A Steganographic method based on Integer Wavelet Transform and Genetic Algorithm

Elham Ghasemi
Islamic Azad University
Science and Research Branch
Tehran, Iran
e.ghasemi@gmail.com

Jamshid Shanbehzadeh
Tarbiat Moalem University
Tehran, Iran
jamshid@tmu.ac.ir

Bahram ZahirAzami
Electrical and Computer Engineering Dept
Ryerson University
Toronto, Ontario, Canada
zahir@ieee.org

Abstract: **This paper proposes a novel steganography scheme based on Integer Wavelet Transform and Genetic Algorithm. The novel scheme embeds data in integer wavelet transform coefficients by using a mapping function based on Genetic Algorithm in an 8x8 block on the cover image. The optimal pixel adjustment process is applied after embedding the message. We employed frequency domain to increase the robustness of our steganography method. Integer wavelet transform avoids the floating point precision problems of the wavelet filter. We use GA and Optimal Pixel Adjustment Process to obtain an optimal mapping function to reduce the difference error between the cover and the stego-image and to increase the hiding capacity with low distortions respectively. Simulation results show that the novel scheme outperforms adaptive steganography technique based on integer wavelet transform in term of peak signal to noise ratio and capacity, 35.17 dB and 50% respectively.**

Keywords- Steganography, Integer Wavelet Transform, Genetic Algorithm, Optimal Pixel Adjustment Process, Peak Signal to Noise Ratio

I. INTRODUCTION

Steganography is the art and science of hiding data in a cover. The cover can be text, audio, image, video, etc. We can divide the data hiding techniques into two groups: spatial and frequency domain [1-3]. The first group is based on embedding message in the Least Significant Bits (LSB) of image pixels. The basic LSB method has a simple implementation and high capacity [4]. However, it has low robustness versus some attacks such as low-pass filtering and compression [5]. A variant of LSB method can be found in [6] that proposes an Optimal Pixel Adjustment Process (OPAP) in which image quality of the stego-image can be improved with low computational complexity. The group finds the frequency coefficients of images and then embeds the messages with them. These hiding methods overcome the robustness and imperceptibility problem found in the spatial domain.

JPEG, a standard image compression technique, employs Discrete Cosine Transform (DCT). Several steganography techniques for data hiding in JPEG have been proposed; such as JSteg [7], JP Hide&Seek [7] and OutGuess [8]. Most recent researches utilize Discrete Wavelet Transform (DWT) because of its wide application in the new image compression standard, JPEG2000. An example is the employment of an adaptive data embedding technique with the use of OPAP to hide data in Integer Wavelet coefficients of the cover image [9].

The application of Genetic Algorithm in steganography can increase the capacity or imperceptibility [10-12]. Fard, Akbarzadeh and Varasteh [11] proposed a GA evolutionary process to make secure steganography encoding on the JPEG images. Rongrong *et al* [12] introduced an optimal block mapping LSB method based on Genetic Algorithm. 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 cover images and, it applies the Optimal Pixel Adjustment Process after embedding the message to maximize the PSNR.

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.

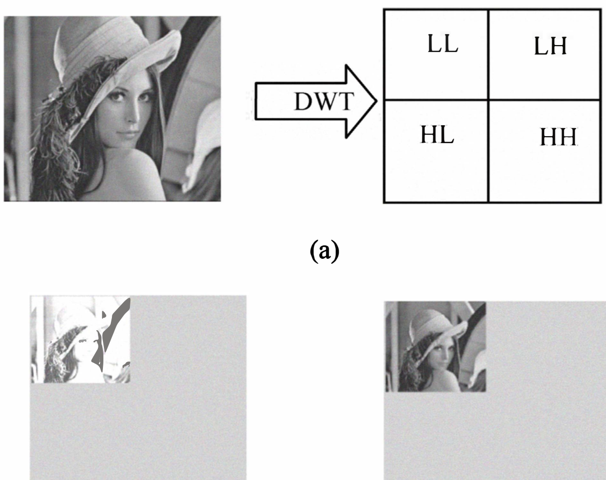
II. THE STEGANOGRAPHY METHOD

In the proposed method, the message is embedded on Integer Wavelet Transform coefficients based on Genetic Algorithm. Then, OPAP algorithm is applied on the obtained embedded image. This section describes this method, and the embedding and extracting algorithms in detail.

978-1-4244-9799-7/11/$26.00 ©2011 IEEE

A. Integer Wavelet Transform

The proposed algorithm employs the wavelet transform coefficients to embed messages into four subbands of two dimensional wavelet transform. To avoid problems with floating point precision of the wavelet filters, we used Integer Wavelet Transform. The LL subband in the case of IWT appears to be a close copy with smaller scale of the original image while in the case of DWT the resulting LL subband is distorted [9] as shown in "Fig. 1".



