Steganography and Steganalysis Based on Digital Image

Ge Huayong*[a,b], Huang Mingsheng[a], Wang Qian[a]
a. School of Information Science and Technology, Donghua University, Shanghai 201620,China
b. Engineering Research Center of Digitized Textile & Fashion Technology,
Ministry of Education, Donghua University, Shanghai 201620, China

***Abstract*—With the rapid development of steganography, steganalysis has advanced quickly. Battle between steganography and steganalysis has become an important issue in information security. Aiming at a commonly used cover media, i.e., digital image, this article reviews steganography and steganalysis based on digital image. Concept and principle of steganography and steganalysis are illustrated. Spatial domain and transform domain embedding methods are generalized. And the recent advances in steganalysis are recapitulated. Then the performance specification of image steganography is discussed. Finally some new trend and problems faced are also discussed.**

Keywords-steganography;steganalysis;digital image; LSB;

cover-image; stego-image;

I. INTRODUCTION

Steganography is an art of hiding communication by embedding message into an innocuous-looking cover media. Using steganography, a secret message is embedded inside a piece of unsuspicious information and sent without anyone knowing the existence of the secret message. Secrets can be hidden inside all sorts of cover information: text, image, audio, video, and so on. Most steganographic utilities hide information inside images, as it is relatively easy to implement. People refer image steganography as the art and science of invisible communication, which is to conceal the very existence of hidden message in digital images. Some facts have motivated active researches and abundant publications in the field of image steganography. For example, images can convey a large of information especially on the internet. Moreover, the nonstationarity of images makes image steganography hard to break. Nowadays, digital image has become an important channel to bear stego information.

Steganography and cryptography are cousins in the spycraft family. Cryptography scrambles a message so it cannot be understood. Steganography hides the message so it cannot be seen. A message in ciphertext, for instance, might arouse suspicion on the part of the recipient while an "invisible" message created with steganographic methods will not. Since steganography is used to hide the occurrence of communication, it has been applied to covert communication, watermarking and fingerprinting that seem to hold promise for copyright protection, tracing source of illegal copies, etc. But it creates a potential problem when this technology is misused for planning criminal activities. Hence, it becomes very essential to distinguish between benign images from anomalous stego-images.

Aiming at detecting secret information hidden in a given image using steganographic tool, steganalysis has been of interest since the end of 1990's. It is fair to say that steganalysis is both an art and a science. The art of steganalysis plays a major role in the selection of features or characteristics to test for hidden message, while the science helps in designing the tests themselves. As more and more techniques of hiding information are developed quickly, the wide-spread availability of tools for the same has led to an increased interest in steganalysis techniques. In the last few years, many new and powerful steganalysis techniques are reported in the literature [1]. Many of these techniques are specific to different embedding methods and indeed have shown to be quite effective in this regard. Research and development of steganography precedes steganalysis, and steganalysis has been forced to catch up. More recently, steganalysis has had some success and steganographers have had to consider the stealthiness of their hiding methods more carefully.

In this paper, we summarize the steganography and steganalysis based on digital image. The rest of this paper is organized as follows. Section II illustrates the concept and principle of steganography and steganalysis. Section III discusses the spatial domain and transforming domain embedding methods of steganography. In section IV, the category and technique of steganalysis based on digital image are introduced. The performance specification of image steganography is discussed in section V. Concluding remarks are given in section VI.

II. PRINCIPLE OF STEGANOGRAPHY AND STEGANALYSIS

With the wide spread of steganographic tool on the internet, the criminals may communicate covertly by steganography, and the research on steganography and steganalysis is becoming more and more important. The modern formulation of steganography and steganalysis is often given in terms of the prisoner's problem. Figure 1 shows the illustration of steganography and steganalysis system, in which Alice and Bob are two inmates who wish to communicate in order to hatch an escape plan. However, all communication between them is examined by the warden, Wendy, who will put them in solitary confinement at the slightest suspicion of covert communication. Specifically, in the general model for

978-1-4244-9306-7/11/$26.00 ©2011 IEEE

steganography, Alice wishes to send a secret message m to Bob. In order to do so, she embeds secret message m into a cover message c by secret key and obtains a stego-object s. The stego-object s is then sent through the public channel. The warden Wendy who is free to examine all messages exchanged between Alice and Bob can be passive or active. A passive warden simply examines the message and tries to determine if it potentially contains a hidden message. If it appears that it does, she takes appropriate action, or else, she lets the message through without alteration. An active warden on the other hand can alter message deliberately, even though she does not see any trace of a hidden message, in order to foil any secret communication that can nevertheless be occurring between Alice and Bob.

