ISSN:0975-9646

A High Capacity and Secured Color Image Steganographic Technique Using Discrete Wavelet Transformation

Juned Ahmed Mazumder, Kattamanchi Hemachandran

Department of Computer Science,
Assam University, Silchar,
Assam, India

Abstract - Steganography is the art and science of concealing information into another medium so that no one apart from the intended sender and receiver can see the information. In this paper an attempt has been made for high capacity and security steganography of color images in the domain of discrete wavelet transformation. We use Discrete Wavelet Transformation (DWT) in the process of steganography so that we can clearly identify the high frequency and low frequency information of each pixel of the image. In our proposed method we applied the Haar-DWT. We have used four different types of image file formats for analysis of our proposed method. In each of the file formats 6389 bytes of message were inserted for the purpose of steganography. For the analysis of the proposed method MSE and PSNR was calculated for each of the cover and stego image and also the RGB histogram of each of the file format was analyzed. The MSE and Capacity are improved with acceptable PSNR compared to the existing algorithm.

Keywords - Cryptography, DWT, MSE, PSNR, Steganography, Steganalysis

I. INTRODUCTION

Steganography can be dated back to 440 BC, where the tale of Demaratus sending a warning by using a wax tablet and Histiaeus using a tattoo on his slave's shaved head were mentioned by Herodotus in the Histories of Herodotus [1], [2]. Steganography is the art and science of concealing information into another medium so that no one apart from the intending sender and receiver can see the information. In other words, Steganography is the process of hiding a secret message within a larger one in such a way that no one can know the presence or contents of the hidden message. Steganography will hide the message so that there is no knowledge of the existence of the message in the place. The term Steganography is forked from the Greek words "steganos" meaning "cover" and "graphia" meaning "writing" defining it as "covered writing" [3]. In this case any digital media can be used as a carrier for the secret information like text, images, audio or video files. But among those most widely used medium is images because it takes advantage of our limited visual perception of colors and also this field expected to grow continually as computer graphics power also grows. The following figure describes the basic process of steganography.

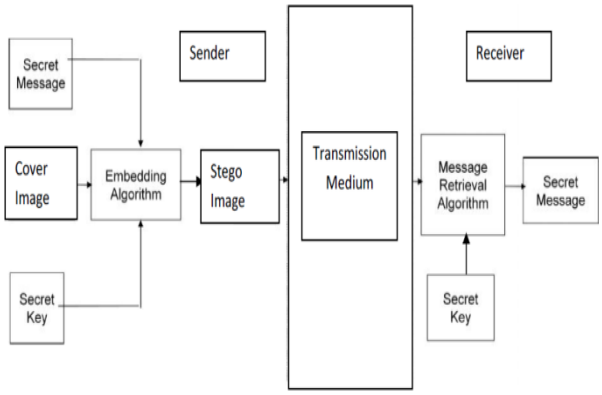

Fig. 1 A Block diagram of a generic steganographic process

The cover medium is any digital medium like image, audio or video. In this paper an attempt has formulate to study the image as cover medium. Generally, there are potentially two basic requirements for a secure image hiding system. First, the secret data embedded in the stego-image should be perceptually invisible. Secondly, the receiver can exactly recover the original data without the knowledge of the cover-image [4].

Liu and Liao [4], proposed a high-performance JPEG steganographic method that adopts the complementary embedding strategy to avoid the detections of several statistical attacks. To show the effectiveness of the proposed method, several statistical attacks are simulated and used to detect the stego-images created by this method. In this method embedding process is integrated with JPEG encoding process. The raw data of the cover-image is first transformed by DCT. The DCT coefficients are then quantized and rounded to the nearest integers.

Tong and Ding [5] proposed a steganographic method in which information is hidden into a publicly accessed color image by a quantization-based strategy, so the transportation of the secret information will not attract the attention of illegal eavesdropper. With this approach, the secret information is embedded in the wavelet domain of every chrominance component, so the hiding capacity is larger than the similar steganography software. The embedded sequence can be reliably extracted without resorting to the original image.

