PERFORMANCE EVALUATION OF QUANTIZED TABLE BASED DATA HIDING

Ms. Rucha R. Shriram[1], Prof. I. I. Mujawar[2]
[1]M.E. (ENTC) Nagesh Karajagi Orchid College of Engineering and Technology, Solapur
[2]H.O.D. (ENTC) Nagesh Karajagi Orchid College of Engineering and Technology, Solapur
Email:[1]shriramrucha@gmail.com, [2] isakmujawar@nkorchidenggmgmt.ac.in

Abstract— Security to data can be provided using encryption techniques or by hiding it into some cover. Data hiding provides a means for covert type of communication. Although its main purpose is hiding the data, the data can be hidden for personal use or can be used to hide secret messages and send it to the intended recipient. Performance of a data hiding technique based on quantization table is evaluated. An image is divided into 16x16 non-overlapping blocks instead of conventional 8x8 block size. Two-dimensional discrete cosine transform is applied on each block. The obtained frequency coefficients are then used to hide the secret data. Different parameters are chosen to evaluate the quality of image produced after hiding the data. Chosen data can be provided security to the extent upto which the opponent is unable of detecting the secret message by perception. The more the stego-image resembles the carrier image, the data hiding technique is more secure.

***Index Terms*—Data Hiding, Discrete Cosine Transform, Images, PSNR, MSE.**

I. INTRODUCTION

Steganography or data hiding is the art and science of covert communication. Data can be provided security by either encrypting it or by hiding it. Encryption involves changing the nature of original data while data hiding techniques maintain the very existence of the data [1]. In the encryption techniques, the data is scrambled using a particular methodology. After scrambling the data it becomes unintelligible for the opponent to read. The opponent can read the data only if he is known of the respective decryption technique. Decryption techniques are those which perform the reverse operation on the data to unscramble it. After decryption original data is obtained back. Encrypting the data provides good security but catches the attention of the opponent [2].

In the ancient time steganographic techniques were employed with a desire to hide messages or some type of information from others. The huge geoglyphs of the Nazca in Peru can be considered as a form of steganography [3]. The geoglyphs can be viewed normally. But many of the images were detected only after the geoglyphs were viewed aerially. In 5th century B.C., Histaiacus shaved the head of a messenger and wrote a secret message on his head. When the messenger's hair grew back, he was sent to deliver the message. The message was retrieved by shaving the messenger's head [4]. Another example of data hiding is the use of Cardano Grill. It is a piece of paper with holes cut in it. When the grill is placed over the text, the intended message can be viewed [5]. These were a few examples that were used in the past to deliver a secret message by hiding it.

Data hiding techniques intend to hide the data within a certain carrier. The carrier can be any of the media files such as an image file, audio file or video file [1]. Such media files are chosen as carrier since they contain redundant data. This redundant data can be manipulated to hide a secret message or any data. After choosing the appropriate carrier, the data is hidden. A stego-object is produced after the data is hidden. A stego-object is nothing but the carrier plus the data. The stego-object so obtained must look alike the original carrier object.

A general description about a data hiding system is given below. Various aspects relative to the data hiding system which define its nature are also mentioned as follows:

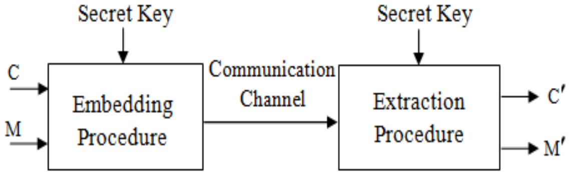

Fig I. General Data hiding process

Fig. I shows a general data hiding process. C is the carrier and M is the message to be hidden. Embedding procedure is nothing but the algorithm used for hiding the message. A secret key can be used during the embedding process. This key should be known to the sender and the intended recipient only. By using a secret key the data can be provided more security as the message can be retrieved only when the secret key is known. At the intended receiver's side extraction procedure is performed in order to retrieve the message back. C′ is the retrieved carrier and M′ is the retrieved message. More the accurate extraction algorithm, exact message is obtained.

There are three different aspects of data hiding: capacity, security and robustness. Capacity refers to the amount of information that can be hidden in the carrier. Security should be provided from an opponent's ability to detect the hidden data. Robustness refers to the amount of modification the stego-object can withstand before the opponent can destroy the hidden data. Data hiding and watermarking are relative to some extent. The primary goal of watermarking is to achieve a high level of robustness so that it should be impossible to destroy the watermark [6]. Data hiding on the other hand strives for high capacity and security. Even slight modifications to the stego-object can destroy it. Modern data hiding techniques attempt to be detectable only if secret data is known namely the key. This holds true according to the Kerchkoff's principle in cryptography, which states that a cryptographic system's security should solely rely on the key [7]. For hidden data to remain undetected, the carrier file must be kept secret. If it is revealed then a comparison between the carrier and the stego-object takes place.