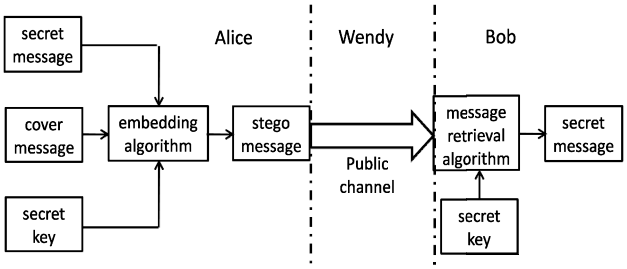

Figure 1. Illustration of steganography and steganalysis system

In pure steganography framework, the technique for embedding the message is unknown to Wendy and shared as a secret between Alice and Bob. However, it is generally considered that the algorithm in use is not secret. Only the key used by the algorithm is kept as a secret between the two parties. This assumption is also known as kerchoff's principle in the field of cryptography. The secret key, for example, can be a password used to seed a pseudo-random number generator to select locations in a cover-object for embedding the secret message.

In this context, steganalysis refers to the techniques that aid Wendy in distinguishing between cover-image and stego-image. It should be noted that Wendy has to make this distinction without any knowledge of the secret key which Alice and Bob may be sharing and sometimes even without any knowledge of the specific algorithm that might be used for embedding the secret message. Hence steganalysis is inherently a difficult problem. However, it should be noted that Wendy does not have to glean anything about the contents of the secret message m, just determining the existence of a hidden message is enough. Generally speaking, steganalysis can be described by a statistical hypothesis model-test. For the test image, using the testing function f : s→ true or false:

$$f(s) = \begin{cases} \text{true or } 1, \text{covert information existing in s} \\ \text{false or } 0, \text{no covert information existing in s} \end{cases} \quad (1)$$

In the steganalysis, a cover-image is probably mistaken for a stego-image and a stego-image is probably mistaken for a cover-image. Therefore there exists false alarm and miss. The aim of steganalysis is to reduce false alarm rate and miss rate.

III. IMAGE STEGANOGRAPHY

There have been a large number of steganography embedding techniques proposed in the literature. These techniques modify the cover-image with different approaches as well as constrains. But all embedding techniques share the important goal of maximizing the capacity of the stego channel. In other words their aim is to embed at highest possible rate while remaining undetectable to steganalysis attack. All the popular data hiding methods can be divided into two major classes: spatial domain embedding and transform domain embedding. Next we will review them.

A. Spatial Domain Embedding

Spatial domain embedding technique is the first technique proposed in the literature. Generally, these techniques operate on the principle of tuning the parameters of the cover-image (e.g., the payload or disturbance) so that the difference between the cover-image and the stego-image is little and imperceptible to the human eyes. Their popularity is derived from their simple algorithmic nature and ease of mathematical analysis. Spatial domain embedding is easy to implement, providing high payload capacity but their robustness is weaker than their counter part.

The most widely known image steganography algorithm is based on modifying the least significant bit layer of images, hence known as the LSB technique[2]. LSB based methods can be divided into two main groups: LSB replacement and LSB matching. In LSB replacement, the LSB bit of cover-image is replaced with secret bits. While in LSB matching, the pixels are randomly incremented or decremented by the secret bits. LSB based techniques pose a difficult challenge to a steganalyst in the passive warden model as it is difficult to differentiate cover-image from stego-image, given the small changes that have been made. Of course with an active warden, such techniques can be easily defeated by randomizing the LSB. In a low resolution (small number of pixels) image with 8 bit color, the effects of manipulating the LSB can cause noticeable shifts in colors. As the resolution and depth of color increase in an image, the impact of manipulating the LSB becomes less noticeable. Thus high resolution images are preferred for use as cover-images.

B. Transform Domain Embedding

Transform domain embedding includes discrete Fourier transform (DFT), discrete cosine transform (DCT) and discrete wavelet transform (DWT)[3]. Regardless of the domain, significant transform coefficients are often selected to mix with secret signal in a way such that information hiding is transparent to human eyes. These transforms may be performed blockwise, or over the entire image. For a blockwise transform, the image is broken into smaller blocks (8 × 8 and 16 × 16 are two popular sizes), and the transform steganography is performed individually on each block.

DCT domain embedding technique is the most popular one, mostly because of the fact that DCT based image format are widely available in public domain as well as the common output format of digital cameras. Embedding in DCT domain is simply done by altering the DCT coefficients, for example by changing the least significant bit of each coefficient. Although

changing the DCT coefficients will cause unnoticeable visual artifices, they do cause detectable statistical changes. Therefore, computer statistical analysis is still promising to detect such a distinction that is difficult for humans to perceive. In order to minimize statistical artifacts left after the embedding process, different methods for altering the DCT coefficients have been proposed, namely Outguess, F5, Model based and Perturbed quantization.