Fridrich and Soukal [6] proposed a matrix embedding technique for large payloads. Matrix embedding is a method for improving embedding efficiency of steganographic schemes. It involves a coding procedure can be applied to most steganographic schemes without any other changes to their embedding mechanism to increase the number of bits embedded using one embedding change. They proposed two new approaches to matrix embedding that are suitable when the embedded message length is close to the embedding capacity. The first approach is based on random linear codes of small dimension. Random linear codes provide good embedding efficiency that is fairly close to the theoretical upper bound for the class of codes of fixed length. Also, their relative embedding capacity densely covers the range of large payloads, which makes these codes suitable for practical applications. The second approach to matrix embedding for large payloads proposed in this paper is based on a family of structured codes-the simplex codes. Structured codes are more computationally efficient and can be used even for relative payloads above 0:9. Their performance for shorter payloads is however not as good as for the random codes.

Nag et al.[7] proposed a gray image steganography based on DWT, where DWT is used to transform original image (cover image) from spatial domain to frequency domain. Firstly two dimensional Discrete Wavelet Transform (2-D DWT) is performed on a gray level cover image of size M × N and Huffman encoding is performed on the secret messages/image before embedding. Then each bit of Huffman code of secret message/image is embedded in the high frequency coefficients resulted from Discrete Wavelet Transform. Image quality is to be improved by preserving the wavelet coefficients in the low frequency sub-band.

Wang et al. [8], proposed a new image steganographic technique capable of producing a secret-embedded image that is totally indistinguishable from the original image by the human eye and the new method overcomes the difficulty of falling-off-boundary problem by using pixel-value differencing and the modulus function.

Reddy and Raja [9], proposed High Capacity and Security Steganography using discrete wavelet transform. The wavelet coefficients of both the cover and payload are fused into single image using embedding strength parameters: alpha and beta. The cover and payload are preprocessed to reduce the pixel range to ensure the payload is recovered accurately at the destination. The use of wavelet in image stenographic model lies in the fact that the wavelet transform clearly separates the high frequency and low frequency information on a pixel by pixel basis. Discrete Wavelet Transform (DWT) is preferred over Discrete Cosine Transforms (DCT) because image in low frequency at various levels can offer corresponding resolution needed.

Ghasemi et al. [10], proposed a method by the application of Wavelet Transform and Genetic Algorithm in a novel Steganography scheme. They employ a genetic algorithm based mapping function to embed data in Discrete Wavelet Transform coefficients in 4x4 blocks on the cover image. The optimal pixel adjustment process is applied after embedding the message. Here frequency domain used to improve the robustness of Steganography and implement Genetic Algorithm and Optimal Pixel Adjustment Process to obtain an optimal mapping function to reduce the difference error between the cover and the stego-image, therefore improving the hiding capacity with low distortions.

Mandal and Das [11], proposed an adaptive steganography based on modified pixel-value differencing through management of pixel values within the range of gray scale pixel value difference (PVD) method is used and check whether the pixel value exceeds the range on embedding positions where the pixel exceeds boundary has been marked and a delicate handle is used to keep the value within the range. In PVD method pixel values in the stego-image may exceed the gray scale range which is not desirable as it may lead to improper visualization of the stego-image. In this paper they introduced a method to overcome this problem. In the proposed method they have used the original PVD method to embed secret data.

II. CLASSIFICATION OF STEGANOGRAPHY

Image steganography primarily classified into two categories. First one is the image based or spatial domain steganography and second is the transform domain or frequency domain steganography. In spatial domain steganography we directly deal with the pixel value of the image that is secret message bit is inserted into the image by modifying the pixel value of the image. The most common and popular spatial domain steganography is the Least Significant Bit (LSB) insertion method. For better understanding LSB method let us consider the following example where we are inserting the simple message"T" into a 3X3 image. The binary value of T is 1010100. Each bit of the secret message is inserted into each least significant bit of the image as follows

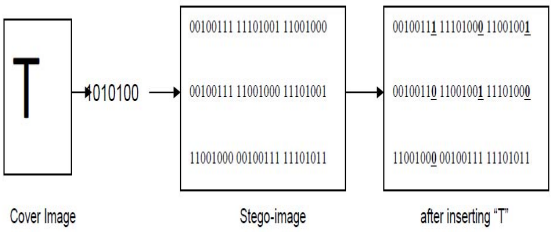

Fig. 2 LSB steganography method

In the stego image the highlighted right most bit of each pixel are replaced for inserting the message "T". The basic LSB method has a simple implementation and high capacity [12]. However it has low robustness and pros to some attacks like low-pass filtering and compression.

