A New Chaos Steganography Algorithm for Hiding Multimedia Data

Mazhar Tayel, Hamed Shawky, Alaa El-Din Sayed Hafez

Electrical Engineering Department, Faculty of Engineering, Alexandria University

Hamed_shawky@yahoo.com, Alaahafez@ieee.org

Abstract— **The paper is devoted to propose a new chaos steganography algorithm for hiding the multimedia data, image, text, or sound. The proposed algorithm based on coordinate the data in the image dimensions using chaos distribution arrangement. The data is embedded with the original image in the pixels least significant bits, so can't appears within the image. When the image received the embedded data is separated and rearranged using the initial condition of the chaos coordination. The algorithm is implemented using Matlab program on three types of data, the first is image data, the second is a text data, and the third is sound signal data. The results show a good hiding for the tested data in the original images with high degree of security if steganalysis is performed on the composed original image.**

***Keywords*— Chaos, Steganography, Image Processing**

I. INTRODUCTION

Some confidential data might be stolen, copied, modified or destroyed by an unintended observer. Therefore, security problems become an essential issue. Encryption is a well-known procedure for secure data transmission. The frequently used encryption methods include RSA and DES algorithms. Although these two methods achieve certain security effects, they make the secret messages unreadable and unnatural. These unnatural messages usually attract some unintended observer's attention. This is the reason a new security approach called "Steganography" arises [1,2]. Steganography literally means covered writing and is the art of hiding secret messages within another seemingly innocuous message, or carrier [3]. With the advent of digital technology, the list of carriers has been made to include, images, text, e-mails, audio and video messages. Some Steganography utilities use secret keys. It can distinguish two kinds of keys: steganographic keys and cryptographic keys [4]. A steganographic key controls the embedding and extracting process. For example, it can scatter the message to be embedded over a subset of all suitable places in the carrier medium. Without the key, this subset is unknown, and each sample used to detect embedding by a statistical attack is a mixture of used and unused places which spoils the result. A cryptographic key is used to encrypt/decrypt the message [5,6]. This paper describes a new method of steganography technique for hiding large volumes of data using digital images, text, and audio as a cover medium. Before embedding the message in the cover medium, a logistic map chaotic encryption technique are used to obtain an undefined (encrypted) message then embedded the message in the cover media by using the proposed steganography method. This will be discussing in the next section in this work.

II. THEORETICAL DESCRIPTION

To provide a Theoretical description of steganography, terms and concepts should first be explained an overview of the different kinds of steganography is given at a later stage. Some definitions common to the steganography field: cover medium, this is the medium in which the data hide for it. Embedded message, this is the hidden message that want to put it in the cover medium. Stego-key, this is represented by some secret information, which is needed in order to extract the embedded message from the stego-medium. Stego-medium, this is the final piece of information that the casual observer can see. We can define this simple formula,

cover medium + embedded message = stego_message.

There are two general approaches to classify steganographic systems. The first approach is based on the type of cover file while the second approach is based on the hiding method or the layout of modification used in the embedding process [7]. These two general classification approaches of steganography are explained as follows, Cover-Type Based Classification, Since many kinds of digital media can be used as cover files of steganography, the first approach of classification breaks down steganography according to the type of the cover file used. However, the properties of these cover files vary from one type to another and these properties control how the secret data can be hidden in these cover files. To this end, knowing the type of cover file can give us an indication or idea where the secret data might be hidden. Mostly, steganographic systems are classified according to the cover file used. Accordingly, different steganography types can be distinguished such as: image, audio, video, text, and HTML steganography [7]. Hiding Method-Based Classification, Regardless of the cover type used for data hiding, steganography can be classified according to the method used to hide secret data. Furthermore, this approach of steganography classification is the most preferred approach in the steganography research community. Accordingly, there

are three ways to hide secret data in cover files: insertion based, substitution-based, and generation-based method [8].

