

A Hybrid Chaos- Fuzzy -Threshold Steganography Algorithm for Hiding Secure 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— Security problem is an essential requirement for data transmission. Although the will known methods of data encryption attain certain security levels, they make the encrypted data unreadable, and unnatural, this attracts some unintended observer attention. So it is required to search for a higher security approach. The present paper introduces a new hybrid steganography algorithm to achieve a higher level of security. The proposed algorithm consists of a compound processing using Chaos – Fuzzy- Thresholding steganography method (CFT). Implementation of the proposed CFT hybrid algorithm shows a higher degree of security level. The hidden data reconstruction is based on an inverse CFT steganography algorithm processing.

Index Terms— Chaos, Steganography, Fuzzy, thresholding

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 compress the image color (or data) range using Fuzzy logic compressor and embed the message in the cover media by using the proposed steganography method. This will be discussing in the next section in this work.

II. THEORITICAL 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,

$$\text{cover medium} + \text{embedded message} = \text{stego_message.}$$

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

Manuscript received April 3, 2012. This work was supported by Electrical Engineering Department, Faculty of Engineering, Alexandria University, Alexandria, Egypt.

Mazhar Basyouni Tayel, Electronic Engineering Professor, Electrical Engineering Department, Faculty of Engineering, Alexandria University, Alexandria, Egypt (e-mail: Mazhar_Tayel@yahoo.com).

Alaa El-Din Sayed Hafez, Affilate Instructor, Electrical Engineering Department, Faculty of Engineering, Alexandria University, Alexandria, Egypt (e-mail: Alaahafez@yahoo.com).

Hamed Shawky, Post Graduate Student (Ph.D), Electrical Engineering Department, Faculty of Engineering, Alexandria University, Alexandria, Egypt (e-mail: Hamed_shawky@yahoo.com).

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. Chaotic Mapping

Chaos is a dynamical system that is extremely sensitive to its initial conditions. It is a deterministic nonlinear system that has random-like behaviours. 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]:

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

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. Fuzzy Logic Compressor

Fig. 1. shows the basic structure of a Fuzzy logic system. The main building units are a fuzzification unit, a Fuzzy logic reasoning unit, a knowledge base, and a defuzzification unit. Defuzzification is the process of converting inferred Fuzzy control actions into a crisp control action. The fuzzy knowledge-base has a rule-base that maps a Fuzzy inputs variable, E , into a Fuzzy output, U [16].

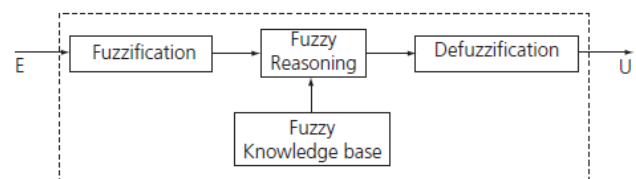


Fig. 1. Basic Structure of a Fuzzy Logic Controller (FLC)

This can be expressed by a linguistic statement such as:

$$E \rightarrow U \text{ (conditions } E \text{ implies condition } U)$$

which may be written as:

$$\text{IF } E \text{ THEN } U.$$

The Fuzzy knowledge-base also has a database defining the variables. A Fuzzy variable is defined by a Fuzzy set, which in turn is defined by a membership function. Fuzzy reasoning is used to infer the output contributed from each rule. The Fuzzy outputs reached from each rule are aggregated and defuzzified to generate a crisp output. In the design of an FL system it is assumed that, a solution exists, the input and output variables can be observed and measured, an adequate solution is acceptable, and a linguistic model can be created based on the knowledge of a human expert. In order to model a system linguistically, one needs to identify the input and output variables of the process to be controlled and define subsets that cover the universe of discourse of each variable and assign a linguistic label to each one. For example, the linguistic variable *speed* may be defined as three Fuzzy subsets: *slow*, *medium*, and *fast* as shown in Fig. 2. Form a rule-base by assigning relationships between inputs and outputs. Determine a defuzzification method to be used to generate a crisp output from the Fuzzy outputs generated from the rule-base.

