file and the secret message is called stego [4, 5]. Digital images are widely used in the field of steganography due to the popularity of digital image transmission via the Internet.

Image steganography is performed using various techniques. These techniques are classified into two main approaches; spatial-domain and transform-domain. The spatial domain steganography techniques refer to methods in which the secret message is hidden directly in pixel values of the cover image. However, the spatial domain techniques have a low robustness against attacks and not secure methods [2, 4, 6-8].

The transform domain refers to the methods that transform the image from the spatial domain to another domain, such as a frequency domain. Then the data is embedded in the transformed image’s pixels rather than direct pixels [7, 9]. Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) are the commonly used transforms techniques to convert the image to frequency domain [5, 8, 9]. Transform domain techniques are more secure and more robust to attacks. However, the size of secret data that can be embedded into a cover image is less than the spatial domain [5, 10]. In addition, in both domains, large size of secret data leads to less quality of stego image.

DWT has been recently used by researchers for image steganography due to its merits. It has a good capacity and high robustness against many image-processing operations such as compression, cropping, blurring, sharpen and noise [10-14]. To get a high capacity of secret data and at the same time a high quality of stego image with DWT, many researchers encode and compress the secret message using various encoding algorithms [15-17]. In this paper, Lempel Ziv Welch (LZW) algorithm is used, because it reduces effectively the redundant data of the secret message. In addition, it is a simple and a lossless compression algorithm. Although, LZW increases the size of secret data, the quality of stego image needs to be enhanced. To do so, Optimal Pixel Adjustment Process (OPAP) is applied to reduce the embedding error between the cover image and stego image. Thus, a new algorithm which combines DWT, LZW and OPAP is proposed in this paper. A large size of secret data with higher quality compared with the state of the art is obtained.

The rest of this paper is organized as follows. Section 2 gives a briefer view of

DWT. An overview of OPAP and LZW are introduced in section 3 and section 4 respectively. The proposed method is described in section 5. Experimental results and discussion are presented in section 6. Finally, section 7 includes the conclusions.

2. Discrete Wavelet Transform (DWT)

The wavelet transform is based on small waves, called wavelets. They come with varying frequencies and limited duration. Let $\varphi(x,y)$ is the low-frequency component called a scaling function. $\psi^H(x,y)$, $\psi^V(x,y)$ and $\psi^D(x,y)$ are the high-frequency components called wavelet functions. The ψ^H measures variations along columns (horizontal edges), ψ^Vresponds to variations along rows (vertical edges) and ψ^D corresponds to variations along diagonals (diagonal edges) [18, 19].
The two-dimensional scaling and wavelet functions are defined as in Equations (1) and (2) [18, 19]:

$$\varphi_{j_0,m,n}(x,y) = 2^{\frac{j_0}{2}}\,\varphi\left(2^{j_0}x - m, 2^{j_0}y - n\right) \qquad (1)$$

$$\psi^i_{j,m,n}(x,y) = 2^{\frac{j}{2}}\,\psi^i\left(2^j x - m, 2^j y - n\right), \quad i = \mathrm{H, V\ or\ D} \qquad (2)$$

Where index *i* identifies the directional wavelets. *j₀* is an arbitrary starting scale. Normally, j_0= zero and $j \geq j_0$. *m* and *n* determine the position of scaling function (φ) and wavelet function (ψ) along the x-axis and y-axis respectively. *x* and *y* are image coordinates. The Two-Dimensional DWT (2-D DWT) of image *f(x,y)* with size *M*×*N* is written using the scaling and wavelet functions as in Equations (3) and (4) [18, 19]:

$$W_\varphi(j_0,m,n) = \frac{1}{\sqrt{M \times N}} \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\ \varphi_{j_0,m,n}(x,y) \qquad (3)$$

$$W^i_\psi(j,m,n) = \frac{1}{\sqrt{M \times N}} \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\ \psi^i_{j,m,n}(x,y), i = H, V\ or\ D \qquad (4)$$

Where $W_\varphi(j_0,m,n)$ coefficients define an approximation of *f(x,y)* at scale j_0. $W^i_\psi(j,m,n)$ coefficients add horizontal, vertical, and diagonal details for scales $j \geq j_0$.

