Paper ID: IOTTSF26

TEXT AND AUDIO DATA HIDING USING LSB AND DCT A REVIEW APPROACH

Snehal Kurane[1], Hrutuja Harke[2], Sheetal Kulkarni[3]

Department of Electronics and Telecommunication
MMCOE, Pune, India.
Snehalkurane46@gmail.com
harkehrutuja@gmail.com
sheetal_k28@yahoo.co.in

Abstract:

Steganography is about to conceal the existence of data within a cover media such as image, text, audio and video. Basically it is a Greek word- steganos means "covering" and graphy means "writing or drawing". Steganography is not only the art of hiding data but also hiding the fact of transmission of secret data. Steganography techniques are categorized into spatial and frequency domain techniques. For hiding secret information in images, there exists variety of techniques like LSB, PVD, MLSB, ISB, DCT, DWT etc. Steganography and cryptography are two sides of a coin. As Cryptography provides privacy while steganography provides secrecy. This paper gives a brief review on the embedding methods which are LSB and DCT for steganography.

Keywords: Steganography, LSB and DCT.

I. Introduction

In today's world communication is the basic need of everyone as the communication is growing people want to secure their private as well as confidential data. As the rise of the internet and multimedia techniques has increasing interest in hiding data in digital media. Early research concentrated on watermarking to protect copyrighted multimedia products such as images, audio, video and text. Sreganography is the art of hiding the fact of transmission of secret data. Many carrier files can be used but digital images are mostly used because of their frequency on the internet. Cryptography scrambles a message so it cannot be understood and steganography hides the message so it cannot be seen.

The term steganography was first come from Greek historian Herodotus in Persian wars. The goal of steganography is to convey the message under cover image, concealing the very existence of information exchange. Image steganography is preferred because altered image with slight variations in its colors will be indistinguishable from the original image by human eye. Vital points in steganography are embedding capacity and quality of the image after embedding the hidden message.

Spatial domain steganographic methods are LSB, MSB and ISB techniques while frequency domain steganographic methods are DCT, DWT, DFT techniques.

II. Types of Steganography

Steganography is classified into four types:
1. Text Steganography
2. Video Steganography
3. Audio Steganography
4. Image steganography

1. Text steganography: Text steganography simply means information is hidden in text files. The text steganography involves anything from changing the formatting of an existing text, for changing the word within the text, to generating random sequences or using context-free grammars to generate readable texts.

2. Video steganography: Video files are generally a collection of images and sounds, therefore, most of the presented techniques on images and audio can be applied to video files. Due to fact that video is a moving stream of images and sounds, the large amount of data that can be hidden inside the video.

3. Audio steganography: Digital sounds are used to embed secret messages and this secret message is embedded by slightly altering the binary sequence of a sound file, this is known as audio steganography.

4. Image steganography: It is a process that hides the secret image behind the cover image in such a way that the presence of the secret image is locked and the cover image appears to be the same.

Steganography Technique:
Steganography consists of two terms that is message and Cover image. Message is the secret data that needs to hide and cover image is the carrier that hides the message in it.

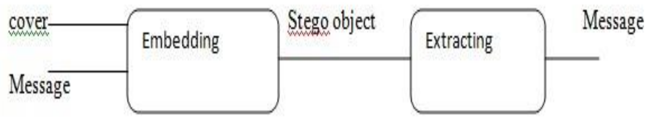


Fig 1. Steganography techniques

Block diagram:

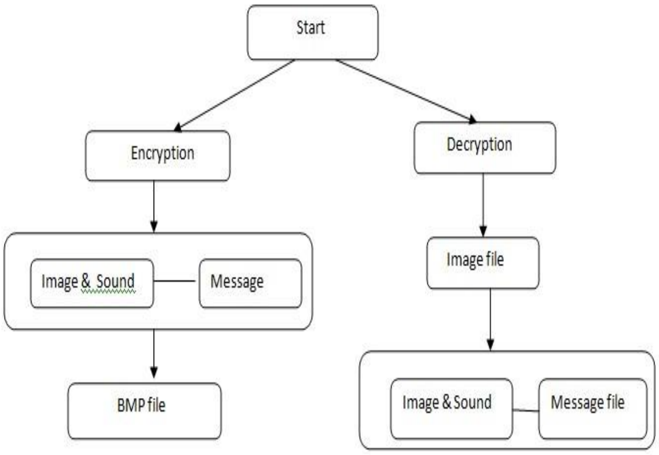


Fig 2. Block diagram of steganography [2].