(b) (c)

Figure 1. (a) Original image Lena and how to analyze in wavelet domain. (b) One level 2DDWT in subband LL (c) One level 2DIWT in subband LL.

Lifting Scheme is one of the techniques on integer wavelet transform. The decomposing filter in integer wavelet transform can be calculated according to [13]:

$$s_{1,k} = \left\lfloor (s_{0,2k} + s_{0,2k+1})/2 \right\rfloor$$
$$d_{1,k} = s_{0,2k+1} - s_{0,2k} \qquad (1)$$

Then the inverse transform can be calculated by:

$$s_{0,2l} = s_{1,l} - \left\lfloor d_{1,l}/2 \right\rfloor$$
$$s_{0,2l+1} = s_{1,l} + \left\lfloor (d_{1,l}+1)/2 \right\rfloor \qquad (2)$$

These equations should be in 2D in order to be applied on images. Simple 2D transform has employed in this paper and it can be computed for an image according to [14]:

$$A_{i,j} = \left\lfloor (I_{2i,2j} + I_{2i+1,2j})/2 \right\rfloor$$
$$H_{i,j} = I_{2i,2j+1} - I_{2i,2j}$$
$$V_{i,j} = I_{2i+1,2j} - I_{2i,2j}$$
$$D_{i,j} = I_{2i+1,2j+1} - I_{2i,2j} \qquad (3)$$

Where I is the original image, A, H, V and D are the low pass, horizontal, vertical and diagonal coefficients. The inverse is given in (4).

$$I_{2i,2j} = A_{i,j} - \left\lfloor H_{i,j}/2 \right\rfloor$$
$$I_{2i,2j+1} = A_{i,j} + \left\lfloor (H_{i,j}+1)/2 \right\rfloor$$
$$I_{2i+1,2j} = I_{2i,2j+1} + V_{i,j} - H_{i,j}$$
$$I_{2i+1,2j+1} = I_{2i+1,2j} + D_{i,j} - V_{i,j} \qquad (4)$$

B. Genetic Algorithm

This paper embeds the message inside the cover with the least distortion therefore we have to use a mapping function to LSBs of the cover image according to the content of the message. We use Genetic Algorithm to find a mapping function for all the image blocks. Block based strategy can preserve local image property and reduce the algorithm complexity compared to single pixel substitution.

- Chromosome Design

In our GA method, a chromosome is encoded as an array of 64 genes containing permutations 1 to 64 that point to pixel numbers in each block. Each chromosome produces a mapping function as shown in "Fig. 2".

60	7	24	....	52	3

Figure 2. A simple chromosome with 64 genes

- GA Operations

Mating and mutation functions are applied on each chromosome. The mutation process causes the inversion of some bits and produces some new chromosomes, then, we select elitism which means the best chromosome will survive and be passed to the next generation.

- Fitness function

Selecting the fitness function is one of the most important steps in designing a GA-based method. Whereas our GA aims to improve the image quality, Pick Signal to Noise Ratio (PSNR) can be an appropriate evaluation test. Thus the definition of fitness function will be:

$$PSNR = 10\log_{10}\frac{M \times N \times 255^2}{\sum_{i,j}(y_{i,j} - x_{i,j})^2} \qquad (5)$$

Where M and N are the image sizes and, x and y are the image intensity values before and after embedding.

C. OPAP algorithm

The main idea of applying OPAP is to minimize the error between the cover and the stego image. For example if the pixel number of the cover is 10000 (decimal number

16) and the message vector for 4 bits is 1111, then the pixel number will change to 11111 (decimal number 31) and the embedding error will be 15, while after applying OPAP algorithm the fifth bit will be changed from 1 to 0, and the embedding error is reduced to 1.

The OPAP algorithm can be described as follows:

Case 1 ($2k-1 < \delta i < 2k$): if $pi' \geq 2k$, then $pi'' = pi' - 2k$ otherwise $pi'' = pi'$;

Case 2 ($-2k-1 < \delta i < 2k-1$): $pi'' = pi'$;

Case 3 ($-2k < \delta i < -2k-1$): if $pi' < 256 - 2k$, then

$pi'' = pi' + 2k$; otherwise $pi'' = pi'$;

P_i, P_i^{δ} and P_i^{δ} are the corresponding pixel values of the i^{th} pixel in the three images; cover, stego and the obtained image by the simple LSB method, respectively. δ_i $(= P_i^{\delta} - P_i)$ is the embedding error between P_i and P_i^{δ} *[6]*. Therefore after embedding *k*-LSBs of P_i with *k* message bits, δ_i will be as follows:

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

D. Embedding Algorithm

The following steps explain the embedding process:

Step1. Divide the cover image into 8x8 blocks.

