Genetic Algorithm based Color Image Steganogaphy using Integer Wavelet Transform and Optimal Pixel Adjustment Process

Medisetty Nagendra Kumar, S. Srividya

Abstract: Information security has become a cause of concern because of the electronic eavesdropping. For hiding secret information in images, there exist a large variety of steganographic techniques. Some of them are more complex than others and all of them have respective strong and weak points. Different applications have different requirements of the steganography technique used. For example, some applications may require absolute invisibility of the secret information, while others require a larger secret message to be hidden. In short, capacity, robustness and invisibility are three important parameters in information hiding and are quite difficult to achieve in a single algorithm. This paper proposes a novel steganography technique for digital colour image which aims at effective retrieval of hidden data in the colour image without significant degradation in the quality of the colour image. The professed methodology utilizes the least significant bits of the three colour channels (Red, Green, Blue) in a given colour image for embedding the secret message based on Integer Wavelet Transform, Genetic Algorithm and Optimal Pixel Adjustment Process (OPAP). The experimental results show that the proposed method is a secure steganographic method that effectively extracts the hidden message with good image quality and provides reasonable hiding capacity as compared to the adaptive methods of gray scale steganography systems.

Keywords- Genetic Algorithm, Histogram Modification, Integer Wavelet Transform, Optimal Pixel Adjustment Process, RGB channels, Steganography.

I. INTRODUCTION

Since the rise of the internet, one of the major concerns of Information Technology and Communication has been the security of information. Cryptography was created as a technique for securing the secrecy of communication. Many different methods have been developed to encrypt and decrypt data in order to keep the message secret. Unfortunately, it is sometimes not enough to keep the contents of a message secret. It may also be necessary to keep the existence of the message secret. The technique used to implement this, is called steganography. Steganography[1] is the art and science of invisible communication.

Steganography focuses on keeping the existence of a message secret. The format of the message can be text, audio, image, video, etc. Once the presence of hidden information is revealed or even suspected, the purpose of steganography is partly defeated. Data hiding techniques are generally divided into two groups: Spatial and Frequency Domain.

Manuscript received October, 2013.

Medisetty Nagendra Kumar, Department of Electronics and Communication Engineering, Chaitanya Institute of Science and Technology, Kakinada, India.

S. Srividya, Department of Electronics and Communication Engineering, Chaitanya Institute of Science and Technology, Kakinada, India.

The first group embeds message in the Least Significant Bit (LSB) of the image pixel [2]. This method is sensitive against attacks such as low pass filtering and compression. But its implementation is simple and its capacity is high. The second group embeds the messages in the frequency coefficients of images. These hiding methods overcome the problem related to the robustness and imperceptibility found in the spatial domain.

JPEG is a standard image compression technique [3]. Several steganography techniques have been implemented to hide data in JPEG images. Most recent researches apply Integer Wavelet Transform [5] due to its wide application in the new image compression standard. An example is the employment of an adaptive data embedding technique with the use of OPAP to hide data in the Integer Wavelet coefficients of the cover image.

The application of Genetic Algorithm in steganography can increase the capacity or imperceptibility of information [4]. This paper proposes a method to embed data in Integer Wavelet Transform coefficients using a mapping function based on Genetic Algorithm in 8x8 blocks on the RGB channels of the cover image [6] and it applies the OPAP after embedding the message to maximize the PSNR. We also used a pseudorandom generator function to select the embedding locations of the integer wavelet coefficients to increase the system security.

The rest of this paper is organized as follows: Section II introduces the proposed algorithm in detail. Section III discusses the achieved results and compares the proposed scheme with the state of the art. Section IV concludes the paper.

II. STEGANOGRAPHIC METHOD

In this section, we shall present the proposed steganographic method for hiding a message in a color image. The proposed methodology uses the principle of embedding message bits in Least Significant Bits (LSB) of Integer Wavelet Transform coefficients according to the best mapping function. Fitness evaluation is performed using genetic algorithm to select the best mapping function. Then, OPAP algorithm is applied on the obtained embedded image. The embedding and extraction procedures of this related scheme would be sequentially described in this section.

