A framework of secured and bio-inspired image steganography using chaotic encryption with genetic algorithm optimization (CEGAO)

A Framework of Secured and Bio-Inspired Image Steganography using Chaotic Encryption with Genetic Algorithm Optimization (CEGAO)

Debiprasad Bandyopadhyay[1], Kousik Dasgupta[1], J. K. Mandal[2], Paramartha Dutta[3], Varun Kumar Ojha[4], Vaclav Snasel[4]

[1]Dept. of CSE, Kalyani Government Engineering College, Kalyani-741 235, India
{debigr8,kousik.dasgupta}@gmail.com
[2]Dept. of CSE, Kalyani University, Kalyani-741 235, India
jkm.cse@gmail.com
[3]Dept. of CSS, Visva-Bharati University, Santiniketan -731 235, India
paramartha.dutta@gmail.com
[4]IT4Innovations & Dept. of Computer Science, VSB – Technical University of Ostrava, Czech Republic
{varun.kumar.ojha,vaclav.snasel}@vsb.cz

Abstract. The two key issues related to steganography techniques are, statistical undetectability and picture quality. Image steganography takes the advantage of limited power of Human Visual System (HVS). The proposed framework offers an approach of secure data hiding technique in digital images. Novel scheme, presented encrypts meaningful secret data using nonlinear dynamics (chaos theory) before embedding into host or cover image. A basic LSB embedding method is used for encrypting data into cover image. Genetic Algorithm based pixel adjustment process is used to reduce the difference of error between the host image and its stego version with low distortions. The results of proposed scheme are compared with other steganographic algorithm using Peak Signal to Noise Ratio (*PSNR*) and Structural Similarity (*SSIM*) index, color frequency test and StirMark analysis.

Keywords: Image Steganography, Spatial Domain, Dynamic System, Chaotic Maps, HVS, Steganalysis.

1 Introduction and Motivation

Steganography, an information hiding method, is used in secret communication. It is about secure transmission of secret information using images, audio, video and other digital media as a cover, embedding the secret information to be sent into the carrier signal, transmitting as an unnoticeable way through public channels, aiming at sending out the information secretly and safely without causing the suspicion of the behavior of hidden message [1]. Steganographic methods can be classified into spatial domain and frequency domain embedding. Steganography in spatial domain is more

challenging, as the change in the image content may leave visually or statistically detectable features [2]. This paper deals with LSB replacing method due to its high embedding capacity and low computational complexity.

Chaos theory was developed by Edward Lorenz in 1972, is a mathematical physics which cuts across traditional scientific disciplines, trying to gather unrelated kinds of wilderness and irregularity into a system. Chaos occurs when a system is very sensitive to initial conditions. Two nearly undistinguishable sets of initial conditions for the same system will result in two final situations that differ greatly from each other. Chaotic systems are mathematically deterministic but nearly impossible to predict [3].

In recent years, use of chaotic system in several fields is noticeable. Easy implementation, more randomness, confidentiality and non-periodicity are the key issues related to chaotic steganographic techniques. K Ganesan et. al. [4] focused on developing algorithm that can be used to hide the secret messages using random number logic. Based on the fractal theory, an optimization technique has been presented in [5] by modifying a chaos optimization algorithm (COA). A hybrid model of chaotic function and cellular automata is presented in [6]. Lifang Yu et al. [7] proposed an improved adaptive steganography method where chaotic parameters are selected by the Genetic Algorithm.

2 Proposed Method (CEGAO)

The proposed scheme offers an approach to encrypt the secret image based on chaos theory before embedding into the cover image. A Genetic Algorithm (GA) based optimization technique is used to reduce the difference error between the host image and its stego version. After extracting the encrypted secret image from the stego image chaotic decryption is performed to get back the actual secret image.

2.1 Preliminary

The logistic map [3] is one of the simplest chaotic maps defined by the quadratic recurrence equation as per (1) below,

$$X_{k+1} = \mu X_k(1 - X_k) \tag{1}$$

Where $0 \leq \mu \leq 4$, $X_k \in [0,1]$, μ is a control parameter or bifurcation parameter. The logistic map stands in chaotic state when $3.5699456 < \mu \leq 4$ [3]. Thus, the sequence $\{X_{k(k=0,1,2,\ldots,N-1}\}$ generated is non-periodic and non-convergent. Two logistic sequences generated from different initial conditions are statistically uncorrelated. Figure 1 illustrates bifurcation diagram of logistic map for different values of the bifurcation parameter.