Many steganography techniques have been proposed during the last few years. These techniques differ in the mechanism or principle being used to hide a secret message. Therefore, there are six categories of steganography techniques: spread spectrum techniques, statistical methods, distortion techniques, cover generation techniques, substitution systems, and transform domain techniques [9]. Spread spectrum communications define as" the process of spreading the bandwidth of a narrowband signal across a wide band of frequencies". In spread spectrum steganography, the frequency domain of the cover file is considered to be a communication channel and the secret message as a signal that is transmitted through it. Since the secret message is spread through a wide frequency band, this technique is relatively robust against stego file modification or message removal [9]. Statistical Techniques embed only one bit of secret data in a cover file. Therefore, it is known as "1-bit" steganography scheme. If "1" is hidden in a cover file, some statistical characteristics (e.g. entropy and probability distribution) of this cover file must be changed significantly to clearly indicate the existence of a message. However, if the hidden bit is "0", the cover file is left unmodified. Therefore, this technique entirely depends on the ability of the receiver to differentiate between changed and intact cover files [7, 9]. Distortion Techniques, which means that a receiver does not need the original cover file to extract the hidden message from the corresponding stego file. However, if a distortion technique is used, the receiver requires the original cover file in order to recover the secret message. For a receiver, the embedded message is the difference between the modified cover file received (the stego file) and the original cover file [6]. Cover Generation Techniques is preventing such kind of detection since only stego files are available and there is no cover files used. The major limitation of this method is the limited stego files which can be generated. Moreover, the generated stego files might be unrealistic files for end users (e.g. an image contains different shapes and colors without any sense or a text without any meaning). Therefore, the main media for such techniques are random-looking images and English text files [6]. Substitution Systems, it is important to find out some areas or data that can be modified without having any significant effects on this cover file. Therefore, a secret message can be embedded by replacing the redundant or insignificant parts of a cover file with secret message bits, without adding any significant noise to this cover file. Generally, digital covers have a large number of redundant bits (e.g. least significant bits (LSB)) [9]. In the substitution technique of steganography, the bits of the secret message substitute the LSB of the bytes of the cover file without causing a drastic change to this cover file. Moreover, the LSB technique is a spatial domain technique since it embeds the secret bits directly in the cover file. Since LSB substitution technique is relatively quick and easy to use, it is the most common technique used for digital steganography and especially with digital images. However, the embedded information using the LSB technique is highly vulnerable for images as a covering media, the LSB of a pixel is replaced with an M's bit. If we choose a 24-bit image as cover, we can store 3 bits in each pixel by modifying the LSBs of R, G and B array. To the human eye, the resulting stego image will look identical to the cover image [11, 12, 13]. Transform Domain Techniques unlike spatial domain techniques (e.g. LSB technique), transform (frequency) domain techniques hide secret data in significant parts of the cover file. Therefore, frequency domain techniques are considered more robust to attacks than spatial domain techniques. Hence, most of robust steganographic systems known today rely on frequency domain techniques. There are many transforms used to map a signal into the frequency domain. Discrete cosine transforms (DCT), discrete wavelet transforms (DWT), and discrete Fourier transform (DFT) are methods used as mediums to embed secret data in digital images. However, when we add a slight noise or secret data to some frequency domain components, it changes the whole image rather than changing only this part of the image. Thus, secret and embedded data will be spread across the entire image and will not be concentrated on one certain area or region [13, 14, 15].

A. Chaos theory

Chaos is a dynamical system that is extremely sensitive to its initial conditions. It is a deterministic nonlinear system that has random-like behaviors. Chaos theory has become a new branch of scientific studies today. Discrete chaotic dynamic systems are used in this system. The implemented map is logistic map, which is one of the simplest form of one dimensional chaotic maps and mathematically its equation can be written as [13]:

$$w_{i+1} = \mu w_i (1 - w_i) \qquad (1)$$

Where w_i is a real value in (0,1), and μ is bifurcation parameter satisfying $0 \leq \mu \leq 4$. The logistic map is chaotic on the condition $0.35699 \leq \mu \leq 4$.

B. Steganographic Algorithm evaluation

In order to make a decision of which steganographic system or technique is better than another an evaluation scheme for steganographic systems is needed. Currently, no standard test or measure is available in order to evaluate the performance or the effectiveness of steganographic systems. However there are some guidelines and general procedures that can be considered when evaluating or designing steganographic systems.

C. Evaluation of capacity

Evaluating the capacity of a Steganography technique means to find out the maximum number of bits that can undetectably be hidden. Increasing the Steganographic

capacity and maintaining an acceptable level of Stego image quality is considered a good contribution.

