Research Inventy: International Journal of Engineering And Science
Issn (e): 2278-4721, Issn (p):2319-6483, PP: 54-62
www.researchinventy.com

A Comparative Study on Image Steganography Techniques

Oyendrila Samanta[1,*], Ruchismita Sahu[2]

[1]Assistant Professor, Department of CSE, Techno International Batanagar, Kolkata, West Bengal, India, Mail id - oyendrila.samanta@tib.edu.in

[2]Assistant Professor, Templecity Institute of Technology and Engineering

Abstract

Steganography is the knack of embedding significant secret material inside different information which is digitally covered. Using steganography, information can be hidden in different embedding mediums, known as carriers. These carriers can be image files, audio files, video files, and text files. In this paper an image file is used as a carrier because of their high frequency on the internet. If it is achieved successfully, the message does not attract attention from attackers and eavesdroppers. Image Steganography is a method of using an image file as a cover object. Steganalysis, an inherently difficult problem of Steganography is the technology that attempts to deterioration the Steganography by sensing the unseen information and by pulling out. In this paper, focus will be given upon recent image steganography and steganalysis techniques that are used for hiding secret messages and will also cover classifications and applications of Image Steganography.

Keywords: *Steganography, carriers, steganalysis.*

I. Introduction

Data. In essence, the modern computing world revolves around this word. In today's world, businesses have started realizing that data has a huge power as it can potentially forecast customer trends, increase sales, and push the organization to newer heights. Within this rapid development in technology and use of data for continuous improvement it has become our topmost concern to secure data. In this modern period, internet offers great convenience in transmitting large amounts of data in different parts of the world. Data sharing is increasing as thousands of messages and data are being transmitted on the internet everyday from one place to another. The protection of data is the primary concern of the sender and it is really import ant that we encrypt our message in a secret way that only the receiver can able to understand.

Steganography also known as imperceptible communication, is a quite earlier art of embedding personal information into other data by using some rules and techniques [1]. It is distinct as the*science and art of concealing* a undisclosed message in different files types, for instance: *digitalimages files*, *digital audio files*, *digital video files*, and *text files* so that illegal users are not capable to see and identify the embedded information.

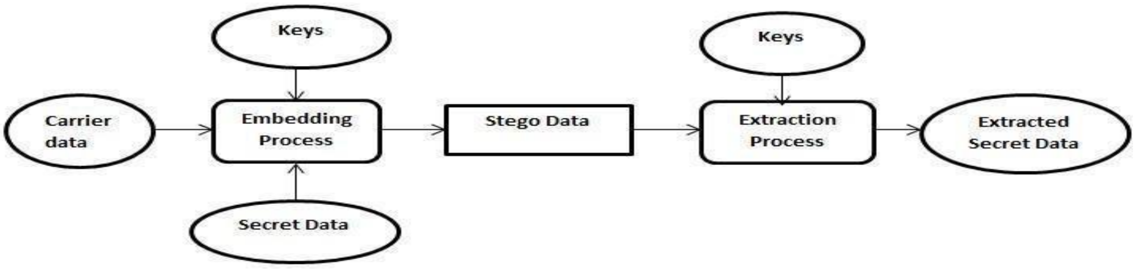


Figure1.1: Steganography Process

A *Steganography* system made up of few apparatuses: *cover − object*, the secret message and the stego-object. In the case of confidential data the safety and security of long-distance communication remains an issue. The need to solve this problem has led to the development of steganography schemes which provides a high level of security.

As displayed in Figure 1.1, the purpose of *steganography* is to conceal-the-message under *cover files*, hiding the very actuality of material exchange.

Definitely, among a *variety of files types*, an "image-steganography" is desired, subsequently the transformed *image* with minor dissimilarities in its *colors* will be impossible to differentiate from the *original image by human eye* [2].

The different types of techniques are used in the Steganography is to hide the messages in the cover images. Figure 1.2 shows two general directions of steganography: protection against detection and protection against removal [3]. Protection against detection uses some ways to embed information invisibly that does not degrade the quality of the original data. Protection against removal supposes that the method should be able to resist to common digital signal processing and noises. Removing the hidden data will definitely reduce the object's quality and its performance will not be functional [4].

