Data Security through more Robust DCT based Methodology and Assessment in terms of PSNR

Ajit Singh[1], Swati Aggarwal[2]

[1]BPS University, School of Engineering and Sciences, Khanpur Kalan, Sonepat, Haryana, India

[2]School of Engineering and Science, BPS University, Khanpur Kalan, Sonepat, Haryana, India

Abstract: ***This paper presents a new frequency domain approach for steganography algorithm.The proposed system proved to be more robust than the existing systems and calculates its PSNR, MSE, CER and Correlation coefficient.This paper also analyse the Least Significant Bit (LSB)spatial domain based Steganography and Discrete Cosine Transform (DCT)frequency domain based Steganography. LSB based Steganography embed the text message in least significant bits of digital picture. Least significant bit (LSB) insertion is a common, simple approach to embedding information in a cover file. Unfortunately, it is vulnerable to even a small image manipulation. DCT based Steganography embed the text message in least significant bits of the Discrete Cosine (DC) coefficient of digital picture.In DCT technique,message is hidden in more significant areas of cover image.An implementation of both these methods and DCT based proposed method with their performance analysis (in terms of PSNR) has been done in this paper. Also different types of attacks such as salt and pepper noise has been applied.In these techniques,a trade-off between the amount of data to be hidden and the robustness of the image will be observed.***

Keywords: Least Significant Bit (LSB), Discrete Cosine Transform (DCT), PSNR, LSB

1. Introduction

In steganography, data can be secured digitally. It not only hides the data but also very existence of communication which is happening. This paper discusses this data hiding application using steganography. The purpose of this paper is to create a user friendly steganography application that allows users to hide private data in image files. Due to the high proliferation of digital images and the high degree of redundancy present in digital images, there is an increased interest in the usage of images as the cover object [1, 2]. The different types of steganography techniques used in this paper are:

A. Spatial Domain Embedding

The best widely known steganography algorithm is based on modifying the least significant bit layer of images, hence known as the *LSB* technique. This technique makes use of the fact that the least significant bits in an image could be thought of random noise and changes to them would not have any effect on the image.

In the LSB technique, the LSB of the pixels is replaced by the message to be sent. The message bits are permuted before embedding, this has the effect of distributing the bits evenly, thus on average only half of the LSB's will be modified. Popular steganographic tools based on LSB embedding vary in their approach for hiding information.

B. Transform Domain Embedding

Another category for embedding techniques for which a number of algorithms have been proposed is the transform domain embedding category. Most of the work in this category has been concentrated on making use of redundancies in the DCT (discrete cosine transform) domain, which is used in JPEG compression. But there have been other algorithms which make use of other transform domains such as the frequency domain. Embedding in DCT domain is simply done by altering the DCT coefficients, for example by changing the least significant bit of each coefficient.

Transform or Frequency Domain Techniques – are independent on image formats and thus can be applied to lossy formats as well. They involve algorithms and tools that manipulate the image by applying transforms such as DCTs and Wavelet Transformations. They hide messages in more significant areas of the cover I mage and may manipulate image properties such as their luminance [3,4].

Hence in these techniques, a trade-off between the amount of data to be hidden. Steganography have to guarantee these requirements [5]:

A. Robustness – the embedded information is said to be robust if its presence can be reliably detected after the image has been modified but not destroyed beyond recognition.
B. Undetectability – embedded information is undetectable if the image with the embedded message is consistent with a model of the source from which images are drawn.
C. Perceptual transparency – it is based on the properties of the human visual system or the human audio system. The embedded information is imperceptible if an average human subject is unable to distinguish between carriers that do contain hidden information and those that do not.
D. Security – the embedding algorithm is said to be secure if the embedded information cannot be removed beyond reliable detection by targeted attacks based on a full knowledge of the embedding algorithm and the detector, and the knowledge of at least one carrier with hidden message.