D. Evaluation of Imperceptibility

Two types of perceptibility can be distinguished and evaluated in signal processing systems, namely fidelity and quality. Fidelity means the perceptual similarity between signals before and after processing. However, quality is an absolute measure of the goodness of a signal to avoid any suspension and therefore detection. Even though the peak signal to nose ratio (PSNR) and the mean square error (MSE) are by definition fidelity metrics, a high quality image entails a large PSNR value and therefore both cover image and Stego-image are very similar and quite undistinguishable. Significantly, "fidelity" is defined as the perceptual quality of Stego files and therefore PSNR and MSE describe how imperceptible the secret message is. PSNR and MSE are the most common and widely-used metrics for Steganographic evaluation. Thus it is very important that there is no visual difference between the cover file and the Stego file. However PSNR values between 20 and 40 dB it is difficult for the human visual recognize any difference between a cover and Stego file if the PSNR value exceeds 36dB.

$$\mathrm{MSE} = \left(\frac{1}{\mathrm{MN}}\right)\sum_{i=1}^{M}\sum_{j=1}^{N}(X_{ij} - \overline{X_{ij}})^2 \qquad (2)$$

$$\mathrm{PSNR} = 10\log_{10}\frac{I^2}{\mathrm{MSE}}\ \mathrm{db} \qquad (3)$$

where X_{ij} is the i^{th} row and the j^{th} column pixel in the original (cover) data. $\overline{X_{ij}}$ is the i^{th} row and the j^{th} column pixel in the reconstructed (stego) data. M and n are the height and the width of the data, I is the dynamic range of amplitude values or the maximum value that a data can take for 8 –bit data; I =255.

III. The Proposed Method

To The proposed steganography algorithm started with increasing the original image pixels from byte to word colour capacity, and then distributes the hidden-image pixel randomly within the lower byte of the cover image pixels using chaos distribution. The stego-image then is ready for transmission. The original image is separated from the received stego-image at the first stage of the receiver. The initial condition of the chaotic random sequence is used to collect the stego-image from the lower byte of the pixels. Then the hidden-image is reconstructed. Figure 1. shows demonstrates the proposed method graphically. The same idea is performed with any kind of data such as text or sound signals.

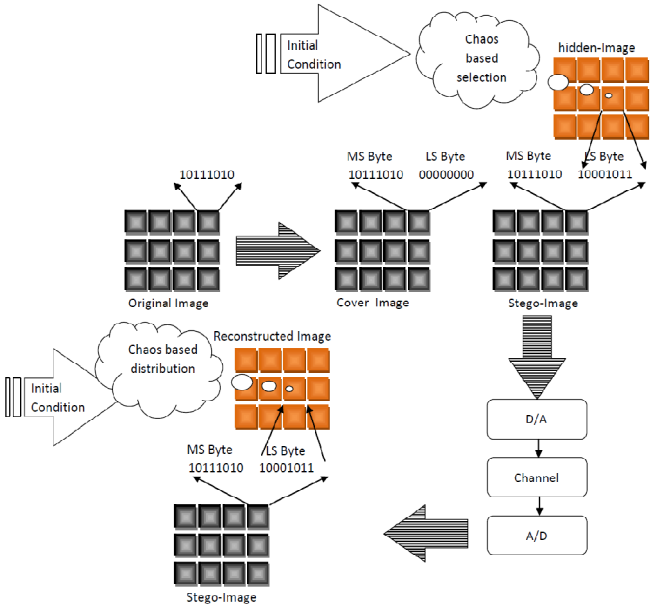
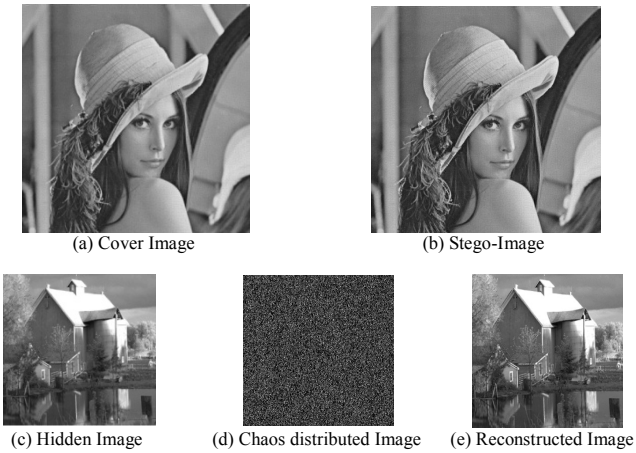


