A Novel Data Hiding Method by Using Chaotic Map and Histogram

R. Haghnazar Koochaksaraei, V. Aghazarian, A. Haroonabadi, and A. Hedayati

***Abstract*—Data hiding is to conceal the existence of secret data and it is considered for more protection of multimedia data. A reversible data hiding method can extract the cover image without any distortion from the stego-image after the hidden data have been extracted. This study tackles a chaotic based reversible data hiding. In this paper first image histogram is employed for detect the pixels which are selected for hiding a bit of secret data, then after a sequence of hiding a bit stream is determined by logistic chaotic map.**

Experimental results show that WICA not only demonstrates superior hiding effect, but also resists various typical attack. The obtained PSNR of the proposed method is approximately 54 which is proven our method excellence.

***Index Terms*—Data hiding, chaotic map, and entropy.**

I. Introduction

Due to the fast growth of multimedia products, and the widespread distribution of digital products on the Internet, protection of digital information from unauthorized copying and distribution is becomingmore important each passing day. To achieve this goal, various algorithms have been proposed for hiding information in images [1]-[5].Information can be hidden in different ways, but the most important, and the most widely used methodintroduced, is the employment of the least valuable bits for hiding the information [6].

There are a few methods of hiding information which are not able to recover hidden information from the cover data after the hiding operation is carried out [7], [8], while many types of information must be completely recoverable to be considered valuable. For this very reason, methods of hiding recoverable information have become very important these days; and many methods have been proposed for this purpose [1]-[3], [9]-[11].

In [4], a method has been introduced for hiding information on the basis of the differences and the means of two neighboring pixels. In this method, the encrypted data is hiddenbetween the differences of the two adjacent pixels. In another research presented in reference number [5], a method called the Codebooks method is introduced in which the differences among several codewords are used to hide the bits of the encrypted data. In [12], a method based on the chaotic model has been used. In this method, each bit of the encrypted data is hidden in the least valuable bit of a pixel of the image the position of which is determined by the chaotic model. In [13] also, a method has been introduced which is based on the histogram of the image; and in which the highest and the lowest frequency of the gray surfaces in the image are employed for hiding the bits of the encrypted data. The major difficulties in most of these presented methods can be studied under two headings: (a) the unbalanced distribution of the bits of the encrypted data on the surface of the cover image, and (b) the reduction in the hiding capacity

In this paper, a new method is introduced in which the chaotic model and the histogram of the image are used. The chaotic feature of the signals of the chaotic model has been extensively used in topics related toimage processing, while it has not been used much in the context of hiding information.

In the method proposed in this article, the histogram of the image is employed for finding the gray surfaces of the pixels of the cover image, each of which pixels covers one bit of the encrypted data. This greatly enhances the capacity for hiding information. On the other hand, the main duty of the chaotic model is the balanced distribution of the bits of the encrypted data on the gray surfaces determined by the histogram.

The rest of the paperis organized as follows: first, chaotic models are briefly described in section 2. In section 3, the proposed method is explained. The tests performed on the proposed method are dealt with in section 4. Finally, the conclusions drawn from the discussion are presented in section 5.

II. The Chaotic Model

Chaotic signals seem like noise, but they are completely definite: if the initial values and the mapping function are known, the same values can be accurately reproduced. The advantages of these signals are studied under the following three headings:

A. Sensitivity to the Initial Conditions

This means that any slight change in the initial values will cause huge changes in the subsequent values of the function – i.e., if there is a small change in the initial values of the signal, the resultant signal will be very different from the initial one.

Manuscript received April 16, 2012; revised June 10, 2012.

R. HaghnazarKoochaksaraei is with Department of Computer Science, Science and Research Branch, Islamic Azad University, Khouzestan-Iran (Roozbeh.haghnazar@gmail.com).

V. Aghazarian, A. Haroonabadi, A.Hedayati are with Department of Computer Science, Central Tehran Branch, Islamic Azad University, Tehran, Iran.

