Analysis and Refinement of Steganography Techniques

Tamanna
Post Graduate Student
Guru Kashi University
Talwandi Sabo, Bathinda

Ashwani Sethi
Dy. Director
Guru Kashi University
Talwandi Sabo, Bathinda

ABSTRACT
Steganography is an art and technique of hiding data in some media i.e image, audio or video file. For hiding data various steganography techniques are used. This paper is an effort to provide comprehensive comparison of these steganography techniques based on different performance metrics such as PSNR, MSE and Embedding capacity. [16] The embedding capacity of Jpeg-Steganography is very less than spatial domain techniques. The spatial domain techniques provide high PSNR, high perceptual quality and high embedding capacity but these not provide robustness. On the other hand transform domain provide robustness while providing very less embedding capacity, low PSNR and low perceptual quality.

Keywords
Steganography, Frequency Domain, Spatial domain, LSB method, PSNR, MSE

1. INTRODUCTION
Steganography, from the Greek, means covered or secret writing, and is a long-practiced form of hiding information. Although related to cryptography, they are not the same. Steganography intent is to hide the existence of the message, while cryptography scrambles a message so that it cannot be understood. More precisely, "The goal of Steganography is to hide message inside other harmless messages in a way that does not allow any enemy to even detect that there is a second secret message present."

Steganography includes vast array of techniques for hiding messages in a variety of media. Among these methods are invisible inks, microdots, digital signatures, covert channels and spread-spectrum communications. Today, thanks to modern technology, Steganography is used on text, images, sound, signal and more. [1]

The advantage of Steganography is that it can be used to secretly transmit messages without the fact of the transmission being discovered. Often, using encryption might identify the sender or receiver as somebody with something to hide. For example, that picture of your cat could conceal the plans for your company's latest technical innovation.

To encode the data we use different techniques like steganography and cryptography. The basic difference between these two as follows:

The word STEGANOGRAPHY combines the Ancient Greek words

1. **Steganos** meaning "Covered", concealed, or protected", and
2. **Graphein** meaning "Writing".

Steganography is the study of embedding and hiding messages in a medium called a covertext. It was used by the Ancient Greeks to hide information about troop movements by tattooing the information on someone's head and then letting the person grow out their hair. Simply put, steganography is as old as dirt. [1]

Cryptography is that we can keep a message a secret by encoding it so that no one can read it. If a good cryptographic cipher is used, it is likely that no one, not even a government entity, will be able to read it.

Watermarking is used to verify the identity and authenticity of the owner of a digital image. It is a process in which the information which verifies the owner is embedded into the digital image or signal. These signals could be either videos or pictures or audios. For example, famous artists watermark their pictures and images. If somebody tries to copy the image, the watermark is copied along with the image. [2]

1.2. Challenges in Steganography
The major challenges of effective steganography are-

1.2.1 Security of hidden Communication- In order to avoid raising the suspicious of eavesdroppers, while evading the meticulous screening of algorithms detection, the hidden contents must be invisible both perceptually and statistically. Steganography techniques should produce high imperceptible Stego-Image.

1.2.2 Size of payload - Unlike watermarking, which needs to embed only a small amount of copyright information, steganography aims at hidden communication and therefore usually requires sufficient embedding capacity. Requirements for higher payload and secure communication are often contradictory. Depending on the specific application scenarios, a tradeoff has to be sought.

1.2.3 Robustness- Stego-image should provide robustness to image processing technique like compression, cropping, resizing etc. i.e. when any of these technique are performed on stego-image, secret information should not be destroyed completely.

There is no technique of steganography which provide all the three properties at high level. There is a trade-off between the capacity of the embedded data and the robustness to certain attacks, while keeping the perceptual quality of the stego-medium at an acceptable level. It is not possible to attain high robustness to signal modifications and high insertion capacity at the same time. [3]

1.3 Types of Attack
While the purpose of steganography is to hide message, there exist several attacks that one may execute to test for steganographic data. The strength of a stegnaographic algorithm depends on its ability to successfully withstand attacks. An attack is dependent on what information is available to the steganlyst. Attacking stegnaographic algorithm is very similar to attacking cryptographic algorithms and similar techniques apply. There are six general

protocols used to attack the use of steganography as pointed out by Katzenbeisser and Petitolas. These are follows:

1.3.1 Stego-only attacks only the steganography medium object is available for analysis.

1.3.2 Known-carrier attack the carrier, that is original cover, and steganography media/object are both available for analysis or are known.