The Inverse Discrete Wavelet Transform (IDWT) is defined as in Equation (5) [18, 19]:

$$f(x,y) = \frac{1}{\sqrt{M \times N}} \sum_{m} \sum_{n} W_{\varphi}(j_0, m, n)\ \varphi_{j_0,m,n}(x,y) + \frac{1}{\sqrt{M \times N}} \sum_{i=H,V,D} \sum_{j=j_0}^{\infty} \sum_{m} \sum_{n} W_{\psi}^{i}(j,m,n)\ \psi_{j,m,n}^{i}(x,y) \quad (5)$$

The 2-D DWT can be implemented using two digital channel filters and a group of down-samplers, as shown in Figure (1) [20]. The digital filters used are Low-Pass Filter (LPF) and High-Pass Filter (HPF). Four different sub-images are obtained by applying First-Level (1-L) 2-D DWT on the image; LL1 is approximation coefficients and LH1, HL1 and HH1 are details coefficients [12]. The most commonly used filters are the Haar Discrete Wavelet Transform (HDWT) [20].

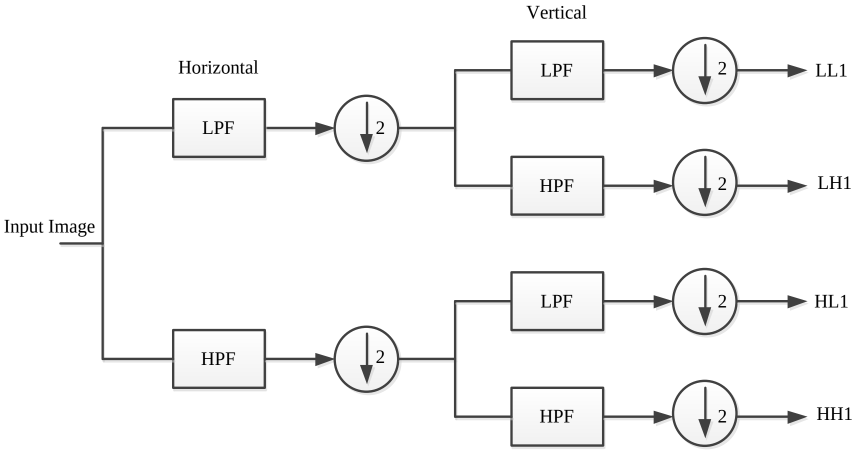


Figure [1]: 2-D discrete wavelet transform

3. Optimal Pixel Adjustment Process (OPAP)

The main aim of OPAP is to minimize the embedding error between cover image and stego image. This leads to improve the quality of stego image. For example, let the 8-bits pixel value of the cover image is 00010000 (decimal 16), and the vector of the four bits secret message is 1111. After embedding the vector secret message bits in LSBs of the cover image pixel, the pixel value is changed to 00011111 (decimal 31) and the embedding error is 15. After applying OPAP algorithm, the fifth bit is changed from one to zero. The pixel value becomes 000**0**1111 (decimal 15), and the embedding error is reduced to one [11, 21].

Let p_i is the pixel value of i^{th} pixel in the cover image.

$\acute{p}_i$ is the pixel value of i^{th} pixel in the stego image that obtain by the direct replacement of the *k*-LSB of p_i with *k*-length of secret message bits.

$\acute{\acute{p}}_i$ is the pixel value of i^{th} pixel in the refined stego image that obtain after OPAP.