The above mentioned requirements are mutually competitive and cannot be clearly optimized at the same time. If we want

to hide a large message inside an image, we cannot require at the same time absolute undetectability and large robustness. A reasonable compromise is always a necessity. On the other hand, if robustness to large distortion is an issue the message that can be reliably hidden cannot be too long. In this paper, a new DCT based steganography algorithm is proposed with implementation on various images. Then, attack is applied such as Salt and Pepper noise and robustness feature has been observed.

2. Method of concealing data in digital image

2.1 Least Significant Bit(LSB)

Least Significant Bit coding is one of the simplest methods for inserting data into digital signals in noise free environments. In a grey level, every pixel consists of 8 bits. One pixel can hence display 2^8=256 variations. The weighting configuration of an 8 bit number is illustrated in Figure 1[6].

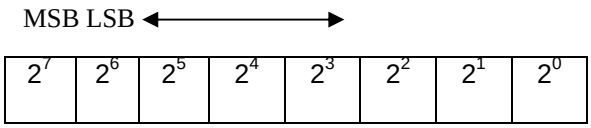

Figure 1: Weighting of an 8-bit pixel

The basic concept of LSB substitution is to embed the confidential data at the rightmost bits (bits with smallest weightings) so that the embedding procedure will not significantly affect the original pixel value. Practically, it can be seen that embedding in the 4th LSB generates example of embedding from the 1st LSB to the 4th LSB is illustrated in Figure 3. more visual distortion to the cover image as the hidden information is seen as “non-natural”[7].

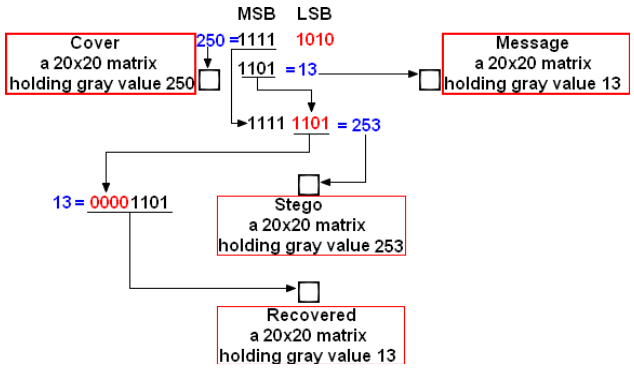

Figure 2: A practical example of embedding from the 1st LSB to the 4th LSB

The mathematical representation for LSB method is:

$$x_i^{'} = x_i - x_i \bmod 2^k + m_i \quad (1)$$

In equation (1), $X_i^{'}$ represents the i th pixel value of the stego-image and X_i represents that of the original cover-image and m_i. represents the decimal value of the i the block in the confidential data. The number of LSBs to be substituted is k. The extraction process is to copy the k-rightmost bits directly. Mathematically the extracted message s represented as

$$m_i = x_i^{'} \bmod 2^k \quad (2)$$

Hence, a simple permutation of the extracted m_i gives us the original confidential data. This method is easy and straightforward. However, when the capacity is greatly increased, the PSNR decreases a lot and hence poor stego-image results. Furthermore, the confidential data might be easily stolen by simply the k-rightmost bits directly [8].

2.1.1LSB Algorithm

Algorithm to embed text message:-
Step 1: Read the cover image and text message which is to be hidden in the cover image.
Step 2: Convert text message in binary.
Step 3: Calculate LSB of each pixels of cover image.
Step 4: Replace LSB of cover image with each bit of secret message one by one.
Step 5: Write stego image.
Step 6: Calculate MSE and PSNR of stego image.
Step 7: Apply different attacks on stego image such as salt and pepper noise. Compute PSNR of noisy stego image.

Algorithm to retrieve text message:-
Step 1: Read the stego image.
Step 2: Calculate LSB of each pixels of stego image.
Step 3: Retrieve bits and convert each 8 bit into character.

