# Steganographic Techniques of Data Hiding using Digital Images

A.Komathi[1] , M.Revathy[2], K.Sivasankari[3],

Mrs. Kavitha Subramani.M.E.,(Ph.D).,

Computer Science and Engineering
Panimalar Engineering College,
Chennai,India

***Abstract*— Steganography is the art of hiding the fact that communication is taking place, by hiding information in other information. Many different file formats can be used, but digital images are the most popular to be used, because of their frequency on the Internet. For hiding those secret information in images there exist a large variety of steganographic techniques, some are more complex than others and all of them have respective strong and weak points. Different applications may have different requirements of the steganography techniques to be used. Some applications may require absolute invisibility of the secret information while some other may require a larger secret message to be hidden. This paper gives an overview of image steganography and its uses and techniques. It also uses to identify the requirements of a good steganographic algorithm and briefly reflects on which steganographic techniques are more suitable for which applications.**

***Keywords: Digital Image Steganography, data to be hidden,*cover-image, stego-image, redundant bits, least significant bit, most significant bit, reversible data hiding.**

## I. INTRODUCTION

Steganography word is of Greek origin and essentially means concealed writing. To protect the transmission data from being intercepted or tampered has led to the development of various steganographic techniques. Though, steganography has been manifested long way back during the ancient Greek times. Greek tyrant Histiaeus in 499 BC shaved the head of his slave and wrote message on his scalp. After the hair grew back, slave was dispatched with the hidden message. Pliny the Elder explained how the milk of the *Thithymallus* plant dried to transparency when applied to paper but darkened to brown when subsequently heated, thus providing the way for hiding information. Giovanni Battista Porta described how to conceal a message within a hardboiled egg by writing on the shell with a special ink. In World War II long sentences of regular letters were used to disguise secret messages. With the various advancement in digital signal processing, use of internet, computing power, steganography has become digital. The data hiding process starts by identifying a cover image's redundant bits. i.e., those can be modified without destroying its integrity. The embedding process then creates stego-image by replacing subset of these redundant bits with the bits of the message to be hidden. In digital image steganography, the secret message is embedded within a digital image called cover-image. Cover-image carrying embedded secret data is referred as stego image.

## II. OVERVIEW OF THE EXISTING SYSTEM

Steganography can be used for wide range of applications such as, in defense organizations for safe circulation of secret data, in some of the military and intelligence agencies, in smart identity cards, where personal details are embedded in the photograph itself for copyright control. In medical imaging patient's details are embedded within image providing protection of information and reducing transmission time and cost, in online voting system so as to make the online election secure and robust against a variety of fraudulent behaviors, for data hiding in countries where cryptography is prohibited, in improving mobile banking security, in tamper proofing so as to prevent or detect unauthorized modifications and other numerous applications.

### *A. Features of Existing System:*

Steganographic techniques have various features which characterizes their strengths and weaknesses. Features include:

*Embedding capacity: It is an amount of* data that can be inserted into the cover-media without deteriorating its integrity.
*Perceptual transparency: It is needed to* avoid suspicion the embedding should occur without significant degradation or loss of perceptual quality of the cover media.
*Robustness*: It indicates the ability of an embedded data to remain intact if the stego-image undergoes various transformations such as scaling, rotation, cropping or compression.
*Tamper resistance: It tells the difficult image to be* to alter or forge a message once it is embedded in a cover-media, such as replacing a copyright mark with the one claiming legal ownership.
*Computational complexity:* It refers to the steganography technique employed for encoding and decoding is another consideration and should be given importance.

*B. Classifications of steganographic techniques:*

Classifications of steganographic techniques based on the types of cover files. Since most of the file formats can also be used for steganography, however, only high degree of redundant bits are preferred. The larger size of both audio and video files makes them less popular as compared to images. The term protocol steganography refers to embedding information within network protocols such as TCP/IP. In Spatial domain, cover-image is first decomposed into bits planes and then least significant bit (LSB) of the bits planes are replaced with the secret data bits. Advantages are high embedding capacity, ease of implementation and imperceptibility of the hidden data. The major drawback is vulnerability to various simple statistical methods. Frequency domain embedding are the techniques, which transforms those cover image into frequency domain, secret data is then embedded in frequency coefficients. Advantages include higher level of robustness against simple statistical analysis. Unfortunately, it lacks high embedding. In the compression domain, secret data is embedded into compression codes of the cover-image which is then sent to the most of the receiver. It is very important where bandwidth requirement is a major concern.