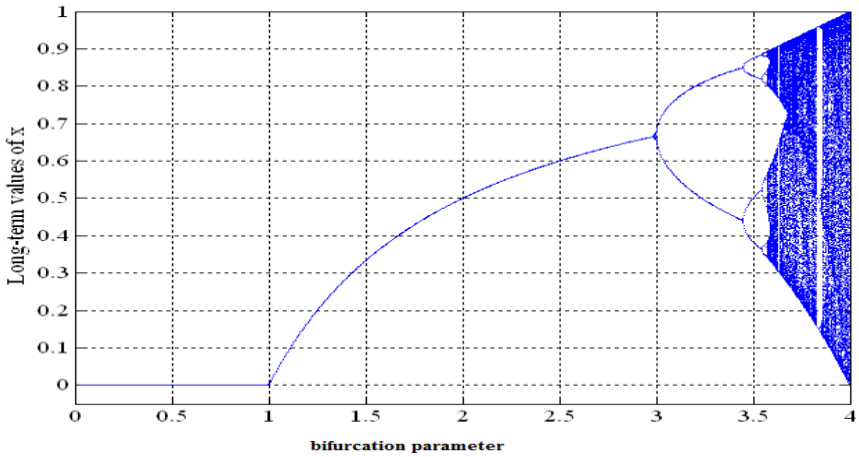


Fig. 1. Bifurcation diagram of the logistic map (Source: http://hyperchaos.files.wordpress.com)

Let i be the 24-bit secret image of size $n \times n$ represented as $x(i,j)$ where $0 \leq i < n$ *and* $0 \leq j < n$. In this paper the secret image is divided into four equal blocks each of size $b \times b$ where $b=n/2$. Thus the total number of pixels in each block is $b \times b$. Each pixel has three 8 bit components R (Red), G (Green) and B (Blue) respectively. So the total number of components (N) in a secret image block is $N = b \times b \times 3$.

Now for the first secret image block using the logistic map as per (1), a chaotic sequence $\{X_{k(k=0,1,2,\ldots,N-1}\}$ of N real numbers is generated with the initial values of μ and X are considered as 3.89 and 0.50 respectively, as the map shows its chaotic behavior when $3.5699456 < \mu \leq 4$ and $X \in (0,1)$.

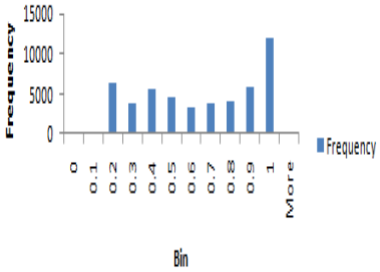
Consider a Lenna image of size 256×256 divided into four equal blocks each of size 128×128 as shown in Figure 3(a). For encryption of the first block a chaotic sequence of 128×128×3 = 49,152 elements is generated using the logistic map as given in (1). The logistic map is iterated 49,152 times. Table 1 shows the elements generated from the logistic map with the initial values of μ and X considered as 3.89 and 0.50 for 20 iterations where actual number of elements generated is 49,152. The grouped frequency distribution of the chaotic numbers with 10 intervals and the corresponding histogram are shown in Figure 2.

Table 1. Elements generated using the logistic map for 20 iterations

0.9725	0.1040	0.3625	0.8990	0.3530	0.8885	0.3853	0.9213	0.2818	0.7873
0.6513	0.8833	0.4008	0.9342	0.2389	0.7073	0.8053	0.6098	0.9255	0.2680

Lower Class Boundary	Upper Class Boundary	Frequency	Cumulative Frequency
0	0.1	0	0
0.1	0.2	6345	6345
0.2	0.3	3707	10052
0.3	0.4	5479	15531
0.4	0.5	4481	20012
0.5	0.6	3415	23427
0.6	0.7	3763	27190
0.7	0.8	4124	31314
0.8	0.9	5905	37219
0.9	1	11933	49152

(a)