DWT domain based embedding technique is quite new, and not as well developed or analyzed as technique which operate on DCT or DFT. But such technique will gain popularity as JPEG2000 compression becomes more popular. Stego Jasper embedding technique based on wavelet operates on JPEG2000 images. Embedding is done by modifying least significant bits of selected wavelet coefficients.

Regardless of the domain, significant transform coefficients are often selected to mix with secret/perturbing signal in a way such that information hiding is transparent to human eyes. For instance, Cheng and Huang have proposed an additive approach to hiding secret information in the DCT and DWT domain [3].

IV. IMAGE STAGANALYSIS

A. Category of Steganalysis

Aiming at detecting secret information hidden in a given image using steganographic tool, steganalysis has been of interest since the end of 1990's. Steganographic attacks consist of detecting, extracting and destroying hidden object of the stego media. There are several types of attacks based on the information available for analysis. Neil Johnson, one of the few experts in the field, classifies attacks in six main categories. They are: 1. Stego-only attack: only the stego-image is available for analysis. 2. Known cover attack: the original cover-image and stego- image are both available. 3. Known message attack: at some point, the hidden message becomes known to the attacker. Analyzing the stego-image for patterns that correspond to the hidden message may be beneficial for future attacks against that system. Even with the message, this may be very difficult and may even be considered equivalent to the stego-only attack. 4. Chosen stego attack: the steganography tool (algorithm) and stego-image are known. 5 Chosen message attack: the steganalyst generates a stego-image from some steganography tool or algorithm from a chosen message. This goal in this attack is to determine corresponding patterns in the stego-image that may point to the use of specific steganography tool or algorithms. 6. Known stego attack: the steganography algorithm is known and both the original and stego-image are available.

Generally speaking, a steganalysis system is considered as broken by an algorithm if the algorithm can decide whether a given image contains a secret message or not. Now the researches of steganalysis mainly focus on detecting the presence of the secret message. With the development of steganography, researchers have made efforts to develop image steganalysis schemes, which is to detect the very presence of hidden message in a given image. According to the targeted steganographic tools, steganalysis can be also broadly classified into two categories[4]: specific steganalysis and universal steganalysis. Specific steganalysis is designed to detect some particular steganalysis. Universal steganalysis is also called blind steganalysis or universal blind steganalysis. It can detect the existence of secret message without any knowledge of the details of steganography algorithms. Therefore it is more applicable and practicable than the specific steganalysis.

B. Steganalysis Technique

There have been a number of steganalysis techniques proposed in the literature, each operating with its own unique approach. In essence steganalysis technique operates by obtaining a set of statistical feature from the input image. Avcibas I first proposed a universal steganalysis algorithm to detect embedded message in images through a proper selection of image quality metrics (IQM) and a multi-variant regression analysis[5]. Then he improved the algorithm and increased the detection precision[6]. Jiang N improved some IQM standards and combined them into a new feature vector[7]. It adopted the supported vector machine (SVM) classifier to distinguish between the original and stego images.

There are many works reporting that high-order statistics are very effective in differentiating stego-images from cover-images. Farid proposed a general steganalysis algorithm based on image higher-order statistics[8]. He modeled the universal steganalysis by supervised learning for the first time and indicated that the supervised learning was effective for detecting stego images without knowing the statistics property of images and steganography methods. In this framework, some sensitive features for data embedding were usually extracted first and then a classifier was chosen to distinguish stego-image from original images. Later Farid H used a wavelet-like decomposition to build higher-order statistical models of natural images. For each image, firstly three-scale quadrature mirror filters (QMF) decomposition was made, then higher-order probability dimensional function (PDF) moments were extracted, finally the FLD classifier was used to detect the test images. Holotyak et al used higher-order moments of the PDF of the estimated stego-image in the finest wavelet level to construct the feature vectors[9]. Shi et al proposed the use of statistical moments of the characteristics functions of the wavelet sub-bands. Since the n-th statistical moment of wavelet characteristics function is related to the n-th derivative of the corresponding wavelet histogram, the constructed 39-dimensional feature vector has proved to be sensitive to embedded data. Usually, the steganalysis algorithm based on the higher-order statistics achieved satisfactory performance on image files, regardless of the underlying embedding algorithm.

However, since the image-embedding method is typically unknown to steganalyst, many researches focused on the design of a blind steganalysis algorithm to detect the presence of steganography independent of the steganography algorithm used. In Geetha's work they employed the higher order statistical features that were collected from a new transform domain, i.e., curvelet transform domain[10]. This work is typically a novel and first attempt of employing these features for steganalysis. Experimental results also proved that their claim was justified.

V. PERFORMANCE SPECIFICATION OF IMAGE STEGANOGRAPHY

There are several important issues to be considered when studying steganographic systems. They are steganographic robustness, capacity, and security. The relationship between them can be expressed by the steganography triangle, which is shown in Figure 2. It represents a balance of the desired characteristics associated with a steganographic method. In order to improve one element, you have to sacrifice one or both of the other two elements. For instance, in order to improve capacity, you sacrifice security. This is logical since inserting hidden information to some degree is the same as tampering with an image. The more you tamper with an image, the more probable that an observer will notice the degradation and suspect something is out of place. Each element is described below.