A. Histogram Modification

In the proposed system, Histogram modification is used to prevent overflow or underflow that occurs when the changed values in integer wavelet coefficients produce stego image pixel values to exceed 255 or to be smaller than 0. This problem was found to be caused by the values near 255 or 0.

The problem can be solved by mapping the lowest pixel value to the value of 15 and the highest pixel value to the value of 240.This procedure modifies the pixel values in the following manner:

$$\begin{aligned} &\text{if} \quad P(i,j) \leq 15 \\ &\quad \text{then} \quad P'(i,j)=15 \\ &\text{else if} \quad P(i,j) \geq 240 \\ &\quad \text{then} \quad P'(i,j)=240 \\ &\text{else} \\ &\quad P'(i,j) = P(i,j) \end{aligned} \tag{1}$$

Where P(i,j) is the actual pixel value in i^{th} row and j^{th} column of the image and P' (i, j) is the modified pixel value in the i^{th} row and j^{th} column of the pixel value.

B. Integer Wavelet Transform

Wavelet domain techniques are becoming very popular because of the developments in the wavelet stream in the recent years. Wavelet transform is employed to convert a spatial domain into frequency domain. The employment 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. Haar wavelet operates by calculating the sums and differences of adjacent elements. This wavelet operates first on adjacent horizontal elements and then on adjacent vertical elements. One nice feature of the Haar wavelet transform is that the transform is equal to its inverse. After each transform is performed the size of the square which contains the most important information is reduced by a factor of 4. The proposed system uses the wavelet transform coefficients to embed messages into four sub-bands LL, LH, HL and HH of two dimensional wavelet transform as shown in the figure1.

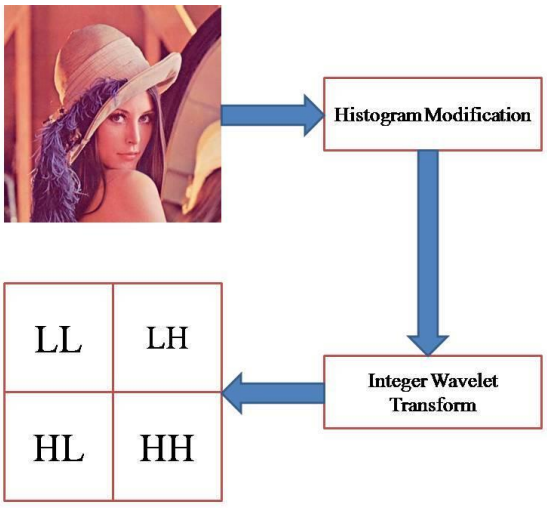


Figure1. Frequency sub-band representation of Original Lena Image.

Hiding data in these regions allow us to increase the robustness while maintaining good visual quality. In discrete wavelet transform, the used wavelet filters have floating point coefficients so that when we hide data in their coefficients, any truncations of the floating point values of the pixels that should be integers may cause the loss of the hidden information which may lead to the failure of the data hiding system [5,9]. To avoid problems with floating point precision of the wavelet filters, we used Integer Wavelet Transform.

