Image Steganography Using Entropy Thresholding Technique

Vivek Ugle[#1], Abdulgani Tamboli[#2], Vaibhav Bansode [#3]
#Department of Electronics and Communication Engineering, Dr. Babasaheb Ambedkar Technological University Lonere

ABSTRACT-Nowadays, many applications are based on the Internet and in some cases it is desired that the communication be secret. The two most important aspects of any steganographic system based on images are the imperceptibility and the capacity of the stego image. In this document, a new steganographic image method is proposed that uses the block level entropy threshold technique. This document evaluates the performance and efficiency of the use of optimized embedding tables within JPEG steganography. After dividing the cover image into non-superimposed 8X8 blocks, the DCT (discrete cosine transform) is calculated and based on the entropy threshold scheme (ET), these blocks are selected for the decision to incorporate information. The secret message is hidden in the valid Entropy Block of the cover image with its average frequency of the DCT coefficients. Finally, a JPEG image is generated. The steganography scheme based on DCT provides greater resistance to image processing attacks, such as JPEG compression, noise, rotation, translation, etc. We obtain that the proposed method has a greater message capacity and that the quality of the stego images of the proposed method is acceptable.

Keywords- **Steganography, Discrete Cosine Transform (DCT), data concealment, entropy threshold, MSSIM**.

I) INTRODUCTION

Steganography is the art and science of invisible communication. This is achieved by hiding the information in other information, thus hiding the existence of the information communicated. In steganography of images, information is hidden only in images. Its main purpose is to hide the occurrence of communication through a public channel. Contrary to cryptography, steganography tends to hide the very existence of the message or any form of communication, while cryptography tries to hide the content of the secret message. The aspect of communication can be hidden by incorporating a secret message in an innocent medium, such as an image, which nobody, except the sender and the recipient, can suspect.

An information hiding system uses two algorithms to communicate: an embedding algorithm to produce the modified cover data that results after the embedding of the secret message and an extraction algorithm to retrieve the image message stego. The digital image containing perceptually irrelevant or redundant information can be used as a cover or messenger to hide secret messages. After embedding the secret message in the image on the cover, you get the image called stego. The hidden data can be of various uses, such as copyright information, titles, time stamps or subtitles of movies, etc. There are two types of steganographic image techniques: spatial domain and Methods based on frequency domain. The schemes of the first type directly incorporate the secret data inside the pixels of the cover image, such as the insertion of the least significant bit (LSB) [1]. The schemes of the second type incorporate the secret data within the cover image that has been transformed, for example DCT (discrete cosine transformation). The DCT coefficients of the transformed cover image are modified according to the secret data [2]. The capacity of the spatial domain schema (the amount of data embedded within a given image) is better than that of the transformation-based schema. However, patterns of frequency domains are more robust than those of spatial domain [4]. In this document, data is entered directly into the image and not into the header. The embedded data are self-detectable. The modified image has a good resemblance to the original image. In JPEG compression, the image is divided into separate blocks of 8x8 pixels, a two-dimensional DCT is applied to each block and then the DCT coefficients of these blocks are encoded. Most steganography techniques used for JPEG images use standard JPEG compression. The cover image is divided into 8x8 pixel non-overlapping blocks to perform DCT and provide compressed images [4] and [6]. Note that the secret image is incorporated into the part of the average frequency of the DCT coefficients of the entropy block valid in our method.

The rest of this paper is organized as follows. Section II will examine various data integration techniques. Section III will propose our data concealment JPEG method. Section IV and Section V propose respectively Data Integration Algorithm and Extraction Algorithm. Section VI will list the experimental results. Finally, the conclusions will be presented in Section VII. The stochaographic approaches to hide data are divided into two types:

1)Spatial / time domain Spatial domain techniques directly embed messages in pixel intensity. The least significant bit (LSB) is the first technique of steganography of the most famous and easy spatial domain. It incorporates the bits of a message in sequence into the LSB of the image pixels, but the problem with this technique is that if the image is compressed, the embedded data can be destroyed [1]. Therefore, there is the fear of damage to the message that could have confidential information. Furthermore, these types of methods are easy to attack using steganography techniques.

2) Transform / frequency domain In this domain, the incorporation of the data is done in the transformation domain with a set of transformation coefficients in the medium frequency bands, since they are better preserved under compression attacks than the high frequency coefficients [7] The transformation domain technique behaves well against attacks such as compression, clipping, etc. And it is imperceptible to the human sensory system. That's why it's more undetectable. In this document, the entropy property of an image block is considered for the decision to embed data. To robustly hide large volumes of data in images without causing significant perceptual impairment, concealment techniques must be adapted to local characteristics within an image. The entropy threshold (ET) method at the block level decides whether or not to incorporate data in each block

(typically 8X8) of the transformation coefficients, depending on the entropy within that block. If a particular block fails this test, we keep it as it is and incorporate the same data into the next block that passes the test.