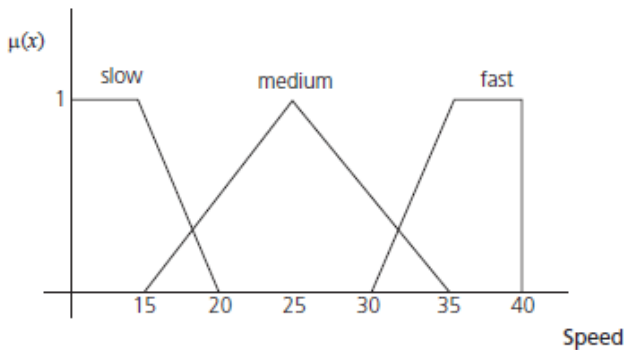


Fig. 2. Possible graphical representation of Fuzzy sets [18]

C. Membership Functions

Discrete and continuous membership functions of a Fuzzy set are intended to capture a person's thinking. Fuzzy membership functions can still be determined subjectively in practical problems based on an expert's opinion. In such a situation one can think of membership functions as a technique to formalize empirical problem solving that is based on experience rather than the knowledge of theory. The expert's way of thinking can be captured either directly or through a special algorithm. Such determination could become more focused by physical measurements if the need arises. Available frequency histograms and other probability data can also help in constructing the membership function. It is important, however, to note that membership function values, or grades of membership, are not probabilities and they do not have to add to 1. Membership construction can be further simplified by selecting their form from the smaller family of the commonly used ones.

D. Fuzzification

Fuzzification is the process of making a crisp quantity Fuzzy. We do this by simply recognizing that many of the quantities that we consider to be crisp and deterministic are actually not deterministic at all; they carry considerable uncertainty. If the

form of uncertainty happens to arise because of imprecision, ambiguity, or vagueness, then the variable is probably Fuzzy and can be represented by a membership function [16].

E. Defuzzification

For a given input, several IF/THEN rules could be launched at the same time. Each rule would have a different strength, because a given input may belong to more than one Fuzzy set, but with different membership values. In general, the output of the Fuzzy reasoning would involve more than two Fuzzy sets; therefore, one can write:

$$F = \bigcup_{i=1}^k F_i \quad (2)$$

Assuming the support of F is $X = \{x_1, x_2, x_3, \dots\}$ then for $x_i \in X$, $F(x_i) = w_i$ indicates the degree to which each is suggested by the rule-base as a good output for the given input. The defuzzification operation is applied on F to determine the best crisp output. Numerous defuzzification methods have been suggested in the literature; however, sometimes different authors name the same method differently. No method has proved to be always more advantageous than the others. The selection of which method to use depends primarily on the experience of the designer [17].

F. Centroid method

This method is also known as the *centre of mass*, or *centre of gravity* method. It is probably the most commonly used defuzzification method. The defuzzified output, x^* is defined by,

$$x^* = \frac{\int \mu_F(x) x dx}{\int \mu_F(x) dx} \quad (3)$$

where the symbol $\int$ denotes algebraic integration. Fig. 3. Represent this method graphically [18].

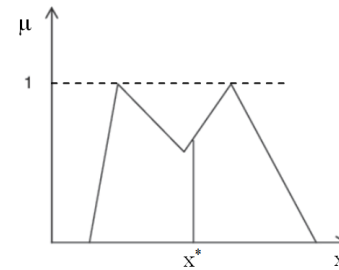


Fig. 3. Centroid Method

G. Centre of largest area method

This method is applicable when the output consists of at least two convex Fuzzy subsets which do not overlap. The result is biased towards a side of one membership function.

III. THE PROPOSED METHOD

The proposed steganography algorithm started with increasing the original image pixels from byte to word color capacity, and then distributes the hidden-image pixel randomly within the lower byte of the cover image pixels using chaos

distribution. Then this stego image processed using Fuzzy logic compressor to invert and compress its color range over a specific threshold color decided by the value of the upper significant bits of the image i.e. 192 to 255. 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 and invert and expand the color range using Fuzzy logic decompressor. 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. Fig. 4. demonstrates the proposed method graphically. The same idea is performed with any kind of data such as text or sound signals.