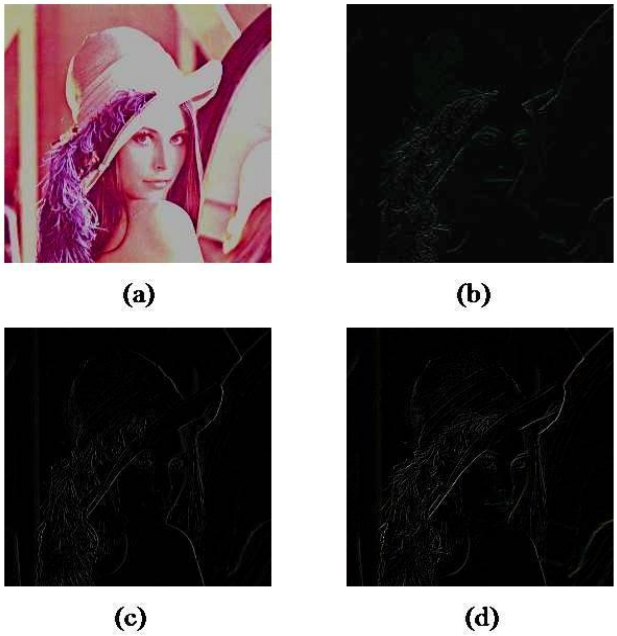


Figure 2. One level 2DIWT decomposition of Lena image in (a) LL Subband (b) LH Subband (c) HL Subband (d) HH Subband.

Integer wavelet transform maps an integer data set into another integer data set. The LL sub band in the case of IWT appears to be a close copy with smaller scale of the original image as shown in the figure 2. But, in the case of DWT, the obtained LL sub-band is distorted to a higher level.

The lifting scheme is for both designing wavelets and performing the integer wavelet transforms. Basically it is worthwhile to merge these steps and design the wavelet filters while performing the wavelet transform. The lifting scheme is an algorithm to calculate wavelet transforms in an effective way. It is also a generic technique to create so-called second-generation wavelets. Lifting scheme allows us to implement reversible integer wavelet transforms. In conventional scheme, it involves floating point operations, which introduces rounding errors due to floating point arithmetic. While in the case of lifting scheme, perfect reconstruction is possible. It allows a faster implementation of the wavelet transform. It requires half number of computations as compared to traditional convolution based discrete wavelet transform.

The decomposing filter in integer wavelet transform can be calculated according to [10]

$$\begin{aligned} D_{1,n} &= S_{0,2n+1} - S_{0,2n} \\ S_{1,n} &= (s_{0,2n} + D_{1,n})/2 \end{aligned} \tag{2}$$

Where $S_{i,n}$ is the n^{th} low frequency and $D_{i,n}$ is the n^{th} high frequency wavelet coefficients at the i^{th} level.

The inverse transform can be calculated by

$$\begin{aligned} S_{0,2n} &= S_{1,n} - (D_{1,n})/2 \\ S_{0,2n+1} &= D_{1,n} + S_{0,2n} \end{aligned} \tag{3}$$

C. Genetic Algorithm

A Genetic Algorithm (GA) is a search technique used in computing to find exact or approximate solutions to optimization and search problems. Genetic algorithms are a basic category of Evolutionary Algorithms (EA) that use techniques galvanized by organic process biology like inheritance, mutation, selection, and crossover. The frequency domain representation of the respective created blocks is estimated by two dimensional Integer Wavelet Transform in order to accomplish 4 sub bands LL, HL, LH, and HH. 1 to 64 genes are generated containing the pixels numbers of each 8x8 blocks as the mapping function. The message bits are embedded in k-LSBs of the IWT coefficients mapped by the mapping function.

• Chromosome Design

In this GA algorithm, a chromosome is encoded as an array of 64 genes containing permutations 1 to 64 that point to pixel numbers in each block as shown in the figure 3.

59	47	1	33	…	41	16	9	60

Figure 3. A simple chromosome with 64 genes

Each pixel in a block is considered as a chromosome as shown in figure 4. Several generations of chromosomes are created to select the best chromosomes by applying the fitness function to replace the original chromosomes. Crossover is applied by randomly selecting two chromosomes and combining them to generate new chromosomes. Once the process of selection, reproduction and mutation is complete, the next block is evaluated.

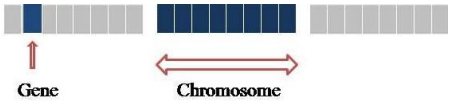


Figure 4. Gene and Chromosome

• GA Operations

Mating and mutation functions are applied on each chromosome. Mutation changes the bit values in which the data bit is not hidden and exchanges any two genes to generate new chromosome. Then we select elitism which means the best chromosome will survive and be passed to the next generation.