(b)

Fig. 2. (a)Frequency distribution of chaotic numbers and corresponding (b) Histogram

The histogram exhibits that there is an asymmetry in the frequency distribution of the chaotic numbers. For asymmetrical distribution, median is a suitable measure of location [8]. So the median [8] of this grouped data is calculated according to the formula given in (2).

$$Median = L + \frac{\frac{N}{2} - cf_b}{f_m} \times w \tag{2}$$

Where L is the lower class boundary of the group containing the median, N is the total frequency, cf_b is the cumulative frequency of the groups before the median group, f_m is the frequency of the median group and w is the group width. The median calculated according to (2) is 0.6305 for the first secret image block. Next the median is considered as threshold T. Thus for each $X_{k(k=0,1,2,\ldots,N-1)} \geq T$ then $B_{k(k=0,1,2,\ldots,N-1)}=1$ otherwise 0. A binary sequence $\{B_{k(k=0,1,2,\ldots,N-1}\}$ generated for the first secret image block for 20 iterations is shown in Table 2.

Table 2. Binary sequence generated in 20 iterations

1	0	0	1	0	1	0	1	0	1	1	1	0	1	0	1	1	1	1	0

Now for each 8 bit component $C_{k(k=0,1,\ldots N-1)}$ of the first secret image block an XOR operation of each bit of the component C_k is done with a single bit of B_k in the binary sequence e.g. if C_0 = 01010000 and B_0 = 1 then $\overline{C_0}$ = 10101111. Where $\overline{C_0}$ is the encrypted component. Repeat this procedure until all such components of the first secret image block is encrypted.

The same procedure is followed for the remaining three blocks of the secret image with different initial conditions of the bifurcation parameter μ. The values of μ considered are 3.90, 3.91, and 3.92 for the remaining three blocks. Finally all the encrypted blocks are merged to form the encrypted secret image. Now each eight bit component of the encrypted secret image is

embedded into the cover image using the base embedding scheme as described in Section 2.3.

2.3 Base Embedding Scheme

In the base scheme first three bits of the secret message are concealed inside three (03) bits of LSB of Red pixel, next three bits in the three (03) bits of LSB of Green pixel. The remaining two bits of secret message are concealed in two (02) bits of LSB of Blue pixel. The particular distribution pattern is taken considering that the chromatic influence of blue to the human eye is more than that of red and green pixels [9].

2.4 Genetic Algorithm Optimization

Finally the quality of the stego image obtained from above is optimized using basic genetic algorithm approach. The first step is determining the initial population by designing the chromosome, fitness function and GA operators.

Chromosome Design: In this step different chromosomal representation of the pixel value of the stego image is obtained. Random selections of candidate solutions are made as initial population. Each of the individual in the initial population has the LSB data same as the encrypted secret data.

Fitness Function: The fitness of each individual in the population is calculated according to the fitness function F as follows.

$$F = w_1 \times p_1 + w_2 \times (1 - p_2) \tag{3}$$

The quality of the stego image is improved in sense of two performance indices MSE (p_1) and HVS deviation (p_2) [10]. MSE measure [11] as per (1) is the most widely accepted statistical image quality feature. The other preferred parameter sis SSIM (structural similarity) [11] index as given as per (6). A good SSIM metric indicates a stego image that is indistinguishable from the original version. The predefined weights w_1 = 0.8, w_2 = 0.2 are used for optimization of the objective function F.

Mutation: Each individual chromosome is evaluated as per (3) and best fitted chromosome is selected twice for mutation with probability 0.05 and the least fitted individual is discarded. The bits of the chromosomes, except the target layers, are changed from ‘1’ to ‘0’ or ‘0’ to ‘1’ depending upon the mutation value. The output of this step results in a new mating pool ready for crossover.