On the other hand in transform domain or frequency domain steganography before imbedding the secret message the cover image is transformed into its frequency domain using one of the suitable methods like Fast Furrier Transform, Discrete Cosine Transform or Discrete Wavelet Transform. In transform domain the JPEG is the most common steganography which uses DCT coefficient for embedding secret information. As we know, the JPEG compression is based on the discrete cosine transform (DCT), and reduces the visual redundancy to achieve good

compression performance. Therefore, the embedding capacity provided by JPEG steganography is relatively smaller than those provided by the other steganographic methods [13] but security in JPEG steganography is very high it is difficult to do steganalysis on stego-images which uses transform domain steganography. An attempt has been made to study the high capacity and security parameters of the steganography using discrete wavelet transformation. So if we use Discrete Wavelet Transform (DWT) we can achieve both of the important parameter for steganography that is Capacity and security, in the next section we will briefly discuss about the DWT.

III. DISCRETE WAVELET TRANSFORMATION

Discrete wavelet transform (DWT) is used to transform the image from its spatial domain into its frequency domain. We use DWT in the process of steganography so that we can clearly identify the high frequency and law frequency information of each pixel of the image. On the other hand if we tell about one dimensional DWT which is a repeated filter bank algorithm and the input is the combination of both the high pass filter and a low pass filter.

The first literature that relates to the wavelet transform is Haar wavelet. It was proposed by the mathematician Alfrd Haar in 1909. However, the concept of the wavelet did not exist at that time. Until 1981, the concept was proposed by the geophysicist Jean Morlet. Afterward, Morlet and the physicist Alex Grossman invented the term wavelet in 1984. Before 1985, Haar wavelet was the only orthogonal wavelet people know [13], [14]. The main feature of DWT is multi-scale representation of function. By using the wavelets, given function can be analyzed at various levels of resolution. The DWT is also invertible and can be orthogonal [13]. Before explaining more about wavelet transformation we have to explain some notation. We are considering an image of N X N as a two dimensional array I with N rows and N columns.

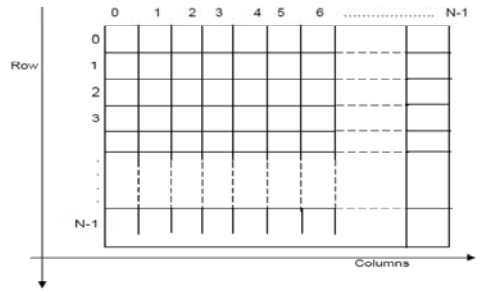

Fig. 3 Image representation as two dimensional arrays I, where the rows are enumerated from top to bottom and the columns from left to right, starting at index zero

In color images each pixel is represented by several color components. Typically there are three of them per pixel. In the RGB color space, e.g., there is one component for red, green, and blue, respectively. Other choices are the YUV color space (luminance and chrominance) and the CMYK color space (cyan, magenta, yellow, black).

In our proposed method the frequency domain transform we applied was the Haar-DWT. There are two operations comprising a 2-dimensional Haar-DWT. One is the horizontal operation and the other is the vertical one. The detailed operations of 2-D Haar-DWT is stated as follows [15]

Step 1: At first, scan the pixels from left to right in horizontal direction. Then, perform the addition and subtraction operations on neighboring pixels. Store the sum on the left and the difference on the right as illustrated in Figure 4. Repeat this operation until all the rows are processed. The pixel sums represent the low frequency part (denoted as symbol L) while the pixel differences represent the high frequency part of the original image (denoted as symbol H).

