UNIVERSITI PUTRA MALAYSIA

DATA HIDING TECHNIQUES IN DIGITAL IMAGES

SALAH RAMADAN ALTHLOOTHI

DATA HIDING TECHNIQUES IN DIGITAL IMAGES

By

Salah Ramadan Althloothi

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Partial Requirements for the Degree of Master of Science

July 2003

To My Parents, Brothers and Sisters

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the partial requirements for the degree of Master of Science

DATA HIDING TECHNIQUES IN DIGITAL IMAGES

By

SALAH RAMADAN ALTHLOOTHI

July 2003

Chairman: **Associate Professor Abd Rahman Ramli, Ph.D.**

Faculty: **Engineering**

With literally millions of images moving on the Internet each year it is safe to say that hiding data in digital image is of real concern to many in the security field. Therefore how to protect secret messages during transmission becomes an important issue. Hiding data provides a good layer of protection on the secret message, so the purpose of this thesis is to study the data hiding techniques in digital images as a new and powerful technology capable of solving important practical problems.

Depending on what information in which form is hidden in the digital images, one can distinguish two types of data hiding techniques, spatial domain techniques, and frequency domain techniques.

In the spatial domain techniques, a digital image serves as a carrier for a secret message. For instance, by replacing the least significant bit of each pixel in the carrier image with the secret message after changing it to stream of bits, the changes to the carrier image will be imperceptible and the secret message will be masked by carrier

image. In this side, two programs had been implemented using MATLAB program to illustrate the main idea involved in least significant technique (low bit encoding), and the other to illustrate the masking technique inside the carrier image.

In the frequency domain techniques, a short message is embedded in the carrier image in a robust algorithm. Robustness means the ability to survive common image processing operations, such as lossy compression, filtering, noise adding, geometrical transformations, etc. So in this technique two programs had been implemented to illustrate the main idea involved in frequency domain, one with the Fast Fourier Transform (FFT), and the other with the Discrete Wavelet Transform (DWT).

After all these studies, one of the algorithms in the masking technique is developed and implemented using JAVA program to embed message into true color image with a good quality and higher capacity. Beside that the carrier images in different techniques were examined by exposing them to common signal processing operations such as image resizing, rotation, histogram equalization, lossy Compression, and Gaussian noise addition to illustrate the characteristic of the data hiding techniques, such as hiding capacity, robustness, undetectability, and perceptual transparency

Finally, it has been shown that the frequency transformation techniques are more robust, and hence suitable for water marking and data hiding purpose. The spatial domain techniques exhibit loss robustness but due to its higher capacity and good quality are perfect for data hiding purposes.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi sebahagian keperluan untuk ijazah Master Sains

TEKNIK PENYEMBUNYIAN DATA DALAM IMEJ DIGITAL

Oleh

SALAH RAMADAN ALTHLOOTHI

Julai 2003

Pengerusi: **Profesor Madya Abd Rahman Ramli, Ph.D.**

Fakulti: **Kejuruteraan**

Kebanjiran berjuta-juta imej yang dihantar melalui Internet pada setiap tahun telah mendapat perhatian yang serius oleh pakar-pakar keselamatan. Persoalan utama ialah bagaimana menggunakan peluang ini untuk menghantar mesej-mesej rahsia menggunakan imej-imej tersebut. Penyembunyian data di dalam imej menyediakan suatu lapisan perlindungan untuk mesej rahsia. Justeru itu, tesis ini mempelajari dan mengkaji teknik-teknik penyembunyian data sebagai suatu teknologi baru yang berpotensi menyelesaikan masalah-masalah praktikal.

Bergantung kepada maklumat yang disembunyikan di dalam imej-imej digital, terdapat dua teknik utama untuk penyembunyian data iaitu domain ruang dan domain frekuensi.

