ROBUST AND SECURED DATA HIDING SCHEMES USING DIGITAL IMAGE STEGANOGRAPHY

By

NAGHAM HAMID ABDUL-MAHDI
(1040210473)

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy (Communication Engineering)

School of Computer and Communication Engineering
UNIVERSITI MALAYSIA PERLIS

(2013)

UNIVERSITI MALAYSIA PERLIS

DECLARATION OF THESIS

Author's full name : Nagham Hamid Abdul Mahdi

Date of birth : 08/September/1977

Title : ROBUST AND SECURED DATA HIDING SCHEMES USING DIGITAL IMAGE STEGANOGRAHY

Academic Session : 2012/2013

I hereby declare that the thesis becomes the property of Universiti Malaysia Perlis (UniMAP) and to be placed at the library of UniMAP. This thesis is classified as :

☐ **CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)*

☐ **RESTRICTED** (Contains restricted information as specified by the organization where research was done)*

☐ **OPEN ACCESS** I agree that my thesis is to be made immediately available as hard copy or on-line open access (full text)

I, the author, give permission to the UniMAP to reproduce this thesis in whole or in part for the purpose of research or academic exchange only (except during a period of _____ years, if so requested above).

Certified by:

SIGNATURE	**SIGNATURE OF SUPERVISOR**
G 1396141	Dr. ABID YAHYA
(NEW IC NO. / PASSPORT NO.)	**NAME OF SUPERVISOR**
Date : July 2013	Date : July 2013

NOTES : * If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization with period and reasons for confidentially or restriction.

Acknowledgements

Praise and thanks to Allah (SWT) who gave me the strength and courage to complete this project. My most special thanks are to my supervisor Dr. Abid Yahya, who supported me throughout the long academic PhD, pursuing journey. His guidance, ideas, encouragement, affable nature, valuable advice, kindness and support were greatly helpful. I also extend my sincere thanks to my co-supervisor Prof. Dr. R. Badlishah Ahmad for his academic support, suggestions, and insights throughout the whole period of the study. My heartfelt thanks are also to the Universiti Malaysia Perlis (UniMAP) for awarding me the Graduate Assistant (GA). In this context, I would like to express my sincere thanks to the staff of the UniMAP; especially the staff members' of the School of Computer and Communication Engineering for their support and eagerness to provide the ideal research environment. Special thanks go to Dr. Osamah M. Al-Qershi from Universiti Sains Malaysia (USM), for his valuable suggestions, ideas, encouragement, kindness and support despite his very hectic and busy schedule. Particular thanks are due to my parents, my brothers and my sister for their daily prayers, giving me the motivation and strength, and for encouraging me to achieve my goals. I am really indebted to them all and words are not sufficed to express the gratitude I owe to them. Special thanks go to all my friends for their motivation, help and support during my academic period. I am indebted to all those namely mentioned above for their friendship and spiritual support that kept me going ahead. Last, but not least, I offer to my family the sincerest words of gratitude for their patience and unshakable faith in me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADR	Accuracy of Correctly Detected Characteristic Regions
AES	Advanced Encryption Standard
AET	Image Adaptive Energy Thresholding
A-MSPU	Adaptive More Surrounding Pixels Using
ANNTS	Artificial Neural Network Technology for Steganography
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BCH	Bose-Chaudhuri-Hochquenghem
BMP	Bitmap Format
BSSIS	Blind Spread Spectrum Image Steganography
bpp	Bit per Pixel
CD	Compact Disc
CDF	Cohen-Daubechies-Feauveau
CDMA	Code Division Multiple Access
CPA	Chosen-Plaintext Attack
CR-BIS	Characteristic Region-Based Image Steganography
dB	Decibel
DCT	Discrete Cosine Transform
DES	Data Encryption Standard
DFT	Discrete Fourier Transforms
DoG	Difference-of-Gaussian
DS/FH	Direct Sequence/Frequency Hopping
DVD	Digital Video Disc

DWT	Discrete Wavelet Transforms
ECC	Error Correcting Coding
FCM	Fuzzy C-Means
GA	Genetic Algorithm
GF	Galois Field
GIF	Graphics Interchange Format
HC-RIOT	Homogenous Connected-Region Interested Ordered
HCSSD	High Capacity and High Security Steganography System
HVS	Human Visual System
ID	Identity Card
IRSS	Improved Robust And Secured Steganography
JPEG	Joint Photographic Experts Group
LDPC	Low-Density Parity-Check
LSB	Least Significant Bit
MB	Model-Based Steganography
MBNS	Multiple Base Notational Systems
MLE	Maximum Length Embeddable
MMS	Multimedia Messaging Service
MSE	Mean Squared Error
MVE	Mean Value of Energy
PNG	Portable Network Graphics
PSNR	Peak Signal to Noise Ratio
PRNG	Pseudo-Random Number Generator
PVD	Pixel-Value Differencing
QF	Quality Factor