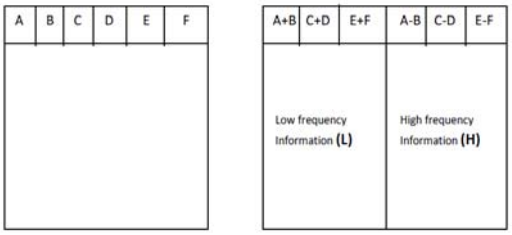

Fig. 4 Horizontal operation

Step 2: Secondly, scan the pixels from top to bottom in vertical direction. Perform the addition and subtraction operations on neighboring pixels and then store the sum on the top and the difference on the bottom as illustrated in Figure 5. Repeat this operation until all the columns are processed. Finally we will obtain 4 sub-bands denoted as LL, HL, LH, and HH respectively. The LL sub-band is the low frequency portion and hence looks very similar to the original image.

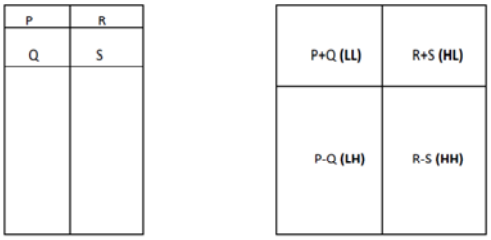

Fig. 5 Vertical operation

The process described above is known as 2-D Haar-DWT. Let us consider an example for better understanding the 2-D Haar DWT of an image.

7	6	3	2
1	0	6	7
4	5	9	3
2	1	3	7

Fig. 6 Pixel by pixel representation of a 4 X 4 image

After applying 2-D Haar-DWT we get the following figure

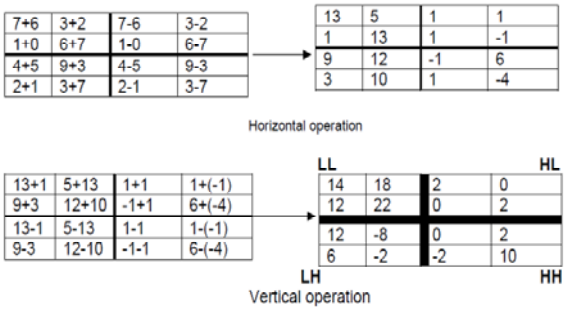

Fig. 7 Example of 2D-Haar DWT of an Image

After applying 2D-Haar DWT on a 4 X 4 image we get the four sub-bands of the image each of size are 2 X 2, these are [16]

- **LL**: approximation area that includes information of the average of the image.
- **HL**: horizontal area that includes information about the vertical edges/details in the image.
- **LH**: vertical area that includes information about the horizontal edges/details in the image.
- **HH**: diagonal area that includes information about the diagonal details, e.g., corners, in the image

After applying 2D-Haar DWT on the image "water_lily.jpg"[21] the following four sub-bands are shown in the fig. 8.

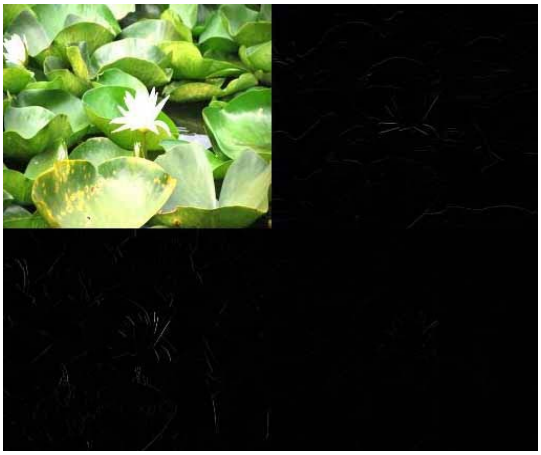

Fig. 8 D-Haar DWT of an image

IV. PROPOSED METHOD

The proposed method is basically in the domain of Haar Discrete Wavelet Transformation. In this method DWT is used to increase the security of the steganography as well as the capacity, we know that after applying DWT we get the four sub-bands of images separating the high frequency and low frequency information and if we insert our secret information in the high frequency coefficients of the sub-bands it is very difficult to detect the steganography in the stego-image. Also among the four sub-bands three of them have high frequency information and in the proposed method we have distributed our secret message among the three high frequency sub-bands. The proposed method is a color image steganography so when we convert a color image (RGB) of size 512 X 512 into its matrix form we get the three classes of the color that is red, green and blue as a result we get the matrix of coefficient of size 512 X 512 X 3. So when we apply DWT to a color image we get the four sub-bands of image each of size 256 X 256 X 3. In this proposed method we have given a secret key for embedding message into the cover image, the exact key is needed for extracting the message from the stego-image. The information of the key is stored in the HL sub-band of the image. The message embedding and extracting procedures are described in the following figure.