II. DATA EMBEDDING TECHNIQUES

The stochaographic approaches to data concealment are divided into two types: 1) spatial / temporal domain and 2) transformation domain. The transformation domain technique works well against attacks such as compression, cropping, etc. and it is imperceptible to the human sensory system. That's why it's more undetectable. Data integration is performed in the transformation domain with a set of transformation coefficients in the medium frequency bands, since they are better preserved in the compression attacks than the high frequency coefficients [7]. In this document, the entropy property of an image block is considered for the decision to embed data. To robustly hide large volumes of data in images without causing significant perceptual deterioration, cloaking techniques must be adapted to local characteristics within an image. It is believed that two methods apply local criteria that determine a large volume of hidden data. The first is the entropy threshold (ET) threshold method that decides whether or not to incorporate the data in each block (usually 8X8) of the transformation coefficients as a function of the entropy within that block. If a particular block fails this test, we keep it as it is and incorporate the same data into the next block that passes the test. The second is the method of inclusion of the selective coefficient , which decides whether or not to incorporate data based on the magnitude of the coefficient.

III. THE PROPOSED METHOD

The media files are large and consume a lot of disk space. So, to reduce the size of the image file, the compression technique is used. Compression works by eliminating redundant data that effectively summarizes the contents of a file to preserve the greatest possible original meaning. For these, there are several image file formats. Of all these file formats, the JPEG file format is the most popular on the Internet due to the reduced image size. Tseng and Chang have proposed a new steganographic method based on JPEG. DCT has been applied for each 8x8 pixel block to improve capacity and control compression ratio [8]. Chang et al. developed a steganographic method based on JPEG and modified the 8x8 quantization table to improve the concealment ability of the Jpeg-Jsteg method. They used the average frequency band for insertion in order to obtain a better concealment capability and an acceptable stego image quality [4]. Since the energy of the image is concentrated in the lower frequency coefficients, the modification of these coefficients can cause a degradation of the quality of the output image. However, the high frequency coefficients will be discarded due to the quantization process. However, the high frequency coefficients will be discarded due to the quantification process [Figure 1]

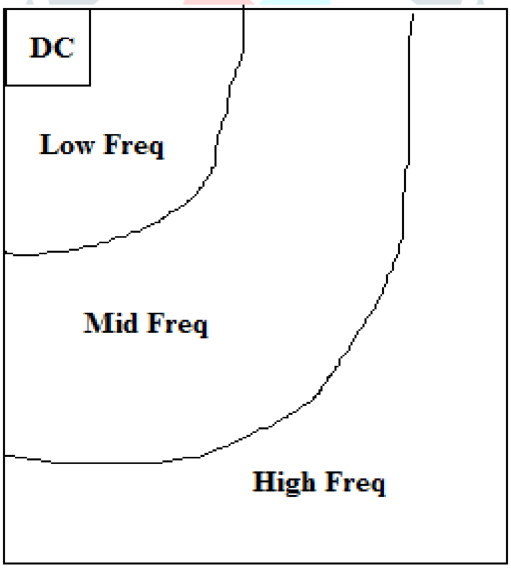


Figure-1 Frequency Distribution in a DCT Block

Figure 2 shows the block diagram of the proposed steganographic model. The proposed method uses the method of preprocessing the JPEG image in the cover image. Partitioning a cover image C in non-overlapping blocks of 8x8 pixels, then we use DCT to transform each block into DCT coefficients.

