High Capacity Image Steganography usingWavelet Transform and Genetic Algorithm

Elham Ghasemi, Jamshid Shanbehzadeh, Nima Fassihi

Abstract— **This paper presents the application of Wavelet Transform and Genetic Algorithm in a novel steganography scheme. We 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. We utilize the frequency domain to improve the robustness of steganography and, we 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. Our Simulation results reveal that the novel scheme outperforms adaptive steganography technique based on wavelet transform in terms of peak signal to noise ratio and capacity, 39.94 dB and 50% respectively.**

Index Terms— **Steganography, Discrete Wavelet Transform, Genetic Algorithm, Optimal Pixel Adjustment Process, Image Processing**

I. INTRODUCTION

Steganography is the art and science to hide data in a cover that it can be text, audio, image, video, etc. Data hiding techniques are generally divided in two groups: spatial and frequency domain [1-3]. The first group embeds message in the Least Significant Bit (LSB) of the image pixel. This method is sensitive against attacks such as low-pass filtering and compression but, its implementation is simple and its capacity is high. For instance, Raja et al [4] exhibited variety of LSB using Optimal Pixel Adjustment Process (OPAP) and enhanced the image quality of the stego-image with low computational complexity. Furthermore, this hiding method improved the sensitivity and imperceptibility problem found in the spatial domain.

The second group embeds the messages in the frequency coefficients of images. These hiding methods overcome the problem related to robustness and imperceptibility found in the spatial domain. JPEG is a standard image compression technique. Several steganography techniques such as JSteg, JP Hide&Seek and OutGuess implemented to hide data in JPEG images. Most recent researches apply Discrete Wavelet Transform (DWT) due to 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 the Integer Wavelet coefficients of the cover image [5]. Raja et al [6] presented a Genetic Algorithm (GA) based steganography in Discrete Cosine Transforms (GASDCT) domain and, GA based Steganography using Discrete Wavelet Transforms (GASDWT). GASDWT has an improvement in bit rate error compared to GASDCT.

The application of GA in steganography can increase the capacity or imperceptibility [7-8]. Fard, Akbarzadeh and Varasteh [7] proposed a GA evolutionary process to make secure steganography encoding on the JPEG images. R. Elshafie, N. Kharma and R. Ward [8] introduced a parameter optimization using GA that maximizes the quality of the watermarked image. This paper proposes a method to embed data in Discrete Wavelet Transform coefficients using a mapping function based on Genetic Algorithm in 4x4 blocks on the cover image and, it applies the OPAP 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 algorithms. Section IV concludes the paper.

II. THE STEGANOGRAPHY METHOD

The proposed method embeds the message in Discrete Wavelet Transform coefficients based on GA and OPAP algorithm and then applied on the obtained embedded image. This section describes this method, and embedding and extracting algorithms in detail.

A. Haar Discrete Wavelet Transform

Wavelet transform has the capability to offer some information on frequency-time domain simultaneously. In this transform, time domain is passed through low-pass and high-pass filters to extract low and high frequencies respectively. This process is repeated for several times and each time a section of the signal is drawn out.

DWT analysis divides signal into two classes (i.e. Approximation and Detail) by signal decomposition for various frequency bands and scales. DWT utilizes two function sets: scaling and wavelet which associate with low and high pass filters orderly. Such a decomposition manner bisects time separability. In other words, only half of the samples in a signal are sufficient to represent the whole signal, doubling the frequency separability.

Haar wavelet operates on data 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. Each transform computes the data energy in relocated to the top left hand corner. Figure 1 shows the image Lena after one Haar wavelet transform.

Fig. 1. The image Lena after one Haar wavelet transform

After each transform is performed the size of the square which contains the most important information is reduced by a factor of 4.

B. Genetic Algorithm

Genetic Algorithm is a technique which mimics the genetic evolution as its model to solve problems. The given problem is considered as input and the solutions are coded according to a pattern. 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 bests, are picked out in every generation. 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:

$$PSNR = 10 \log_{10} \frac{M \times N \times 255^2}{\sum_{ij}(y_{i,j} - x_{i,j})^2} \qquad (1)$$

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

A solution to the problem is translated into a list of parameters known as chromosomes. These chromosomes are usually displayed as simple strings of data. In the first step, several characteristics are generated for the pioneer generation randomly and the relevant proportionality value is measured by the fitness function. The next step associates with the formation of the second generation of the society which is based on selection processes via genetic operators in accordance with the formerly set characteristics. A pair of parents is selected for every individual. Selections are devised so that to find the most appropriate component. In this way, even the weakest components enjoy their own chance of being selected and local solutions are bypassed. In the current study, Tournament method has been exploited.

The contents of the two chromosomes which enter the generation process are interacted to produce two newborn chromosomes. In this approach two of the bests are mixed to give a superb one. In addition, during each process, it is likely for a series of chromosomes to undergo mutations and breed a succeeding generation of different characteristics.

C. Embedding Algorithm

The following steps explain the embedding process:

Step1. Divide the cover image into 4x4 blocks.