DOI: 10.7763/IJIMT.2012.V3.312

B. The Apparently Random Behavior

Compared to the producers of the natural random numbers in which the string of the random numbers produced cannot be reproduced, the methods used in producing random numbers in algorithms based on chaotic models allow the reproduction of the same random numbers, provided that the initial values and the mapping function are known.

C. Definite Operation

Although chaotic models appear to be random, yet they are completely definite: if the mapping function and the initial values are known, a set of values can be produced (apparently without any order in their production) in order to be used in the reproduction of those same initial values. Equation 1 shows the signal known as the Logistic Map, which is one of the most famous signals having chaotic behavior:

$$X_{n+1} = rX_n(1 - X_n) \quad (1)$$

In this equation, X_n assumes a value in the interval [0, 1]. This signal, due to fact that the r parameter is divided into three different intervals, shows three different chaotic behaviors which, assuming that $X_0 = 0.3$, can be described as follows:

1) If r ϵ [0, 3], then the signal behaves somewhat chaotically in the first 10 iterations and becomes stable after the tenth iteration [9] seeing Fig.1 (a).
2) If r ϵ [3, 3.57], then the signal will behave somewhat chaotically in the first 20 iterations and, after the 20th iteration, varies between two stable values [9] seeing Fig. 1(b).
3) If r ϵ [3.57, 4], then the behavior of the signal will generally be chaotic [9] seeing Fig. 1 (c).

Considering what has been said above, and because in this article a completely chaotic model is needed for hiding information , the chaotic signal Logistic Map with the initial values of $X_0 = 0.3$ and r ϵ [3.57,4] is used.

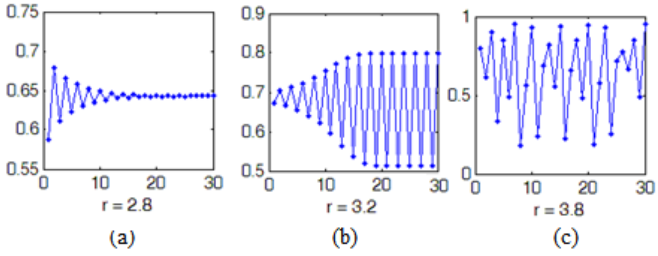

Fig. 1. The chaotic behavior of the Logistic Map signal with X0= 0.3. a) r∈ [0, 3] ,b)r ∈ [3,3.57],c) r∈ [3.57,4]

III. The Proposed Method

In the proposed method, the most frequent gray surfacein the image, which can be seen at the tip of the histogram diagram, is used for hiding the bits of the encrypted data. In the process of hiding information, first the most frequent gray surface and the gray surface with zero frequency are found. For example, Fig. 2 (a) represents the matrix of an imaginary image, and Fig. 2 (b) indicates the histogram of this image in which the most frequent gray surface and the gray surface with zero frequency are shown.

4	0	5	2	6
5	6	4	2	0
7	1	0	6	6
6	5	6	4	7
0	0	2	5	6

Fig. 2. (a) The gray surface of the image in the interval [0, 7],

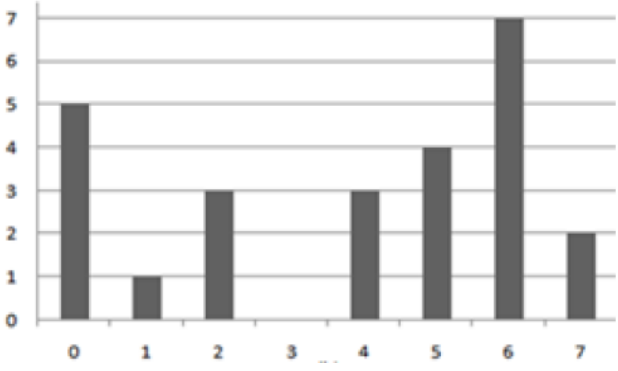

Fig. 2. (b) The histogram of the image in 2 (a).