## III. SPATIAL DOMAIN STEGANOGRAPHIC TECHNIQUES

The most direct way to represent pixel's color is by giving an ordered triplet of numbers: red (R), green (G), and blue (B) that gives particular color. The other way is to use a table known as palette to store the triplet, and put a reference into the table for each pixel. The spatial domain based steganography technique use LSB algorithm for embedding/extraction of data.

*A. EzStego Data Hiding:*

EzStego data hiding scheme was given by the person called Machado. In this method palette is first sorted by luminance to minimize the perceptual distance between consecutive colors. EzStego then embeds the secret data into the LSB of the indices pointing to the palette colors. This approach works quite well in gray scale images and may work well in images with most of the related colors. The most common drawback is, since luminance is a linear combination of colors R, G, and B (Luminance = 0.299 R + 0.587 G + 0.144 B), occasionally colors with similar luminance values may be relatively far from each other. Other drawbacks are the ease of extraction of hidden data, dependency of stego-image quality on number of palette colors, and ease of detection of presence of data using simple statistical histogram analysis. Fridrich proposed a palette modification scheme for hiding data. In this method, both the cost of removing an entry color in a palette and the benefit of generating a new one to replace it are calculated. If the maximal benefit exceeds the minimal cost then the entry colour is to be replaced. His method mostly reduces the distortion of the carrier images, but mostly suffers with the low embedding capacity as EzStego does. Cheng, *et al*. proposed high embedding capacity technique that can hide 1 bit to 8 bits per pixel, and has no distortion in contrast to EzStego. High capacity data hiding algorithm based on relevance of adjacent pixels difference was given by Ren, *et al.* Ren's method guarantees the better quality of image after hiding mass information.

*B. S-Tools, Hide & Seek, StegoDos, White Noise Storm, and other techniques:*

S-Tools by Andy Brown reduce the number of colors from 256 to 32 while maintaining the image quality. Instead of simply going with adjacent colors as EzStego does, S-Tools manipulate the palette to produce colors that have a difference of one bit. As compared to EzStego, non-linear insertions in S-Tools method make the presence and extraction of secret data more difficult and achieve better results in terms of visual perceptibility. StegoDos works only with 320 X 200 pixels image and involves much effort in encoding and

decoding of the secret message. White Noise Storm includes encryption to randomize the bits within an image and suffers with the problem of using large cover file. Younes, *et al.* proposed a method in which data is inserted into LSB of each byte within the cover-image in encrypted form. Mandal proposed a method with minimum deviation of image fidelity resulting high quality stego-image with better embedding capacity.

### *C. Bit Plane Complexity Segmentation Steganography:*

Bit plane complexity segmentation steganography (BPCS) was introduced by Kawaguchi, *et al.* It is based on the simple idea that the higher bit planes can also be used for embedding information. In BPCS, each block is decomposed into bit-plane. The LSB plane would be a binary image consisting of the LSB of each pixel in the image and so on. In each segmented bit-plane its complexity is analyzed and based on a threshold value block is divided into 'informative region' and 'noise-like region' and the secret data is hidden in noise regions without degrading image quality. BPCS provides high embedding capacity and least degradation of the cover-image as compared to traditional LSB manipulation techniques. Maya, *et al.* uses variance of image block as a parameter for complexity measure. Prime advantages achieved are high embedding capacity and robustness against noise as compared to BPCS technique.

### *D. INFORMATION THEORY-BASED DATA HIDING:*

Hadhoud, *et al.* proposed a technique based on entropy calculation. In this method entropy of the '4' most significant bits (MSBs) are calculated first which contains most detail of each pixel. If the entropy is > 2 then it inserts '4' bits into the '4' LSBs, if not then the entropy of the '5' MSBs is calculated. If it is > 2 then it inserts '3' bits into the '3' LSBs, if not then it inserts '2' bits into '2' LSBs. Flowchart for entropy based data hiding is shown in **Fig 6.** This method provides high embedding and high level of image transparency.

