Scientific Journal Impact Factor (SJIF): 1.711

e-ISSN: 2349-9745
p-ISSN: 2393-8161

International Journal of Modern Trends in Engineering and Research
www.ijmter.com

A Review paper on Steganography Techniques

Dhara Rana
Computer Dept, MBICT, Anand, Gujarat 388120, India

Abstract—In today's world the art of sending & displaying the hidden information especially in public places, has received more attention. And so it has to face many challenges. Therefore, different methods have been developed so far for hiding information in different cover media. Steganography is the art and science of invisible communication which takes place between two different entities. This is done by hiding the information in other information. It is the way of hiding the existence of the communicated information. Steganography is often confused with cryptography because the two are similar in the way that they both are used to protect confidential information. Steganography is different from cryptography. Cryptography focuses on keeping the contents of a message secret, while steganography focuses on keeping the existence of a message secret [4]. This paper intends to give an overview of security systems with a main concentration on steganography, its uses and techniques.

Keywords-steganography; steganalysis; discrete cosine transformation (DCT); discrete wavelet transformation (DWT); model-based steganography (MB)

I. INTRODUCTION

With the development of computer and expanding its use in different areas of life and work, the issue of information security has become increasingly important. One of the grounds discussed in information security is the exchange of information through the cover media. To this end, different methods such as cryptography, steganography, coding, etc have been used. The method of steganography is among the methods that have received attention in recent years [1]. The main goal of steganography is to hide information in the other cover media so that other person will not notice the presence of the information.

The word steganography is derived from the Greek words "*stegos*" meaning "cover" and "*grafia*" meaning "writing" defining it as "covered writing" [1]. The strength of steganography can thus be amplified by combining it with cryptography. Two other technologies that are used for information hiding are watermarking and fingerprinting [5]. In watermarking, information hidden in objects is usually a signature to signify origin or ownership for the purpose of copyright protection [6]. With fingerprinting on the other hand, different, unique marks are embedded in distinct copies of the carrier object that are supplied to different customers. This enables the intellectual property owner to identify customers who break their licensing agreement by supplying the property to third parties [5]. In watermarking and fingerprinting the fact that information is hidden inside the files may be public knowledge – sometimes it may even be visible – while in steganography the imperceptibility of the information is crucial [4].

Most steganography jobs have been carried out on images, video clips, texts, music and sounds .Nowadays, using a combination of steganography and the other methods, information security has improved considerably. Steganalysis is used to detect the presence of steganography.

Hiding information into a media requires following elements [7].

- The cover media(*C*) that will hold the hidden data
- The secret message (*M*), may be plain text, cipher text or any type of data
- The stego function (*Fe)* and its inverse (*Fe-1*)
- An optional stego-key (*K*) or password may be used to hide and unhide the message.

The stego function operates over cover media and the message (to be hidden) along with a stego-key (optionally) to produce a stego media (*S*). The schematic of steganographic operation is shown below.

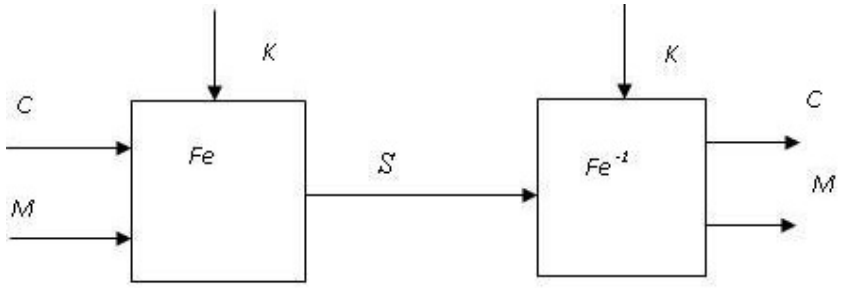

Figure 1. The Steganographic operation

Steganography and Cryptography are great partners in spite of functional difference. It is common practice to use cryptography with steganography. Table 1 shows a comparison of different techniques for communicating in secret [8].