Figure 1. The proposed steganography algorithm

IV. Results

To verify the effectiveness of the proposed method, it has been implemented on two different test images with two hidden-images. Figure 2., 3. (a) shows the original cover image before processing, (b) demonstrate the stego-image after embed the hidden image chaotically within its pixels, (c) illustrate the hidden image before processing, (d) the chaos based distributed hidden image, and (e) shows the reconstructed image after reconstruction from the stego-image. The mean square error (MSE) for the first image is 7.315e+003 which achieve 57.6872 dB peak signal to noise ratio (PSNR), MSE for the second image is 2.2374e+004 which achieve 52.8321 dB peak signal to noise ratio (PSNR). These values of PSNR are acceptable for hiding the image data and can't notice by inspection. The same algorithm is applied to text data shown in Figure 4. and gives 5.7401e+003 MSE and achieved PSNR is 58.7404 dB. The sound test data MSE is 6.3252e+003 and PSNR is 58.3189 dB.



(a) Cover Image (b) Stego-Image

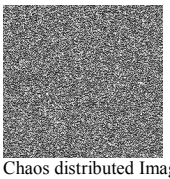
(c) Hidden Image (d) Chaos distributed Image (e) Reconstructed Image

Figure 2. stages of test image (1)

(a) Cover Image

(b) Stego-Image

(c) Hidden Image



(d) Chaos distributed Image

(e) Reconstructed Image

Figure 3. Stages of test image (2)

Humans accept input from five sense organs and senses: touch, smell, taste, sound, and sight in different physical formats (and even the sixth sense as mystics tell us) [1]. By some incredible process, not yet fully understood, humans transform input from these organs within the brain into the sensation of being in some reality. We need to feel or be assured that we are somewhere, in some coordinates, in some place, and at some time. Thus, we obtain a more complete picture of an observed scene than would have been possible otherwise (i.e., using only one sense organ or sensor). The human activities of planning, acting, investigating, market analysis, military intelligence, complex art work, complex dance sequences, creation of music, and journalism are good examples of activities that use advanced data fusion (DF) aspects and concepts [1] that we do not yet fully understand. Perhaps, the human brain combines such data or information without using any automatic aids, because it has a powerful associative reasoning ability, evolved over thousands of years.

(a) Stego-text

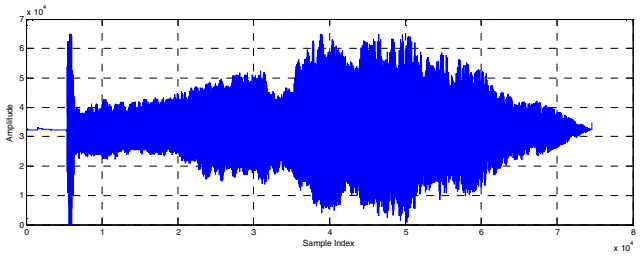
Animals recognize their changing environment by processing and evaluation of signals, such as data and some crude information, from multiple directions and multiple sensors sight, sound, touch, smell, and taste [1]. Nature, through the process of natural selection in very long cycles, has evolved a way to integrate the information from multiple sources and sensing organs to