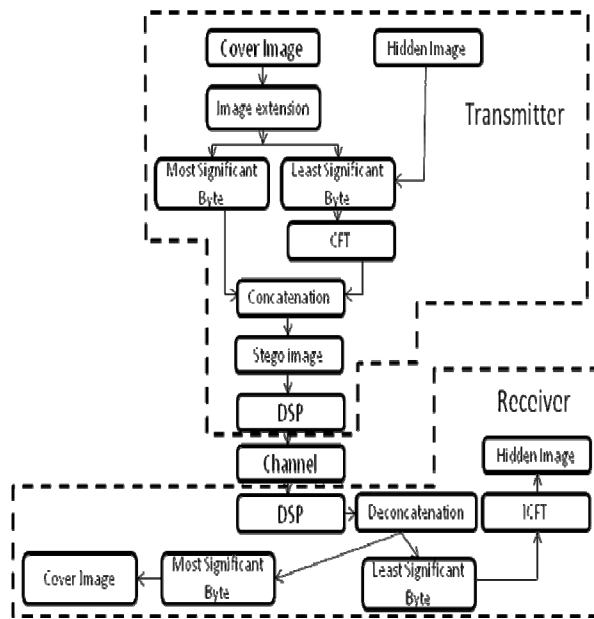


Fig. 4. The proposed Steganography algorithm

IV. RESULTS

The input and output membership functions of the proposed Fuzzy logic compressor (FLC) and Fuzzy logic Decompressor (FLD) is shown in Fig. 5. Fig. 6 demonstrate the transfer characteristics of the compressor and decompressor. To verify the effectiveness of the proposed method, it has been implemented on two different test images with two hidden-images. Fig. (7,8), (a) shows the original cover image before processing, (b) illustrate the hidden image before processing, (c) the chaos based distributed hidden image, (d) the compressed image (e) demonstrate the stego-image after embed the hidden image chaotically and compressed within its pixels, and (f) shows the reconstructed image after reconstruction from the stego-image. The mean square error (MSE) for the first image is $5.9143e+004$ which achieve 48.16 dB peak signal to noise ratio (PSNR), MSE for the second image is $5.2445e+004$ which achieve 49.13 dB peak signal to noise ratio (PSNR). These values of PSNR are acceptable for hiding the image data and can't notice by inspection.