TABLE 1
COMPARISON OF SECRET COMMUNICATION TECHNIQUES

Secret Communication Techniques	Confidentiality	Integrity	Unremovability
Encryption	Yes	No	Yes
Digital Signatures	No	Yes	No
Steganography	Yes/No	Yes/No	Yes

II. Types of Information Security

Following Figure 2 shows the different areas of information security. Cryptography is used to protect the confidential message from unintentional receivers. Information hiding is divided into two major areas. One is steganography and the other one is document marking. Document marking is used for the purpose of proving the ownership of the document. Document marking can be done either by watermarking or fingerprinting. In watermarking all of the instances of an object are “marked” in the same way. The kind of information hidden in objects when using watermarking is usually a signature to signify origin or ownership for the purpose of copyright protection [6]. With fingerprinting on the other hand, different, unique marks are embedded in distinct copies of the carrier object that are supplied to different customers. This enables the intellectual property owner to identify customers who break their licensing agreement by supplying the property to third parties [5]. This paper intends to give an overview of different algorithms used for steganography.

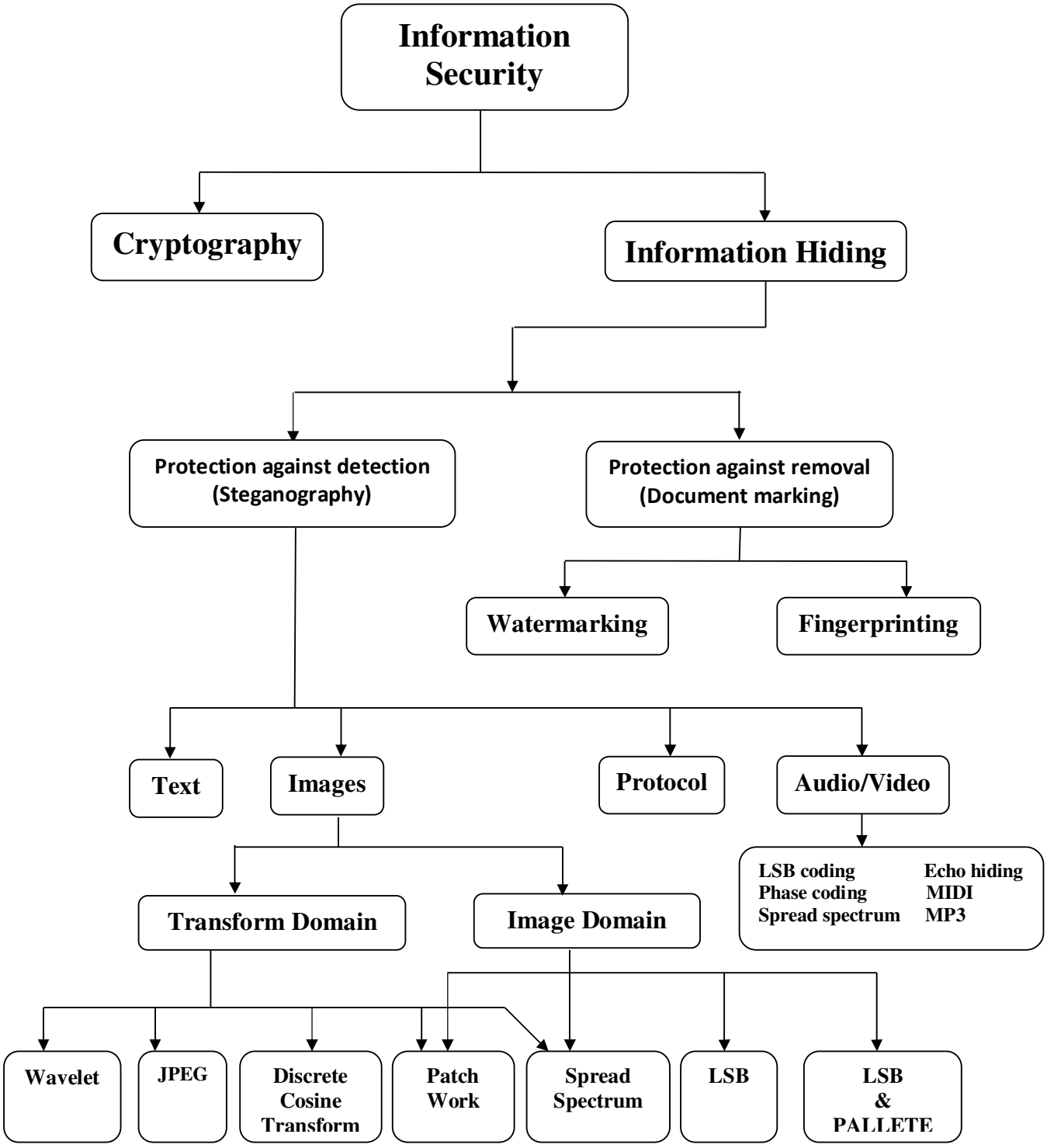

Figure 2. Toxonomy of Information Security

III. Steganography Media/Cover used for digitally embedding a message

- Texts
- Images
- Audio/video
- Protocol

Historically the most important method of steganography is hiding information in texts. Text steganography using digital files is not used very often since text files have a very small amount of redundant data. Most widely used cover for steganography in now a day is digital image because of having large amount of redundant bits present in the digital representation of images. In audio steganography, message is embedded into digitized audio signal. To hide information in audio files similar techniques are used as for image files. Masking is one different technique unique to audio steganography, which exploits the properties of the human ear to hide information unnoticeably. A faint, but audible, sound becomes inaudible in the presence of another louder audible sound [1]. This property creates a channel in which to hide information. Although nearly equal to images in steganographic potential, the larger size of meaningful audio files makes them less popular to use than images [9]. The term protocol