III. Review on literature

Steganography is classified into two categories:
1. Spatial Domain
2. Frequency Domain

1. Steganography in Spatial Domain:

Spatial domain method offers number of techniques which are LSB, MLSB, and ISB etc. Out of which we are using LSB. We choose this technique because Vandana Ladwani proposed human eye is not that much sensitive to perceive changes in the image achieved by altering the least significant bit. Therefore an altered image with slight variations in its colors will be indistinguishable from the original by a human eye, just by looking at it [1]. Kurak proposed a technique in which one image can be hidden in another image by replacing the LSB of the cover image by MSB of hidden image [1].

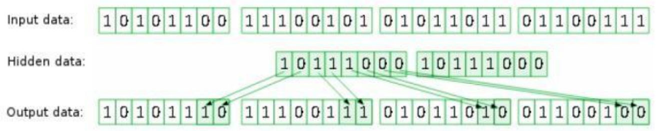

Input data: 10101100 11100101 01011011 01100111
Hidden data: 10111000 10111000
Output data: 10101110 11100111 01011010 01100100

Fig 3. LSB with one least significant bit

Ravindra Reddy, Roja Ramani, Akanksha Kaushal and Veenita Chaudhary gives the basic concept of LSB substitution is to hide the secret message behind the cover image at the LSB so that embedding procedure doesn't have any impact on the original pixel value greatly[2][7]. LSB varies according to the number of bits in an image. Ravindra Reddy, Roja Ramani gave this point nicely, for 8 bit image first the secret message should be converted into binary and then each least significant bit i.e 8th bit of that byte is replaced with the least significant bit or last bit of that 8 bit image. For 24 bit image, "colors of each component like RGB are changed" [2].
According to the Ravindra Reddy, Roja Ramani the compression in BMP is lossless as compare to JPEG, as JPEG is lossy compression technique thus LSB is very effective for BMP images [2]. LSB substitution is also possible for GIF format, but the major drawback of GIF image is whenever the LSB of message is changed, there will be some detectable changes in the colored cover images which are not feasible. So the problems can be solved by using GIF gray scale images as the gray scale images contains only 256 shades i.e it includes 8 bits and it is very hard to detect[2].

Mathematical representation of LSB [7]:

$$\mathbf{Xi' = xi - (xi \bmod 2^k) + mi} \qquad \mathbf{(1)}$$

Where,
Xi'- i th pixel value of stego image.
xi - original cover image.
mi – decimal value of the i th block in the secret data.

For reconstruction of LSB of secret message following equation if used

$$\mathbf{mi = xi \bmod 2^k} \qquad \mathbf{(2)}$$

Simple permutation of extracted mi gives original data.

Algorithm to embed the text message [4]:

1. Read the cover image and text message.
2. Convert the text message into binary.
3. Calculate LSB of each pixel of cover image.
4. Replace LSB of cover image with each bit of secret message.
5. Write stego image.

Algorithm to embed the text message:

1. Read the stego image.
2. Calculate LSB of each pixels of stego image.
3. Retrieve bits and convert each 8 bit into character.

2. Steganography in Frequency Domain:

The Fourier transform is mostly used tool which decomposes the image into its sine and cosine components. Frequency domain method offers number of techniques which are DCT, DWT and DFT, out of which we are using DCT. DCT stands for Discrete Cosine Transform; it transforms a signal or image from the spatial domain to frequency domain. Nikolaos, Constanntinos said that DCT separates the image into low, middle and high frequency coefficients [4].

Nikolaos, Constanntinos proposes that "embedding the image in a middle frequency band does not scatter the secret information to most visual important parts of the image i.e the low frequency and also it do not over expose them to removal through compression and noise attacks where high frequency components are targeted"[4].

Akanksha Kaushal described in her paper that cover image is divided into Small Square of k x k pixel commonly in square of 8 x 8 pixels to compress the cover image using DCT. "For k values, we have more compression but lower quality" [7].