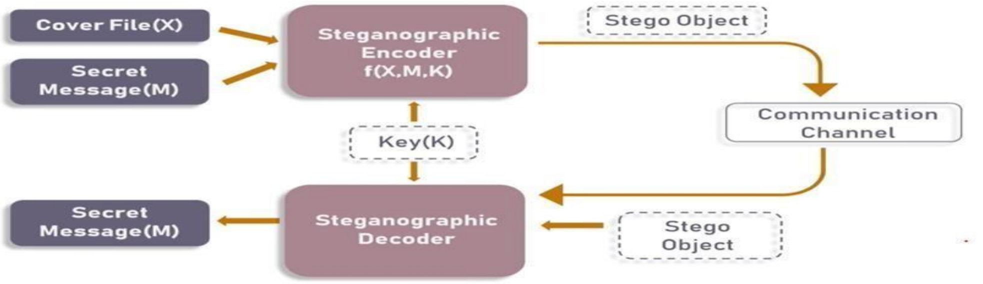


Figure 1.2: A Typical Steganography Technique

1.1 Terminologies of Steganography

- **Secret Message**: This is information which needs to be hidden into some suitable digital media [5].

- **Cover Message**: It is the carrier of message such as image, audio, video or other digital media. It is mainly the object in which the data is to be hidden.
- **Stego-object**: The cover-object carrying the secret message is known as stego-object.
- **Stego-key**: Key used for encrypting and decrypting the secret message.
- **Encoding**: The secret text message is encrypted using an encryption key.

- **Decoding**: The Stego-Image is fed into the decoder which uses a decryption algorithm to provide the original cover and the secret message as output.

1.2 Cryptography and Steganography

Cryptography made up of *Krypto's* means unseen and *Graphene* means *Writing*. *Steganography* word is self-possessed of two-Greek-words, namely: *Stegano and Graphy*. Theword *Stegano means a Covered*, whereas *Graphy* means Script. Therefore, *steganography* means a Covered Writing. The main transformation between *Steganography and cryptography* is that *cryptography* focuses on keeping

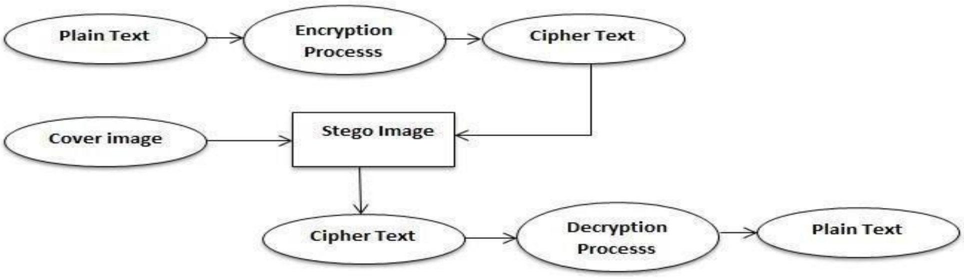


Figure 1.3: Combination of Cryptography and Steganography

the contents of a message secret whereas steganography focuses on keeping the existence of a message secret [6]. The strength of steganography can thus be amplified by combining it with cryptography. Steganography and cryptography are both ways to protect information from unwanted parties but neither technology alone is perfect and can be compromised.

Reasons for using Digital Image as cover-media for steganography:
The reasons for using digital image as cover-media for steganography are:

a. It is the most widely used medium.
b. Takes benefit of the limited-visual-perception of*colors*.
c. Digital images are made up of pixels.
d. The arrangement of pixels makes up the image's '*raster data*'.
e. 8-bit and 24-bit images are common
f. The larger the image size, the more information can be hidden