2.1.2 PSNR (Peak Signal to Noise Ratio)

How do we determine the quality of a digital image? Human eyes perception is the fastest approach. However, although this criterion is effective in general, the results may differ from person to person. To establish an objective criterion for digital image quality, a parameter named PSNR (Peak Signal to Noise Ratio) is defined as follows:

$$PSNR = 10 \log_{10} 255^2 / MSE \quad (3)$$

where MSE stands for the mean squared error between the cover image and stego image. The mathematical definition for MSE is:

$$MSE = (1/M \times N) \sum_{i=1}^{M} \sum_{j=1}^{N} (a_{ij} - b_{ij})^2 \quad (4)$$

In Equation (4), a_{ij} means the pixel value at position (i, j) in the cover-image and b_{ij} means the pixel value at the same position in the corresponding stego-image. The calculated PSNR usually adopts dB value for quality judgment. The larger the PSNR, the higher the image quality is (which means there is only little difference between the cover-image and the stego-image). On the contrary, a small dB value of PSNR means there is great distortion between the cover-image and the stego-image.

2.2 Discrete Cosine Transform (DCT)

In order to obtain a better performance, the stego message is transformed to the frequency domain. It can separate the image into high, middle and low frequency components [9].

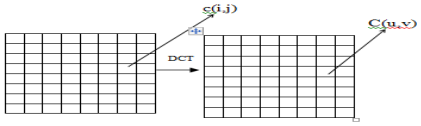

Figure 3: Discrete Cosines Transform of an Image

The general equation for a 1D (N data items) DCT is defined by the following equation:

$$c(u) = a(u)\sum_{x=0}^{N-1} f(x)\cos\left[(2x+1)u\pi / 2N\right] \quad (5)$$

for $u = 0,1,2.....N-1$

The general equation for a 2D (*N* by *M* image) DCT is defined by the following equation:

$$c(u,v) = a(u)a(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cos\left[(2x+1)u\pi / 2N\right].\cos\left[(2y+1)v\pi / 2N\right] \quad (6)$$

for $u, v = 0,1,2,3.....N-1$

Here, the input image is of size N X M. c(i, j) is the intensity of the pixel in row i and column j; C(u,v) is the DCT coefficient in row u and column v of the DCT matrix. Signal energy lies at low frequency in image; it appears in the upper left corner of the DCT. Compression can be achieved since the lower right values represent higher frequencies, and generally small enough to be neglected with little visible distortion [10, 11].

DCT is used in steganography as- Image is broken into 8×8 blocks of pixels. Working from left to right, top to bottom, the DCT is applied to each block. Each block is compressed through quantization table to scale the DCT coefficients and message is embedded in DCT coefficients [12].

2.2.1DCT Based Steganography

Algorithm to embed text message:-
Step 1: Read cover image.
Step 2: Read secret message and convert it into binary.
Step 3: The cover image is broken into 8×8 block of pixels.
Step 4: Transform each block to spatial frequency domain via DCT.
Step 5: Embed stream of bits in cover image by choosing random variable.
Step 6: Inverse DCT is applied to each non-overlapping block and stego image is obtained.
Step 7: Calculate MSE and PSNR.
Step 8: Apply different attacks on stego image such as salt and pepper noise. Compute PSNR of noisy stego image.

Algorithm to retrieve text message:-
Step 1: Read stego image.
Step 2: Stego image is broken into 8×8 block of pixels.
Step 3: DCT is applied to each block.
Step 4: Convert data into message vector. Compare it with original message.