(b) Hidden text

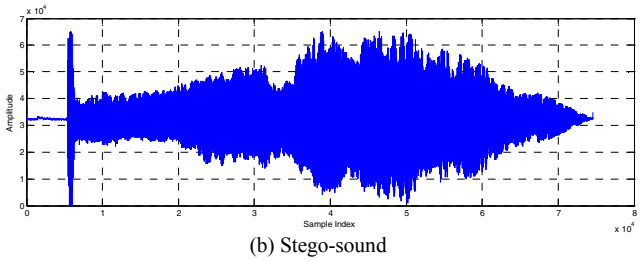
Figure 4. Simulated Test Text

V. CONCLUSIONS

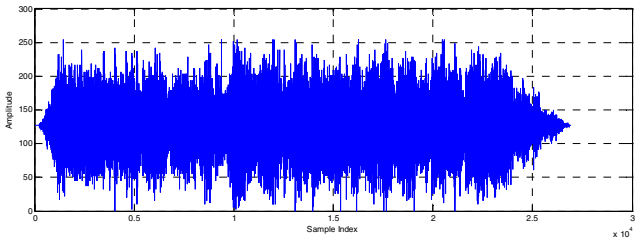
The paper proposes a new chaos steganography algorithm for hiding the multimedia data, image, text, or sound. The data is embedded with the original image in the pixels least significant bits, so can't appears within the image. The algorithm is implemented using Matlab program on three types of data, the first is image data, the second is a text data, and the third is sound signal data. The results show a good hiding for the tested data in the original images with high degree of security if steganalysis is performed on the composed original image.



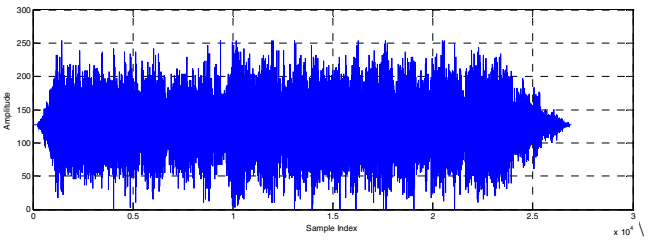
(a) Original cover sound



(b) Stego-sound



(c) Hidden sound



(d) Reconstructed sound

Figure 5. Simulated Test Audio Signal

REFERENCES

[1] *Uma Devi.G MS by Research – CSE" Steganography-Survey on File Systems" by Research – CSE by Research – CSE 19 October 2006.*

[2] *S.K.Bandyopadhyay, Debnath Bhattacharyya, Poulumi Das, S. Mukherjee, Ganguly, "A Tutorial Review on Steganography", IC3 Noida, pp. 106-114, August 2008.*

[3] *Richerd A. Mllin "An Introduction to Cryptography" Second Edition Discrete Mathematics and its application Series Editor Kenneth H. Rosen 2007 by Taylor & Francis Group, LLC www. copyright.com (http://www.copyright.com/).*

[4] *Abbas Cheddad, Joan Condell, Kevin Curran and Paul Mc Kevitt " Digital Image Steganography: Survey and Analysis of Current Methods", BT48 7JL, Northern Ireland, United Kingdom,2009.*

[5] *Peter Wayner "disappearing cryptography information hiding: steganography and watermarking" Published by Elsevier Inc, Copyright 2009.*

[6] *Send Lawyers, Guns, and Money" Introduction to Steganography" Revised July 6, 2009.*

[7] *Katzenbeisser, S. & Petitcolas, F. A. P. "Information Hiding Techniques for Steganography and Digital Watermarking", Artech House2000.*

[8] *Cole, E. "Hiding in Plain Sight: Steganography and the Art of Covert Communication", Indiana, John Wiley & Sons Inc (2003).*

[9] *Kipper, G. "Investigator's Guide to Steganography", Florida, CRC Press LLC (2004).*

[10] *Adel Almohammad, " Steganography-Based Secret and Reliable Communications: Improving Steganographic Capacity and Imperceptibility" thesis submitted for the degree of Doctor of Philosophy, Brunel University, August, 2010.*

[11] *Yogendra Kumar Jain "A Novel Image Steganography Method with Adaptive Number of Least Significant Bits Modification Based on Private Stego-Keys" International Journal of Computer Science and Security (IJCSS) Volume 4, Issue 1 Publishing Date: 30-03-2010.*

[12] *Rabah, K. Steganography "The Art of Hiding Data. Information Technology", Journal, 3, 245-269 (2004).*

[13] *Nuno Roma, Leonel Sousa"A tutorial overview on the properties of the discrete cosine transform for encoded image and video processing "an international journal in Signal Processing February 18, 2011, www.elsevier.com/locate/sigpro.*

[14] *Kh. Manglem Singh, S. Birendra Singh and L. Shyam Sundar Singh, "Hiding Encrypted Message in the Features of Images", IJCSNS, VOL. 7, No.4, April 2007.*

[15] *S. Praveen Kumar, K. Anusha, R.Venkata Ramana" A Novel Approach to Enhance Robustness in Steganography Using Multiple Watermark Embedding Algorithm" International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307 (Online), Volume-1, Issue-1, March 2011.*