RBGC	Reflected Binary Gray Code
RGB	Red Green and Blue
RS-Code	Reed–Solomon Code
SIFT	Scale-Invariant Feature Transform
SINR	Signal to Interference Plus Noise Ratio
SNR	Signal to Noise-Ratio
SS	Spread Spectrum
SSIS	Spread Spectrum Image Steganography
SURF	Speeded-Up Robust Features
TIFF	Tagged Image File Format
VBs	Valid Blocks
VCs	Valid DCT Coefficients
WMF	Windows Metafile
XOR	Exclusive-Or-Operation
YASS	Yet Another Steganographic System

SKIM-SKIM DATA TERSEMBUNYI MANTAP DAN TERJAMIN MENGGUNAKAN IMEJ DIGITAL STEGANOGRAFI

ABSTRAK

Penghantaran data melalui jaringan awam seperti Internet memerlukan peningkatan dalam keselamatan komunikasi data; terutamanya menggunakan pemindahan dokumen sangat sensitif. Teknik-teknik steganografi telah diperkenalkan dan dibangunkan untuk menyediakan keselamatan untuk aplikasi ini. Pada dasarnya, matlamat steganografi adalah bukan sahaja untuk menghalang musuh dari menyahkod mesej tersembunyi, tetapi juga untuk mencegah musuh daripada mengesyaki kewujudan komunikasi tersembunyi. Ia bukan sahaja untuk menggantikan kriptografi, tetapi memperbaiki keselamatan menggunakan ciri-ciri kekaburan. Jika seseorang yang mencurigakan muncul semasa menggunakan teknik steganografi, teknik ini akan mengalahkan tanpa mengira sama ada mesej didedahkan atau tidak. Dalam kajian ini, dua teknik steganografi untuk imej digital telah dikaji. Algoritma pertama menyediakan sistem steganografi yang baru dan cekap dikenali juga sebagai Characteristic Region-Based Image Stefanography (CR-BIS) atau Ciri-ciri Wilayah Berdasarkan Imej Steganografi. Ia menggabungkan kedua-dua kemantapan teknik Speeded-Up Robust Features (SURF) atau Ciri-ciri Kemantapan Dipercepat dan Discrete Wavelet Transform (DWT) atau Ubahan Wavelet Diskret untuk mencapai ciri-ciri wilayah Steganografi serentak. Ia mengelakkan penyembunyian data di seluruh imej dengan memilih ciri kawasan bagi proses pembenaman secara dinamik. Apa-apa cara pemilihan wilayah yang dinamik akan meningkatkan keselamatan data terbenam. Keputusan eksperimen menunjukkan bahawa CR-BIS yang disediakan imej-imej stego dengan kualiti persepsi yang baik; ini telah ditunjukkan oleh nilai yang diperoleh dari high Peak Signal to Noise Ratio (PSNR) atau Isyarat Puncak tinggi kepada Nisbah Bunyi, sehingga 48.30 dB. Algoritma kedua, iaitu, Improved Robust and Secured Steganography (IRSS) atau Steganografi yang Diperbaik Kemantapan dan Bercagar, merupakan pembaikan algoritma Mali et al. yang mempunyai kecacatan kebolehpercayaan kerana beberapa data tidak boleh diambil pada fasa pengekstrakan. IRSS telah mengatasi masalah kehilangan maklumat melalui mengamalkan konsep peta pembenaman. Selain itu, ia telah dibuktikan secara eksperimen bahawa IRSS mengatasi yang asal dari segi kualiti imej-stego; ini telah dibuktikan oleh nilai PSNR yang dicapai antara (37.28-39.74) dB. Dalam dua algoritma yang dicadangkan, rahsia maklumat terbenam boleh dipulihkan dengan betul tanpa merujuk kepada ‘cover’ imej yang asal. Di samping itu, penambahan pembaikan kemantapan kemasukan sistem stego oleh Kod Pembetulan Ralat (ECC) dan menambah bit lebihan mesej rahsia yang terbenam, ditakrifkan dan dinilai, untuk mencadangkan kaedah mengalakkan kemantapan yang bersesuaian untuk algoritma yang dicadangkan. Sepertimana ECC, Reed-Solomon Kod (RS-Kod) telah digunakan untuk menghasilkan bit pembetulan bersamaan dengan bilangan bit yang terhasil oleh faktor lebihan khusus. Sebagai kesimpulan daripada keputusan eksperimen bahawa RS-kod meningkatkan kemantapan algoritma CR-BIS lebih daripada penambahan bits lebihan. Sebaliknya, menambah bit lebihan kepada mesej mempunyai kesan yang lebih baik pada kemnatapan algoritma IRSS.