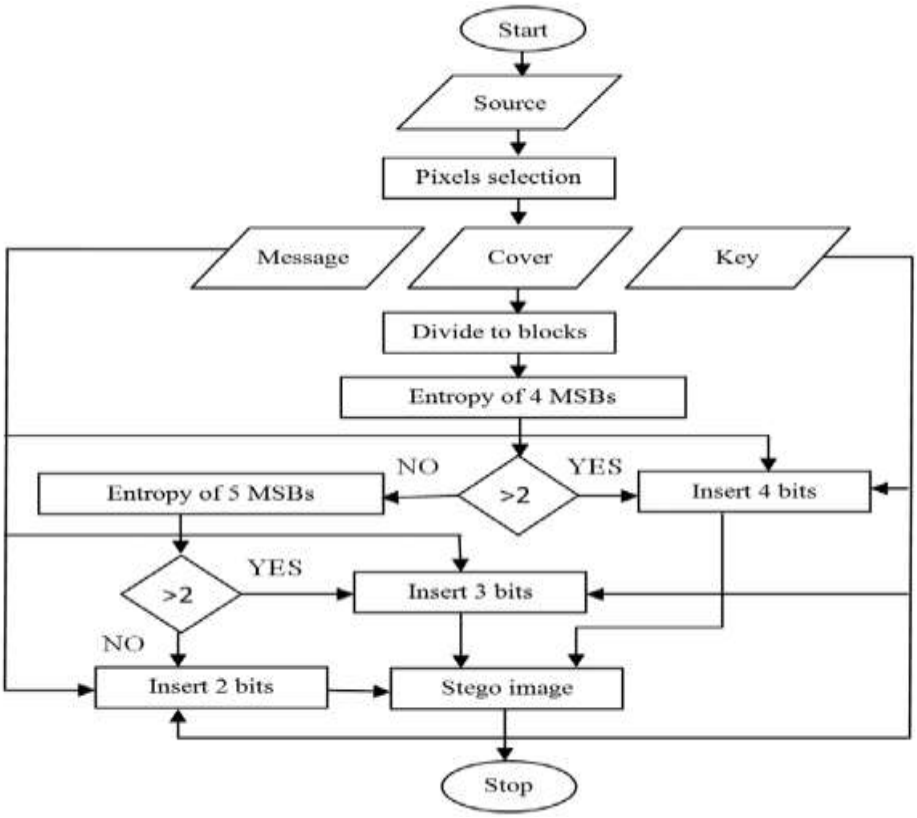


Fig 6: Entropy- based Data hiding

## IV. DRAWBACKS OF THE EXISTING SYSTEM

- Message is very much hard to recover if the image is subject to attack such as translation and rotation.
- Significant damage to picture appearance.
- Message easily lost if the picture is subject to compression such as JPEG**.**

## **V.** . OVERVIEW OF THE PROPOSED SYSTEM **VI.**

Today steganography is mostly used on computers with digital data being the carriers and networks being the high speed delivery channels. Steganography differs from cryptography in the sense that where cryptography focuses on keeping the contents of a message to be secret, whereas steganography mainly focuses on keeping the existence of a message secret. Steganography and cryptography are both ways to protect information from unwanted parties but neither technology alone is perfect and that can be compromised. When the presence of hidden information is revealed then the purpose of steganography is mostly defeated. The important strength of steganography can be amplified by combining it with cryptography. There are two technologies that are closely related to steganography are watermarking and finger printing.

In watermarking all of the instances of an object are "marked" in the usual way. The various kind of information hidden in the objects when using watermarking is usually a signature to signify origin or ownership for the purpose of copyright protection. 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 of an owner to identify customers who break their licensing agreement by supplying the property to third parties.

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 . A successful attack on a steganographic system consists of an adversary observing that there is information hidden inside a file, when a successful attack on a watermarking or fingerprinting system would not be to detect the mark, but to remove it.

### *A. ADVANTAGES OF THE PROPOSED SYSTEM*

- Original image is same as the altered image.
- Hard to detect as a fundamental image and message.
- Not susceptible to most of the attacks such as rotation and translation.