Step2. Find the frequency domain representation of blocks by 2D Integer Wavelet Transform and get four subbands LL1, HL1, LH1, and HH1.

Step3. Generate 64 genes containing the pixels numbers of each 8x8 blocks as the mapping function.

Step4. Embed the message bits in 4-LSBs IWT coefficients each pixel according to mapping function.

Step5. Fitness evaluation is performed to select the best mapping function.

Step6. Apply Optimal Pixel Adjustment Process on the image.

Step7. Calculate inverse 2D-IWT on each 8x8 block.

Extraction Algorithm.

The extraction algorithm consists of four steps as follows:

Step1. Divide the cover image into 8x8 blocks.

Step2. Extract the transform domain coefficient by 2D IWT of each 8x8 block.

Step3. Employ the obtained mapping function in the embedding phase and find the pixel sequences for extracting.

Step4. Extract 4-LSBs in each pixel.

III. EXPERIMENTAL RESULTS

The proposed method is applied on 512x512 8-bit grayscale images "Baboon" and "Lena". The simulation is implemented on 2.5GHZ Core 2 Duo processor, 4GB RAM and Windows Vista OS and Matlab7.6. The messages are generated randomly with the same length as the maximum hiding capacity. Table I shows the stego image quality by PSNR as described in Eq. (5). Human visual system is unable to distinguish the grayscale images with PSNR more than 35 dB. This paper embedded the messages in the 4-LSBs and received a reasonable PSNR. "Fig. 3" shows the images after and before embedding.

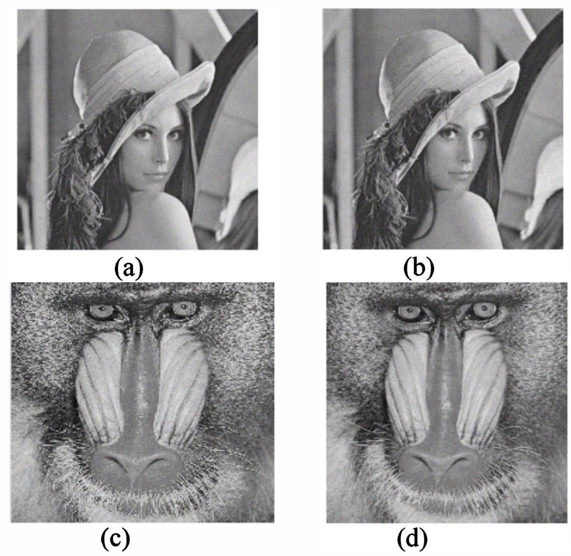


Figure 3. (a)Lena image after embedding 3-LSBs. (b) Lena image after embedding 4-LSBs. (c) Baboon image after embedding 3-LSBs (d) Baboon image after embedding 4-LSBs.

We can increase the capacity up to 5-LSBs. Table I shows the capacity and the PSNR of the proposed method for 4-LSBs and 5-LSBs.

TABLE I. COMPARISON OF CAPACITY AND PSNR FOR 4-LSBS AND 5-LSBS

Cover image	Hiding Capacity (bits)	PSNR (dB)
Lena	1048576 (4-LSBs)	35.17
	1310720 (5-LSBs)	29.08
Baboon	1048576 (4-LSBs)	36.23
	1310720 (5-LSBs)	29.49

Table II summarizes the results for two images, Lena and Baboon, and compares the PSNR and the capacity of new method to the one in [9].

According to the obtained results in this paper the performance analysis for Lena as a cover image is shown in "Fig. 4".

TABLE II. COMPARISON OF CAPACITY AND PSNR OBTAINED FROM OUR METHOD AND THE PROPOSED METHOD IN [9].

Cover image	Method	Max. H. C. (bits)	Max H.C. (%)	PSNR (dB)
Lena	Proposed method	1048576	50%	35.17
	Adaptive steganographic based IWT[9]	986408	47%	31.8
Baboon	Proposed method	1048576	50%	36.23
	Adaptive steganographic based IWT[9]	1008593	48%	30.89

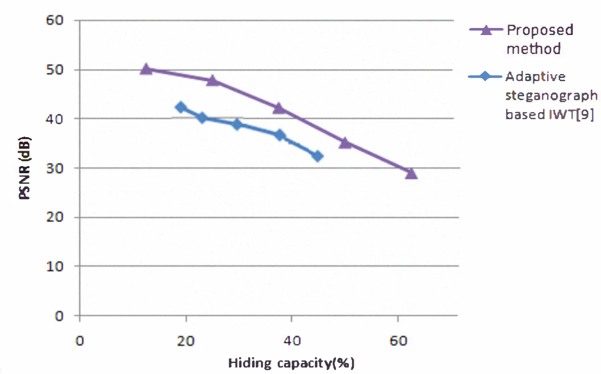


Figure 4. Comparison performance analysis of Lena for two methods

"Fig. 5" shows that when the size of message increases, the histogram tends to be smoother. We can see that there is no significant change in stego image histogram for 4-LSBs Lena image, thus it is robust against some statistic attacks.