• Fitness function

The fitness function evaluates every candidate solution most of which are chosen randomly. Evolution begins from a completely random set of entities and is repeated in subsequent generations. The most suitable, and not the best, are picked out in every generation. Our GA aims to improve the image quality. Peak Signal to Noise Ratio (PSNR) can be an appropriate evaluation test.

Thus the definition of fitness function will be:

$$PSNR=10\log_{10}(M \times N \times 255^2)/ \sum_{i,j}(y_{i,j} - x_{i,j})^2 \quad (4)$$

Where M x N is the size of the image, x is the image intensity value of the cover image and y is the image intensity value of the Stego image.

D. OPAP algorithm

The proposed system uses optimal pixel adjustment algorithm, while taking into consideration that each modified coefficient stays in its hiding capacity range where each value of L (number of bits that can hide) is calculated according to the absolute value of the wavelet coefficients. Any significant change in this value will produce different value of L to be calculated at the receiver.The main idea of applying OPAP is to minimize the error between the cover and the stego image. OPAP (Optimal Pixel Adjustment process) is applied where $(k+1)^{th}$ bit of every pixel is modified if the modified version seems to give better results and thus contributing to a decrease in the MSE value. For example, if a binary number 11000 (decimal number 24) is changed to 11111 (decimal number 31) because its three LSB's were replaced with embedded data. The difference from the original number is 7. This difference in the original value is called the embedding error. By adjusting the fourth bit from a value of 1 to a value of 0, the binary number now becomes 10111 (decimal number 23) and the embedding error is reduced to 1 while at the same time preserving the value of the three embedded bits. The algorithm depend on calculating the difference (Δ_i) between original value P(x, y) and the modified value P'(x, y)

$$\Delta_i (x, y) = P_i'(x, y) - P_i(x, y) \quad (5)$$

After calculating the (Δ_i), the algorithm modifies the changed value in the following manner:

Case 1: $(-2^k < \Delta_i < -2^{k-1})$

if $P_i'(x, y) < 256 - 2^k$

then $P_i'(x, y)^* = P_i'(x, y) + 2^k$

else $P_i'(x, y)^* = P_i'(x, y)$

Case 2: $(-2^k \leq \Delta_i \leq 2^{k-1})$

$P_i'(x, y)^* = P_i'(x, y)$

Case 3: $(2^{k-1} < \Delta_i < 2^k)$

if $P_i'(x, y) \geq 2^k$

then $P_i'(x, y)^* = P_i'(x, y) - 2^k$

else $P_i'(x, y)^* = P_i'(x, y)$ (6)

E. Embedding Algorithm

The following steps explain the embedding process:

Step1. Read the cover image file into a two dimensional decimal array to handle the file data more easily.

Step2. Apply histogram modification to avoid overflow or underflow.

Step3. Divide the cover image into 8x8 blocks.

Step4. Find the frequency domain representation of blocks by 2D Integer Wavelet Transform and get four sub-bands LL1, LH1, HL1, and HH1.

Step5. Generate 64 genes containing the pixels numbers of each 8x8 blocks as the mapping function.

Step6. Embed the message bits in 4-LSBs IWT coefficients each pixel according to mapping function.
Step7. Fitness evaluation is performed to select the best mapping function.
Step8. Apply Optimal Pixel Adjustment process on the image.
Step9. Calculate inverse 2D IWT on each 8x8 block.

F. Extraction Algorithm

At the receiver end the extraction algorithm is used to obtain the secret message. The extraction algorithm consists of four steps as follows:
Step1. Divide the image into 8x8 blocks.
Step2. Extract the transform domain coefficient by 2D IWT of each 8x8 block.
Step3. Employ the obtained function in the embedding phase and find the pixel sequences for extracting.
Step4. Extract 4-LSBs in each pixel.