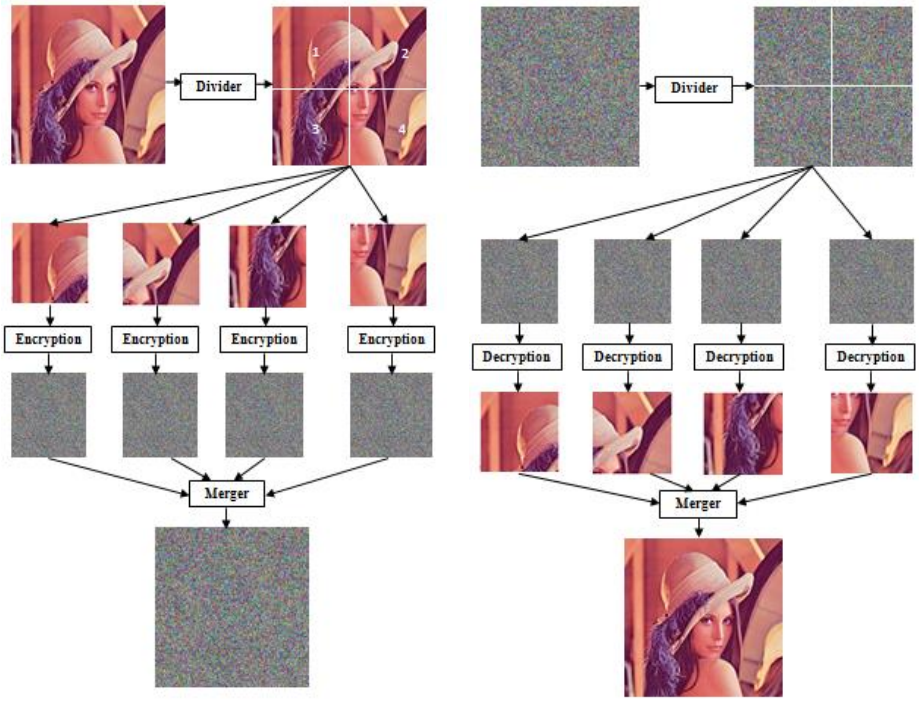
Crossover: A random single point crossover is chosen and portion lying on one side of crossover site is swapped with the other side. Thus a new pair of individuals is generated.

The steps Mutation and Crossover are repeated iteratively till either we get a chromosome having pixel value closest to the original value or maximum number of iterations is completed. Finally the optimized stego image is obtained.

2.5 Extraction and Chaotic Decryption

A reverse method is applied for decoding where all the bits are extracted from LSB of the RGB pixels of the stego image in the order 3, 3, 2 respectively [9]. The encrypted secret image so obtained of size $n \times n$ is divided into four equal blocks as shown in Figure 3(b). A chaotic sequence of N real numbers is generated as $\{X_{k(k=0,1,2,\ldots,N-1}\}$, considering the same initial values of μ and X as in encoding. Next the median is considered as threshold value T. Now for each $X_{k(k=0,1,2,\ldots,N-1)} \geq T$ then $B_{k(k=0,1,2,\ldots,N-1)} = 1$ otherwise 0. Thus a binary sequence $\{B_{k(k=0,1,2,\ldots,N-1)}\}$ is generated. Now for each encrypted 8 bit component $\overline{C_k}$ of the encrypted secret image part XOR each bit of the component with a single bit $B_{k(k=0,1,2,\ldots,N-1)}$ in the binary sequence e.g. if the extracted encrypted component $\overline{C_0}$ =10101111 and B_0 =1 then then the decrypted component C_k will be 01010000.

This procedure is repeated until all the encrypted eight bit components of the secret image block are decrypted. For decryption of the remaining three blocks the same procedure is followed with different initial conditions of the logistic map as assumed during encoding. All the decrypted secret image blocks are merged to form the original secret image of size $n \times n$.



(a) (b)

Fig. 3. (a) Chaotic encryption process of the secret image Lenna before embedding and (b) Chaotic decryption process of the encrypted Lenna after extraction

A simulation environment is implemented using visual C++ 2013 as IDE and OpenCV 2.4.6.0 as the graphics library. The proposed technique is applied on various 24-bit JPEG images as cover images. Three of them are "Baboon", "Peppers", and "Tiffany" of size 512 × 512. A 24-bit image Lenna.png of size 256 × 256 is taken as the secret image. The experimental results (as given in Table 3) are compared with the existing C-LSB [12] algorithm in terms of PSNR and SSIM. The SSIM measure given in (6), measures the perceptual similarity between the original image and its stego version. The quality measure is measured by PSNR [11] as per (4).

$$\text{PSNR} = 10log_{10}\frac{L^2}{MSE} \tag{4}$$

Where L is peak signal level for a gray scale image and it is taken as 255. The value of MSE [11] is computed as per (5).

$$\text{MSE}= \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W}(P_{ij} - S_{ij})^2 \tag{5}$$

Where H and W are height and width and P_{ij} represents the original image and S_{ij} represents corresponding stego image. Whereas, the similarity measure between two images x and y is measured by Structural Similarity (SSIM) index [11] as follows.