## VII. ALGORITHM AND DESIGN

AES is based on a design principle known as a substitution-permutation network, combination of both substitution along with permutation, and it is fast in both software and hardware. Unlike DES, AES does not use a Feistel network. AES is a variant of Rijndael which consist of fixed block size-128 bits, and a key size - 192, or 256 bits. By contrast, the Rijndael specification *per se* is specified with block and key sizes that may be any multiple of 32 bits, both with a maximum of 256 and a minimum of 128 bits.

The Advanced Encryption Standard (AES) is defined in each of

- FIPS PUB 197: *Advanced Encryption Standard (AES)*
- ISO/IEC 18033-3: *Information technology — Security techniques — Encryption algorithms — Part 3: Block ciphers*

*A. ARCHITECTURE DIAGRAM*

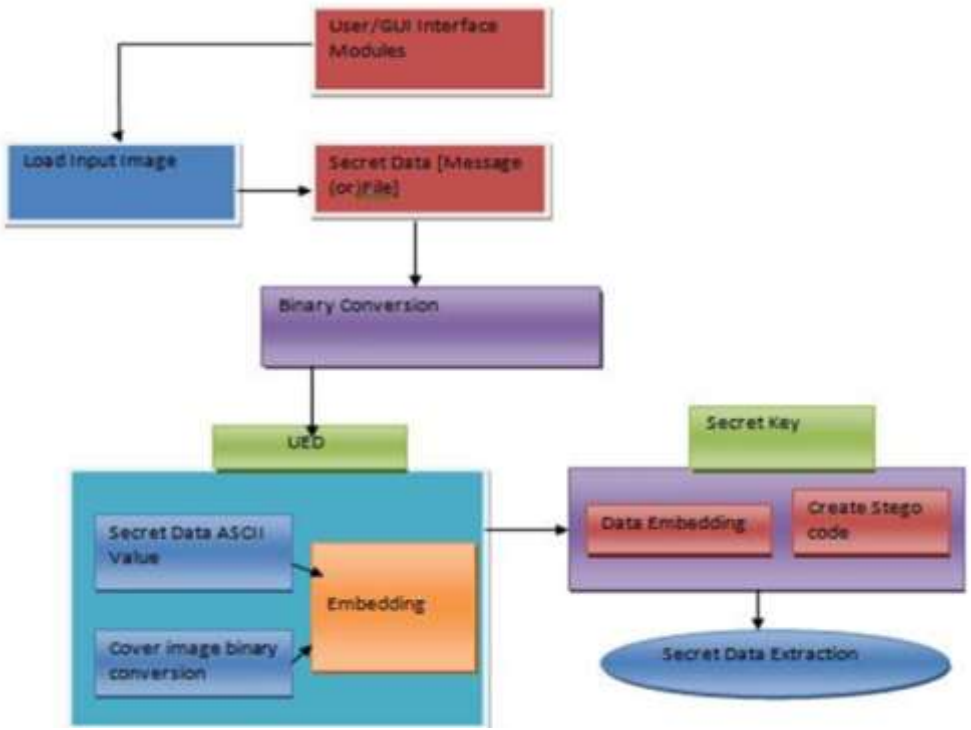


Fig 6: Entropy- based Data hiding

## VII. IMPLEMENTATION DETAILS

### *A. Embedding a message:*

This module is used to embed a message within files like image, audio and video. Here we have to specify the master file and the output file. The process of embedding information during JPEG compression results in a stego image with a high level of invisibility, though embedding takes place in the transform domain. JPEG is the most popular image file format on the Internet and the image sizes are too small because of compression, thus making it the least suspicious algorithm to use. However, the process of the compression is a very mathematical process, making it more difficult to implement. Suggested applications: The JPEG file format can be used for most applications of steganography, but is especially suitable for images that have to be communicated over an open systems environment.

*B. Embedding a file & Compression:*

Module is used to embed files like image, audio and video. Here we have to specify the master file and the output file. To compress an image into specified JPEG format, first the RGB color representation is converted into a YUV representation. In this representation the Y component indicates the luminance (or brightness) and the U and V components stand for chrominance (or color). According to research the human eye is more sensitive to changes in the brightness (luminance) of a pixel than to changes in its color. This fact is exploited by the JPEG compression by down sampling the color data to reduce the size of the file. The color components (U and V) are halved in horizontal and vertical direction, thus decreasing the file size. JPEG does this by dividing all the values in a block by using a quantization coefficient. Hence, the results are rounded to integer values and the coefficients are encoded using Huffman coding to further reduce the size.