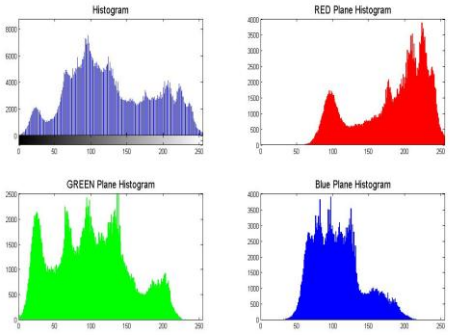
III. EXPERIMENTAL RESULTS

In the proposed method three 512 × 512 digital color images of Lena, Elephant, Peppers and one 512 × 512 gray scale image of Lena has been taken as cover images and tested for maximum hiding capacity. The program was implemented using Matlab 7.8.0 running on 2.27 G Hz i3 core processor under Windows 7. The effectiveness of the stego process proposed has been studied by calculating MSE and PSNR for all the three color images using the proposed method and the results are as shown as follows.

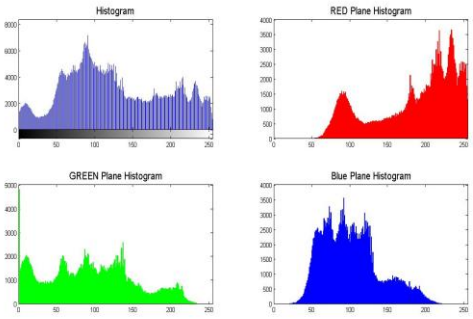
Figure 5. (a) Cover image of Lena (b) Lena image after Embedding 4-LSBs.

Figure. 5 show that there is no much difference observed visually in the quality of the image before and after embedding.

The same can be analyzed using histograms shown in the Figure 6. It is observed that there is no significant change in the stego image histogram of Lena when compared to the cover image histogram of Lena and further analyzing it using R,G,B channel histogram also shows no significant change.



(a)



(b)

Figure 6. (a) Histogram of Cover image of Lena (b) Histogram of Lena image after Embedding 4 LSBs

The same image and Histogram analysis is made on the Elephant and Peppers images as shown in the Figure 7 and Figure 8 and no noticeable changes of higher significance are found in the analysis. Thus this makes the proposed method more secure and robust against susceptible attacks.

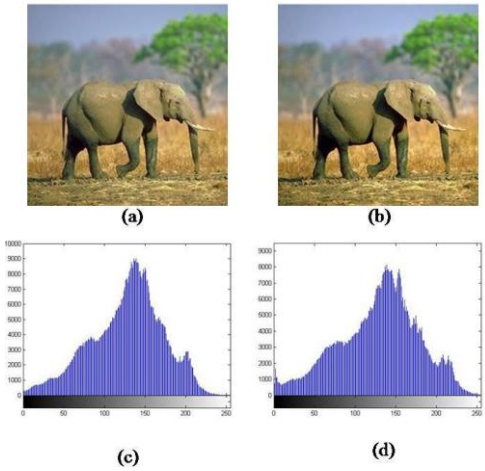


Figure 7. (a) Cover image of Elephant. (b) Elephant image after Embedding 4 LSBs (c) Histogram of Elephant. (d) Histogram of Elephant after Embedding 4 LSBs

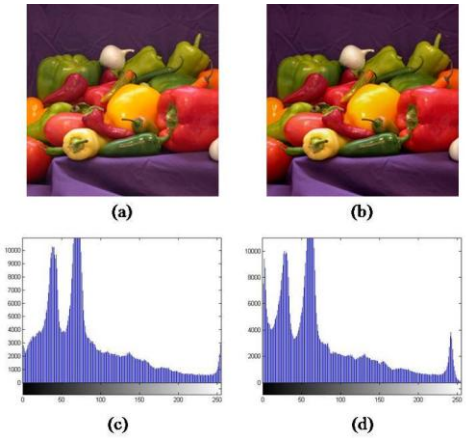


Figure 8. (a) Cover image of Peppers. (b) Peppers image after Embedding 4 LSBs (c) Histogram of Peppers (d) Histogram of Peppers after Embedding 4 LSBs.