As can be seen in Fig. 2 (a), the gray surface 6 is the most frequent, while and the gray surface 3 has zero frequency. The rules of embedding the bits of the encrypted data are determined based on the values of the pixels of the image. On the basis of the value of each pixel, three different positions can be assumed for the pixels of the image:

1) Pixels whose gray surfaces lie in between those of the most frequent gray surface and the gray surface with zero frequency (group 1 pixels);
2) Pixels whose gray surfaces lie outside the interval of the most frequent gray surface and the gray surface having zero frequency (group 2 pixels);
3) The pixels whose gray surfaces are equal to that of the pixel which has the most frequent gray surface (group 3 pixels). After classifying the pixels under study, the following steps are taken for hiding the bit string:

- Group 1 pixels

The pixels in this group approach the pixel with zero frequency.

- Group 2 pixels

The pixels in this group do not change at all.

- Group 3 pixels

After finding group 1 pixels, a number starting from zero is assigned to the position of each of the pixels. For example, for Fig. 2 we will have the following sequence (Fig. 3)

$6_{1.5}$	$6_{2.2}$	$6_{3.4}$	$6_{3.5}$	$6_{4.1}$	$6_{4.3}$	$6_{5.5}$
0	1	2	3	4	5	6

Fig. 3. Assigning a number to the most frequent pixel

The chaotic model is used to select one of the pixels. To do this, we need to produce an initial value to start the operation of the chaos model. This initial value is determined by using an 80-bit key as follows:

$$K = K_0, K_1, \ldots, K_9(ASCII) \quad (2)$$

In this key, K_i signifies an 8-bit block. The mentioned key will be converted into the binary form (Equation 3).

$$K = \left\{ \begin{array}{c} K_{01}, K_{02}, K_{03}, K_{04}, K_{05}, K_{06}, K_{07}, \\ K_{08}, \ldots\ldots\ldots K_{91}, K_{92}, K_{93}, \\ K_{94}, K_{95}, K_{96}, K_{97}, K_{98}(Binary) \end{array} \right\} \quad (3)$$

In equation 3 will represent the j^{th} bit of the i^{th}block; and, using equation 4, the value of X_0 is obtained in the interval [0, 1].

$$X_0 = \left\{ \begin{array}{l} B_{01} \times 2^{79} + B_{02} \times 2^{78} + \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ B_{11} \times 2^{71} + B_{12} \times 2^{70} + \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ + B_{n7} \times 2^1 + B_{r8} \times 2^0 \end{array} \right\} / 2^{80} \quad (4)$$

After finding the initial value of the Logistic Map chaotic function, the first (X_1) value in the interval [0, 1] is determined. This value is quantized in the interval [0, $n-1$] by using equation 5.

$$Position = round(X_n \times (n-1)) \quad (5)$$

In this equation, n indicates the length of the array shown in Fig. 3. The number produced in this stage is the number of one of the elementsof this array. With the determination of the first gray surface for hiding the first bit of the encrypted data, the remaining bits of the encrypted data are hidden on the basis of the following two rules:

1) If the value of the bit of the encrypted data is 1, the gray surface found using the chaos model does not change at all; 2) If the value of the bit of the encrypted data is zero, the gray surface found by using the chaos model approaches the gray surface having zero frequency (in the above example, 6 is converted to 5, because the gray surface having zero value is 3).

To better understand this process, we hide one bit of the encrypted data with the value of 10 in Fig. 2. In this example, it is assumed that the value of r is 3.99 and the initial value of X_0, according to equations 2, 3, and 4, is 0.55.