II. Several Techniques of Image Steganography

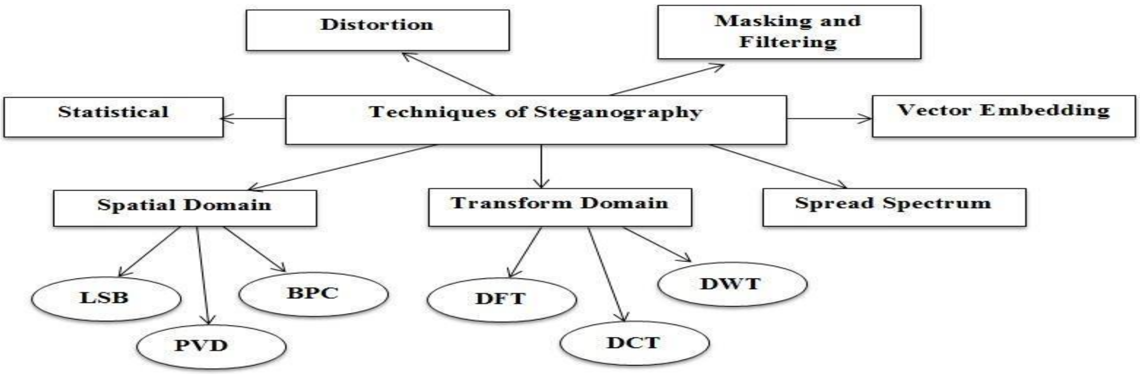


Figure 2.1: Techniques of Steganography

2.1 Spatial Domain based Steganography

The term spatial domain refers to the image plane itself and approaches in this domain are based on direct manipulation of pixels of an image. Spatial domain methods, also known as substitution techniques are procedures that operate directly on pixels and are aggregate of pixels composing an image.

In the spatial domain approach, the secret message is embedded directly into the pixels of a cover image. This method involves modification of the secret data in the spatial domain of the cover image. There are several kinds of Spatial Domain Techniques, some of which are discussed below.

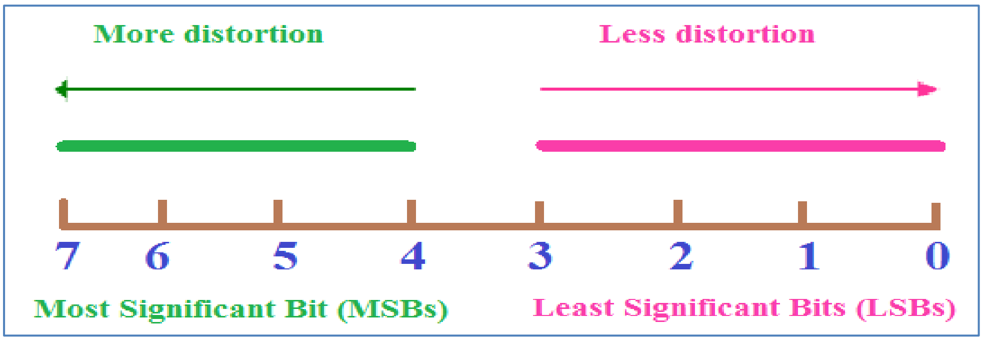


Figure 2.2: One byte representation of a pixel with integer to binary conversion

2.1.1 LSB based Steganography

Least Significant Bit Substitution (LSB) [7] is the most simple and commonly used steganographic technique. The main objective of this technique involves the embedding of the secret information at the bits which having the least weightage so that it will not influence the value of original pixel and also the most significant bits of the pixel of that image can be hidden. This method often works with raster images, presented in a format without compression (e.g. *.gif, *.bmp). This file formats are preferred because they offer "lossless" compression. But, other image formats are used as cover image as well [8]. For example, the least significant bit of the binary number 10111001, is the far right 1 where the secret information is stored.