Teknik domain ruang menggunakan imej digital sebagai pembawa kepada mesej rahsia. Sebagai contoh, penukaran bit yang paling kurang signifikan pada setiap piksel

di dalam imej pembawa dengan mesej rahsia selepas penukaran kepada aluran bit, aktiviti penukaran tidak dapat dikesan dan mesej rahsia akan ditopeng oleh imej pembawa. Dua program telah dibangunkan dengan menggunakan perisian MATLAB untuk menunjukkan idea utama yang digunakan di dalam teknik paling tidak signifikan (pengekodan bit rendah) dan satu lagi untuk menunjukkan teknik topeng di dalam imej pembawa.

Teknik domain frekuensi pula menempelkan mesej pendek di dalam imej pembawa menggunakan algoritma yang kukuh. Kekukuhan bermakna kebolehan menangani operasi biasa pemprosesan isyarat seperti penukaran saiz, pemusingan, penyamaan histogram, pemampatan hilang dan penambahan hingar Gaussian. Dua program telah dibangunkan untuk menunjukkan idea utama yang digunakan oleh domain frekuensi iaitu Penukaran Pantas Fourier (FFT) dan Penukaran Anak Isyarat Terputus (DWT).

Melalui pembelajaran yang dilakukan, salah satu algoritma di dalam teknik topeng telah dibangunkan menggunakan pengaturcaraan JAVA untuk menempelkan mesej ke dalam imej warna sebenar. Hasil yang didapati adalah baik dan mempunyai keupayaan yang lebih besar. Selain daripada itu, imej-imej pembawa di dalam teknik-teknik yang berbeza dikaji dengan menggunakan operasi biasa pemprosesan isyarat. Pengkajian ini dilakukan untuk membandingkan ciri-ciri teknik penyembunyian data seperti keupayaan penyembunyian, kekukuhan, kesukaran pengesanan, dan ketelusan pemahaman.

Akhirnya, telah ditunjukkan bahawa teknik penukaran domain ferkuensi adalah lebih kukuh dan sesuai untuk tujuan penandaan dan penyembunyian data. Teknik domain ruang adalah kurang kukuh berbanding teknik domain frekuensi. Ia juga mempunyai kualiti yang agak baik dan keupayaan yang lebih baik untuk penyembunyian data.

AKNOWLEDGEMENTS

First of all, the author would like to express his utmost thanks and gratitude to Almighty Allah for giving him the ability to finish this thesis successfully.

The author gratefully wish to express his profound appreciation and gratitude to his supervisor, Assoc. Prof. Abd Rahman Ramli, for his supervision, support, and constructive suggestion throughout the duration of the project until it turn to real success.

The author also would like to indebted the members of his supervisory committee, Puan. Roslizah Ali and Puan. Wan Azizun, for their affectionate guidance, prompt decision and valuable assistance during this period.

Appreciation also to the assistance rendered by the respective lectures, staffs, technicians of engineering faculty for providing the facilities required for undertaking this project. Also the author would like to seize this opportunity to thank all the friends in the multimedia lab.

The author would like to thank his family for the encouragement and supporting, also to all the individuals at the Department of Computer and Communication System Engineering for their cooperation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The art of information hiding has been around nearly as long as the need for covert communication. Steganography, hiding the information, arose early on as an extremely useful method for covert information transmission. The word itself is derived from a Greek phrase meaning *"covered writing,"* and some of the first recorded examples of steganography come from the ancient Greeks. One ancient history tells the account of a king who, wishing to send a secret message across unfriendly borders, shaved the head of a trusted servant and tattooed a message on his bare scalp. Once the slave's hair grew back, the message was concealed from prying eyes and could be safely delivered to its destination. This technique works so well, that it was employed by German spies even as late as the 20th century. The use of invisible inks, made readable only through a special developing process, qualifies as another classic example of steganography in practice. This information-hiding technique was used extensively during both World Wars, as was concealing information on microdots, pictures the size of a printed period made through special photographic techniques (Fridrich, 1998).

While the technology for information transmission has changed much with the advent of the digital age, the need for concealing information is as present today as it has ever been. In recent years, much attention has been paid to steganographic applications using digital images. Considering the frequency with which such images are transmitted on a daily basis, they serve as perfect containers for hidden messages.