$\delta_i = \acute{p}_i - p_i$ is the embedding error between p_i and $\acute{p}_i$ according to the simple LSB embedding. Therefore, the embedding error is as in Equation (6):

$$-2^k < \delta_i < 2^k \qquad (6)$$

The OPAP algorithm can be described as in Equations (7), (8) and (9) [21]:

Case 1: If $(2^{k-1} < \delta_i < 2^k)$ and $(\acute{p}_i \geq 2^k)$, then $(\acute{\acute{p}}_i = \acute{p}_i - 2^k)$; else $(\acute{\acute{p}}_i = \acute{p}_i)$. (7)

Case 2: If $(-2^{k-1} \leq \delta_i \leq 2^{k-1})$, then $(\acute{\acute{p}}_i = \acute{p}_i)$. (8)

Case 3: If $(-2^k < \delta_i < -2^{k-1})$ and $(\acute{p}_i < 256 - 2^k)$, then $(\acute{\acute{p}}_i = \acute{p}_i + 2^k)$; else$(\acute{\acute{p}}_i = \acute{p}_i)$. (9)

4. LZW

LZW method is a simple and a lossless compression algorithm. It is used effectively to reduce the size of the secret message. LZW does not need the dictionary table during the decompression process [15, 22, 23]. The secret message is encoded and compressed using LZW to reduce the size of the secret message. This process increases the hidden capacity and enhances the quality of stego image.

5. The proposed method

The proposed method consists of two algorithms; embedding algorithm and extraction algorithm. In the embedding algorithm, the cover image is decomposed by using 2-D HDWT into four sub-bunds. Besides, the secret message is encoded and compressed using LZW. After that, the encoded secret message bits are embedded into k-LSBs of cover image coefficients. Then, The OPAP is applied to decrease the difference error between the cover image and the stego image. Finally, the inverse HDWT is applied to obtain the stego image.

In the extraction algorithm, the stego image is decomposed by using 2-D HDWT into four sub-bunds. Then, the encoded secret message is extracted from k-LSBs of stego image coefficients. Finally, the LZW decompression is applied to obtain the secret message.

a. Embedding Algorithm

The process of the proposed embedding algorithm is described as follows:

Input: Cover image, Secret message

Output: Stego image

Steps:

1. Read the standard cover image.
2. Apply 2-D HDWT for the cover image.
3. Read the secret message.
4. Apply LZW algorithm to the secret message.
5. Embed encoded secret message bits in k-LSBs of cover image coefficients.
6. Apply OPAP to stego image coefficients.
7. Apply inverse 2-D HDWT to obtain the stego image.

The flowchart of the embedding algorithm is illustrated in Figure (2).

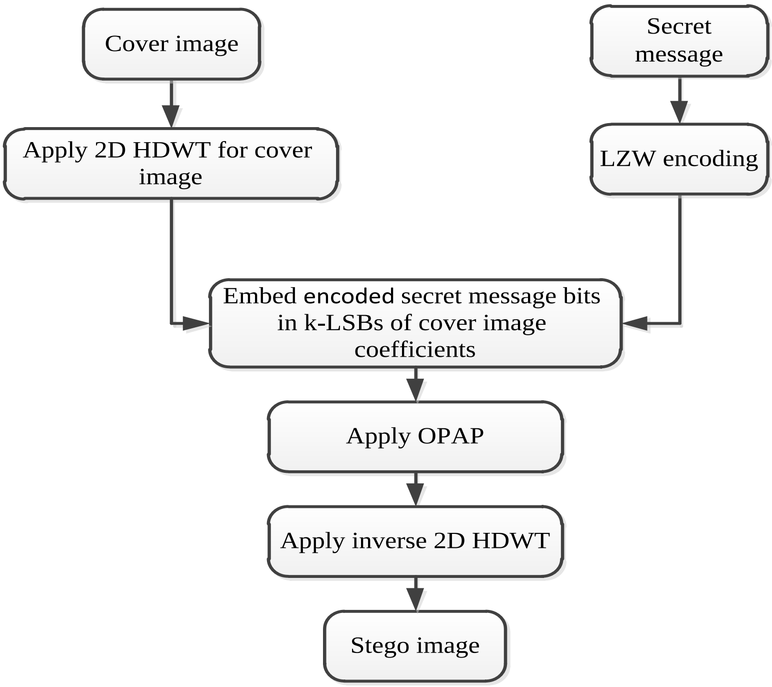

Cover image
Secret message
Apply 2D HDWT for cover image
LZW encoding
Embed encoded secret message bits in k-LSBs of cover image coefficients
Apply OPAP
Apply inverse 2D HDWT
Stego image

Figure [2]: Flowchart of the embedding algorithm

b. Extraction Algorithm

To extract the secret message from the stego image, the same steps of embedding algorithm are applied. The following steps describe the extraction algorithm.