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

Step3. Generate 16 genes containing the pixels numbers of each 4x4 blocks as the mapping function.

Step4. Embed the message bits in k-LSBs DWT coefficients each pixel according to mapping function. For selecting value of k, images are evaluated from k=3 to 6. K equal to 1 or 2, provide low hiding capacity with high visual quality of the stego image and k equal to 7 or 8, provide low visual quality versus high hiding capacity.

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

Step6. Apply Optimal Pixel Adjustment Process on the image.

Step7. Calculate inverse 2D-HDWT on each 4x4 block.

D. Extraction Algorithm

The extraction algorithm consists of four steps as follows:

Step1. Divide the cover image into 4x4 blocks.

Step2. Extract the transform domain coefficient by 2D HDWT of each 4x4 block.

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

Step4. Extract k-LSBs in each pixel.

III. EXPERIMENTAL RESULTS

The proposed method is applied on 512x512 8-bit grayscale images "Jet", "Boat", "Baboon" and "Lena". 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. (1). Human visual system is unable to distinguish the grayscale images with PSNR more than 36 dB [5]. This paper embedded the messages in the k-LSBs, from k=3 to k=6 and received a reasonable PSNR. Table I shows PSNR for variant value of k.

Table I presents the results and we can see that for k equal to 4 or 5, we obtain the highest hiding capacity and reasonable visual quality. Therefore, we take k equal to 4 as the number of bits per pixel.

Cover image	PSNR			
	K=3	K=4	K=5	K=6
Lena	46.83	39.94	32.04	24.69
Jet	51.88	45.20	37.45	29.31
Boat	48.41	40.44	31.17	23.60
Baboon	47.32	40.34	32.79	24.80

Figure 2 shows the original cover images along with their histogram analyze which will be used later to compare it with the ones of the resulting stego image to test for imperceptibility.
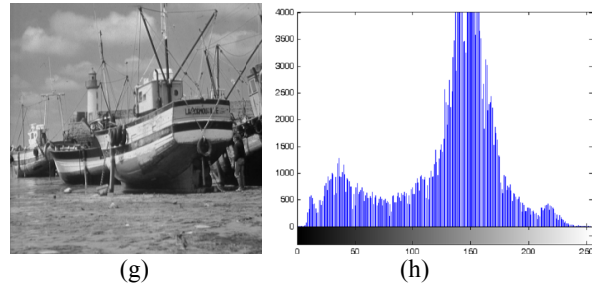

(a) (b) (c) (d) (e) (f)

(g) (h)

Fig. 2. Four cover images used in system simulation and their corresponding histogram (a)cover image Lena (b) Lena histogram(c) cover image Jet (d) Jet histogram(e) cover image Baboon (f) Baboon histogram (g) cover image Boat (h) Boat histogram