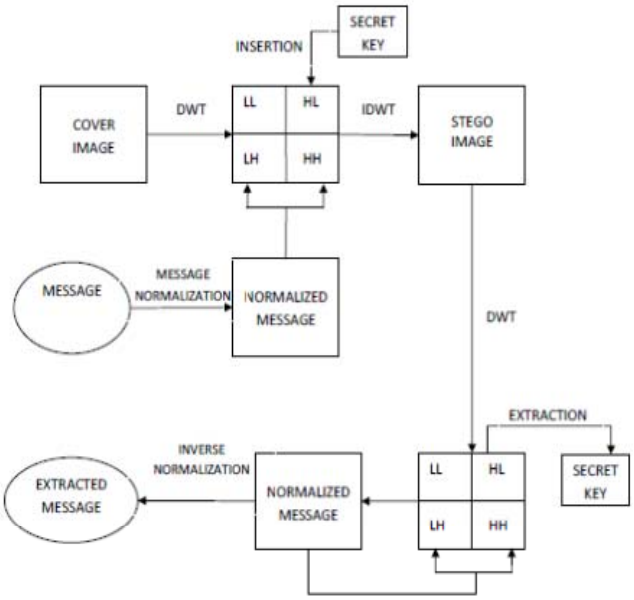

Fig. 9 Message embedding and extraction process of proposed method

A. *Message embedding algorithm*

Step1: Apply 2D-Haar DWT of the cover image, where we get the four sub-bands which separate the high and low frequency information.

Step2: Calculating the length of the message (n) and save it to the sub-bands HL.

Step3: Convert the each character of the message into its ASCII format.

Step4: Normalizing the ASCII format of the message using the following formula

$$\text{Normalized Message} = \frac{\text{ASCII value of each character of the message}}{\text{Length of the message (n)}}$$

Step5: Divide the Normalized message into two parts so that first part is to be inserted in the sub-band LH and the second part of the message in the sub-band HH.

Step6: Message is inserted into all the color components of LH and HH sub-bands that are Red, Green and Blue color components. Message is inserted starting from the last column of each of the color components from top to bottom depending upon the length of the message.

Step7: If the message length is larger than the No. of rows of each of the color components of each of the sub-bands then rest of the normalized message will go to the second last column of the each component, in this way all the normalized message is to be distributed as described in fig.10.

Step6: Insert the secret key in the HL sub-band.

Step8: At the last step we have to take the Inverse Discrete Wavelet Transformation (IDWT) so that we can get the stego-image.

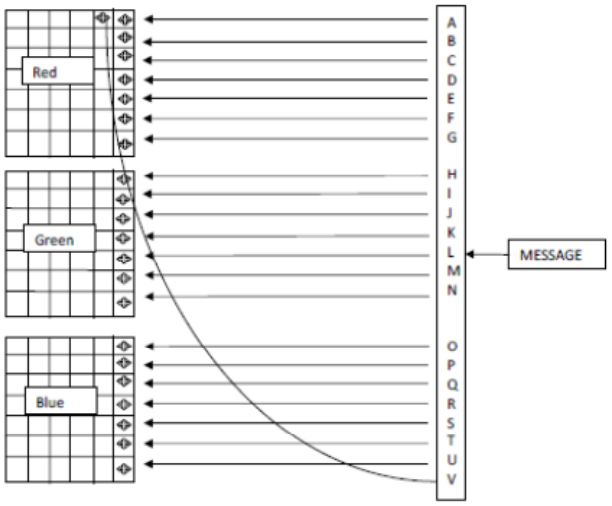

Fig. 10 Message embedding process in each color components Red, Green and Blue

B. Message extraction algorithm

Step1: Apply 2D-Haar DWT to the stego-image so as to separate all the four sub-bands that are LL, HL, LH and HH

Step2: Ask the user to inserting the secret key (k).