ROBUST AND SECURED DATA HIDING SCHEMES USING DIGITAL IMAGE STEGANOGRAPHY

ABSTRACT

Transmitting data over a public network such as the Internet necessitates increasing the security of data communications; especially with the highly sensitive document transfer. Steganography techniques have been introduced and developed to provide security to these applications. Fundamentally, the steganography goal is not only to hinder the adversary from decoding a hidden message, but also to prevent an adversary from suspecting the existence of covert communications. It does not replace cryptography but rather improves the security using its obscurity features. If one's suspicious is raised while using a steganography technique, the goal of the latter will be defeated regardless whether or not a plaintext is revealed. In this research, two steganography techniques for digital images were developed. The first algorithm provides a new and efficient steganographic system, called Characteristic Region-Based Image Steganography (CR-BIS). It combines both the robustness of Speeded-Up Robust Features technique (SURF) and Discrete Wavelet Transform (DWT) to achieve characteristic region steganography synchronization. It avoids hiding data in the whole image by dynamically selecting characteristic regions for the process of embedding. Such a dynamic manner of region selection increases the security of embedded data. The experimental results showed that CR-BIS provided stego-images with a good perceptual quality; this was indicated by the obtained high Peak Signal to Noise Ratio (PSNR) values, up to 48.30 dB. The second algorithm, namely, an Improved Robust and Secured Steganography (IRSS), is an improvement of Mali et al.'s algorithm, which has a reliability defect as some data cannot be retrieved at the extraction phase. IRSS has overcome the problem of information loss via adopting the concept of the embedding map. Besides, it has been proved experimentally that IRSS outperformed the original one in terms of stego-image quality; this was demonstrated by the achieved PSNR values, which were between (37.28-39.74) dB. In the two proposed algorithms, the embedded secret information can be correctly recovered without referring to the original cover-image. In addition, improving the robustness of the stego-system by Error Correcting Codes (ECC) insertion and adding redundancy bits to the secret embedded message is defined and evaluated, in order to suggest an appropriate robustness enhancing method for the proposed algorithms. As a popular ECC, Reed-Solomon Code (RS-Code) was used to produce correction bits equal to the number of bits produced by a specific redundancy factor. It has been concluded from the experimental results that RS-code improved the robustness of CR-BIS algorithm more than the addition of redundancy does. On the other hand, adding redundancy bits to the message has a much better effect on the robustness of IRSS algorithm.

CHAPTER ONE

INTRODUCTION

1.1 Introduction

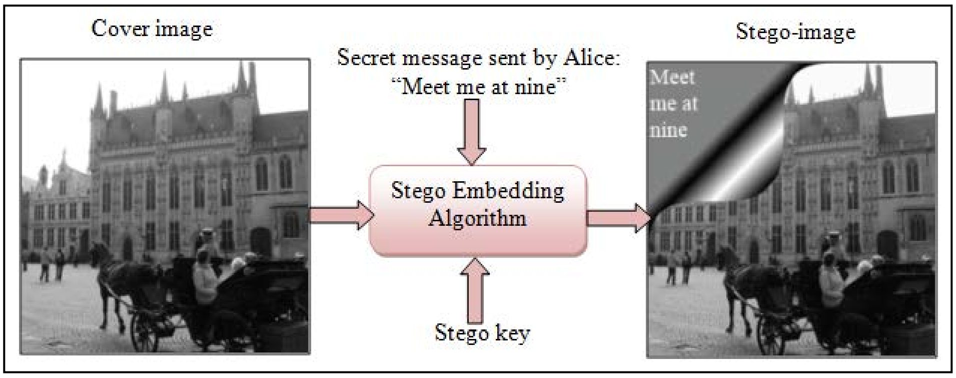
With the development in digital communication technologies and the rapid growth of network bandwidth, the Internet has become a popular channel for transmitting various data, such as, audio, video, image and text, in digital form. Many techniques have been proposed and developed for providing a secure transmission of data. A common approach to provide the secure environment for important data transmission is the use of cryptographic techniques (Ling, 2005).