Figure 3 shows images for k equal to 4 that there is no significant change in stego image histogram for 4-LSBs images, thus it is robust against some statistic attacks.
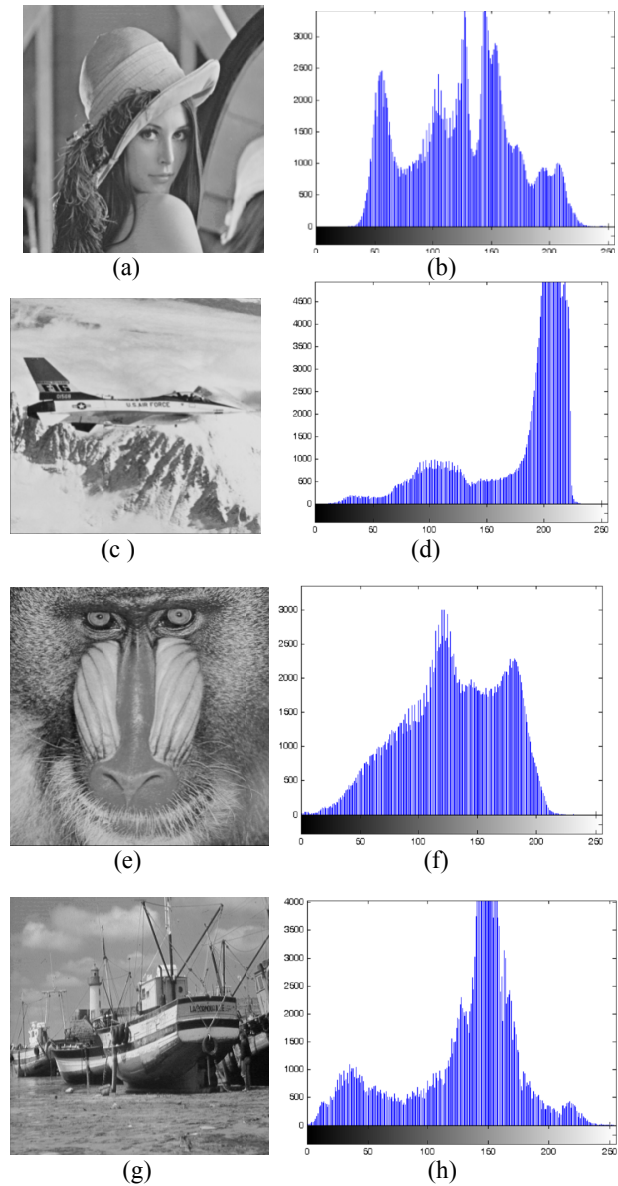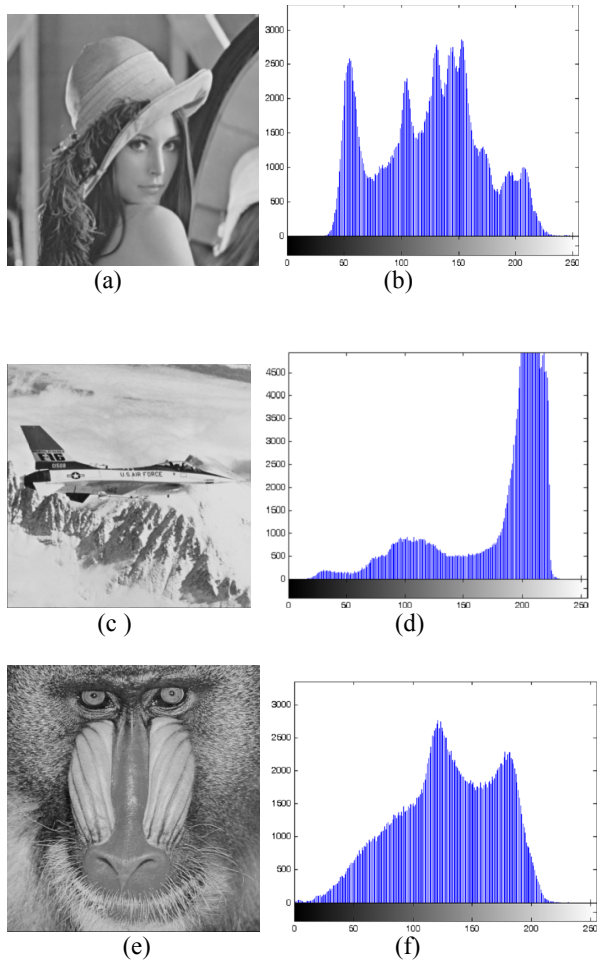

(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 3. Output stego image of k=4 for embedding data and their corresponding histograms (a)Lena image (b) Lena histogram (c) Jet image (d) Jet histogram (e) Baboon image (f) Baboon histogram (g) Boat image (h) Boat histogram

IV. CONCLUSIONS

In this research, we introduced a novel steganography technique to increase the capacity and the imperceptibility of the image after embedding. GA employed to obtain an optimal mapping function to lessen the error difference between the cover and the stego image and use the block mapping method to preserve the local image properties. Also we applied the OPAP to increase the hiding capacity of the algorithm in comparison to other systems. However by this method, the computational complexity is high, our results show that capacity and imperceptibility of image have increase simultaneity. Also, we can select the best block size to reduce the computation cost and to increase the PSNR using optimization algorithms such as genetic algorithm.

TABLE II

COMPARISON OF HIDING CAPACITY ACHIEVED AND THE OBTAINED PSNR BETWEEN OUR PROPOSED METHOD AND METHODS IN [5], [9] AND [10].

Cover image	Method	Hiding Capacity (bit)	Hiding Capacity (%)	PSNR (DB)
Lena	Proposed Method	1048576	50%	39.94
	Adaptive [5]	986408	47%	31.8
	HDWT[10]	801842	38%	33.58
	DWT [9]	573550	27.34%	44.90
Baboon	Proposed Method	1048576	50%	40.34
	Adaptive [5]	1008593	48%	30.89
	HDWT[10]	883220	42%	32.69
	DWT [9]	573392	27.34%	44.96
Jet	Proposed Method	1048576	50%	45.20
	DWT [9]	573206	27.33%	44.76
Boat	Proposed Method	1048576	50%	40.44
	DWT [9]	573318	27.33%	44.92

REFERENCES

[1] H. Inoue, A. Miyazaki, T. Katsura, "An Image Watermarking Method Based On the Wavelet Transform." Vol. 1,pp. 296-300. Aug 2002.

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

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

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

[5] 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.

[6] 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.

[7] 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.

[8] D. R. ElShafie, N. Kharma, R. Ward, "Parameter Optimization of an embedded watermark Using a Genetic Algorithm," International Symposium on Communications, Control and signal Processing, pp 1263-1267,Jun 2008.

[9] P. Chen, H. Lin, "A DWT Based Approach for Image Steganography." International Journal of Applied Science and Engineering,Vol. 4, No. 3, pp. 275-290, 2006.

[10] B. Lai and L.Chang, "Adaptive Data Hiding for Images Based on Haar Discrete Wavelet transform," Lecture Notes in Computer Science, Vol 4319, 2006.