Input: Stego image
Output: Cover image, Secret message
Steps:

1. Read the stego image.
2. Apply 2-D HDWT for the stego image.
3. Extract the encoded secret message bits from k-LSBs of stego image coefficients.
4. Apply LZW decompression to obtain the secret message.

The flowchart of the extraction algorithm is illustrated in Figure (3).

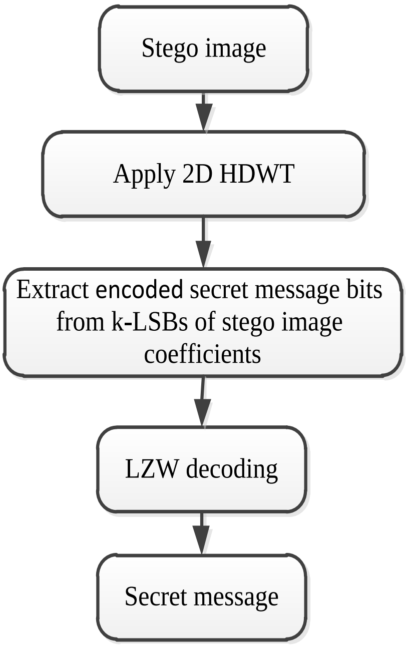

Stego image
Apply 2D HDWT
Extract encoded secret message bits from k-LSBs of stego image coefficients
LZW decoding
Secret message

Figure [3]: Flowchart of the extraction algorithm

6. Experimental Results and Discussion

The proposed algorithm has been tested by using four standard gray scale images as a cover image. These standard gray scale images are “Lena”, “Jet”, “Baboon”, and “Boat” with size 512×512 pixels. Regarding the secret messages, two types of messages are used. The first one is a standard gray scale image, “Cameraman” with size 512×256 pixels. The second secret message is a text. The size of the secret message is 1048576 bits, and the hidden capacity is 50%. The four standard cover images are illustrated in Figure (4) and the two secret messages are shown in Figure (5). The parameters that used to measure the performance of stego image are:

a) **Peak Signal to Noise Ratio (PSNR):** It measures the average cumulative squared errors between the stego image and the cover image. It is defined as in Equation (10) [24]:

$$PSNR = 10\ log_{10} \frac{M \times N \times (2^8 - 1)^2}{\sum_{x=1}^{M} \sum_{y=1}^{N} \left(C_{xy} - S_{xy}\right)^2} \qquad (10)$$

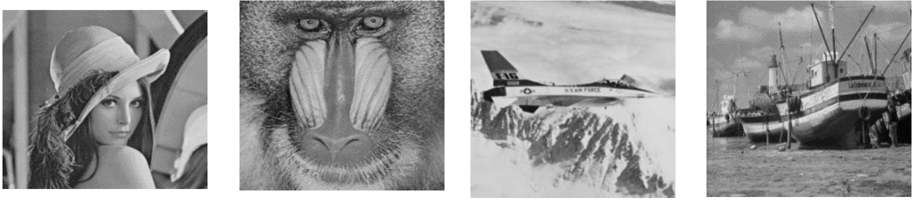
Where: *x* and *y* are the image coordinates. *M* and *N* are the dimensions of the image. C_{xy} is the cover image. S_{xy} is the stego image.

b) **Hidden Capacity (HC):** It measures the maximum size of data that is embedded in a cover image [8]. It is defined as in Equation (11).

$$HC = \frac{S_b}{C_b} \qquad (11)$$

Where: S_b is the number of secret message bits that are hidden. C_b is the number of cover image bits.

c) **Histogram analysis:** It measures the change in the stego image corresponding to cover image. The small difference between cover image and stego image shows more quality and resistance.