This procedure workings by substituting some of the evidence ina given $pixel\ with\ information$from the $data$ in the $image$. Since A $LSB\ algorithm\ replaces$ the $most - right\ bits$ of a cover-files-bytes. In situation a $bit\ of\ the\ cover\ image\ C(m,\ n)$ is alike to the $bit\ of\ a\ secret\ message$ that to be entrenched, $C(m,\ n)$ stay untouched, or $C(m,\ n)$ is set to bit of a $secret\ message\ (SM)$[9]. For example, the letter $'C'$ is an $ASCII\ code$, which is 01000011 in binary and 67 in decimal and bits-of-image pixels before inserting a $secret\ message$ are:

$Pixel1$: 1111100**0** 1100100**1** 0000001**1**
$Pixel2$: 1111100**0** 1100100**1** 0000001**1**
$Pixel3$: 1111100**0** 1100100**1** 00000011

Least-Significant-Bit $(LSB) algorithm$pelts bits of letter 'A', which are **01000001**, into $image\ pixels\ to\ produce$:

$Pixel1$: 1111100**0** 1100100**1** 0000001**0**
$Pixel2$: 1111100**0** 1100100**0** 0000001**0**
$Pixel3$: 1111100**1** 1100100**1** 00000011

The $least\ significant\ bit\ (in\ other\ words, the\ 8th\ bit)$ of few or all of the $bytes$ inside an imageis altered to a bit-of -secret communication. In a $24 - bit\ image$, a bit of each of the red, $green$ and blue color components can be used, since they are each represented by a byte.

2.1.2 Palette based Steganography

In this image steganography technique the bits of secret message will be hidden inside the cover image using the stretched palette of the cover image. Palette based image enables 8 bits per pixel or less to look almost as good as 24 bits per pixel. Rather than each pixel in the image having all three RGB colors (one 8-bit red, one 8-bit green and one 8-bit blue) each pixel contains one 8-bit number that indexes into the 256-color lookup table, which contains the RGB values [8]. This algorithm consists of color-quantization and dithering. Color-quantization selects the palette of the image by truncating all colors of the original raw, 24-bit image to a finite number of colors. Dithering is used for apparent increasing of color depth that uses the integrating properties of the human visual system and creates the illusion of additional colors by trading space resolution for color depth.

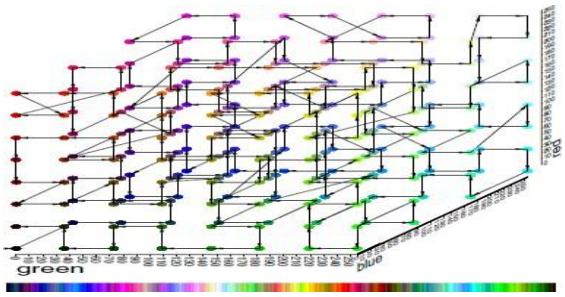
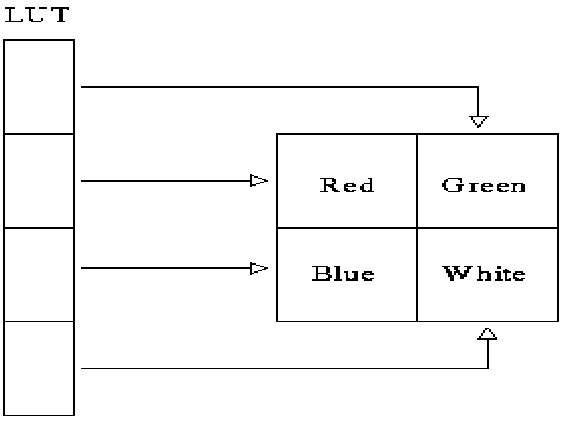


Figure 2.3: Sorting of color in palette as used in EzStego method



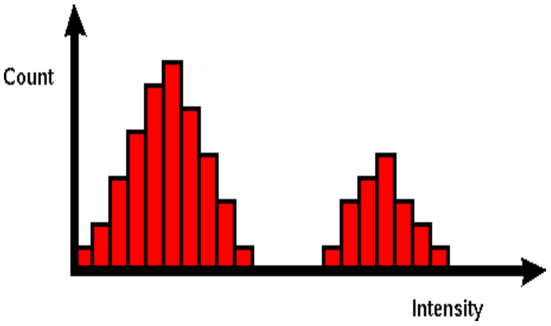


Figure 2.4: Intensity Histogram **Figure 2.5: Common Names:** Histogram

One common approach is the *popularity algorithm*, which creates a histogram of all colors and retains the 256 most frequent ones. The 256 values of the colormap are divided into four sections containing 64 different values of red, green, blue and white. As can be seen in Figure2.5, a 2×2 pixel area is grouped together to represent one composite color, each of the four pixels displays either one of the primary color or white. In this way, the number of possible color is increased from *256* to 64^4.