Although digital steganography is still a young field, a variety of techniques have been developed to implement the hiding of information within digital images.

With literally millions of images moving on the internet it is safe to say that digital image steganography is of real concern to many in the IT security field. Digital images could be used for a number of different types of security threats. The interest in data hiding techniques has increased with the recent activity in digital copyright protection schemes. One way to protect the ownership of a digital image is to secretly embed data in the content of the image identifying the owner (Eugene et al, 2001).

1.1 Problem Statement

A typical digital steganographic encoder is shown in Figure 1.1. The message is the data that the sender wishes to remain confidential and can be text, images, audio, video, or any other data that can be represented by a stream of bits. The cover or host is the medium in which the message is embedded and serves to hide the presence of the message.

The message embedding technique is strongly dependent on the structure of the cover, and in this thesis covers are restricted to being digital images. It is not required that the cover and the message have homogeneous structure. For example, it is possible to embed a recording of Shakespeare's lines (an audio stream message) inside a digital image cover as Johnson had done (Johnson and Jajodia, 1998).

The image with the secretly embedded message produced by the encoder is the *stego-image*. The stego-image should resemble the cover image under casual inspection and analysis. In addition, the encoder usually employs a stego-key, which ensures that only recipients who know the corresponding decoding key will be able to extract the message from a stego-image (Eugene et al 2001).

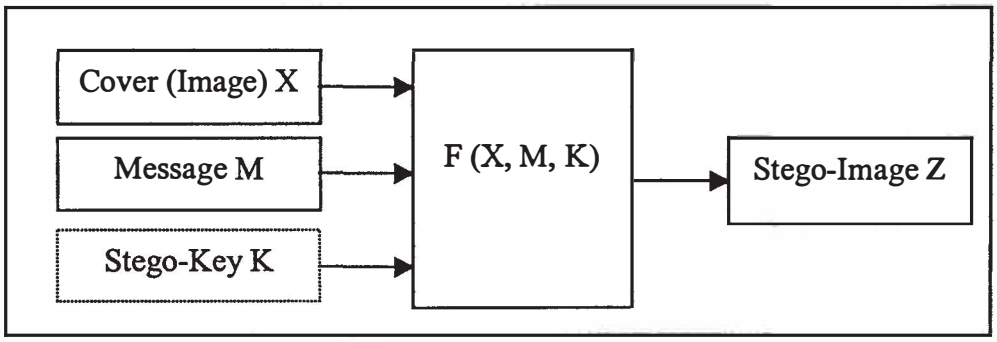

Figure 1.1: Steganographic Encoding

Recovering the message from a stego-image requires the stego-image itself and a corresponding decoding key if a stego-key was used during the encoding process. The original cover image may or may not be required; in most applications it is desirable that the cover image not be needed to extract the message. Steganography is not the same as cryptography. In cryptography, the structure of a message is changed to render it meaningless and unintelligible unless the decryption key is available. Cryptography makes no attempt to disguise or hide the encoded message. Steganography does not alter the structure of the secret message, but hides it inside a cover. It is possible to combine the techniques by encrypting a message using cryptography and then hiding the encrypted message using steganography. The resulting stego-image can be transmitted without revealing that secret information is being exchanged. Furthermore, even if an attacker were to defeat the steganographic

technique and detect the message from the stego-image, he would still require the cryptographic decoding key to decipher the encrypted message (Eugene et al 2001).

1.2 Objectives

The objectives of this thesis are:

- To study the use of data hiding techniques in digital images, and to get general understanding of Data Hiding Techniques in different domains.
- To investigate and implement different methods using MATLAB for data hiding techniques.
- To develop and implementing a method that covertly embeds data (text, image) into true color container digital images using JAVA program.
- To identify the robust method against a series of images attacks such as noise addition, translation, lossy compression, rotation, etc.