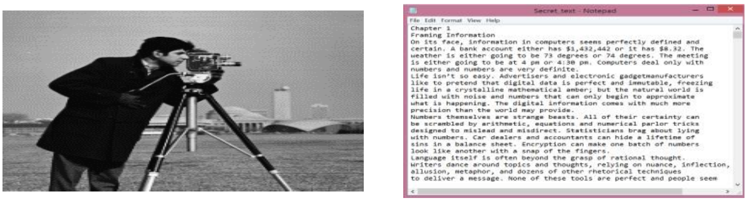
a) Cover image “Lena”

b) Cover image "Baboon"

c) Cover image “Jet”

d) Cover image “Boat”

Figure [4]: Four standard cover images



```
Chapter 1
Framing Information
On its face, information in computers seems perfectly defined and
certain. A bank account either has $1,432,442 or it has $8.32. The
weather is either going to be 73 degrees or 74 degrees. The meeting
is either going to be at 4 pm or 4:30 pm. Computers deal only with
numbers and numbers are very definite.
Life isn't so easy. Advertisers and electronic gadgetmanufacturers
like to pretend that digital data is perfect and immutable, freezing
life in a crystalline mathematical amber; but the natural world is
filled with noise and numbers that can only begin to approximate
what is happening. The digital information comes with much more
precision than the world may provide.
Numbers themselves are strange beasts. All of their certainty can
be scrambled by arithmetic, equations and numerical parlor tricks
designed to mislead and misdirect. Statisticians brag about lying
with numbers. Car dealers and accountants can hide a lifetime of
sins in a balance sheet. Encryption can make one batch of numbers
look like another with a snap of the fingers.
Language itself is often beyond the grasp of rational thought.
Writers dance around topics and thoughts, relying on nuance, inflection,
allusion, metaphor, and dozens of other rhetorical techniques
to deliver a message. None of these tools are perfect and people seem
```

a) Secret message “Cameraman”

b) Secret message "Text"

figure [5]: Two secret messages

In this experiment, the proposed algorithm has been compared with the algorithms [25] and [26]. Table (1) presents a comparison between the results of the proposed algorithm and the existing recent algorithm. PSNR and HC are used to evaluate the performance of different algorithms. It can be noted from Table (1) that using LZW with OPAP gives a visible improvement of PSNR. This is because LZW reduces the effects of redundant data in the secret message and the OPAP minimizes the embedding error between the cover image and the stego image. This leads to enhance the quality of stego image.

Table [1]: Comparison of PSNR and HC between the proposed algorithm and the recent algorithms

Cover image	Algorithm	Max. HC (bits)	Max. HC (%)	Secret message	
				Image	Text
				PSNR (dB)	PSNR(dB)
Lena	Proposed algorithm	1048576	50%	40.47	46.38
	LSB and GA [25]	1048576	50%	32.22	-
	GASDCT [26]	80000	3.8%	44.78	-
Baboon	Proposed technique	1048576	50%	39.83	44.36
	LSB and GA [25]	1048576	50%	32.53	-
	GASDCT [26]	80000	3.8%	44.92	-
Jet	Proposed technique	1048576	50%	40.34	46.37
Boat	Proposed technique	1048576	50%	40.21	45.82
	GASDCT [26]	80000	3.8%	44.86	-

Table (1) shows explicitly the proposed algorithm is better than [25] algorithm in term PSNR. The algorithm in [26] has high PSNR, but the HC for this algorithm is lower than the proposed algorithm. Therefore, the increasing of PSNR value is expected. However, this does not mean that its results is better than the results of the proposed algorithm. This is because the evaluation is considered when both criteria, PSNR and HC are used. To make sure, the proposed algorithm is tested with the same HC of [26] (3.8%). The PSNR values obtained are 53.98 with “Lena”, 47.54 with “Baboon” and 51.60 with “Boat”.

The comparison of HC and PSNR between the proposed algorithm and the recent algorithms is represented graphically to investigate the results more clearly as illustrated in Figure (6) and (7).