2.1.3 Pixel Value Differencing (PVD):

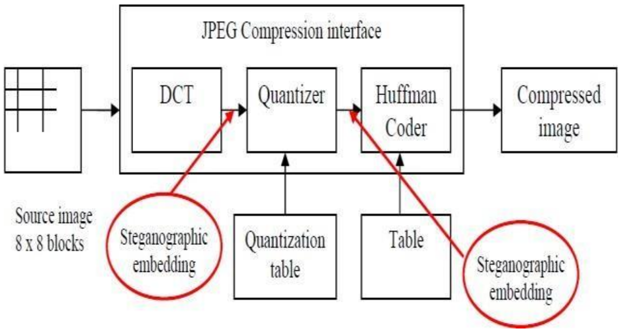
Pixel-value differencing (PVD) scheme selects two subsequent pixels and creates a quantization range table which can able to find out the payload by the difference value between the subsequent pixels. Thus it provides high imperceptibility to the stego image. Besides, it offers the advantage of passing on a large number of payloads, while still sustaining the regularity of an image characteristic after data embedding.

2.2 Frequency Domain based Steganography

There are many transforms used to map a signal into the frequency domain such as *Discrete cosinetransform* (*DCT*), *Discrete wavelet transform* (*DWT*), *Discrete Fourier transform* (*DFT*), Vector Embedding, Spread Spectrum, Statistical and Distortion technique, File and Palette embedding, Image Generation Technique and Image Element Modification Technique etc. Some of the techniques are discussed below.

2.2.1 Discrete Cosine Transform (DCT) based Steganography

It is a widely used JPEG (Joint Photographic Experts Group) image compression technique for converting the signal or image from spatial domain to the frequency domain. Such technique basically applies lossy compression in images and thus they form an image with some loss in bits. In order to compress an image into JPEG format, the RGB color representation is first converted to a YUV representation space and each color plane is partitioned into non-overlapping 8 x 8 blocks of pixels. In this representation the Y component corresponds to the luminance and the U and V components correspond to chrominance. Here the image is broken into 8×8 pixel blocks and these pixel blocks are transformed into 64 DCT.



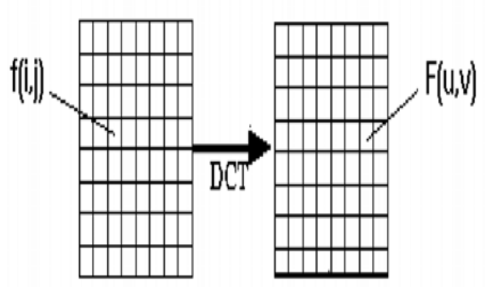


Figure 2.6: DCT based Steganography during JPEG **Figure 2.7:** Process of DCT compression process

A modification of a single DCT coefficient will affect all 64 image pixels in that block. Each DCT coefficient F (u,v) of an 8x8 block of image pixels f(x,y) is given as stated by:

$$F(u,v) = \frac{1}{4}C(u)C(v)\Big[\sum_{x=0}^{7}\sum_{y=0}^{7}(x,y)*^{(2x+1)}\frac{Cos^{(2x+1)\pi}}{16}\Big]\ldots\frac{\ldots\ldots\ldots}{16} \qquad (2.1)$$

Where C(u)=1/√2 when u =0 and C(u) = 1 otherwise. C(v)=1/√2 when v= 0 and C(v) = 1 otherwise.
In this case x, y, u, vϵ\{0,1,…, 7\} and f(x,y) is the particular pixel color space component.

2.2.2 Discrete Wavelength Transform (DWT) based Steganography

A $discrete\ wavelet\ transform$ (DWT) is used to translate as $Signal$ from $spatial$ $domain$ to $frequency\ domain$. It represents an image as a sum of wavelet functions (wavelets) with different locations and scales and any decomposition of an image into wavelets involves a pair of waveforms: one to represent the high frequencies corresponding to the detailed parts of an image (wavelet function) and one for the low frequencies or smooth parts of an image (scaling function).