II. LITERATURE SURVEY

T. Morkel, J.H.P. Eloff and M.S. Olivier mentioned different styles of data hiding as to which digital files can be used to hide the data efficiently. Various digital files such as images, audio and video files can be used to hide the data [8]. C. Chan and L. M. Cheng gave a data hiding technique by simple LSB substitution. The effect of substituting k-LSB bits of the carrier image is also stated [9]. G. Liu, Z. Zhang and Y. Dai presented an LSB matching method for image data hiding [10]. H. Kobayashi, Y. Noguchi, and H. Kiya proposed a data hiding technique in frequency domain using the discrete cosine transform. [11]. C. C. Chang, T. S. Chen and L. Z. Chung proposed a method using DCT to hide the data in images. A modified quantization table different from the default JPEG quantization table was used to reduce the distortion in the reconstructed image [12]. H. W. Tseng and C. C. Chang embedded the secret data into compressed JPEG images. The high frequency coefficients were used to hide the data rather than low and mid-frequency coefficients [13]. Another frequency domain data hiding technique includes the use of wavelet transform as stated by A. Al-Ataby and F. Al-Naima [14]. After the data is hidden a stego-image is obtained which should look alike the cover image. To evaluate the image quality of the stego-image various parameters such as Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR), Structural Content (SC), Maximum Hiding Capacity (MHC) and Time required [15].

Section I gives the introduction of data hiding. The literature survey is included in the Section II. Detailed description of the various data hiding systems is given in Section III. Proposed work is explained in Section IV. Part A of Section IV states the embedding and extraction procedure of data hiding. Part B of Section IV lists the evaluation parameters. Section V states the result and conclusion.

III. IMAGE DATA HIDING

Images can be used to hide the data. They contain more redundant data than the text files which can be manipulated to hide a secret message. Data is embedded in a digital image using an embedding algorithm. A secret key can be used in the embedding procedure to increase the security. After hiding the data into the carrier image, the stego-image should be identical to the carrier image so that the opponent should not notice the existence of any hidden data.

Image data hiding can take place in two domains broadly: Spatial domain and Transform domain. In the spatial domain, redundant bits of the carrier image are manipulated to hide the data. Spatial domain data hiding techniques are comparatively simple to implement. The image formats most suitable for spatial domain data hiding techniques are lossless. Few spatial domain techniques are explained below:

A. Least significant bit (LSB) substitution

It is a simple LSB substitution method. An n-bit secret message is embedded into the k-rightmost LSBs of the carrier image. Firstly secret message is rearranged to form a conceptually k-bit virtual image. Then a subset of pixels from the carrier is chosen. The embedding procedure is completed by replacing the k-LSB bits of the carrier image [9].

TABLE I.
WORST PSNR FOR K = 1-5 BY SIMPLE SUBSTITUTION

k	1	2	3	4	5
PSNR (dB)	48.13	38.59	31.23	24.61	18.30

Table I tabulates the worst PSNR for some k = 1-5. It can be seen that the image quality of the stego-image is degraded drastically when k ≥ 4 [9]. PSNR values falling below 30dB indicate a fairly low quality (i.e., distortion caused by embedding can be obvious); however, a high quality stego-object should strive for 40dB and above.

B. Least significant bit (LSB) matching

In LSB matching method, if the LSB bit of the carrier image is similar to the LSB bit of the message then the LSB bit of the carrier is not changed. If the message bit is different, then the carrier image bit is substituted. LSB matching provides small amount of changes in the pixel values maintaining the perception quality of the stego-image [10].

In the transform domain data hiding, the image is first transformed from spatial domain into frequency domain. These techniques hide the data into significant areas of the carrier image. This makes it more robust than spatial domain data hiding. Many transform techniques do not depend upon the image format . Thus, the hidden message might not get lost after operations like compression, cropping, rotation, resizing etc.

C. Discrete cosine transform (DCT)

One of the transform domain techniques use the discrete cosine transform. The DCT transforms spatial domain representation of an image into frequency domain. A grouping of 8x8 pixel blocks after taking DCT on them provide 64 DCT coefficients each. These coefficients are quantized. The quantized coefficients can be used to hide the data. After the embedding procedure is over the obtained image is further passed through a lossless compression technique and the stego-image is obtained.