2.3 Proposed Robust Image Stegnography Methodology

The Steganography has to guarantee these four requirements i.e. robustness, undetectability, perceptual transparency and security. We are looking for a robust embedding method. In order to find one, a measure of robustness must be defined. An embedding method may be considered robust if the embedded message can be extracted after an image has been manipulated without being distorted. The embedding algorithm must be tested against the different types of attacks (Salt and Pepper noise have been used in this dissertation) in order to determine how much an image can be manipulated before the message is destroyed. If we want to hide a large message inside an image, we cannot ensure at the same time absolute undetectability and large robustness. On the other hand, if robustness to large distortion is an issue, the message that can be reliably hidden should not be too long. Based on the same embedding capacity, our proposed method improves both image quality in terms of PSNR and CER (Character Error Rate.) [13, 14]

2.3.1Proposed DCT based Steganography

This work includes the following steps:

Algorithm to Embed Text Message
Step1: Read the cover image and text message which is to be hidden in the cover image.
Step2: Convert text message in binary.
Step3: Obtain a block from the image and check either block_Num is equal to zero 0 or Tot_Bit is equal to 1.
Step4: Now if block_Num is not equal to zero and Tot_Bit is not equal to one. Compute the DCT of 8*8 blocks and as per data bit i.e. 0 and 1.
Step5: Select two strength random variables K_1 and K_2. Add the value of main diagonal of DCT's AC co-efficient with either K_1 or K_2 depending upon data bits.
Step6: Read next data bit and obtain next block from the image.
Step7: If Block_Num=M*N/8*8 or Tot_Bit=1, then Message has been Embedded successfully.
Step8: Obtain the Stego Image.
Step9: Calculate the Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) of the Stego image and Calculate percentage of the error occurred in the recovered data in terms of CER (Character Error Rate).
Step10: Calculate the effect of noise (such as Gaussian noise) by varying the variance on the recovered data in terms of CER.

Algorithm to Retrieve Text Message
Step 1: Obtain Stego Image and random variables K_1 and K_2.
Step 2: Read a block from the Stego Image and obtain its 8*8 DCT. Compute the Correlation between the off main diagonal DCT's with both K_1 and K_2.
Step 3: If Correlation (off main diagonal DCT, K_1 of the blocks is greater than Correlation (off main diagonal DCT, K_2 of the block, then the message bit is 1 or else 0.Similarly get the data bits of all the 8*8 blocks of stego image. Convert the data bits to message vector 'M'. Compare it with the original message vector 'M'.

3. Performance and Results

Comparative analysis of LSB and DCT based steganography has been done on basis of parameters like PSNR. Peak signal to noise ratio is used to compute how well the methods perform. PSNR computes the peak signal to noise ratio, in decibels, between two images. This ratio is used as a quality measurement between two images. If PSNR ratio is high then images are best of quality [15,16]

3.1 LSB Based Steganography

We hide the data"hello how are u."(12 characters) inside the 225*225 true lena image. The original image showed in Figure 6(a) and stego-image in the Figure (b). The MSE between stego- image and original image is 0.00011826 and PSNR between stego-image and original image 83.8036

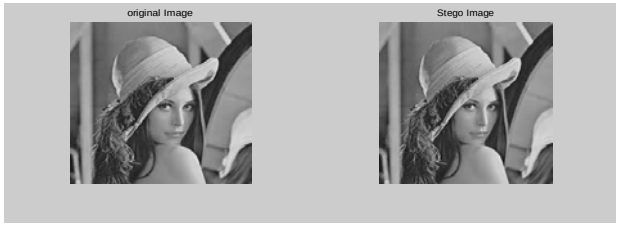

Figure 6: (a) Original image (b) Stego-image

If we want to hide a large message inside an image, we cannot require at the same time absolute undetectability and large robustness. A reasonable compromise is always a necessity. We hide the data"hello how are u." inside the 225*225 true lena image. The MSE between stego- image and original image is 0.00043457 and PSNR between stego-image and original image is 76.8494. The effect of salt and pepper noise on same image is shown in Figure 7.In this case MSE between stego and original image is 2.3981and PSNR between stego- image and original image is 22.693.When the Capacity is greatly is increased, the PSNR decreases a lot and hence poor stego-image quality results.