Haar-DWT is the simplest form of DWT in which the low frequency wavelet coefficients are generated by averaging the two pixel values and high frequency coefficients are generated by taking half of the difference of the same two pixels. Haar wavelet is not continuous, therefore not differentiable and is used to convert spatial domain image to wavelet domain. In this method, a spatial domain image is transformed into a HDWT-based frequency domain image and then the high frequency coefficients are used to embed the secret data. This method provides a high hiding capacity and a good stego-image quality.

2.2.1 Discrete Fourier Transform (DFT) based Steganography

It is the maximum significant discrete-transform which is used to bring out $Fourier$ $analysis$ in numerous-practical-applications. It is the $purely\ discrete\ transform$ where a limited list of equally-spaced samples of a function is converted into the list of coefficients of a finite amalgamation of complex sinusoids ordered by their frequencies. It is a very wild procedure compared with others. This algorithm is also known as $Fast\ Fourier\ Transform$ i.e.FFT and it generates the same result as of the DFT by using the $Inverse\ Discrete\ FourierTransform$ $(IDFT)$.

III. Image Steganalysis

Whether an image consist of secret data or not, it is measured by Steganalysis technique.

3.1 Type-of-Attacks

There are $six\ general\ protocols$ used to attack the use of $Steganography$ as pointed- out-by Katzenbeisser and Petitolas (2000) [10].

3.2 Image based Steganalysis Techniques

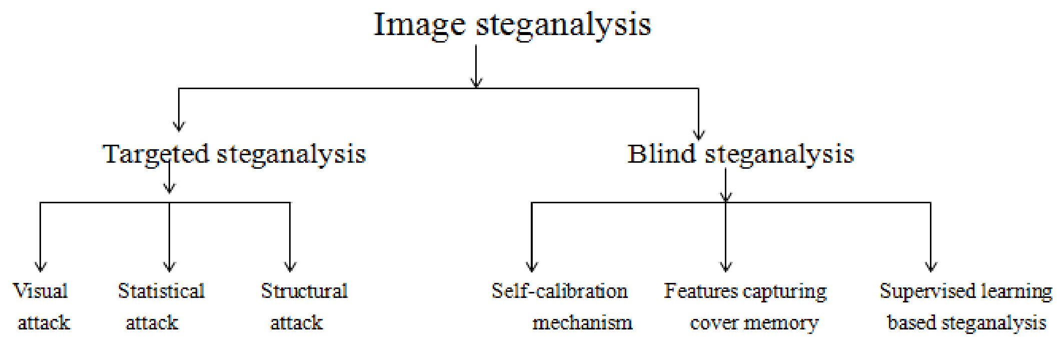


Figure 3.1: Classification of image Steganalysis Techniques

3.2.1 Targeted Steganalysis

(a) *Visual Attacks*

It is the *simplest form of steganalys is* that involves examining the *stego − image* with the naked-eye to discovery out any-kind-of misrepresentation.

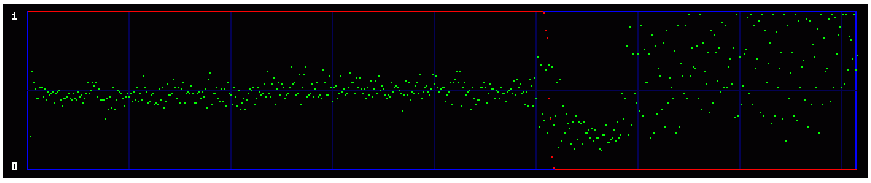
(b) *Statistical Attacks*

In this type of attacks, the statistical analysis of the images by some mathematical formula is performed to detect the presence of hidden data. Statistical tests try to reveal whether an *image* has been amended by formative *ge's statistical properties deviate from a norm*. Chi-square Analysis is one of such attack that belongs to *statistical attack*.