Kobayashi's method embedded only one secret data bit into one 8x8 DCT block. The embedded one bit binary data is replaced with the k^{th} quantized DCT coefficient through zigzag scanning. Secret data bit was embedded into high frequency components rather than low and mid frequency components. Advantage of this method is, the high frequency components often become zeroes after quantization so there is no need of changing the values of coefficients if the data to be hidden is zero. Another advantage is that the high frequency components are more visually resistant to noises than low frequency components [11].

Jpeg-Jsteg is a data hiding tool based on JPEG. In Jpeg-Jsteg the secret messages are embedded n LSB of quantized DCT coefficients whose values are not 0, 1 or -1. The image is divided into non-overlapping blocks of 8x8 pixels and DCT is applied on each block to transform each block into DCT coefficients. The DCT coefficients are scaled according to the quantization table. The secret message is encrypted and then embedded into quantized DCT coefficients. After embedding the secret message, Jpeg-Jsteg uses Huffman coding, Run Length coding and Differential Pulse Code Modulation of JPEG entropy coding to compress each block. The disadvantage of this method is its limited hiding capacity. Jpeg-Jsteg modifies the quantized DCT coefficients in the low frequency part. Therefore, the image quality of Jpeg-Jsteg is degraded, especially when cover image undergoes a high compression ratio [6].

Limitations of Jpeg-Jsteg were overcome by Chang's method wherein the image is divided into non-overlapping blocks of 8x8 pixels and DCT is applied on each block. The DCT coefficients are scaled using a modified quantization table which is different from the default JPEG quantization table. A modified quantization table was used to avoid distortion in the reconstructed image. The secret data was hidden in the middle frequency part of the quantized DCT coefficient. Two least significant bits (LSB) were replaced by the secret data bits instead of one LSB. This increased the hiding capacity. After embedding the secret data, each block was compressed using JPEG entropy encoding [12].

Tseng proposed a method wherein the secret data was embedded into a compressed JPEG image rather than into the quantized DCT coefficients as mentioned in [6], [11] and [12]. Entropy decoding is applied to the JPEG compressed image. Then the secret data bit is embedded into the last ac coefficient. This method provided acceptable image quality and adjustable embedding capacity [13].

D. Discrete wavelet transform (DWT)

This method pre-adjusts the original cover image in order to guarantee that the reconstructed pixels from the embedded coefficients would not exceed its maximum value and hence the message will be correctly recovered. Then, it uses wavelet transform to transform both the cover image and the hidden message. Wavelet transform allows perfect embedding of the hidden message and reconstruction of the original image. DWT as expected due to the ability of wavelet transform to compress data and introducing sparsity, hence increasing the capacity or payload of the data hiding process. Wavelet transform has been used extensively in the last few years in the image processing field, ranging from noise suppression or de-noising to image coding and compression [14].

IV. PROPOSED WORK

Part A states the description of proposed quantized table based data hiding algorithm. Part B states the parameters required for performance evaluation of the data hiding algorithm.

A. Quantized Table based Data Hiding

Embedding procedure for hiding data is shown in fig. II.

Input: A cover image C and a message M.

Output: A stego-image S.

1. Choose a cover image of size NxN pixels.
2. Partition the cover image into non-overlapping blocks $[C_1, C_2, C_3, \ldots\ldots C_{N/16xN/16}]$ .Each C_i contains 16x16 pixels.
3. Use DCT to transform each block Ci into DCT coefficient matrix X_i where $X_i[a,b] = DCT(C_i[a,b])$, where $1 \le [a,b] \le 16$ and $X_i[a,b]$ is the pixel value in C_i.
4. Use modified 16x16 quantization table Q to quantize each X_i and the result after quantization can be given by $Y_i[a,b] = truncate(X_i[a,b]/P[a,b])$.
5. Encrypt message M to be hidden. The resultant message after encryption is represented by $\overline{L} = [L_1, L_2, L_3 .....L_m]$ where L_i is a secret message bit and m is the length of $\overline{L}$.
6. Select $Y_i[a,b]$ to hide $\overline{L}$ respectively where [a,b] corresponds to upper left quantized DCT coefficients.
7. Entropy coding is applied to compress each block $Y_i[a,b]$.
8. All the blocks $Y_i[a,b]$ after step 7 are combined to generate the stego-image S.