*C. Retrieving message from a Master file:*

This module is used to retrieve a message which is embedded on a image, audio and video files. Here we have to specify password and the stegno file. It involves retrieving the embed message from the file

independent of the file format. Whenever the message has been retrieved it has to be converted into original message or file. This can be done by reading the embedded data from the master file. The read data will always be in the bytes format. Hence, this message has been converted into the suitable output file format. Embedding data, to be hidden requires two files. The first is the innocent looking image that will usually hold the hidden information, which is said to be the cover image. The second file is the message- the information to be hidden. A message may be plain text, cipher text, other images, or anything that can be embedded in a bit stream, when combined, the cover image and the embedded message make a stego- image.

### *D. Retrieving embedded file from a Master file:*

This module is used to retrieve a file which is embedded on a image, audio and video files. Here we have to specify password and the stegno file. Most steganographic software neither supports nor recommends using JPEG images, but recommends instead the use of lossless 24-bit images such as BMP. The next best alternative to 24-bit images is 256- color or gray scale images. It mostly found on the Internet are GIF files. In 8-bit color images such as GIF files, each of the pixels is represented by a single byte, and each pixel nearly points to a color index table (a palette) with 256 possible colors. The pixels value ranges between 0 and 255. The software simply paints the indicated color on the screen at the selected pixel position. Many steganography experts recommend the use of images featuring 256 shades of graph. Gray scale images are usually preferred because the shades change very gradually from byte to byte, and less the value changes between palette entries.

## VIII. *CONCLUSION*

The proposed system involves the hiding of encrypted information behind the cover image rather than hiding the original information. So although the case happens whenever the message is hidden behind the cover image the third person can able to get only the encrypted format, since the key by which the message encrypted is not known.

## REFERENCES

[1] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," IEEE Trans. Inf. Forensics Security, vol. 6, no. 3, pp. 920–935, Sep. 2011.
[2] A. Westfeld, "F5—A steganographic algorithm," in Proc. 4th Inf. Hiding Conf., vol. 2137. 2001, pp. 289–302.
[3] J. Fridrich, T. Pevný, and J. Kodovský, "Statistically undetectable JPEG steganography: Dead ends challenges, and opportunities," in Proc. 9th ACM Workshop Multimedia Security, Dallas, TX, USA, Sep. 2007,pp. 3–14.
[4] Y. Kim, Z. Duric, and D. Richards, "Modified matrix encoding technique for minimal distortion steganography," in Proc. 8th Inf. Hiding Conf.,vol. 4437. Jul. 2006, pp. 314–327.

[5] V. Sachnev, H. J. Kim, and R. Zhang, "Less detectable JPEG steganography method based on heuristic optimization and BCH syndrome coding," in Proc. 11th ACM Workshop Multimedia Security, Sep. 2009,pp. 131–140.
[6] T. Filler and J. Fridrich, "Design of adaptive steganographic schemes for digital images," Proc. SPIE, vol. 7880, p. 78800F, Jan. 2011.
[7] C. Wang and J. Ni, "An efficient JPEG steganographic scheme based on the block entropy of DCT coefficients," in Proc. IEEE ICASSP, Kyoto,Japan, Mar. 2012, pp. 1785–1788.
[8] J. Kodovský and J. Fridrich, "Calibration revisited," in Proc. 11th ACM Workshop Multimedia Security, New York, NY, USA, Sep. 2009,pp. 63–74.
[9] J. Kodovský, J. Fridrich, and V. Holub, "On dangers of overtraining steganography to incomplete cover model," in Proc. 13th ACM Workshop Multimedia Security, New York, NY, USA, Sep. 2011,pp. 69–76.
[10] V. Holub and J. Fridrich, "Digital image steganography using universal distortion," in Proc. 1st ACM Workshop Inf. Hiding Multimedia Security,2013, pp. 59–68.