steganography refers to the technique of embedding information within messages and network control protocols used in network transmission [10]. In the layers of the OSI network model there exist covert channels where steganography can be used. An example of where information can be hidden is in the header of a TCP/IP packet in some fields that are either optional or are never used.

3.1. Plaintext Steganography

3.1.1. Using selected characters of the plaintext

A series of integer number (Key) is sent to the recipient by the sender and there is a prior agreement that the secret message is hidden within the respective position of subsequent words of the cover text. For example the series is '1, 1, 2, 3, 4' and the cover text is **"Today whether is very beautiful"**. So the hidden message is **"Twsru"**. A "0" in the number series will indicate a blank space in the recovered message. The word in the received cover text will be skipped if the number of characters in that word is less than the respective number in the series (Key) which shall also be skipped during the process of message unhide.

3.1.2. Use of extra white space characters of plaintext

A number of extra blank spaces are inserted between consecutive words of cover text. A number of white space corresponds to a certain value. This numbers are mapped to a hidden message through an index of a lookup table. For example extra five spaces between adjacent words indicate the number "5" which subsequently indicates a specific text of a look-up table which is available to the both communicating parties as a prior agreement.

3.1.3. XML

XML file is used for data hiding due to its larger size. Data can be hidden in XML by using the different tags as allowed by the W3C. For example both of these image tags are valid and could be used to indicate different bit settings

Stego key:

<img></img> -> 0
<img/> -> 1

In this way a piece of XML like the following could be used to encode a simple bit string.

Stego data:
<img src="foo1.jpg"></img>

<img src="foo4.jpg"></img>

That XML stores the bit string 0110. Another way of hiding data is by using the space inside a tag. Once again the following XML code is used as the key while the code after is an example of how it could be used to store a string:

Stego key:
<tag>, </tag>, or <tag/> -> 0
<tag >, </tag >, or <tag /> -> 1
Stego data:
<user ><name>User1</name ><id >01</id></user>
<user><name >User2</name><id>02</id ></user >

The XML data in this case stores the bit strings 101100 and 010011.