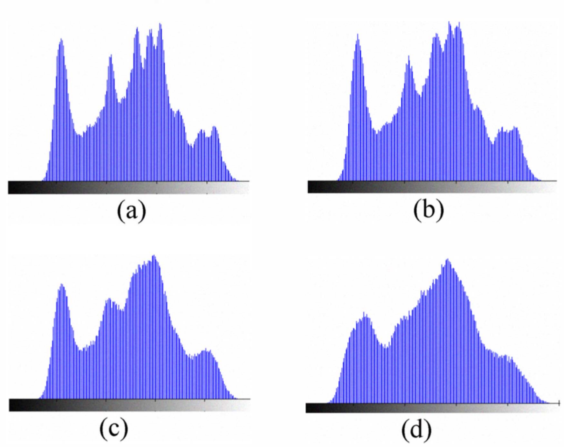


Figure 5. (a) Histogram of Lena image before embedding (b) Histogram of Lena image after embedding 3-LSBs (c) Histogram of Lena image after 4-LSBs (d) Histogram of Lena after embedding 5-LSBs.

IV. CONCLUSIONS

This paper presented a novel technique to increase the capacity and the imperceptibility of the image after embedding. GA is employed to obtain an optimal mapping function to reduce the error difference between the cover and the stego image and use the block mapping method to preserve local image properties and to reduce the algorithm complexity, and then applied the Optimal Pixel Adjustment Process to increase the hiding capacity of the algorithm in comparison to other systems. The drawback of this method is the execution time that can be the subject of our future studies to select the best block size to reduce the computation cost and to increase the PSNR using optimization algorithms such as genetic algorithm.

REFERENCES

[1] A. Z. Tirket, R.G. Van Schyndel, C. F. Osborne, "A digital watermark," Proceedings of ICIP, Austin Texas, Vol. II, 1994, pp. 86-90,1994

[2] W. Bender, N. Morimoto, "Techniques for data hiding," IBM Sys. J. 35(3/4) (1996) 313-336.

[3] K. L. Chung, C.H. Shen, L. C. Chang, "A novel SVD and VQ–based image hiding schem," Pattern Recognition Let. 22(9) 1051-1058 July 2001.

[4] N. Wu and M. Hwang, "Data hiding: current status and key issues," International Journal of Network Security, vol4, No.1, pp. 1-9, Jan. 2007.

[5] W. Chen, "A comparative study of information hiding schemes using amplitude, frequency and phase embedding," PhD thesis, National Cheng Kung University, Taiwan, May 2003

[6] C. K. Chan and L. M. Chang, "Hiding data in images by simple LSB substitution," Pattern Recognition, pp. 469-474, Mar. 2004.

[7] N. Provos, P. Honeyman, "Hide and Seek: an introduction to steganography," IEEE Computer Society, pp. 32-44, May-June 2003.

[8] N. Provos, "Defending against statistical steganalysis," In Proc. Of10th Usenix Security Symp, Usenix Assoc, pp. 323-335, 2001.

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

[10] 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," International Conference on Information System Security Vol. 4812, pp, 51-63. 2007.

[11] A.M. Fard, M.R Akbarzadeh and A. F Varasteh. "A new genetic algorithm approach for secure JPEG steganography," International Conference on Engineering of Intelligence Systems, pp 1-6, 2006.

[12] Ji. Rongrong, Yao. Hongxun, L. Shaohui and W. Liang, "Genetic algorithm based steganography optimal block mapping method for LSB substitution," International Conference on Information Hiding and Multimedia Signal Processing, pp, 215-218, Dec 2006.

[13] A. R. Calderbank, I. Daubechies, W. Sweldens and B. Yeo., "Wavelet transforms that map integers to integers," Applied and Computational Harmonic Analysis, vol. 5, no. 3, pp. 332-369, July 1998.

[14] G M. K. Ramani, E. V. Prasad, S. Varadarajan, "Steganography using BPCS to the integer wavelet transformed image" IJCSNS, Vol. 7, No. 7, pp. 293-302, July 2007.