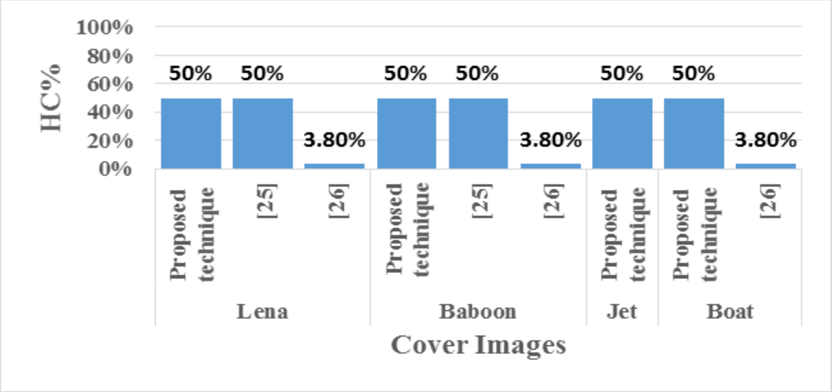
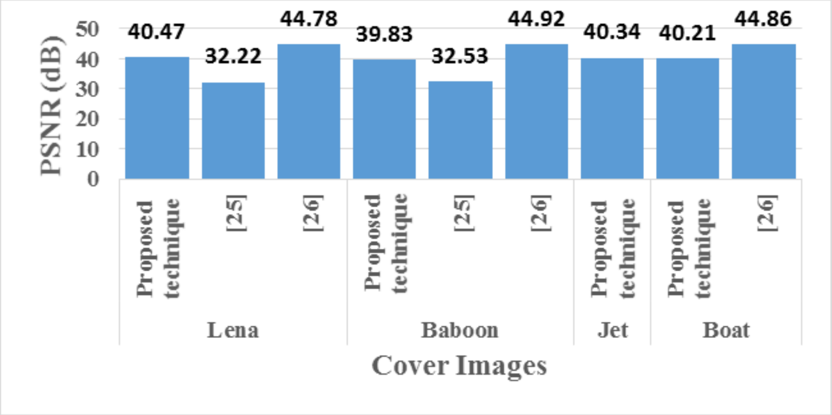
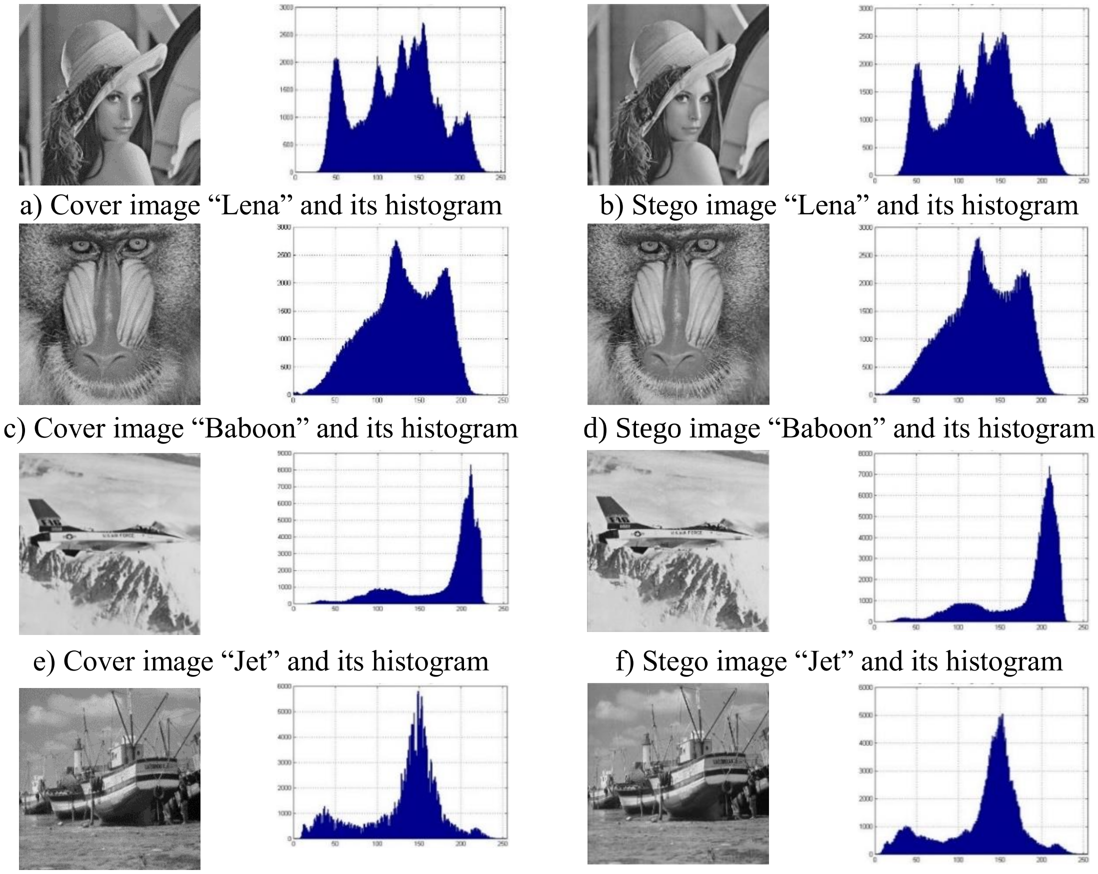