3.2. Image Steganography

3.2.1. Transform Domain

- **JPEG Compression**

In images there are two types of compression: lossy and lossless. Lossy compression removes the details that are too small for the human eye to differentiate, resulting in close approximation of the original image. JPEG uses lossy compression. Lossless compression never removes information from the original image but represents the data in mathematical formulas. Thus it maintains the integrity of the original image and the decompressed image output is bit-by-bit identical to the original image input. To compress an image into JPEG format, the RGB colour representation is first converted to a YUV representation. The Y component refers to the luminance (or brightness) and the U and V components refers to chrominance (or colour).The next step is the actual transformation of the image. For JPEG, the Discrete Cosine Transform (DCT) is used, but similar transforms are for example the Discrete Fourier Transform (DFT). These mathematical transforms convert the pixels in such a way as to give the effect of "spreading" the location of the pixel values over part of the image [14]. The DCT transforms a signal from an image representation into a frequency representation; by grouping the pixels into 8 × 8 pixel blocks and transforming the pixel blocks into 64 DCT coefficients each [15]. A change in a single DCT coefficient will affect all 64 image pixels in that block. The next step is the quantization phase of the compression. JPEG does this by dividing all the values in a block by a quantization coefficient. The results are rounded to integer values and then Huffman coding is used to encode the coefficients to further reduce the size [14].

- **JPEG Steganography**

The major JPEG steganography methods are JSteg/JPHide, F5,OutGuess,MB and YASS.

JSteg/JPHide: JSteg functions to hide the secret data in a cover image by simply exchanging the LSBs of non-zero quantized DCT coefficients with secret message bits. The quantized DCT coefficients, already used to conceal secret message bits in JPHide, are selected randomly by a pseudo-random number generator. JPHide, on the other hand, tends not only to modify the LSBs of

the selected coefficients, but it can also switch to a process where bits of the second least-significant bit-plane are likely to be worked out.

F5: F5 algorithm reduces the absolute value of the coefficient by one if it needs modification. In addition to embedding message bits into randomly chosen DCT coefficients, the F5 algorithm employs matrix embedding that reduces the number of changes necessary for hiding a message of a certain length.

OutGuess: There are two stages representing the embedding process of OutGuess. The first of which is that OutGuess embeds secret message bits along a random walk into the LSBs of the quantized DCT coefficients while skipping 0s and 1s.Soon after modifications are made to the coefficients which are already left during embedding, to make the global DCT histogram of the stego image match that of the cover image.

MB: Model-based steganography (MB) can be defined as a general framework for conducting both steganography and steganalysis by simply using a statistical model of the cover media. The MB method for JPEG images is capable of having high message capacity while remaining secure against many first-order statistical attacks [16].

YASS: Yet another steganographic scheme (YASS) belongs to JPEG steganography, but does not conceal data in JPEG DCT coefficients directly. Instead, an input image in the spatial domain is divided into blocks with a fixed large size, called big blocks (or B-blocks). A later stage is to randomly select within each B-block, an 8 × 8 sub-block known as embedding host block (or H-block). Then via using error correction codes, secret data is encoded and embedded in the DCT coefficients of the H-blocks. Finally, the entire image is compressed and distributed as a JPEG image after inversing DCT on the H-blocks [16].

- **Wavelet Transform**

Spatial domain information is converted to the frequency domain information by the Wavelets transform. Wavelet transform clearly partitions the high-frequency and low-frequency information on a pixel by pixel basis that's why the wavelets are used in the image steganographic model.

- **Discrete Cosine Transform**

DCT is also one way of hiding data. In this the image is split up into 8 x 8 squares. Then these squares are transformed via a DCT, which outputs a multi dimensional array of 63 coefficients. A quantizer rounds each of these coefficients, which essentially is the compression stage as this is where data is lost. Small unimportant coefficients are rounded to 0 while larger ones lose some of their precision. After this result will be an array of streamlined coefficients, which are further compressed via a Huffman encoding scheme or similar. Finally decompression is done via an inverse DCT.