The image quality in terms of PSNR,MSE and maximum hiding capacity of the color images Lena, Elephant and Peppers as per the results obtained from the proposed method are tabulated in the Table I.

TABLE I. Capacity, MSE and PSNR values obtained from proposed method.

Cover Image	Max.H.C (bits)	Max.H.C (%)	MSE	PSNR (db)
Lena	1048576	50%	22.11	34.68
Elephant	1048576	50%	15.88	36.12
Peppers	1048576	50%	23.48	34.42

As the proposed method also well applicable to the Gray scale images, gray scale image of Lena has been taken and the results are compared against the adaptive steganographic [5] using IWT of gray scale image. Stego image quality visual observation and Histogram analysis for Gray scale image Lena in the figure 9 shows that the visual quality of the images having the secret data is well preserved by the algorithm, there is no introduction of noticeable visible artifact by the adjustment.

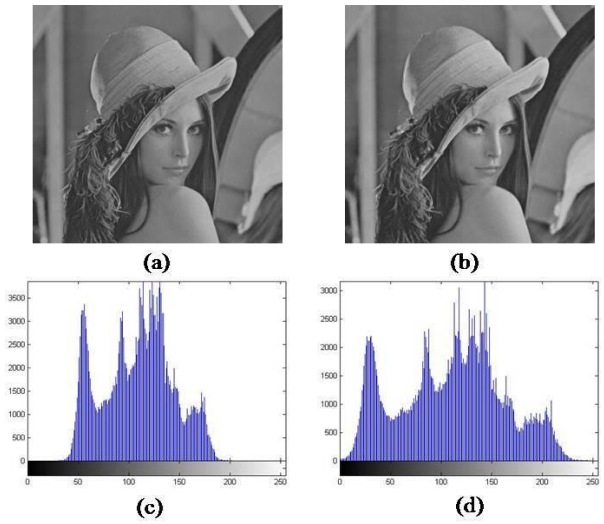


Figure 9. (a) Gray scale cover image of Lena (b) Gray Scale image of Lena after Embedding secret message.(c)Histogram of Lena before embedding message(d) Histogram of Lena after embedding message.

Table II summarizes the results for Gray scale Lena image and compares the Capacity and PSNR values of the proposed method as against the adaptive method [5]

TABLE II. Comparison of Capacity and PSNR values obtained from proposed method and adaptive method [5]

Cover Image	Method	Max.H.C (bits)	Max.H.C (%)	PSNR (db)
Lena	Proposed method	1048576	50%	32.47
Lena	Adaptive Method	986408	47%	31.80

A performance analysis graph is plotted as shown in the figure for the Gray scale Lena image in order to compare the results of proposed method with the adaptive method [5].

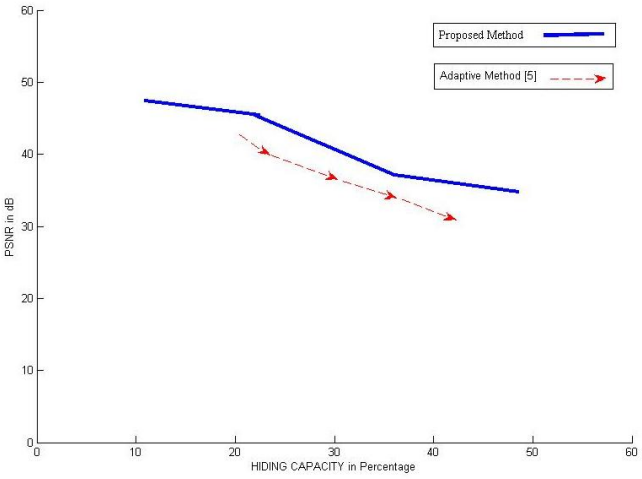


Figure 10. Analysis of Performance of Gray Scale Lena for Proposed and Adaptive Methods

Thus the results are satisfactory in achieving the objective of preservation of visual quality of stego image with improved hiding capacity using color image as well as the gray scale image. However color images are more flexible to maintain more hiding capacity as the secret messages can be embedded in three different channels where as it is only one channel in the case of gray scale image.