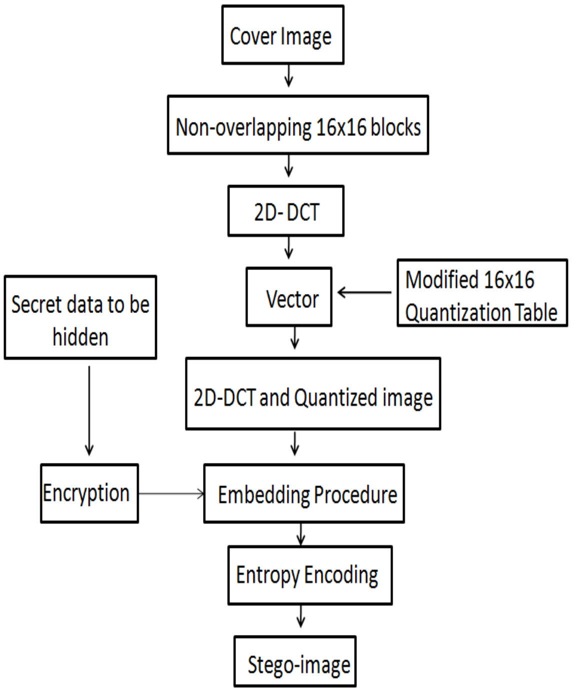

Fig. II. Embedding procedure

Extraction procedure for retrieving the secret data is shown in fig. III.

Input: Stego-image S.

Output: Hidden message M.

1. Entropy decoding is performed on the stego-image S first to decompress it.
2. Extraction procedure is performed to retrieve the encrypted secret message bits and reconstruct the original secret message M.
3. Further to obtain the cover image back, dequantize and apply IDCT.

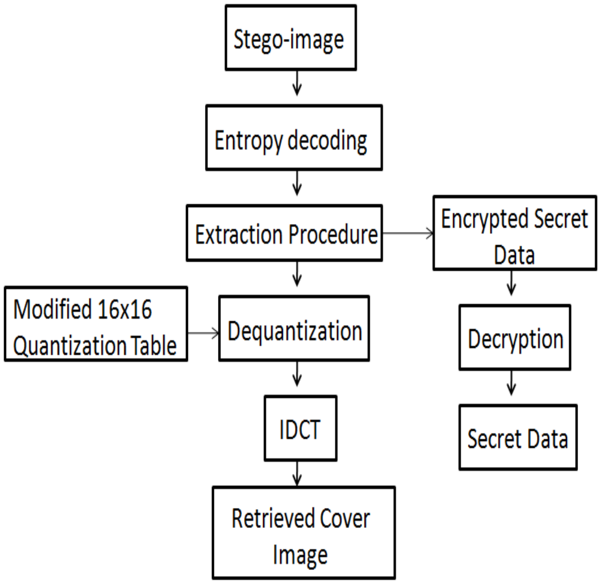

Fig. III. Extracting procedure

B. Evaluation Parameters

The stego-object obtained should look alike the original cover image. The parameters mentioned below are used to evaluate the performance of the data hiding algorithm:

1. Mean Square Error (MSE)

It measures a difference between two images and the result is a degree of similarity or strength of the error signal between the two images. MSE increases as the compression ratio increases [15]. It is used mostly due to its simplicity in calculation. MSE is defined by equation (4.1)

$$MSE = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}[C(i,j)-S(i.j)]^2$$

(4.1)

where C(i,j) is the intensity value of the pixel in the cover image, S(i,j) is is the intensity value of the pixel in the stego-image and MxN is the dimension of the image.

2. Peak Signal to Noise Ratio (PSNR)

PSNR is a measure of peak error between the distorted or compressed image and the original image. It can also be defined as the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation [15].

$$PSNR = 10\log_2 \frac{(M)^2}{(MSE)} \qquad (4.2)$$

where M is the maximum intensity of the pixel and MSE is the mean square error. Higher the PSNR more the stego-image resembles the cover image [15].

3. Structural content (SC)

It estimates the similarity between the structures of two images that is, between the cover image and the stego-image. If structural content is spread at 1, then the stego-image is of better quality. Large value of structural content means that, the stego-image does not resemble closely to the cover image [15].

$$SC = \sum_{i=1}^{M}\sum_{j=1}^{N}[C(i,j)]^2 / \sum_{i=1}^{M}\sum_{j=1}^{N}[S(i,j)]^2 \qquad (4.3)$$

where C(i,j) is the cover image and S(i,j) is the stego-image both having dimensions MxN.

4. Maximum Hiding Capacity (MHC)

Maximum hiding capacity refers to the number of bits of the cover image that can be used to hide the data without any drastic changes in the perception quality of the image.

$$MHC(bits) = \frac{PxCxXxY}{MxN} \quad (4.4)$$

where P is the number of LSB bits substituted of each ac coefficient of the pixel, C is number of ac coefficients, XxY is size of image and MxN is the block size chosen to divide the image into number of blocks.