Step3: Extract the secret key from the HL sub-band which was inserted at the time of stego-image.

Step4: If the secret key (K) match with the saved key in the sub-band HL, then go to step 4 otherwise to step 8.

Step5: Find the length of the message from the sub-bans HL so that we can calculate from which column to which row we can extract the coefficient of the color components of each of the sub-bands LH and HH.

Step6: The inverse normalization is carried out by multiplying each of the extracted coefficients with the length of the message.

Step7: Convert each of the extracted coefficients into character format to form the message.

Step8: Display the message that the secret key did not match.

V. EXPERIMENTAL RESULT AND PERFORMANCE ANALYSIS

In this section we have presented the experimental result and evaluated the performance of the proposed method. The proposed method is implemented in the MATLAB and the operating system used is windows 7. We have used four different types of file formats for the analysis of our proposed method. The images cover and stego shown in the fig. 11. We have tested our proposed method by inserting message of size 6389 bytes to each of the four file formats. The test images before and after steganography are given bellow.

Fig. 11 Cover and stego Images

Before analyzing the result of the proposed method we have to define some parameter by which we can understand the difference between the original and the manipulated image. Any processing applied to an image may cause an important loss of information or quality [19]. In the experimental phase we have used the parameter PSNR (Peak Signal to Noise Ratio) for calculating the difference between the cover image and stego-image. Given a reference image f and a test image g, both of size $M \times N$, the PSNR between f and g is defined by [19]:

$$\mathrm{PSNR(f,g)} = 10\log_{10}\frac{255^2}{MSE(f,g)} \qquad (1)$$

Where

$$MSE(f,g) = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f_{ij} - g_{ij}\right)^2 \qquad (2)$$

The PSNR value approaches infinity as the MSE approaches zero; this shows that a higher PSNR value provides a higher image quality. At the other end of the scale, a small value of the PSNR implies high numerical differences between images. The following table shows the PSNR and MSE value of the four file formats used in this proposed method.

Table 1
Table shows the PSNR and MSE value of the four file formats used in this proposed method

Image File Name	File Format	No. Bytes Inserted	PSNR	MSE	No. Bytes Extracted
water_lily	JPEG	6389	57.143	0.125517	6389
lena	PNG	6389	53.983	0.259832	6389
eaves	BMP	6389	54.379	0.237235	6389
earth	TIFF	6389	57.109	0.126510	6389

Besides the analysis of PSNR and MSE value of test images we have analyzed the histogram of each of the images before and after steganography. An image histogram refers to the probability mass function of the image intensities. This is extended for color images to capture the joint probabilities of the intensities of the three color channels.