Cryptography transforms data into seemingly meaningless bits, called cipher text, by using a sophisticated and robust algorithm. This will help only the intended receivers recover the original messages using a cryptographic key. For those, who do not have a key, the encrypted messages will appear as a stream of meaningless codes (Bender, Gruhl, Morimoto, & Lu, 1996). To overcome the weakness of cryptography, steganographic techniques are proposed to camouflage the existence of the hidden data in such a way that no one away from the sender and the intended receiver even knows that there is a hidden message (Koppola, 2009). Unlike the other forms of communications, the main purpose of steganography is defeated when the communication between the sender and the receiver is detected. Therefore, the first requirement of a steganographic system is its undetectability. In other words, a steganographic system is considered insecure, if anyone can differentiate between the cover-objects and stego-object (Kharrazi, 2006).

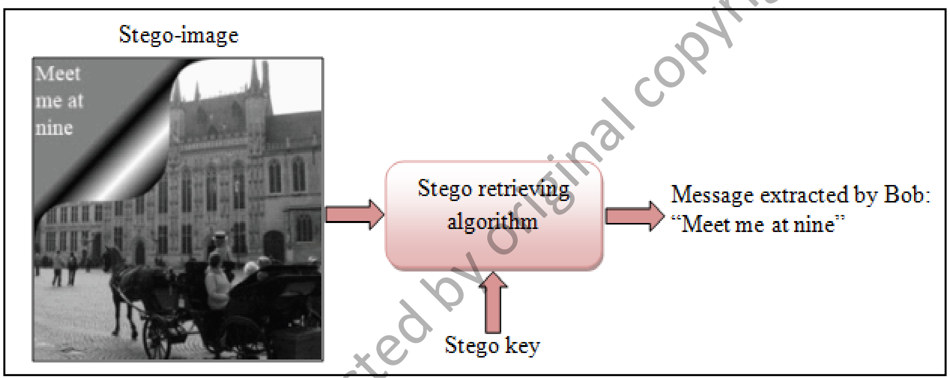
Recently, digital images are popularly used as a cover-object to covey the secret information. Owing to the high propagation of digital images and the high degree of redundancy they showcase, there is an increased interest in the use of images as a cover object in steganography. A lot of research work has been reported on the techniques of hiding data in images (Chang, Chen, & Lin, 2004; CHANG, LIN, & HU, 2002; Cheddad, 2009; Chi-Kwong & Cheng, 2001; Kharrazi, 2006; Koppola, 2009; Ling, 2005; Ran-Zan, Chi-Fang, & Ja-Chen, 2000). There are specific terms that are commonly used by the information hiding committees. Throughout this thesis, the term ‘cover image’ is used to describe the image selected for hiding the secret data. The image with the embedded information is characterized as ‘stego-image’. In addition to the expression 'stego key’, which is a parameter used to restrict other parties from extracting the secret message from stego-image. However, the processing of an image and the efforts of statistical analysis needed for breaking steganography algorithms are known as ‘steganalysis’ or ‘attacks’.

The classical model for modern steganography was proposed by Simmons in 1984 as the prisoners’ problem (Simmons, 1984). An example is illustrated in Figure 1.1 (Katzenbeisser, 2000). In this example, Alice and Bob are arrested for a crime and are thrown into different cells. The two prisoners would like to develop an escape plan; however, all communications between them are monitored by a warden named Wendy known to Alice and Bob as the adversary. Wendy will not let the prisoners communicate through encryption or any other means that make the communication noticeable. To avoid alerting Wendy of any covert message, an ideal way of communication is used to hide the stego message within a cover file, such as an image. Figure 1.1 illustrates the prisoners' problem where Alice places a hidden message, "Meet me at nine", and Bob is able to reconstruct the message with a shared stego key. Note that the difference between the cover image and the stego image is visually unobservable. Wendy is unaware that the picture sent by Alice contains the secret escape plan, i.e., the stego message.

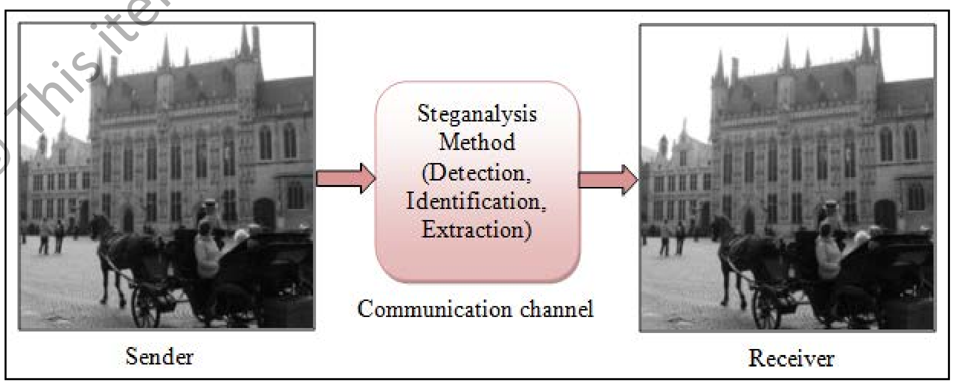
The focus of the current research is on the design of data hiding techniques used for transmitting secret data where digital images are selected as the cover-media. In the proposed techniques, the emphasis is placed on enhancing the robustness and imperceptibility. Moreover, error-free recovery of the embedded secret data without referring to the original cover-media is required.