The first pixel for hiding the first bit is situated in element 6 of the array, the value of which, according to Fig. 3, is $6_{5,5}$; and since the first bit of the encrypted data is equal to 1 , this value does not change. The second bit is hidden as follows:

$$X_2 = 3.99 \times 0.987525 \times (1 - 0.987525) = 0.0491543,$$

$$\text{Position} = \text{round}\,(0.0491543 \times (7 - 1)) = 0$$

In this iteration, the gray surface in the first element of the array, which is $6_{1,5}$ is selected. Since the second value in the bit string is equal to zero, the gray surface in this element approaches the value of the gray surface with zero frequency: the number of the element $6_{1,5}$ changes from 6 to 5.

If we take all the above rules which concern the three groups of pixels into consideration, the final matrix in this example changes as shown in Fig. 4.

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

Fig. 4. The final image after hiding the encrypted bit string.

This process of finding pairs of maximums and minimums in the histogram of the image is continued and the bits of the encrypted data are hidden in the mentioned method.

IV. Empirical Results

In this section,the extent of the storage capacity and the security of the proposed method are discussed. To carry out the intended tests, a number of standard gray surfaces with the dimensions of 128×128 are employed.

A. The Capacity of Hiding Information in Images

In this test, we intend to calculate the number of bits of the encrypted data that can be hidden in the image. According to what has been said in section 3, the number of bits of the encrypted data that can be hidden in the image is equal to the number of pixels having the maximum frequency.

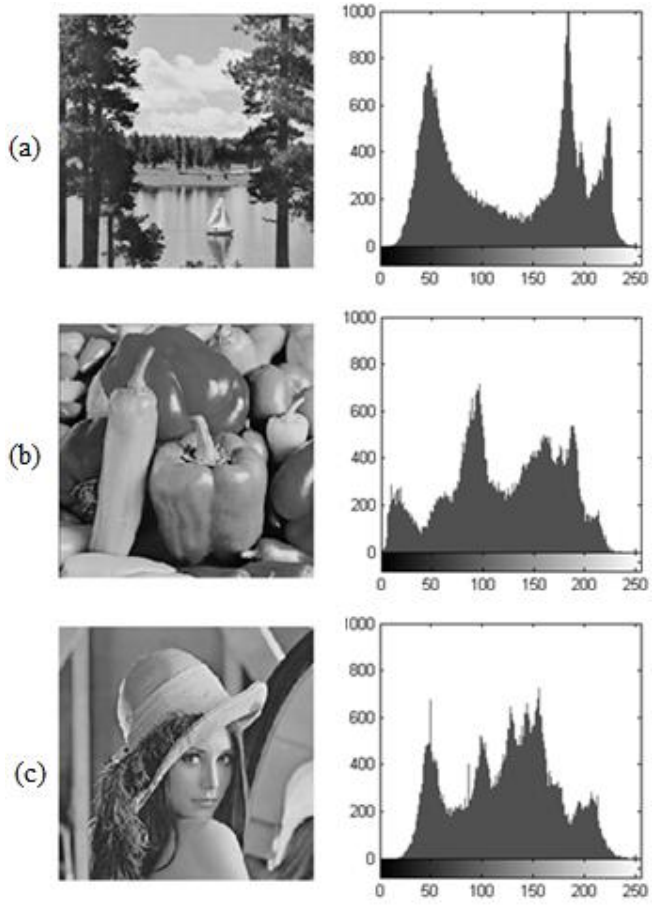

Fig. 5. (a) The image of a boat and its histogram, (b) the image of Peppers and its histogram, (c) the image of Lena and its histogram.

It can be seen in Fig. 5 that 2650,1974, and 2120 bits of the encrypted data can be hidden in the images of the Boat,Peppers, and Lena, respectively.