I *Chi − squareAnalysis*

Westfeld and Pfitzmann, in 2000 [11] used a Chi-square (χ2) test to determine whether the color frequency distribution in an image matches a distribution that shows distortion from embedding data with the probability of statistics under the condition that the dispersals occurrences of the color guides before embedding and afterward inserting are equivalent. They increased the-sample-size and applied the trial at a continuous situation. *Chi − Square Analysis* calculates the avg LSB and builds a table of *frequencies and Pair*of Values; it collects the data from these two-tables and achieves a *chi − square* test. It measures *the theoretical vs. calculated*difference of population. Here is the output of a chi-square analysis on a Steganographed image:

Red=*Chi − Square* result, Green=Average LSB, Blue Horizontal=1KB of data



The red line, a value close to 1, indicates that there is Steganography within the first 5KB of the image. The graph also shows that the LSB's are quite similar during that range, which further supports the conclusion. This attack cannot detect patterns or *Steganography* on very complex *images* with loads of *noise* than 1 can detect through visualization of the Enhanced LSB's.

3.2.2 Blind Steganalysis

It is an approach of detecting secret message embedded into a file even when it is not certain how the evidence might have been surrounded. It works differently to targeted steganalysis as it does not need prior information about *details of the embedding operations*.Some of the methods that belong to the blind steganalysis schemes are Self Calibration Mechanism, Features capturing cover memory, Supervised learning based steganalysis etc.

IV. Performance Evaluation of Different Schemes Explored in this paper

The parameters upon which the steganography is measured is discussed below:

4.1 Measure of Steganographic Capacity

Capacity is the most important parameter since the size of the secret information has direct impact on a steganographic system. Evaluating the capacity of a steganography technique is the maximum number of bits that can be hidden in a given cover image with an insignificant probability of detection by an adversary.

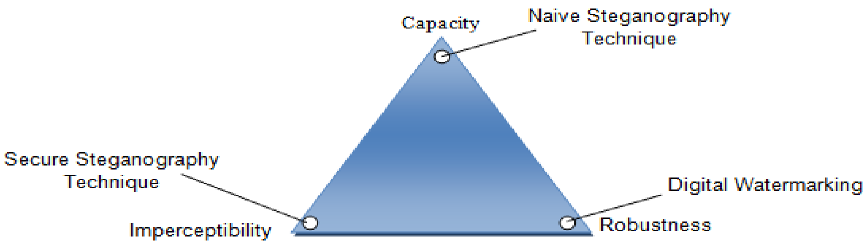


Figure4.1: Competing factors in steganographic systems

Cole and Krutz,(2003) stated that '*the more data you can hide, the better the technique*'. Therefore, designing a steganography technique should take into consideration how to increase the amount of secret data that can be embedded without affecting the properties of stego- image.

4.2 Measure of Robustness

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, and image filtering.

4.3 Measure of Imperceptibility

The invisibility of the embedded information is the first and foremost evaluation criteria, since the strength of image steganography lies in its ability to go unnoticed to human eye. If any changes to the bits of image lead to visual distortion which becomes noticeable then the overall objective of the steganographic method fails. On the other hand, if the level of invisibility is high in image steganography algorithms then the overall objective of the approach is fulfilled. Thus, for better evaluation and comparison it is therefore necessary to consider the perceptibility of the resultant image in the evaluation process. Chang *et al.*, (2002) [14] stated that '*The better quality the stego-image has, the more secure the steganography system will be*'.

Generally, there are two primary ways to measure image quality: objective quality methods(automated) and subjective quality methods (human based). The higher the quality of stego-images, the larger the imperceptibility of the steganographic system.

PSNR and MSE:

Nowadays, the most popular and common distortion measures used to evaluate the quality of images in the field of image processing is the peak signal-to-noise ratio (PSNR) and Mean squared error(MSE).The PSNR measures the similarity between two images (how two images are close to each other) and are usually measured in decibels (dB) while the MSE measures the statistical difference in the pixel values between the original and the reconstructed images and measured in percentage. Greater the $PSNR$ value indicates the enhanced quality of $image$ i.e. less distortion. PSNR is the ratio of the maximum signal to noise in the stego-image.