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1) \times (2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1) \times (\sigma_x^2 + \sigma_y^2 + c_2)} \tag{6}$$

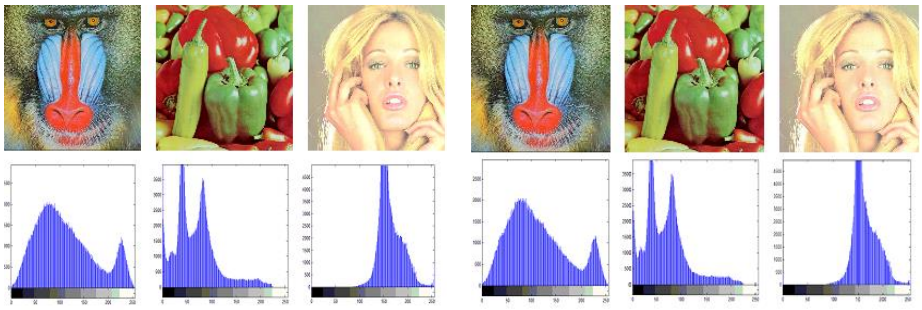
Where μ_x and μ_y are the average of x and y, σ_{xy} is the covariance of x and y, σ_x^2 and σ_y^2 are the variance of x and y. Here $c_1 = (k_2L)^2$ and $c_2 = (k_2L)^2$ where $k_1 = 0.01$, $k_2 = 0.03$ by default and $L = 255$ for a gray scale image.

Table 3. Comparative Study of the proposed technique CEGAO with the C-LSB [11] technique

Cover Image	Method	PSNR	SSIM		
			R (Red) Channel	G (Green) Channel	B (Blue) Channel
Baboon	C-LSB[11]	46.76	0.93	0.97	0.95
	CEGAO	48.19	0.99	0.99	0.99
Peppers	C-LSB[11]	45.81	0.92	0.94	0.98
	CEGAO	47.59	0.99	0.99	0.99
Tiffany	C-LSB[11]	46.53	0.98	0.97	0.96
	CEGAO	48.55	0.99	0.99	0.99

3.1 Resistance against Visual Attacks

For proposed scheme the stego images are visually indistinguishable from their cover counterparts. The original, stego version and their corresponding histograms for the blue components are shown in Figure 4. The results for the other components are similar.



(a) (b)

Fig. 4. (a) Cover images (baboon, peppers and tiffany) and their corresponding histogram (b) Stego images and their corresponding histogram

3.2 Resistance to Statistical Tests

The objective of Steganalysis is to develop some standard statistical tests to determine if the stego image contains any hidden message. Color frequency test is one of the popular first order statistical tests. Here the results of the test are shown empirically.

Color Frequency Test. According to Westfield and Pfitzmann [13] the probability of embedding is given as per (7) below,

$$p = \frac{1}{2^{\frac{v}{2}}\Gamma(\frac{v}{2})}\int_0^{\chi^2_{k-1}} t^{\frac{v}{2}-1}e^{-\frac{t}{2}} \quad dt \tag{7}$$

Where v is the degrees of freedom and $v = k - 1$ is the distinct color categories. The quantity $\chi^2 = \sum_i \frac{(y_i - y_i^*)^2}{y_i^*}$ follows chi-square distribution. The expected distribution y_i^* compared against the observed distributiony_i. In the proposed scheme the probability is computed by taking a sample of 1% of the total number of pixels. The sample is increased by 1% in each iteration till it went up to 100%. In the experiment it is found that the probability of embedding yielded approx 0 for every sample.

StirMark Analysis. Any steganographic scheme should resist some standard tests to establish strength and robustness. The test has been done using StirMark 4.0 [14] and it shows good results. A sample of the results is shown in Table 4.