Figure [6]: Comparison of HC between the proposed algorithm and the recent algorithms



Figure[7]: Comparison of PSNR between the proposed algorithm and the recent algorithms

The histogram analysis is used to show the imperceptibility of stego image. Figure (8) shows the comparison between the histograms of the cover images and the stego images. The histogram analysis shows a small difference between the histograms of the both images as illustrated in Figure (8). This means more resistance for image steganography algorithm and more imperceptibility of stego image.



a) Cover image “Lena” and its histogram

b) Stego image “Lena” and its histogram

c) Cover image “Baboon” and its histogram

d) Stego image “Baboon” and its histogram

e) Cover image “Jet” and its histogram

f) Stego image “Jet” and its histogram

g) Cover image “Boat” and its histogram

h) Stego image “Boat” and its histogram

Figure [8]: Histograms analysis for cover images and stego images

7. Conclusion

In this paper, an image steganography algorithm is proposed based on HDWT, LZW and OPAP to increase the quality of stego image and the capacity of secret data. The LZW algorithm is employed to reduce the effects of redundant data in the secret message. This improves the hidden capacity. The OPAP is applied to decrease the difference error between the stego image and cover image. This enhances the quality of stego image. The results of the proposed algorithm show clear improvement in capacity and visual quality. The improvement in PSNR is about 3 to 9 dB.

8. References

[1] H. A. G. Al-Jbara, L. B. M. Kiah, and H. A. Jalab, "Increased capacity of image based steganography using artificial neural network," in International Conference on Fundamental and Applied Sciences (AIP Conf. Proc.), pp. 20-25, 2012.

[2] G. Chugh, "Image steganography techniques: A review article," Acta Technica Corvininesis - Bulletin of Engineering, vol. 6, no. 3, pp. 97-104, 2013.

[3] G. Kipper, Investigator's guide to steganography, Auerbach Publications a CRC Press, 2004.

[4] P. Rajkumar, R. Kar, A. K. Bhattacharjee, and H. Dharmasa, "A comparative analysis of steganographic data hiding within digital images," International Journal of Computer Applications, vol. 53, no. 1, pp. 1-6, 2012.

[5] S. Singh and A. Singh, "A review on the various recent steganography techniques," International Journal of Computer Science and Network (IJCSN), vol. 2, no. 6, pp. 142-156, 2013.

[6] V. Banoci, G. Bugar, and D. Levicky, "A novel method of image steganography in DWT domain," in 21st International Conference Radioelektronika, pp. 1-4, 2011.