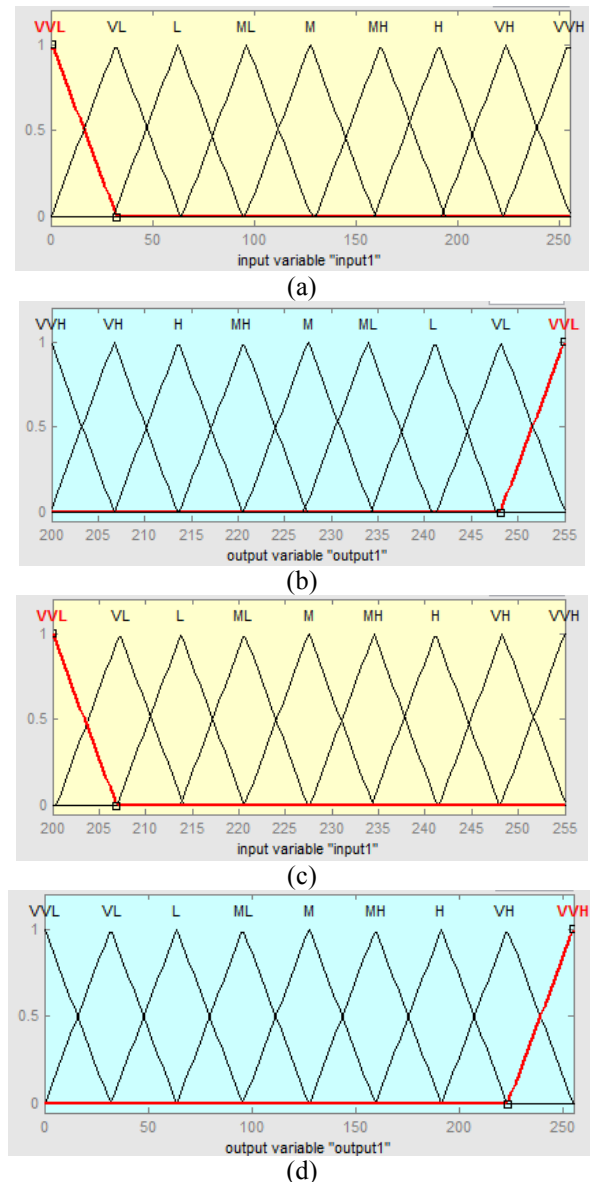


Fig. 5. Membership functions (a) input of the FLC, (b) output of the FLC, (c) input of the FLD, and (d) output of the FLD

V. CONCLUSION

The paper introduced a new hybrid chaos-Fuzzy-thresholding (CFT) steganography algorithm for hiding the multimedia data, image, text, or sound. The algorithm coordinates the data in the image dimensions using chaos distribution arrangement and performs a negative color compression using Fuzzy logic Scheme. The data is embedded with the original image in the pixels least significant bits, so can't appears within the image. The received image is separated and color expanded then rearranged using the initial condition of the chaos coordination. The algorithm is implemented using Matlab program for two test images. 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.

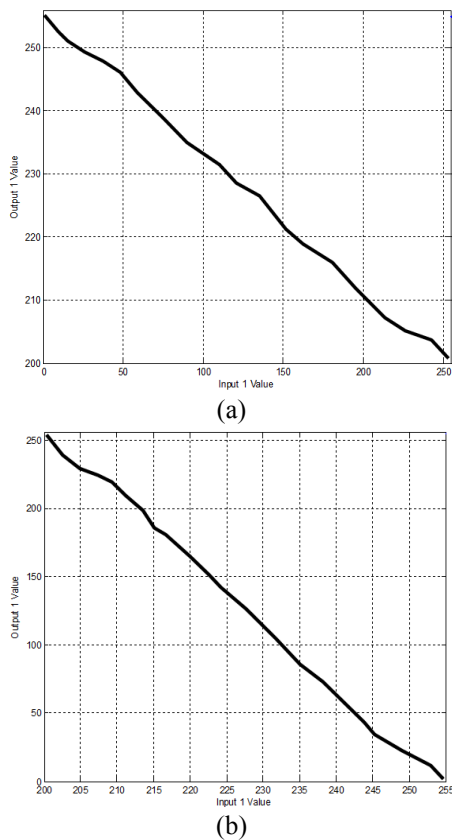


Fig. 6. Fuzzy logic transfer characteristics (a) T.C. of the FLC, and (b) T.C. of the FLD

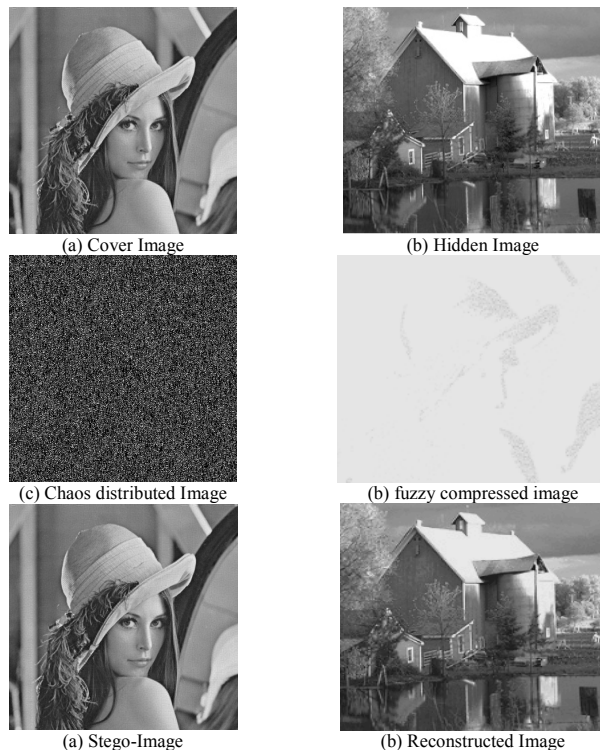


Fig. 7. stages of test image (1)