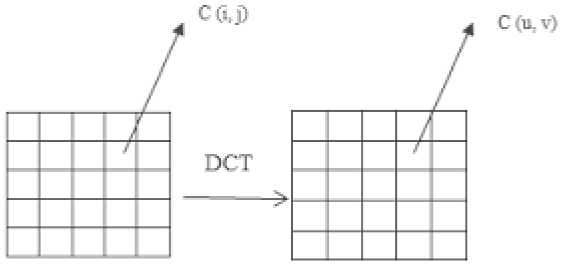


Fig 4. DCT transform

Let I(x,y) denotes an 8 bit cover image with x = 1,2,3……m and y = 1,2,3……n. This m x n cover image is divided into 8 x 8 blocks and 2 D DCT is performed on each of L = m x n /64 blocks [7].

So the equation of DCT for 2D image is

$$C(u,v) = a(u)a(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\ cos\left[\frac{(2x+1)u\pi}{2N}\right] cos\left[\frac{(2y+1)u\pi}{2N}\right]$$

For u= 0, 1……7 and v = 0, 1…...7 (3)

$$where\ \ a(k) = \begin{Bmatrix} \frac{1}{\sqrt{2}} & for\ K = 0 \\ 1 & otherwise \end{Bmatrix}$$

Mathematical equation for IDCT is

$$F(x,y) = \sum_{u=0}^{N-1}\sum_{v=0}^{N=1} C(u)C(v)f(u,v)cos\left[\frac{(2x+1)u\pi}{2N}\right]cos\left[\frac{(2y+1)u\pi}{2N}\right]$$

(4)

For having higher security, the secret information is encrypted first and encrypted ASCII value is converted in binary form [4]. Because of its advantages, such as high compression ratio, small bit error rate, good information integration ability, good synthetic ability, synthetic effect of calculations, DCT is highly preferred in many applications.

Algorithm to embed the text message [4]:

1. Read cover image.
2. Read secret message and convert it into binary.
3. The cover image is broken into 8 x 8 block of pixels.
4. Working from left to right, top to bottom then subtract 128 in each block of pixel.
5. DCT is applied to each block.
6. Each block is compressed through quantization table.
7. Calculate LSB of DC coefficient and replace with each bit of secret messages.
8. Write stego image.

To retrieve text message:

1. Read stego image.
2. Stego image is broken into 8 x 8 block of pixels.
3. Working from left to right, top to bottom then subtract 128 in each block of pixel.
4. DCT is applied to each block.
5. Each block is compressed through quantization table.
6. Calculate LSB of each DC coefficient.
7. Retrieve and convert each 8 bit into character.

3. Quantization:

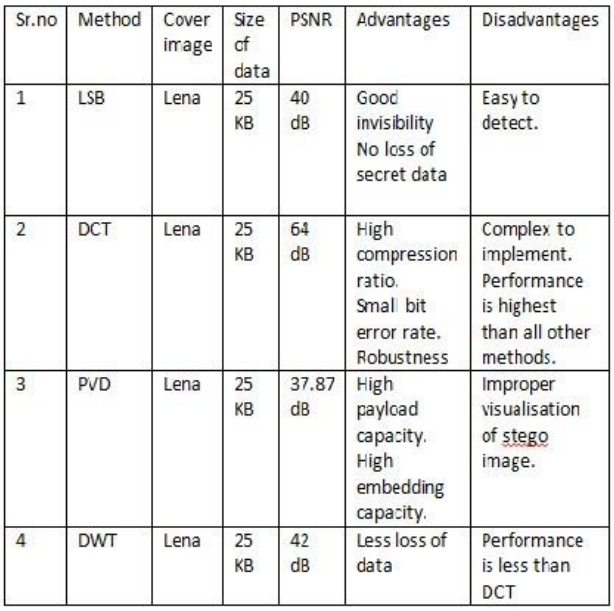
Quantization, involved in image processing, is a lossy compression technique achieved by compressing range of values to a single quantum or relatively small discrete value when the number of discrete symbols in a given stream is reduced, the stream becomes more compressible. For example, reducing the number of colors required to represent a digital image makes it possible to reduce its file size.