5. *Time Required*

This parameter is evaluated to determine the time taken by the algorithm for hiding the data. If the block size chosen is less then time required for algorithm execution is less. If the block size is on higher side then the time required for processing each block is more.

V. RESULTS AND CONCLUSION

TABLE II.
HIDING CAPACITY OF VARIOUS METHODS

Method	Image Size	Number of LSBs replaced	Hiding capacity (bits)
Kobayashi	512x512	1	4096
Jpeg-Jsteg	512x512	2	51094
Chang	512x512	2	212992
Expected Proposed	512x512	2	278528

As seen from table II., the size of cover image chosen is 512x512. Kobayashi's method places one LSB from the cover image to hide the secret data bit and cover image is divided into blocks of 8x8 pixels. The hiding capacity using eq. (4.4) results to 4096 bits. Similarly the data hiding tool Jpeg-Jsteg replaces two LSBs from the cover image resulting to a hiding capacity of 51094 bits. Further the hiding capacity increases by using Chang's method as the number of coefficients available for hiding the data are more than Jpeg-Jsteg's method. The hiding capacity of the proposed method is more than rest all methods since the coefficients available for hiding the data are more. Discrete Cosine Transform is chosen as its characteristic of good energy compaction is advantageous. A 512x512 image is divided into blocks of size 16x16 pixels. Using a 16x16 modified quantization gives 136 coefficients from each 16x16 block. Two LSBs are replaced to hide the secret data of each quantized coefficient resulting to an expected hiding capacity of 278528 bits.

VI. REFERENCES

[1] J. Fridrich and M. Goljan, “Practical Steganalysis—State of the Art,” Proc. SPIE Photonics Imaging 2002, Security and Watermarking of Multimedia Contents, vol. 4675, SPIE Press, 2002, pp. 1–13.

[2] F.A.P. Petitcolas, R.J. Anderson, and M.G. Kuhn, “Information Hiding—A Survey”, Proc. IEEE, vol. 87, no.7, 1999, pp. 1062–1078.

[3] J. C. Judge, “Steganography: Past, Present, Future”, Information Security Reading Room, SANS Institute 2001.

[4] M. Conway, “Code Wars: Steganography, Signals Intelligence, and Terrorism”, Working Papers in International Studies Centre for International Studies Dublin City University, Working paper 6 of 2008.

[5] “Classical Steganography, Cardano Grille”, URL: http://library.thinkquest.org/27993/crypto/steg/classic1.shtml.

[6] N. Provos and P. Honeyman, “Hide and Seek: An Introduction to Steganography”, IEEE Security & Privacy, 1540-7993/03/$17.00, 2003.

[7] A. Kerckhoffs, “La Cryptographie Militaire (Military Cryptography),” J. Sciences Militaires (J. Military Science, in French), Feb. 1883.

[8] T. Morkel, J.H.P. Eloff and M.S. Olivier, “An Overview of Image Steganography”, in Proceedings of the Fifth Annual Information Security South Africa Conference (ISSA2005), Sandton, South Africa, June/July 2005.

[10] C. Chan and L. M. Cheng, “Hiding data in images by simple LSB substitution”, Pattern Recognition, pp. 469-474, Mar. 2004.

[11] G. Liu, Z. Zhang and Y. Dai, “Improved LSB-matching Steganography for Preserving Second-order Statistics”, Journal Of Multimedia, Vol. 5, No. 5, October 2010.

[11] H. Kobayashi, Y. Noguchi, and H. Kiya, “A method of extracting embedded binary data

from JPEG bitstreams using standard JPEG decoder", Proceedings of IEEE International Conference on Image Processing, Vancouver, BC, Canada, Vol. 1. pp. 577–580, 10–13 Sept. 2000.

[12] C. C. Chang, T. S. Chen, L. Z. Chung, "A steganographic method based upon JPEG and quantization table modification", Information Sciences, Vol.141, pp.123-138, 2002.

[13] H. W. Tseng, C. C. Chang, "Steganography using JPEG-compressed images", The Fourth International Conference on Computer and Information Technology, Wuhan: IEEE Computer Society Press, pp.12-17, 2004.

[14] Ali Al-Ataby and Fawzi Al-Naima, "A Modified High Capacity Image Steganography Technique Based on Wavelet Transform", The International Arab Journal of Information Technology, Vol. 7, No. 4, October 2010.

[15] M. Ahmet, Eskicioglu and Paul and S. Fisher, "Image Quality Measures and Their Performance", IEEE Transactions On Communications, Vol. 43, No. 12, December 1995.

[16] S. Venkatraman, A. Abraham and M. Paprzycki, "Significance of Steganography on Data Security", Proceedings of the International Conference on Information Technology: Coding and Computing, 2004.