3.2.2. Spatial Domain

- **LSB**

When hiding the message bits in the image using LSB algorithms, there are two schemes, namely sequential and scattered. The LSBs of the image, in the sequential embedding scheme are replaced by the message bits, whereas in the case of the scattered embedding scheme, the message bits are randomly scattered throughout the image using a random sequence to control the embedding sequence [12].

- **LSB and Palette based images**

GIF images are palette based images.GIF images are indexed images where the colours used in the image are stored in a palette, sometimes referred to as a colour lookup table. In palette based image steganography, one change in the LSB of a pixel will result in a completely different colour since the index to the colour palette is changed. One possible solution is to sort the palette so that the colour difference between consecutive colours is minimized [12].

3.2.3. Image or Transform Domain

- **Patch Work**

Patchwork is a statistical technique that uses redundant pattern encoding to embed a message in an image. The algorithm adds redundancy to the hidden information and then scatters it throughout the image. A pseudorandom generator is used to select two areas of the image (or patches), patch A and patch B. All the pixels in patch A is lightened while the pixels in patch B are darkened. In other words the intensities of the pixels in the one patch are increased by a constant value, while the pixels of the other patch are decreased with the same constant value [6]. The contrast changes in this patch subset encodes one bit and the changes are typically small and imperceptible, while not changing the average luminosity [12].

- **Spread Spectrum**

In spread spectrum techniques, hidden data is spread throughout the cover-image making it harder to detect [4].

3.3. Audio/Video Steganography

3.3.1. LSB coding

Analog audio signal will be converted to digital binary sequence by sampling technique which is then followed by the quantization. In this approach binary equivalent of secret message will replace the LSB of binary sequence of each sample of digitized audio file. For example if we want to hide the letter 'A' (binary equivalent **01100101**) to an digitized audio file where each sample is represented with 16 bits, then LSB of 8 consecutive samples (each of 16 bit size) is replaced with each bit of binary equivalent of the letter 'A'[11].

3.3.2. Phase coding

Human Auditory System (HAS) can't recognize the phase change in audio signal as easy it can recognize noise in the signal. The phase coding method exploits this fact. This technique encodes the secret message bits as phase shifts in the phase spectrum of a digital signal, achieving an inaudible encoding in terms of signal-to- noise ratio [11].

3.3.3. Spread spectrum

- **DSSS:**

In this approach Direct-sequence spread-spectrum transmissions multiply the data being transmitted by a "noise" signal. This noise signal is a pseudorandom sequence of 1 and −1 values, at a frequency much higher than that of the original signal, thereby spreading the energy of the original signal into a much wider band.

- **FHSS:**

Frequency-hopping spread spectrum pseudo-randomly retunes the carrier, instead of adding pseudo-random noise to the data, which results in a uniform frequency distribution whose width is determined by the output range of the pseudo-random number generator.

3.3.4. Echo hiding

In this method the secret message is embedded into cover audio signal as an echo. Three parameters of the echo of the cover signal namely amplitude, decay rate and offset from original signal are varied to represent encoded secret binary message. Video files are generally consists of images and sounds, so most of the relevant techniques for hiding data into images and audio are also applicable to video media. In the case of Video steganography sender sends the secret message to the recipient using a video sequence as cover media [11].

3.3.5. MIDI

MIDI files can also be used to hide information due to the revival this format has had with the surge of mobile phones, which play MIDI ring tones. There are also techniques which can embed data into MIDI files easily. MIDI files are made up of a number of different messages. Some of these messages control the notes you hear while others are silent and make up the file header or change the notes being played [13].

3.3.6 MP3

MP3 is also one way of hiding the data. The hidden data is encoded in the parity bit of this information. As MP3 files are split up into a number of frames each with their own parity bit, a reasonable amount of information can be stored. To retrieve the data it is required to uncompress the MP3 file and read the parity bits as this process is done [13]. This is an effective technique which leaves little trace of any distortions in the music file.