IV. CONCLUSION

In this paper we have proposed a novel stenographic technique based on Genetic Algorithm to embed the secret data in the digital color image with high security, imperceptibility, and robustness. Benefited from the effective optimization, a good balance between the security and image quality is achieved. The employed embedding process depends completely on the nature of the pixels which is not predictable. This makes it completely adaptive and random because the nature of the pixels cannot be controlled; it is inherent of an image. Genetic Algorithm is employed to obtain an optimal mapping function to reduce the error occurrence between the cover and the stego image. Optimal Pixel Adjustment Process is employed to increase the hiding capacity of the algorithm in comparison to other hiding systems. One of the two drawbacks observed in this method is that the high execution time for selecting the best block and the other is that MSE is not uniformly distributed over all the three RGB channels. Our future work will focus on overcoming the above drawbacks. First one may be solved by optimizing genetic algorithm and the second may be solved by using channel indicator.

REFERENCES

[1] N.Provosn and P. Honeyman, Hide and seek: An introduction to steganography, IEEE Security Privacy Mag.,1 (3) (2003) 32–44/
[2] C.K. Chan, L.M. Chen, Hiding data in images by simple LSB substitution, Pattern Recognition 37 (3) (2004) 469–474.
[3] W. Tseng and C. C. Chnag, "High capacity data hiding in jpegcompressed images," Informatica, vol. 15, no. I, pp. 127-142,2004.
[4] K. B. Raja, Kiran Kumar. K, Satish Kumar. N, Lashmi. M. S, Preeti.H, Venugopal. K. R. and Lalit. M. Patnaik "Genetic algorithm based steganography using wavelets," nternational Conference on Information System Security Vol. 4812, pp, 51-63. 2007.
[5] El Safy, R.O, Zayed. H. H, EI Dessouki. A, "An adaptive steganography technique based on integer wavelet transform,"ICNM

International Conference on Networking and Media Convergence, pp 111-117,2009.

[6] M.H. Lin, Y.C. Hu, C.C. Chang, Both color and gray scale secret images hiding in a color image, International Journal of Pattern Recognition and Artificial Intelligence 16 (2002) 697–713.

[7] T. Liu, Z.D. Qiu, A DWT-based color image steganography scheme, in: Proceedings of International Conference on Signal Processing, vol.2, Beijing, China, August 2002, pp. 1568–1571

[8] S. Lee, C.D. Yoo and T. Kalker, "Reversible image watermarking based on integer-to-integer wavelet transform," IEEE Transactions on Information Forensics and Security, Vol. 2, No.3, Sep. 2007, pp. 321-330.

[9] Xuan, J. Zhu, Y. Q. Shi, Z. Ni, and W. Su., "Distortionless data hiding based on integer wavelet transform," IEE Electronic Letters, 38(25): 1646--1648, Dec. 2002.

[10] Elham Ghasemi, Jamshid Shanbehzadeh,Bahram Zahir Azamid, "An Steganographic method based on Integer Wavelet Transform and Genetic Algorithm", 978-1-4244-9779-7/11/$26.00 ©2011 IEEE.

[11] Steganography software for windows, http://members.tripod.com/steganography/stego/software.html.

[12] S. C. Chu, H. C. Huang, Y. Shi, S. Y. Wu, and C. S. Shieh, Genetic watermarking for zerotree-based applications. Circuits, Systems, and Signal Processing, vol. 27, no. 2, pp. 171-182, 2008.

[13] J. Fridrich, M. Goljan, and R. Du, Detecting lsb steganography in color, and gray-scale images, IEEE MultiMedia, pp. 22{28, 2001.