B. The Ratio of Signal to Noise

The ratio of signal to noise, which is used for measuring the quality of the Stego – Image, is called the PSNR of the image and it is calculated using equation 6:

$$PSNR = 10 \times log_{10}^{\frac{255^2}{MSE}} \qquad (6)$$

In which MSE, the mean square error, is obtained from the following equation:

$$MSE = \frac{1}{128 \times 128} \sum_{i=0}^{128} \sum_{j=0}^{128} (H_{i,j} - H'_{i,j})^2 \qquad (7)$$

In this equation, $H_{i,j}$and $H'_{i,j}$ represent the values of the gray surfaces of the pixel before and after the pixel in which the bit of the encrypted data is hidden.

To perform this test, we have used the images in Fig. 5 as the cover image, and we have also employed a bit string which has 1000 bits. The results obtained for the proposed method are shown in Table I.

In another part of this test, the ratio of the change in the PSNR of the image to the number of bits of the encrypted data is investigated. The results of this investigation are shown in the diagram in Fig. 6.

TABLE I: THE PSNR OF THE PROPOSED METHOD

Cover Image	PSNR
Boat	49.28
Peppers	51.14
Lena	51.94

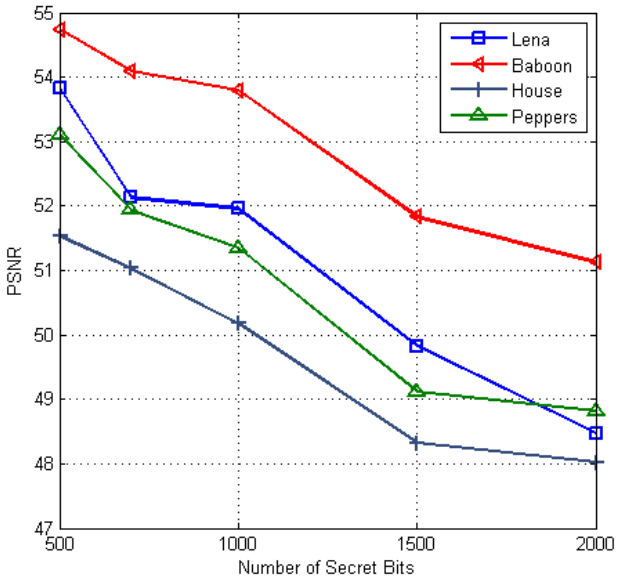

Fig. 6. The ratio of the change in the PSNR of the image to the number of bits in the encrypted data

As can be seen in the figure, the reduction in PSNR is very slight, as compared with the increase in the number of bits of the encrypted data; and this suggests that the quality of the image remains almost constant when the number of bits of the encrypted data increases.

C. The Resistance of the Proposed Method

One of the outstanding advantages of the proposed method is its resistance against common attacks in this domain. The balanced distribution of the bits of the encrypted data all over the cover image, which is carried out by the chaos model called Logistic Map, protects many bits of the encrypted data from attacks such as image truncation. For example, Fig. 7 shows a truncation attack on image 5-c after our method was used to hide 1000 bits of the encrypted data. Results of this test show that almost 132 bits of the encrypted data were destroyed in this attack, which means that about 87% of the bits of the encrypted data were preserved.

Fig. 7. Lena after 25% cutting

This test was conducted, as follows, on several different images using our proposed method and the methods introduced in references [14] and [15]: 1000 bits of the encrypted datawere hidden in different images and 25% of the images experienced a truncation attack. The results are shown in Table II.

TABLE II: PERCENTAGE OF BITS OF THE ENCRYPTED DATA DESTROYED

	PBRDH [14]	PCHS [8]	HCDA1
Lena	35%	39%	21%
Peppers	18%	24%	25%
Boat	47%	39%	27%

As can be seen in Table II, in most cases the proposed method performs better than the other two methods. The reason for this better performance is that in the proposed method the bits of the encrypted data are distributed in a balanced way on the surface of the cover image, while in the other two methods they are placed sequentially in the image.

V. CONCLUSIONS

In this article, a new method for hiding data in an image is proposed in which the information in the histogram and the features of the chaos model are used to hide the bits of the encrypted data in the image.