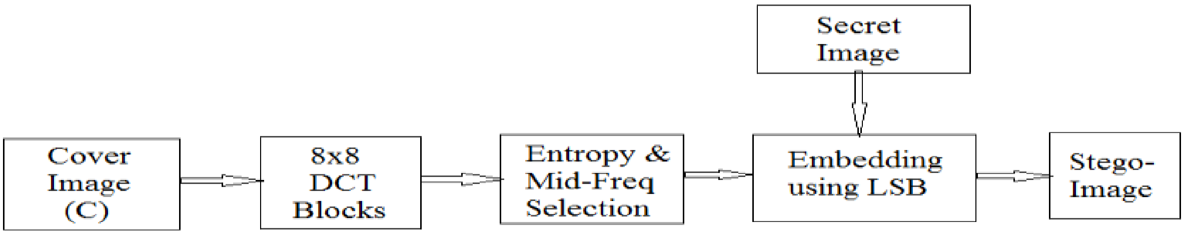


Figure-2.Block diagram of Proposed Embedding Model

To find the embedding region, first calculate the entropy (E) of all 8x8 DCT blocks. The entropy formula is shown in Eq. 1. E = | Xi, j | ^ 2i, j (i, j) ≠ (0,0) - (1) Then calculate the mean entropy (ME) of all block entropy. Now, to obtain the valid block (VB), compare the entropy of each block with the mean entropy (ME). To embed information, we select the block that has an entropy greater than the average entropy value and is called a valid block. Another step is insertion using the matrix K. The average coefficients of K Matrix are those of one and the others are zero, this is because our secret message will be integrated into the average frequency part of the quantized DCT coefficients. In the matrix k there are 26 coefficients located in the central part that are configured to be one.

$$K=\begin{bmatrix} 16 & 11 & 10 & 16 & 1 & 1 & 1 & 1 \\ 12 & 12 & 14 & 1 & 1 & 1 & 1 & 55 \\ 14 & 12 & 1 & 1 & 1 & 1 & 69 & 56 \\ 14 & 1 & 1 & 1 & 1 & 87 & 80 & 62 \\ 1 & 1 & 1 & 1 & 68 & 109 & 103 & 77 \\ 1 & 1 & 1 & 64 & 81 & 104 & 113 & 92 \\ 1 & 1 & 78 & 87 & 103 & 121 & 120 & 101 \\ 1 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

Figure-3.Quantization Matrix K

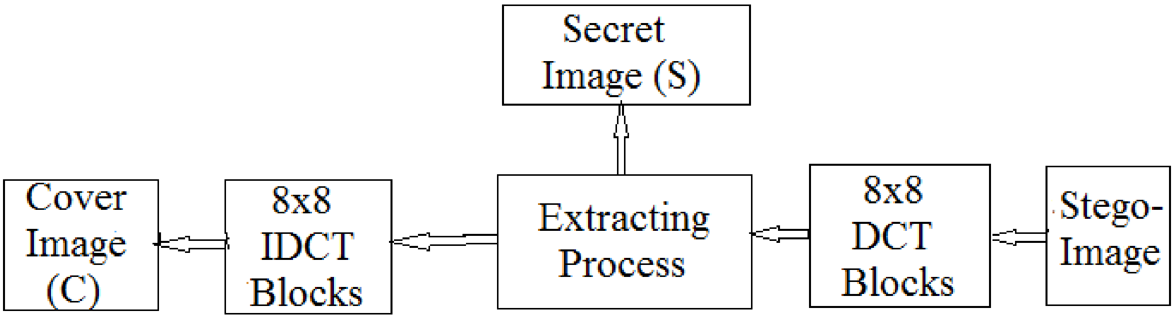
Here K [0.4], K [0.5], K [0.6], K [0.7], K [1.3], K [1.4], K [1.5], K [1,6], K [2,2], K [2,3], K [2,4], K [2,5], K [3,1], K [3,2], K [3,3], K [3,4], K [4,0], K [4,1], K [4,2], K [4,3], K [5,0], K [5,1], K [5,2], K [6,0], K [6,1] y K [7,0]. Here K [a, b] is the value of the row and the element of the column of K [4]. the secret image S will be integrated into the medium frequency part of the DCT coefficients for the valid block Vb. After embedding the secret image, the IDCT is taken to get a Stego image that will look like the image on the cover. The procedure for embedding and extracting secret messages in a cover image for a JPEG-based steganographic approach is described below:

IV. DATA EMBEDDING ALGORITHM

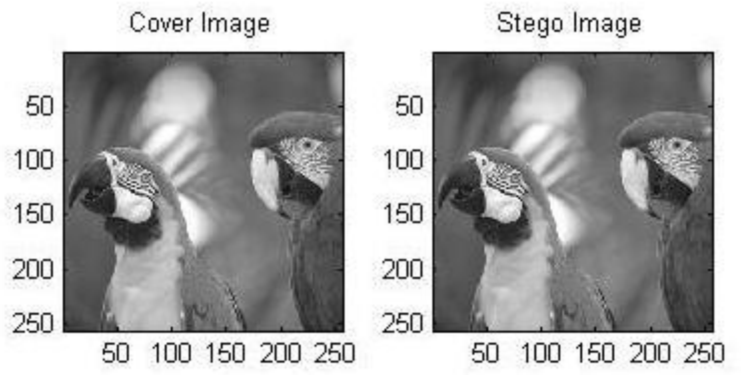
1) First you need to embed the incoming secret image file. 2) The RGB color representation if it first becomes a YUV representation. This step is mandatory if the image on the cover is colored. For JPEG, if you use DCT, however, it can be used to transform the similarities, for example, the DFT (discrete Fourier transform). 3) Read the cover image C (i, j) and divide it into separate blocks of 8 × 8 pixels. Each block indicates B. Now, for each block B (8, 8) in the interval (0-255), it calculates the discrete transformation cosine (DCT), producing 64 DCT coefficients. Both the DCT coefficient of the pth block as Cp (i), 0 ≤ i ≤ 63, p = 1 ... T, where "T" is the total number of blocks in the image. Then click on the CC coefficient of each zero block, which is the first coefficient of each block of 8 x 8 DCT C (1,1). 4) Calculate the entropy (E) of each block of 8 X 8 as, E = Σ || Xij || 2 (i, j) ≠ (0,0) 5) Calculates the mean entropy (ME) of all the blocks in T. Now, to obtain the valid block (VB), it compares the entropy of each block with the mean entropy (ME). To embed information, we select the block that has an entropy greater than the average entropy value and is called a valid block [8]. 6) A further step consists in incorporating using the K matrix as follows, Within each valid block, we chose the coefficients as, select VB [a, b] to hide the Secrets S data, respectively, where [a, b] is equal to [0,4], [0,5] , [0, 6], [0.7], [1,3], [1,4], [1,5], [1,6], [2,2], [2,3], [2,4], [2,5], [3,1], [3,2], [3,3], [3,4], [4,0], [4,1], [4, 2], [4, 3], [5,0], [5,1], [5,2], [6,0], [6,1] and [7,0], respectively. Each VB [a, b] incorporates a secret bit using the LSB technique inside it. 7) Calculate the transformation of the reverse discrete cosine (IDCT) of all the blocks. Blocks that are not used for embedding are taken as is in the stego image. At the end, add the CC coefficient of each block to the respective block and organize all the blocks to get the image, which is called image stego Stego (i, j).

V. DATA EXTRACTION ALGORITHM

When the secret data is extracted, the previous embedding steps are repeated for the stego image. Extract 26 bits of average frequency information (BI) of the valid coefficient of all valid blocks in the zigzag order of the stego image. So we order the extracted BI. Finally decode the image information and acquire the image file of secret input S.



VI. SIMULATION RESULTS



An experiment was conducted to evaluate the efficiency of our method. A 256x256 pixel cover image was used. To measure the quality of image encoding and compression, the criteria of the peak-to-noise signal (PSNR) and the mean square error (MSE) criteria were used. The PSNR for the proposed algorithm was 61.15. Furthermore, the improvement in robustness provided by this algorithm was considerably high.

VII. CONCLUSION

In this review, the secret message is incorporated into the average frequency part of the DCT coefficients. The analysis based on DCT steganography was performed on the basis of parameters such as PSNR. Grayscale images were used for experiments. The peak signal-to-noise ratio is used to calculate the performance of the methods. PSNR calculates the peak-to-noise ratio, in decibels, between two images. This relationship is used as a measure of quality between two images. If the PSNR report is high, the images are of better quality. Our experimental results show that the proposed method provides an acceptable image quality and an extremely robust system with a high security system.

REFERENCES

[1] Wang, R.Z., C.F. Lin and J.C. Lin. Image hiding by optimal substitution of LSB and genetic algorithm. Recognition of patterns, 34, pp. 671-683, 2001.
[2] Y. K. Lee and L.-H. Chen, "Steganographic model of high-capacity images", in IEE procedures on vision, image and signal processing, June, 147 (3), 2000, p. 288-294.
[3] Adel Almohammad, Gheorghita Ghinea, Robert M. Hierons. JPEG steganography: an assessment of the performance of the quantification tables International conference on networks and applications of advanced information 2009.
[4] Chang C.C., T.S. Chen and L.Z. Chung A steganographic method based on JPEG and modification of the definition table. Information Sciences, 141, pp. 123-138, 2002.
[5] ISO DIS 10918-1 "Including Digital and Continuous-Tone Still Image Coding (JPEG)", RECOMMENDATION T.81 of the CCITT.
[6] Q. Li, C. Yu and D. Chu, "A method for understanding the bases in the sign insignia and diffuse classification", The World Congress on Control and Intelligent Automation, 2006. WCICA 2006, 21 - 23.2006 de June, pp.10050-10053.
[7] Mansi Subhedar, Gajanan Birajdar, "Entropy thresholding technique at the block level for the concealment of adaptive data from high volume images" ICGST-GVIP Magazine, Volume 11, Number 3, June 2011.
[8] H.W.Tseng and C.C.Chang. Steganography using compressed JPEG images. The Fourth International Conference on Informatics and Information Technology, ILC '04, 14-16 September, pp: 19-17, 2004.