The quantization is achieved by having a quantization matrix as below:

$$Q = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}.$$

In this review paper, we adopt DCT and LSB embedding method for compression of image. There are number of techniques used for compression which are PVD, DWT, MLSB, and ISB out of this we are using particularly LSB and DCT based on the advantages, disadvantages and PSNR values. On the basis of all these parameters it seems like that LSB and DCT are more preferable in terms of maintaining

Secrecy and providing better image quality. Comparisons of these Methods are shown in below table.



Sr.no	Method	Cover image	Size of data	PSNR	Advantages	Disadvantages
1	LSB	Lena	25 KB	40 dB	Good invisibility No loss of secret data	Easy to detect.
2	DCT	Lena	25 KB	64 dB	High compression ratio. Small bit error rate. Robustness	Complex to implement. Performance is highest than all other methods.
3	PVD	Lena	25 KB	37.87 dB	High payload capacity. High embedding capacity.	Improper visualisation of stego image.
4	DWT	Lena	25 KB	42 dB	Less loss of data	Performance is less than DCT

Table 1. Comparison of different methods

Applications:

1. Confidential communication and secret data storing.
2. Protection of data alteration.
3. Access control system for digital content distribution.
4. Media data base system.
5. Steganography is used in Military applications.

IV. Proposed work

We are going to propose steganography using 2D images for hiding text as well as audio to attain secrecy about confidential data. For this we are using embedding and extracting algorithm based on LSB and DCT substitution method with secret key. In our approach, we are going to compress the image using DCT and MP3 stego software is for audio compression. Because of this process, the secret message is hidden within the cover media i.e the image. DCT based steganography gives higher quality in terms of PSNR rather than LSB based steganography. Using PSNR, we compare the original and stego image. At the receiver side, if the receiver knows the secret key or password, then only the receiver retrieve the code, otherwise the system will get locked and security of system will be protected.

V. Conclusion

While reading all these research papers, we understand that LSB is a technique which is very simple but easy to detect and DCT method is complex and little bit lossy but it provides Higher security than LSB. The original and stego images are compared on the basis of their PSNR values. PSNR is a measure of representing the quality of image. If the image has PSNR value less than 30 dB then the used image is of poor quality and if PSNR value is greater than 30dB then the image is adoptable to any type of compression.

References

1. Vandana M. Ladwani, Srikanta Murthy K. "A new approach to securing images", International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), Bangalore, India.' January 2015, 2319-5940.
2. Ravindra Reddy, Roja Ramani A, "The process of encoding and decoding of image steganography using LSB algorithm", IJCSET November 2012, 2231-0711.
3. Champakamala. B.S, Padmini. K, Radhika. D. K "Least Significant Bit algorithm for image steganography", IJACT 2319-7900 Department of TCE, Don Bosco Institute of Technology, Bangalore, India.
4. Constanntinos Patsakis, Nikolaos G. Aroukatos "LSB and DCT Steganographic Detection using Compressive Sensing", Journal of Information Hiding and Multimedia Signal Processing, January 2014, 2073-4213.
5. Mrs. Kavita Kadam, Ashwini Koshti, Priya Dunghav, "Steganography using Least significant Algorithm", International Journal of Engineering Research and Applications, Pune University, May-June 2012, 2248-9622.
6. R. Amirtharajan, R. Akila, P. Deepikachodavarapu, "A Comparative Analysis of Image Steganography", International Journal of Computer Applications, Tamil nadu University, May 2010, 0975-8887.
7. Akanksha Kaushal, Vineeta Chaudhary, "Secured Image Steganography Using Different Transform Domains", International Journal of Computer Applications, ECE Department Ujjain, India September 2012, 0975-8887.
8. Shahana T, "An Enhanced Security Technique For Steganography using DCT and RSA", International Journal of Advanced Research in Computer Science and Software Engineering, University of Calicut, Kerala, India, July 2013, 2277-128X.
9. Shiksha, Vidhu Kiran Dutt, "Steganography: The art of Hiding Text in Image using Matlab", International Journal of Advanced Research in Computer Science and Software Engineering, University of Hissar, India September 2014, 2277-128X.