Moreover, PSNR and MSE are defined as follows:

Moreover, PSNR and MSE are defined as follows:

$$PSNR=10_{log_{10}}\ (\frac{255^2}{MSE})\ ....(5.1)$$

Where mean square error (MSE) is a measure used to quantify the difference between the cover-image I and the stego (modified) image$I$`.If the image has a size of M *N then

$$MSE=\frac{1}{M*N}\sum^{M}_{i=1}\sum^{N}_{j=1}|(\iota,)-I(\iota,\jmath)|\ ....(5.2)$$

For color images, PSNR is similarly defined as follows:

$$PSNR=10_{log_{10}}\ (\frac{255^2}{MSE_{AVG}})................(5.3)$$

Where the MSE for color images is defined as follows:

$$MSE_{AVG}=\frac{MSE_R}{3}+\frac{MSE_G}{3}+\frac{MSE_B}{3}\ ....(5.4)$$

Where MSE_R, SE_G and MSE_B are the MSE of red, green, and blue components respectively.

4.4 Bit Error Rate (BER)

The $hidden\ information$ can be magnificently recovered from the communiqué channel. It must be ideal but for the real communication channel, the error comes while retrieving hidden information and this is measured by BER. It is the ratio of the number of errors to the total no of bits sent in an image.

V. Conclusion

$Steganography$ is a legitimately new research idea there are continuous progressions in the computer field, suggesting advancements in the field of steganography as well. It is likely that there will quickly be more effective and more progressive procedures for $Steganalysis$. A hopeful advancement is the enhanced sensitivity to minor messages. Knowing how problematic it is to sense the presence of a equally huge text file within an $image$, imagine how

challenging it is to detect even “one or two” sentences entrenched in an $image$! In the forthcoming, it is anticipated that the technique of $Steganalysis$ will advance so that it will turn out to be easier to sense even minor messages within an $image$.

References

[1]. Jin-Suk Kang, Yonghee You, and Mee Young Sung, "Steganography using block-based adaptive threshold," 22nd International Symposium on Computer and Information Sciences, ISCIS2007-Proceedings, pp.1-7,December2007.

[2]. E.Nandhini, M.Nivetha, S.Nirmala and R.Poornima, "MLSB Technique Based 3D Image Steganography Using AES Algorithm,"J.Recent,vol.3, no.1,p.2936,2016.

[3]. Li Bai, Saroj Biswas, and Erik P. Blasch, "An estimation approach to extract multimedia information in distributed steganographic images," 2007 10th International Conference on InformationFusion,July2007.

[4]. A.Martin, G.Sapiro and G.Seroussi, "Is image steganography natural?, "IEEE Transactions on Image Processing, vol.14,no.12,pp.2040-2050, December 2005.

[5]. Mukesh Garg, "An Overview of Different Type of Data Hiding Scheme in Image usingSteganographicTechniques,"ISSN:2277 128XInternational Journal of Advanced Research in Computer Science and Software Engineering Research Paper, vol. 4, no. 1,January2014.

[6]. Huaiqing Wang and Shuozhong Wang,"Cyber Warfare: Steganography vs. Steganalysis," Communications of the ACM, vol. 47,no.10, pp.76-82, October 2004.

[7]. C.KChan and L.M.Chang, "Hiding data in image by simple LSB substitution, "Pattern Recognition, vol.37, pp.469-471, 2003.

[8]. SK Bandyopadhyay and IKMaitra,"An alternative approach of steganography using reference image," arXiv preprintarXiv:1007.1233,2010.

[9]. S.Kurane, H.Harke and S. Kulkarni, "TEXTANDAUDIODATAHIDINGUSINGLSBAND

[10]. DCT A REVIEW APPROACH," Natl. Conf."Internet Things Towar. a Smart Futur."Recent Trends Electron.Commun,2016.

[11]. S.C. Katzenbeisser and F. Petitcolas, "Principles of Steganography," Information Hiding Techniques for Steganography and Digital Watermarking Ed. London: Artech House, pp.43-78,2000.

[12]. Andreas Westfeld and Andreas Pfitzmann, "Attacks on Steganographic Systems," International Workshop on Information Hiding, pp. 61-76, 1999.