3.4. Protocol Steganography

3.4.1. Covert channel communication using 'Flags' field

If sender and recipient both have a prior knowledge of Maximum Transfer Unit (MTU) of their network then they can covertly communicate with each other using DF flag bit of IP header.

3.4.2. Covert channel communication using 'Identification' field

The '16-bit identification field' in Ipv4 header is used to identify the fragmented packed of an IP datagram. If there is no fragmentation of datagram, then this Identification field can be used to embed sender specified information [11].

3.4.3. Covert channel communication using ISN(initial sequence number)field

The large 32-bit address space of the Sequence Number field can be used for covert channel. The sender will craft a TCP/IP packet where the secret binary message can be embedded over the Sequence Number field and the passively listening receiving party will then extract the data. Source Port and Checksum in UDP header Code field in ICMP header is also good candidate for Cover item for Network Steganography.

IV. CONCLUSION

This paper reviewed the several steganographic techniques with the use of different cover media such as text, images, audio or video files and protocols. Image steganographic techniques in both spatial domain and transform domain are also presented. Also different image file formats have different methods of hiding information. Audio and video files are also used for secure information exchange. Different fields of the IP datagram can also be used for data hiding. Recently images are the most popular to be used as the cover media. It is required to exchange the confidential information between two entities now a days on a regular basis so the research for a strong steganographic is continuous process.

REFERENCES

[1] Moerland, T. ,"Steganography and Steganalysis", Leiden Institute of Advanced Computing Science, www.liacs.nl/home/ tmoerl/privtech.pdf

[2] Silman, J., "Steganography and Steganalysis: An Overview", SANS Institute, 2001

[3] Jamil, T., "Steganography: The art of hiding information is plain sight", IEEE Potentials,1999

[4] Wang, H & Wang, S, "Cyber warfare: Steganography vs. Steganalysis", Communications of the ACM, October 2004

[5] Anderson, R.J. & Petitcolas, F.A.P., "On the limits of steganography", IEEE Journal of selected Areas in Communications, May 1998

[6] Marvel, L.M., Boncelet Jr., C.G. & Retter, C., "Spread Spectrum Steganography", IEEE Transactions on image processing, 8:08, 1999

[7] Bret Dunbar ,"Steganographic Techniques and their use in an Open-Systems Environment", The Information Security Reading Room, SANS Institute 2002

[8] Shashikala Channalli, Ajay Jadhav, "Steganography-An Art of Hiding Data", International Journal on Computer Science and Engineering Vol.1(3), 2009, 137-141

[9] Artz, D., "Digital Steganography: Hiding Data within Data", IEEE Internet Computing Journal, June 2001

[10] Ahsan, K. & Kundur, D., "Practical Data hiding in TCP/IP", Proceedings of the Workshop on Multimedia Security at ACM Multimedia, 2002

[11] Soumyendu Das, Subhendu Das, Bijoy Bandyopadhyay, Sugata Sanyal, "Steganography and Steganalysis: Different Approaches"

[12] T. Morkel , J.H.P. Eloff, M.S. Olivier**, "**An overview of image steganography"

[13] Jonathan Cummins, Patrick Diskin, Samuel Lau and Robert Parlett, "Steganography and digital watermarking"

[14] Currie, D.L. & Irvine, C.E., "Surmounting the effects of lossy compression on Steganography", 19th National Information Systems Security Conference, 1996

[15] Krenn, R., "Steganography and Steganalysis", http://www.krenn.nl/univ/cry/steg/article.pdf

[16] B. Li, J. He, J. Huang, and Y.Q. Shi. "A survey on image steganography and steganalysis." Journal of Information Hiding and Multimedia Signal Processing. 2(2), [On line], pp. 142-172. Available: http://bit.kuas.edu.tw/~jihmsp/2011/vol2/JIH-MSP-2011-03-005.pdf [Dec., 2011].