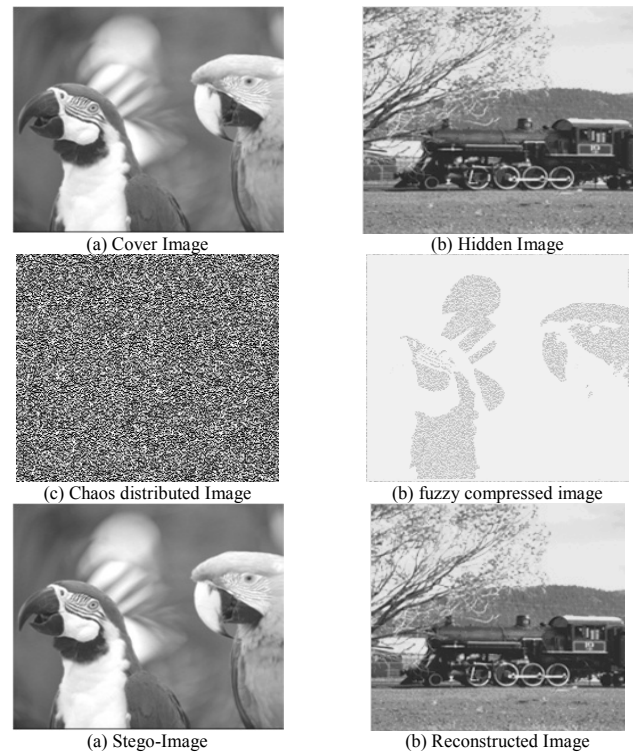


Fig. 8. Stages of test image (2)

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 House 2000.
- [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.
- [16] Timothy J. Ross," Fuzzy Logic with Engineering Applications" John Wiley & Sons Ltd, United Kingdom, 2010.
- [17] Gang Feng," Analysis and Synthesis of Fuzzy Control Systems A Model-Based Approach," CRC Press Taylor & Francis Group, USA, 2010.
- [18] S. N. Sivanandam, S. Sumathi and S. N. Deepa," Introduction to Fuzzy Logic using MATLAB," Springer Pub., New York, USA, 2007.



Mazhar Basyouni Tayel is professor in Faculty of Engineering, Alexandria University, Alexandria, Egypt. He holds B.Sc. in Electronics and Communications from Faculty of Engineering, Alexandria University. He also holds M.Sc. and Ph.D. in Electrical Engineering from Faculty of Engineering, Alexandria University. He taught many technical courses in Electrical and Electronic (Analog and Digital) system design and Implementation, works as System Engineer for more than 20 years, teach up to 30 undergraduate, postgraduate subjects, supervising more than 75 thesis, publish more than 150 papers in different international conference, Forums and Journals.



Alaa El-Din Sayed Hafez is an affiliate instructor in Faculty of Engineering, Alexandria University, Alexandria, Egypt. He was born in Cairo, Egypt in 1967 and became a member of IEEE in 2000, a Senior member of IEEE in 2011. He holds B.Sc. in Electronics and Communications from Faculty of Engineering, Alexandria University, M.Sc. in Electronics and Communication from Arab Academy for Science and Technology and Maritime Transport, Alexandria. He also holds M.Sc. and Ph.D. in Electrical Engineering from Faculty of Engineering, Alexandria University. He received many technical courses in Surveillance Radar design and Implementation, work as Radar System Engineer for more than 5 years, teach up to 20 undergraduate subjects, supervising more than 12 thesis, publish more than 50 papers in different international conferences and journals.



Hamed Shawky Zied is a Post Graduate Student (Ph.D.), Alexandria University, Alexandria, Egypt and became a Member of IEEE in 2012. He was born in Minoufia, Egypt in 1973. He holds B.Sc. in Electronics and Communications from Faculty of Engineering, Alexandria University, M.Sc. in Electrical Engineering from Faculty of Engineering, Alexandria University. He received many technical courses in electronic engineering design and Implementation.