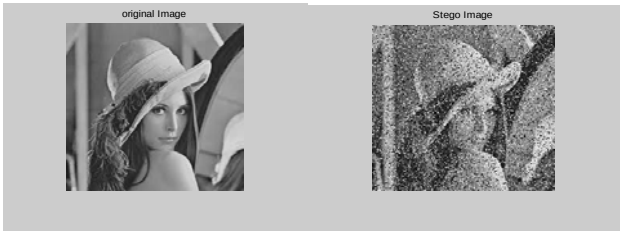

Figure 7: Effect of salt and pepper noise

3.2 DCT based Stegnography

We hide the data "hello how are u."(12 characters) inside the 255*225 true Lena image. The original image shown in Figure 8 (a) and stego-image in the Figure (b). The MSE between stego- image and original image is 0.0011195 and PSNR between stego-image and original image 77.6407.

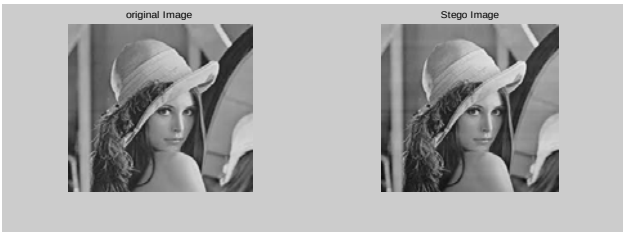

Figure 8: (a) Original image (b) Stego-image

The effect of salt and pepper noise on same image is shown in Figure 8.In this case, MSE between stego and original image is .0066224 and PSNR between stego- image and original image is 69.9207.When the Capacity is greatly increased; the PSNR is within acceptable limits.

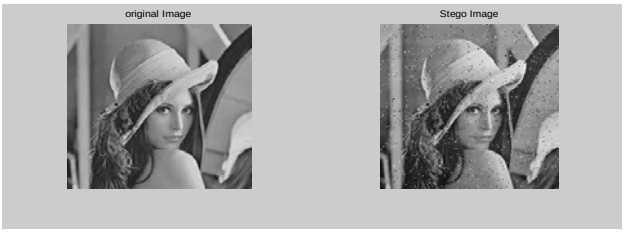

Figure 9: Effect of salt and pepper noise

3.3 Proposed DCT based steganography

We hide the data "hello how are u."(12 characters) inside the 255*225 true Lena image. The original image shown in Figure 10 (a) and stego-image in the Figure (b). The MSE between stego- image and original image is 0.47023 and PSNR between stego-image and original image 51.4077.

Figure 10 : (a) Original image (b) Stego-image

The effect of salt and pepper noise on same image is shown in Figure 11.In this case, MSE between stego and original image is .50521 and PSNR between stego- image and original image is 52.0962.When the Capacity is greatly increased; the PSNR is within acceptable limits.

Figure 11: Effect of salt and pepper noise

Table 1: Survival and PSNR Calculation

IMAGE	*Method*	*MSE*	*PSNR*
Lena	LSB	0.00043457	76.8494
Lena	DCT	0.0011195	77.6407
LENA	PROPOSED	.47023	51.4077

After applying salt and pepper noise:

IMAGE	*METHOD*	*MSE*	*PSNR*
Lena	LSB	2.3981	22.693
Lena	DCT	.0066224	69.9207
Lena	Proposed	.50521	52.0962

So we can say there is a tradeoff between Capacity, Undetectability and Robustness. As the PSNR is decreased but the robustness is increased.

4. Conclusions

LSB based steganography embed the text message in LSB of cover image. DCT based steganography embed the text message in LSB of DC coefficients.This paper proposed and implements LSB and DCT based steganography and computes PSNR ratio. PSNR is the peak signal to noise ratio, in decibels, between two images. This ratio is used as a quality measurement between two images. If PSNR ratio is high then images are better of quality. Comparison of LSB based, DCT based and proposed stego images using PSNR ratio shows that PSNR ratio of LSB based steganography scheme is high as compared to DCT based steganography scheme for all types of images.Proposed DCT based steganography scheme works perfectly with minimal distortion of the image quality as compared to LSB based steganography scheme. Even though the amount of secret data that can be hidden using this technique is very small as compared to LSB based steganography scheme still,Proposed DCT based steganography scheme is recommended because of the minimum distortion of image quality.