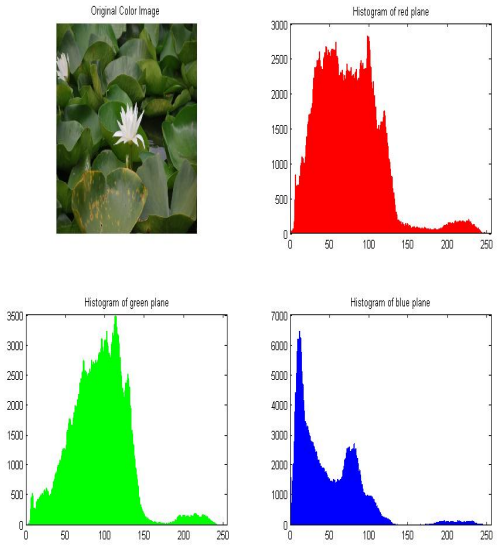

Fig. 12 Histogram of cover image water_lily.jpg

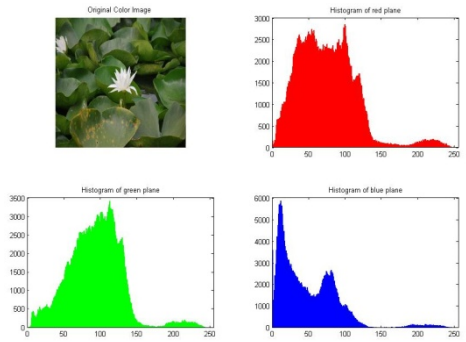

Fig. 13 Histogram of stego-image water_lily.jpg

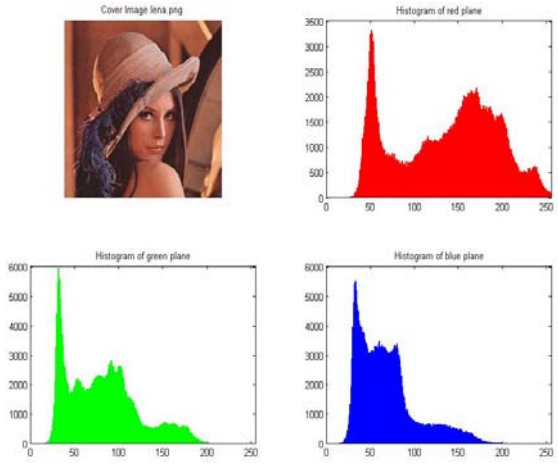

Fig. 14 Histogram of cover image lena.png

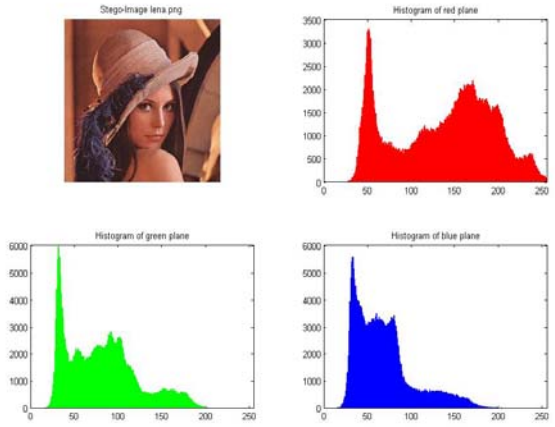

Fig. 15 Histogram of stego-image lena.png

CONCLUSION

Nowadays encryption of confidential data and communications is an increasingly important part of doing business. But on the other hand steganography can take data confidentiality to a whole new level, since it hides messages in ordinary-looking data files, making the very existence of the messages practically undetectable. Although steganography is not a new field and has played a critical part in secret communication throughout history, few people understand exactly how it works today. There exists various algorithms for steganography but still we are concerned about the capacity and security of steganography. In this paper an attempt has been made to study the high capacity and security steganography of color images in the domain of discrete wavelet transformation. The secret message is normalized depending upon the length of the message and the wavelet coefficient of the cover image is obtained by applying discrete wavelet transform. The security of the proposed algorithm is increased as the only high frequency coefficient of the cover image is replaced by the normalized message. The capacity of the proposed algorithm is increased as coefficient of all three color component that is red, green and blue of each of the high frequency sub-band of the cover image are considered. The MSE and Capacity are improved with acceptable PSNR compared to the existing algorithm.

REFERENCES

Journals

[1] Wahab, Ainuddin Wahid, Johann A. Briffa, Hans Georg Schaathun, and Anthony TS Ho. "Conditional probability based steganalysis for JPEG steganography." In *2009 International Conference on Signal Processing Systems*, pp. 205-209. IEEE, 2009.

[2] Petitcolas, Fabien AP, Ross J. Anderson, and Markus G. Kuhn. "Information hiding-a survey." *Proceedings of the IEEE* 87, no. 7 (1999): 1062-1078.

[3] S. Katzenbeisserm and F.A.P. Petitcolas, "Information Hiding Techniques for Steganography and Digital Watermarking" *Artech House, Norwood, MA, 2000.*

[4] Liu, Chiang-Lung, and Shiang-Rong Liao. "High-performance JPEG steganography using complementary embedding strategy." *Pattern Recognition* 41, no. 9 (2008): 2945-2955.

[5] LIU Tong, QIU Zheng-ding "A DWT-based color Images Steganography Scheme" *IEEE International Conference on Signal Processing*, vol. 2, pp.1568-1571, 2002.