[7] S. Sharda and S. Budhiraja, "Image steganography: A review," International Journal of Emerging Technology and Advanced Engineering (IJETAE), vol. 3, no. 1, pp. 707-710, 2013.

[8] R. O. El Safy, H. H. Zayed, and A. El Dessouki, "An adaptive steganographic technique based on integer wavelet transform," in International Conference on Networking and Media Convergence (ICNM), pp. 111-117, 2009.

[9] A. Tiwari, S. R. Yadav, and N. Mittal, "A review on different image steganography techniques," International Journal of Engineering and Innovative Technology (IJEIT), vol. 3, no. 7, pp. 121-124, 2014.

[10] S. Atawneh and P. Sumari, "Hybrid and blind steganographic method for digital Images based on DWT and chaotic map," Journal of Communications, vol. 8, no. 11, pp. 690-699, 2013.

[11] E. Ghasemi, J. Shanbehzadeh, and N. Fassihi, "High capacity image steganography based on genetic algorithm and wavelet transform," in Intelligent Control and Innovative Computing: Springer, pp. 395-404, 2012.

[12] A. A. Abdelwahab and L. A. Hassaan, "A discrete wavelet transform based technique for image data hiding," in 25th National Radio Science Conference (NRSC), pp. 1-9, 2008.

[13] S. Bhattacharyya and G. Sanyal, "A robust image steganography using dwt difference modulation (DWTDM)," International Journal of Computer Network and Information Security (IJCNIS), vol. 4, no. 7, pp. 27-40, 2012.

[14]H. Motamedi and A. Jafari, "A new image steganography based on denoising methods in wavelet domain," in 9th International ISC Conference on Information Security and Cryptology (ISCISC), pp. 18-25, 2012.

[15]S. Goel, P. Kumar, and R. Saraswat, "High capacity image steganography method using LZW, IWT and modified pixel indicator technique," International Journal of Computer Science and Information Technologies (IJCSIT), vol. 5, no. 3, pp. 3759-3763, 2014.

[16]A. Nag, S. Biswas, D. Sarkar, and P. P. Sarkar, "A novel technique for image steganography based on DWT and Huffman encoding," International Journal of Computer Science and Security (IJCSS), vol. 4, no. 6, pp. 561-570, 2011.

[17]K. Pol, "Image steganography based on DWT using Huffman LWZ encoding," International Journal of Engineering and Technical Research (IJETR), vol. 2, no. 3, pp. 100-103, 2014.

[18]S. Mallat, A wavelet tour of signal processing, 2nd ed., Academic press, 1999.

[19]R. C. Gonzalez and R. E. Woods, Digital image processing, 3rd ed., Prentice Hall Upper Saddle River, NJ:, 2007.

[20]P.-Y. Chen and H.-J. Lin, "A DWT based approach for image steganography," International Journal of Applied Science and Engineering, vol. 4, no. 3, pp. 275-290, 2006.

[21]C.-K. Chan and L. M. Cheng, "Hiding data in images by simple LSB substitution," Pattern Recognition, vol. 37, no. 3, pp. 469-474, 2004.

[22]M. Bashardoost, G. B. Sulong, and P. Gerami, "Enhanced LSB image steganography method by using knight tour algorithm, Vigenere encryption and LZW compression," International Journal of Computer Science Issues (IJCSI), vol. 10, no. 2, pp. 221-225, 2013.

[23]S. W. Smith, The scientist and engineer's guide to digital signal processing, 1997.

[24]M. Kutter and F. A. P. Petitcolas, "Fair benchmark for image watermarking systems," in Electronic Imaging '99. Security and Watermarking of Multimedia Contents, pp. 226-239, 1999.

[25]M. Mohamed, F. Al-Afari, and M. A. Bamatraf, "Data hiding by LSB substitution using genetic optimal key-permutation," International Arab Journal of e-Technology, vol. 2, no. 1, pp. 11-17, 2011.

[26]A. Khamrui and J. K. Mandal, "A genetic algorithm based steganography using discrete cosine transformation (GASDCT)," Procedia Technology, vol. 10, pp. 105-111, 2013.