References

[1] Popa, et. al, "An Analysis of SteganographicSystem", in Proc. of the International Conference on Computer and Communication Engineering, pp. 978-983, 2008.

[2] Cheddad, J. Condell, K. Curran and P. Mc Kevitt,Securing information content using new encryption method and steganography, in: Proceedings of the 3rd IEEE International Conference on Digital Information Management, University of East London, UK, 13- 16 Nov. 2008, pp. 563-568

[3] M..M. Amin, M. Salleh, S. Ibrahim, M.R Katmin, "Steganography Using Least Significant Bit (LSB)", Malaysian Science And Technology Congress 2002 (MSTC2002), pp. 19-21 September 2002.

[4] R. Radhakrishnan, K. Shanmugasundaram, and N. Memon, "Data masking: A secure covert channel paradigm," IEEE Multimedia Signal Processing, St Thomas, US Virgin Islands, 2002.

[5] Cachin, "An Information-Theoretic Model for Steganography", in Proc. 2nd Information Hiding Workshop, Vol. 15, pp. 306-318, 1998.

[6] V. Lokeswara Reddy,Dr. A. Subramanyam,Dr.P. Chenna Implementation of LSB Steganography and its Evaluation for Various File Formats, Int. J. Advanced Networking and Applications Volume: 02, Issue: 05, Pages: 868-872 (2011)

[7] Ker,"A Improved detection of LSB steganography in grayscale images".Proc. 6th Information Hiding Workshop. Springer LNCS, vol. 3200, pp. 97-115, 2004.

[8] J. Fridrich, M. Goljan and R. Du, "Detecting LSB Steganography in Color and Gray Scale Image", IEEE Multimedia, vol.8, no. 4, pp. 22-28,2001.

[9] KokSheik Wong, Xiaojun Qi, and Kiyoshi Tanaka,"A DCT based Mod4 Steganography " Method‖ Signal Processing 87, 1251-1263, 2007

[10] Manikopoulos, Y.Q. Shi, "Detection of block DCT-based Steganography in gray-scale images" in Proc. of Sixth Indian Conference on Computer Vision Graphics & Image Processing, pp. 7709-7713, 2002.

[11] F. S. Abed, N. Abdul and A. Mustafa, "A proposed Technique for Information Hiding Based on DCT", International Journal of Advancements in Computing Technology, Vol. 2, pp.184-188, 2010

[12] Abbas Cheddad, Joan Condell, Kevin Curran and Paul Mc Kevitt "Digital Image Steganography:Survey and Analysis of Current Methods" , 1992.

[13]M. Ramkumar and A.N. Akansu, "Some Design Issues For Robust Data hiding Systems", IEEE Transactions on Information Theory, Vol. 51, pp. 334-33, 2005.

[14] A.Nag, S. Biswas, D. Sarkar and P.P. Sarkar,' A novel technique for image steganography based on Block-DCT and Huffman Encoding' in International Journal of Computer Science and Information Technology, Vol. 2, No. 3, June 2010.

[15] Constantine Manikopoulos, Yun-Qing Shi,'Detectionof block DCT-based Steganography in gray scale images' in the proceedings of Sixth Indian Conference on Computer Vision, Graphics & Image Processing, 0-7803-7713, 2002 IEEE.

[16] Dr. Fadhil Salman Abed, Nada Abdul Aziz Mustafa,'A proposed Technique for Information Hiding Based on DCT', in International Journal of Advancements in Computing Technology Volume 2,Number 5, pp184-188 December 2010.

[17]Venkatraman.S, Ajith Abraham 'Significance of Steganography on Data Security',In the Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'04) pp 364-367 2004 IEEE.