[6] Fridrich Jessica and Soukal David, "Matrix Embedding for Large Payloads" *SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents* , vol. 6072, pp. 727-738. 2006.

[7] Nag Amitava, Biswas Sushanta, Sarkar Debasree, Sarkar Partha Pratim, "A Novel Technique for Image Steganography Based on DWT and Huffman Encoding" *International Journal of Computer Science and Security, (IJCSS), Volume (4): Issue (6) (2011): 561-570*

[8] Wang, Chung-Ming, Nan-I. Wu, Chwei-Shyong Tsai, and Min-Shiang Hwang. "A high quality steganographic method with pixel-value differencing and modulus function." *Journal of Systems and Software* 81, no. 1 (2008): 150-158.

[9] Reddy, HS Manjunatha, and K. B. Raja. "High capacity and security steganography using discrete wavelet transform." *International Journal of Computer Science and Security (IJCSS)* 3, no. 6 (2009): 462.

[10] Ghasemi, Elham, Jamshid Shanbehzadeh, and Nima Fassihi. "High capacity image steganography using wavelet transform and genetic algorithm." In *International MultiConference of Engineers and Computer Scientists*, vol. 1. 2011.

[11] Mandal, J. K., and Debashis Das. "Steganography Using Adaptive Pixel Value Differencing (APVD) of Gray Images Through Exclusion of Overflow/Underflow." *arXiv preprint arXiv:1205.6775* (2012).

[12] Anderson, Ross J., and Fabien AP Petitcolas. "On the limits of steganography." *Selected Areas in Communications, IEEE Journal on* 16, no. 4 (1998): 474-481.

[13] Kociołek, Marcin, Andrzej Materka, Michał Strzelecki, and Piotr Szczypiński. "Discrete Wavelet transform–Derived features for digital Image texture Analysis." In *Proc. of Interational Conference on Signals and Electronic Systems*, pp. 163-168. 2001.

[14] Rafael C. Gonzalez and Richard E. Woods, "Digital Image Processing" *Second Edition, Prentice Hall Upper Saddle River, New Jersey 07458*

[15] Chen, Po-Yueh, and Hung-Ju Lin. "A DWT based approach for image steganography." *International Journal of Applied Science and Engineering* 4, no. 3 (2006): 275-290.

[16] Audithan, S., and R. M. Chandrasekaran. "Document Text Extraction from Document Images Using Haar Discrete Wavelet Transform." *European Journal of Scientific Research* 36, no. 4 (2009): 502-512.

[17] Phad Vitthal, S., S. Bhosale Rajkumar, and R. Panhalkar Archana. "A Novel Security Scheme for Secret Data using Cryptography and Steganography." *International Journal* 4 (2012).

[18] Wang, Shen, Bian Yang, and Xiamu Niu. "A secure steganography method based on genetic algorithm." *Journal of Information Hiding and Multimedia Signal Processing* 1, no. 1 (2010): 28-35.

[19] Horé, Alain, and Djemel Ziou. "Image quality metrics: PSNR vs. SSIM." In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pp. 2366-2369. IEEE, 2010.

[20] Mazumder, Juned Ahmed, and K. Hemachandran. "Review Of Different Techniques Used In Recent Steganography Researches." *International Journal of Engineering* 1, no. 8 (2012).

Website

[21] http://www.dreamstime.com/free-photos

AUTHORS

Juned Ahmed Mazumder received his Master of Science in Computer Science (5 years integrated course) degree with first class in 2011 from Department of Computer Science, Assam University, Silchar, where he is currently doing his Ph.D. His research interest includes Image Processing, Steganography, Neural Network and Data Security.

Prof. K. Hemachandran is associated with the Department of Computer Science, Assam University, Silchar, since 1998. Currently he is serving as the Head of the Department in the Department of Computer Science, Assam University, Silchar. He obtained his M.Sc. Degree from Sri Venkateswara University, Tirupati and M.Tech and Ph.D Degrees from Indian School of Mines, Dhanbad. His areas of research interest are Image Processing, Software Engineering and Distributed Computing.