The histogram of the image is used to find the gray surfaces of the pixels of the cover image in each of which one bit of the encrypted data should be hidden. The main application of the chaos model is for the balanced distribution of the bits of the encrypted data in the gray surfaces determined by the histogram. This balanced distribution increases the resistance of the proposed method against different types of inflicted damages. As was also shown in the section on empirical results, the proposed method performs better than similar ones. Moreover, the value obtained for PSNR using the proposed method is 54, which value suggests the desirable efficiency of the proposed method.

REFERENCES

[1] L. Kamstra and H. J. A. M. Heijmans, "Reversible Data Embedding into Images Using Wavelet Techniques and Sorting," *IEEE Transactions on Image Processing*, vol. 4, no. 12, pp. 2082–2090, 2005.

[2] C. C. Chang, T. D. Kieu, Y. C. Chou, "Reversible Information Hiding for VQ Indices Based on Locally Adaptive Coding," *J. Vis. Commun. Image Represent*, vol. 20, no. 1, pp. 57–64, 2009.

[3] C. H. Yang and Y. C. Ling, "Reversible Data Hiding of a VQ Index Table Based on Referred Counts," *J. Vis. Commun. Image Represent*, vol. 20, no.6, pp. 399–407, 2009.

[4] J TIAN, "Reversible Data Embedding Using a Difference Expansion," *IEEE Trans. Circuits Syst. Video Technol*, vol. 13, no. 8, pp. 890–896, 2003

[5] C.C. Chang and C. Lut, Reversible Index-Domain Information Hiding Scheme Based on Side-Match Vector Quantization, *J. Syst. Soft*, vol. 79, no. 8, pp. 1120–1129, 2006.

[6] F. J. Neil and J. Sushil, "Exploring Steganography: Seeing the Unseen," *Computer Practices*, pp. 26–34, 1998.

[7] N. Yu, L. L. Cao, W. Fang, and X. L. Li, "Practical Analysis of Watermarking Capacity," in *Proc. IEEE International Conference on Communication Technology*, 2003, pp. 1872–1877.

[8] P. Tsai, Y. C. Hu, and C. C. Chang, "A Progressive Secret Reveal System Based on SPIHT Image Transmission," *Signal Processing: Image Commu- nication*, vol. 19, no. 3, pp. 285–297, 2004.

[9] Z. Ni, Y. Q. Shi, N. Ansari, and W. Su, "Reversible Data Hiding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 3, pp. 354–361, 2006.

[10] C. L. Tsai, H. F. Chiang, K. C. Fan, and C. D. Chung, "Reversible Data Hiding and Lossless Reconstruction of Binary Images Using Pair-Wise Logical Computation Mechanism," *Pattern Recognition*, vol. 38, no. 11, pp. 1993–2006, 2005.

[11] L. Kamstra, H. J. A. M. Heijmans, "Reversible Data Embedding into Images Using Wavelet Techniques and Sorting," *IEEE Transactions on Image Processing*, vol. 4, no. 12, pp. 2082–2090, 2005.

[12] R. Enayatifar, F. Mahmoudi, and K. Mirzaei, "Using the Chaotic Map in Image Steganography," *Int. Conference on Information Management, Kuala Lumpur, Malaysia*, pp.491-495.

[13] C. H. Yang and M. H. Tsai, "Improving Histogram-Based Reversible Data Hiding by Interleaving Predictions," *IET Image Process*, vol. 4, Iss. 4, pp. 223–234, 2010.

[14] H. W. Tseng and C. P. Hsieh, "Prediction-Based Reversible Data Hiding," *Inf. Sci.*, vol. 179, pp. 2460–2469, 2009.

[15] P. Tsai, Y. C. Hub, and H. L.Yeh, "Reversible Image Hiding Scheme Using Predictive Coding and Histogram Shifting," *Signal Processing*, vol. 89, pp. 1129–1143, 2009.