[14] J. Fridrich, M. Goljan, and D. Hogea. Attacking the outguess, Proc. of ACM Workshop Multimedia and Security, 2002.

[15] J. Fridrich, M. Goljan, and D. Hogea, Steganalysis of jpeg images: Breaking the f5 algorithm, Proc. of ACM Workshop on Multimedia and Security 2002, 2002.

[16] D. E. Goldberg, The genetic algorithms in search, optimization and machine learning, Addison- Wesley, 1989.

[17] C. T. Hsu, J. Wu, and L. Hidden, Digital watermarks in images, IEEE Trans. Image Processing, pp. 58-68, 1999.

[18] H. C. Huang, C. M. Chu, and J. S. Pan, The optimized copyright protection system with genetic watermarking, Soft Computing, vol. 13, no. 4, pp. 333{343, 2009.

[19] H. C. Huang, J. S. Pan, Y. H. Huang, F. H. Wang, and K. C. Huang, Progressive watermarking techniques using genetic algorithms, Circuits, Systems, and Signal Processing, vol. 26, no. 5, pp. 671{687, 2007.

[20] E. Kawaguchi and R. O. Eason, Principle and application of bpcs-steganography, Proc. Of SPIE:Multimedia Systems and Applications, pp. 464{472, 1998.

[21] A. R. S. Marcal and P. R. Pereira, A steganographic method for digital images robust to rs steganal- ysis, Lecture Notes in Computer Science, pp. 1192{1199, 2005.

[22] N. Provos, Steganography detection with stegdetect, http://www.outguess.org/detection.php.

[23] A.Westfeld, F5-a steganographic algorithm, Proc. of the 4th International Workshop on Information

[24] Hiding, Lecture Notes in Computer Science,2137.Springer-Verlag, pp. 289{302, 2001.

[25] A. Westfeld and A. P_tzmann, Attacks on steganographic systems, Proc. of Information Hiding- Third International Workshop, 1999.

[26] A. Westfeld and A. P_tzmann, Attacks on steganographic systems, Lecture Notes in Computer Science, pp. 61{76, 1999.

[27] D. C. Wu and W. H. Tsai, A steganographic method for images by pixel-value di_erencing, Pattern Recognition Letters, pp. 1613{1626, 2003.

[28] X. Zhang and S. Z. Wang, Statistical analysis against spatial bpcs steganography, Computer-Aided Design & Computer Graphics, pp. 395{406, 2003.

[29] X. Zhang and S. Z. Wang, Vulnerability of pixel-value di_erencing steganography to histogram analysis and modi_cation for enhanced security, Pattern Recognition Letters, pp. 331{339, 2004.

[30] Bruice Schneier, Applied Cryptography Protocols, Algorithm and Source Code in C. Second edition. Wiley India edition 2007

[31] S. Katzenbeisser, F.A.P. Petitcolas, Information Hiding Techniques for Steganography and Digital Watermarking, Artech House, Norwood, MA, 2000.

[32] W. Bender, D. Gruhl, N. Morimoto, A. Lu, ―Techniques for data hiding‖ IBM Syst. J. 35 (3&4) (1996) 313–336.

[33] G. J. Simmons, "The prisoners' problem and the subliminal channel" in *Proc. Advances in Cryptology (CRYPTO '83)*, pp. 51-67

[34] R.Amirtharajan and R.John Bosco Balaguru. ―Constructive Role of SFC & RGB Fusion versus Destructive Intrusion‖ .*International Journal of Computer Applications* 1(20):30–36

[35] R.Amirtharajan and Dr. R. John Bosco Balaguru, ―*Tri-Layer Stego for Enhanced Security – A Keyless Random Approach*‖ - IEEE Xplore, DOI, 10.1109/IMSAA.2009.5439438.

[36] Abbas Cheddad, Joan Condell, Kevin Curran, Paul Mc Kevitt, Digital image steganography: Survey and analysis of current methods Signal Processing 90 (2010) 727–752