Table 4. StirMark Analysis of the proposed scheme on peppers image (size: 512 X 512)

Tests	Factor	Cover Image	Stego Image
Self Similarities	1	34.2465	34.2462
Self Similarities	2	48.7681	48.7602
PSNR	10	37.6924	37.7050
PSNR	20	33.3307	33.3395
AddNoise	20	9.6329	9.6384
AddNoise	40	8.2113	8.2126
Small Random Distortions	1.00	15.6120	15.5954
Small Random Distortions	1.05	15.3870	15.2895
ConvFilter	1.00	12.1887	12.1879
ConvFilter	2.00	-5.2074	-5.2074
MedianCut	3.00	29.4185	29.4179

4 Conclusion

A secure Genetic Algorithm based image steganography technique GACIS has been proposed in this paper using the concept of non-linear dynamic system (chaos). This paper shows that the proposed algorithm obtained better results than an existing approach C-LSB which results in good visual imperceptibility of the stego image with low distortions. The proposed technique is applied to JPEG files as cover images, however it can work with any other formats. Further work focuses on improving the efficiency of the algorithm by finding the best parameters for the logistic map using soft computing techniques.

Acknowledgement

….

References

1. Rudramath, P.R., Madki, M.R.: High Capacity Data Embedding Technique using Improved BPCS Steganography. International Journal of Scientific and Research Publications, vol. 2, issue 7, 1-4, (2012).
2. Paul, Goutam, Davidson, Ian, Mukherjee, Iman, Ravi, S.S.: Keyless Steganography in Spatial Domain Using Energetic Pixels. In: ICISS 2012, LNCS 7671, Springer, pp. 134-148 (2012).
3. Gleick, James: CHAOS, The Amazing Science of the Unpredictable. Vintage Books (1998).

4. Ganesan, K., Venkatalakshmi, B., Moothy , R.K.: Steganography Using Enhanced Chaotic Encryption Technique. In: International Conference of Communication Systems, (2004).
5. Tavazoi, M.S., Haeri, Mahammad: An optimization algorithm based on chaotic behavior and fractul nature. Journal of Computational and Applied Mathematics 206, 1070-1081, (2007).
6. Alirezaanejad, Mehdi, Enayatifar, Rasul: Steganography by using logistic map function and cellular automata. Research Journal of Applied Sciences, Engineering and Technology, 4991-4995, (2012).
7. Yu, Liafang, Zhao, Yao, Ni, Rongrong, Li, Ting: Improved Adaptive LSB Steganography Based on Chaos and Genetic Algorithm. EURASIP Journal on advances in signal Processing, vol. 2010,1-6, (2010).
8. Das, N.G.: Statistical Methods. Tata-McGraw-Hill Education, 1st edition, (2008).
9. Dasgupta, Kousik, Mandal, J. K., Dutta, Paramartha: Hash Based Least Significant Bit Technique For Video Steganography (HLSB). International Journal of Security, Privacy and Trust Management (IJSPTM), vol. 1, no. 2, 1-11 (2012).
10. Dasgupta, Kousik, Mandal, J. K., Dutta, Paramartha: Optimized Video Steganography using Genetic Algorithm (GA). In: Procedia Computer Science 00(2013), 1-8 (2013).
11. Kutter, M., Petitcolas, F. A. P.: A Fair Benchmark For Image Watermarking Systems. Electronic Imaging ’99. Security and Watermarking of Multimedia Contents, The International Society for Optical Engineering, vol. 3657, 1-14 (1999).
12. Bandyopadhyay, Debiprasad, Dasgupta, Kousik, Mandal, J.K., Dutta, Paramartha: A Novel Secure Image Steganography Method Based on Chaos Theory In Spatial Domain. International Journal of Security, Privacy and Trust Management (IJSPTM), vol. 3, no. 1, 11-22 (2014).
13. Westfield, A., Pfitzmann, A.: Attacks on Steganographic Systems. In: Pfitzmann, A. (ed.) IH 1999. LNCS, Springer, Heidelberg, vol. 1768, 61-76 (2000).
14. StirMark Benchmark, http://www.petitcolas.net/fabien/watermarking/stirmark