The purpose of this thesis is to study data hiding in digital imagery as a new and powerful technology capable of solving important practical problems. The field of data hiding in digital image is relatively very young and is growing at an exponential rate. Data hiding is a highly multidisciplinary field that combines image and signal processing with cryptography, communication theory, coding theory, signal compression, and the theory of visual perception.

By embedding almost invisible signals in images, one makes the images "intelligent" in the sense that the hidden message can carry information about the content of the image (thus protecting its integrity), additional information about the

author of the image, or other useful data related to the image. In another application, one can achieve a very secure mode of communication by embedding messages into the noise component of digital images. This way, the very presence of communication becomes hidden (Eugene et al 2001).

1.3 Scope of the thesis

In this thesis we concentrate on the following:

- Concept and Study:
  - ✓ Several data hiding techniques, including their mechanism, implementation and computational efficiency will be studied. The relative merits and demerits of the approaches used will be presented.
- Investigation and implementation:
  - ✓ Several algorithms are implemented using MATLAB program to illustrate the main idea involved in low bit encoding, frequency domain encoding and encoding in the wavelet domain.
- Develop and implementation:
  - ✓ One of the algorithms is developed and implemented using JAVA program to illustrate the characteristic of the data hiding techniques such as (Hiding Capacity, Quality, and Perceptual Transparency).

1.4 Thesis Organization

This thesis includes five chapters. The first chapter summarized the research plan by stating the problem, objectives, motivation and scope of the thesis. Chapter II

discusses the origin of steganography and how it has historically evolved to its present-day computer applications. Chapter III describes the methodology used to meet the research objectives followed by Chapter IV's research analysis and results. Finally, Chapter V presents the conclusions, and research recommendations with future work in this field.

CHAPTER II

LITERATURE REVIEW

Data hiding, a form of steganography embeds data into digital image for the purpose of identification, annotation, and copyright (Bender et al, 1996). Several constraints affect this process: the quantity of data to be hidden and carrier image, the type of the carrier image, and the degree to which the data must be immune to interception, modification, or removal by any attack (Johnson et al, 2001).

An overview of data hiding techniques in digital images is presented in this chapter. In particular (next chapter) it will describe the use of Steganography to hide information in a digital image. Steganography is related to cryptography and is the basis for many of the data hiding techniques currently being developed. The interest in data hiding has raised with the recent activity in digital copyright protection schemes. One way to protect the ownership of a digital image is to secretly embed data in the content of the image identifying the owner (Eugene et al, 2001).

2.1 Definition of Data Hiding

In cryptography the message is broken into smaller units using a known or pre-determined algorithm or key. It may be a substitution, additive, transposition, transcription, etc. The science of cryptography is used for message encryption technology, digital signatures and private and public key cryptosystems used in digital certificates. The purpose of cryptography is to make messages unintelligible so that those who do not posses secret keys cannot recover the messages. In cryptography, the structure of a message is changed to render it meaningless and

unintelligible unless the decryption key is available. Cryptography makes no attempt to disguise or hide the encoded message.

Steganography conceal the existence of a hidden communication. The secret message to be transmitted is camouflaged in a carrier so that its detection becomes difficult. Information related to the sender and the receiver of the message also can be hidden this way. Steganography does not alter the structure of the secret message, but hides it inside a cover image. It is possible to combine the techniques by encrypting a message using cryptography and then hiding the encrypted message using steganography.

A digital watermark is meshed into a file, audio, image or video, to identify the copyright information for the intellectual property rights of an item, as the traditional watermark on paper identified its creator or a currency's authenticity. This is accomplished, in the simplest terms, by inserting a specific pattern of bits into the file in such a way that they are invisible or visible to the human eyes. The digital watermark must be robust enough to remain constant when a file is compressed, or otherwise changed.

Digital watermarks are not steganography by definition. The difference is primarily one of intent. Steganography conceals information; watermarks extend information and become an attribute of the cover image. In steganography, the object of communication is the hidden message to be inserted and extracted from the cover image without any error. In digital watermarks, the object of communication is the