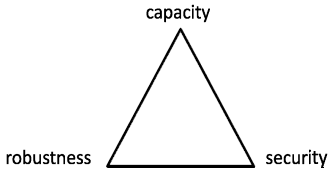

Figure 2. The steganography triangle

Robustness refers to an embedded message's ability to survive either deliberate attack by a suspecting third party or the random corruption of noise during some phase of the transmission process. If a secret message is able to survive when a carrier image moderately degraded, then the steganographic method is said to be very robust. However, it is most desired that the embedded content be fragile so as to reduce the possibility that an interceptor would be able to reassemble the embedded message.

Capacity refers to the maximum number of bits which could be embedded in the image, while the obtained stego-image remains undetectable and visually intact. The cover-image used to create a stego-image is acting as an information channel with which the embedded message is transferred. Like any other information channel, an important property of a stego channel is its capacity. Shanon defines the capacity of an information channel as the maximum achievable rate, with which error free transmission could be achieved. But capacity of steganography channel has a number of additional constraints. Firstly the stego channel needs to be undetectable by definition. In other words the statistical properties of the stego and cover image need to be indistinguishable. The second constraint on the capacity of stego channel is that the stego channel should preserve the properties of the cover channel.

The fundamental characteristic of steganography is its ability to offer a means of communication without suspicion. Security is the ability of an embedding carrier to remain undiscovered. The whole purpose of steganography, unlike other forms of communication, is defeated by the detection of communication between the sender and the receiver. Therefore, the first requirement of a steganographic system is its undetectability. In other words, a steganographic system is considered to be insecure, if the warden is able to differentiate between cover-image and stego-image. Cachin has defined a steganography technique to be ε-secure if the relative entropy of the probability distribution of cover-object and stego-object is less than or equal to ε. A steganography technique is perfectly secure if ε is zero. It is demonstrated that there do exist steganographic techniques that are perfectly secure. However, it should be noted that classical definition of steganography is statistical and not perceptual.

VI. CONCLUSION

Steganography, especially combined with cryptography, is a powerful tool which enables people to communicate secretly. With the rapid development of digital technology and internet, steganography has advanced a lot over the past years. Accordingly the steganalysis developed quickly. In this paper, we present an overview on steganography and steganalysis based on digital image.

However, steganography and steganalysis are still at an early stage of research. With the rapid development of steganography, steganalysis are facing new challenge. There are several problem needed to be investigated in steganography and steganalysis based on digital image. Notion of security and capacity for steganography needs to be investigated deeply. Steganography and corresponding steganalysis using image models needs to be further investigated. How do you choose good cover-image for given stego message? What kinds of images are good for using as cover-images? Therefore steganography and steganalysis need to make effort on it.

REFERENCE

[1] S.Z.Wang, X.P.Zhang, "Recent advance in image-based steganalysis research," Chinese journal of computers,vol. 32,pp. 1247-1263,July 2009

[2] A.Yadollahpour, H. M. Naimi, "Attack on LSB steganography in color and grayscale images using autocorrelation coefficients," European Journal of Science Research,vol.31,pp. 172-183,Feburary 2009

[3] Q. Cheng ,T.S. Huang, "An additive approach to transform-domain information hiding and optimum detection structure, " IEEE Transaction on Multimedia,Vol.3, pp.273-284,March 2001

[4] L. Xiangyang, L.Fenlin,Y.Chunfang, W.Daoshun, "Image universal steganalysis based on best wavelet packet decomposition," Sci China Inf Sci, vol. 53,pp.634-647,March 2010

[5] I. Avcibas, N. D. Memon, "Steganalysis of watermarking techniques using image quality metrics," Procedings of SPIE, Vol.4314, pp.523-531,2001

[6] I. Avcibas, N. D. Memon, "Steganalysis using image quality metrics," IEEE Trans Image Process,vol,12, pp. 221-229,2003

[7] N. Jiang, J. Wang, "A new method for blind image steganalysis, "J Beijing University Posts Telecommun, vol,29, pp.1-4,2006

[8] H. Farid, "Detecting hidden message using higer-order statistical models," IEEE on signal processing, vol, 2, pp. 905-908, 2002

[9] T. Holotyak, J. Fridrich, "Blind statistical staganalysis of additive staganography using wavelet higher order statistics," Lecture notes in computer science,pp.273-274,2005

[10] S. Geetha, K.Kamaraj, "Passive staganalysis based on higher order image statistics of curvelet transform," International Journal f Automation and Computing,vol,7,pp.531-542,April 2010