(a)



(b)



(c)

Figure 1.1: The prisoners' problem (a) Steganography Embedding System Used by Alice (b) Steganography Retrieval System Used by Bob (c) Steganalysis System Developed by the Warden Wendy.

1.2 Fundamental Requirements for Steganography

The performance of a steganographic system can be measured using several properties (I. J. Cox, Miller, Bloom, Fridrich, & Kalker, 2008):

- Imperceptibility (undetectability) of the data shows how difficult it is to determine the existence of a hidden message. This parameter is the first and the primary requirement; it represents the ability to avoid human eyes detection. However, the techniques that do not alter the image in such a way to be perceptible to the human eye may still be able to alter the image in a way that it is detectable by the statistical tests. Truly secure steganographic techniques should be undetectable neither by the human eye nor by the statistical attacks (Amirtharajan & Rayappan, 2012; Bahi, Couchot, & Guyeux, 2012). To assess the level of imperceptibility, the visual difference between the original cover-image and the stego-image is calculated. By comparing the original cover-image and the final stego-image, the visual difference is determined and then, the imperceptibility level is specified (Ling, 2005).

- Robustness refers to how well the steganographic system resists the extraction of hidden data. It is the second parameter that measures the ability of the steganographic technique to survive the attempts of removing the hidden information. Such attempts include, image manipulation (like cropping or rotating), data compression, noise, and image filtering (Bahi et al., 2012).

- Payload capacity represents the maximum amount of information that can be safely embedded in a work without having statistically detectable objects. The more data bits to be hidden in the cover-image, the higher embedding capacity will be achieved. In general, imperceptibility is not proportional to the embedding capacity. When the embedding capacity increases, the imperceptibility level decreases and vice versa. Accordingly and by using this parameter, small amounts of data could be hidden without being detected by the human eye. Larger amounts of information, on the other hand, may detect artifacts by the Human Visual System (HVS) or by statistical tests (G. Chen, Zhang, Chen, Fu, & Wu, 2012; C. F. Lee & Huang, 2012).

- Reliability is the most important parameter that characterizes the applicability of the stego-system based on error free recovery for the hidden secret information. In other words, the feasible stego-system should allow its users to be able to retrieve the hidden information without any loss i.e. with 100% recovery. This vital factor should be taken into consideration in designing an applicable and accurate information hiding system in addition to the other three factors.

It is noteworthy that it is impossible to obtain the highest degree of robustness and the maximum embedding capacity with an acceptable level of imperceptibility at the same time. Therefore, a compromise must be made between robustness, imperceptibility and the embedding capacity. For different applications, the acceptable balance between these three constraints is different, depending on the nature of the requirements of the application (Ling, 2005).

1.3 Problem Statement

Ensuring safety and security of long-distance communication is a critical problem. This is particularly important in the case of confidential data storage and transmission in a public network, like the Internet. The security of such data communication, which is necessary and vital for many applications nowadays, has been a major concern and an ongoing topic since Internet is by design open and public in nature (Ling, 2005). Many techniques have been proposed for providing a secure transmission of data. Data encryption and information hiding techniques have become popular and they usually complement each other (Shankar, Sahoo, & Niranjan, 2012). The main problem is that, once you encrypt a file, even with a strong encryption algorithm, it looks like a random stream of bytes. In computer world, random bytes are very rare, and it is very easy to detect in a flow of trillions structured bits (Marwaha, 2010). The present research work in this thesis has identified the following problems in the present image steganography schemes:

- The majority of the existing methods presuppose that flexibility to noise, compression, and other image processing manipulations are not necessary in the steganographic context, which obviously are not tailored to steganography applications where flexibility, robustness and security are required (Cheddad, Condell, Curran, & Kevitt, 2010).
- In most of the current image steganography techniques, the process of information hiding modifies almost all the cover components. Such a process may negatively affect the visual quality of an image and increase the possibility