1.3.3 Known-message attack in this case, the hidden message is known and can be compared with stego-object/medium.

1.3.4Chosen-stego attack the stegnaographic medium/object and tool (algorithm) are both available for analysis.

1.3.5 Chosen-message attack here a chosen message and steganography tool (or algorithm) is used to create steganography media for future analysis and comparisons.

1.3.6 Known-steganography attack the secret message, steganography mediums/ object and the stegnaographic tool (algorithm) are known and available for analysis.

1.3.7 Destroy everything attack this type of attack aims in destroying the message completely and the attacker might not even try to retrieve the message.

1.3.8 Random tweaking attacks here small changes in the files are added so that the message will be unreadable.

Add new information in some case the attackers might use the same technique of data hiding to embed a new message into the stego-file. The original message might be overwritten.

1.3.9 Reformat attack a common way to destroy the information hidden in a file is by changing the file format. This type of attack can produce a lot of damages to the hidden message.

1.3.10Compression attack the attacker might compress the file which might result in the total loss of the secret message embedded in the file.

The attacks presented above discuss ways to destroy the hidden message. But for all such case, the attack should be on the suspected image. It might also be a case that an attack can be performed on an innocent image that does not contain any secret data. Based on this certain attacks are implemented in steganography to evaluate of the image contains hidden data.

1.4 The Classical Prisoner Problem

The best way to explain the purpose of using steganography is the Prisoner Problem as illustrated in figure1.1. Consider two prisoners Alice and Bob, and Wendy, the warden. Alice and Bob want to devise a plan for escaping and they have to communicate with each other in the presence of Wendy, If Wendy is a passive warden, and she would not interfere in their communication. If she is active, then she would interfere and try to extract or modify the confidential message or simply stop the communication. Hence Alice and Bob have to communicate with a shared secret key in such a way that Wendy would not be able to decipher their plan without the key "Kerckhoffs principle". This could be done by the technique of steganography.

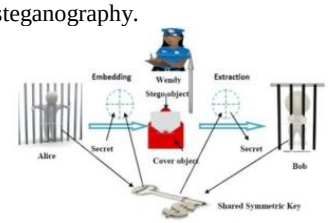

Fig 1: Simmons prisoners' problem

2. PROBLEM FORMULATION

In these methods explained above some problem are there:

2.1 In intermediate method: there is no criteria specified how to select the modifiable fit from a set of bits. It can be selected on the basis of modulus of no of bit.

2.2 In color based method: intensity of a particular color is selected and its occurrences are found and no of these occurred pixels are used to hide data. But RGB colors has different capacity to hide the data so which color is to be selected and how?

2.3 Shape based method: a shape is considered and only under the area of that shape hiding data is suggested.

But amount of data to be hiding depends upon the area of shape, and image size where the area has to be drawn.

The successful data hiding efficiency can be calculated on the basis of

- **SNR (Signal Noise Ratio)**
- **PSNR (Peak signal Noise Ratio)**
- **MSE (Mean Square Error)**

But all these factors also have limitations like:

SNR: depends on the image transmission i.e. if SNR is > than 1:1 then the image is less distorted

PSNR: depends upon maximum power of a signal and its min value. It may high dynamic. The square of S value is calculated rather than considering actual value.

MSE: it depends upon image intensity that may be high or low for any color[5][11].

But due to all these limitations, we cannot conclude which method is best or not. So the problems I had found can be solved one by one after concentrating on these methods. For eg: In intermediate bit method, there is no protocol defined that How to be replaced bit is selected. We can find the length of binary data and calculate the mid of binary data string and select the bit.

3. RELATED WORK

3.1 LSB Substitution:

This is the most common method used. In this type, the data to be hidden is inserted into the least significant bits of the pixel information. Increase or decrease of value by changing the least significant but doesn't change the appearance of the image, such that the result stego-image looks exactly same as the cover image. [4][16]

Embedding Algorithm:
Input: cover image, key, secret message
Procedure:
Step I. Convert the secret message into bit stream (Length L)
Step II. Generate L number of pseudo random number using seed key
Step III. Calculate the non-collide L pixel positions in the cover image
Step IV. While complete bit stream not embedded
{ Replace LSB of pixel denoted by ith pixel position, with secret bit
Insert pixel into cover image.
}
End
Output: Stego-image

On the receiver side, first of all the pixel positions are calculated in the same way with the use of the same key. Then secret bit-stream is formed by the LSBs of these pixels. The extraction algorithm is as below:

Extraction Algorithm:
Input: Stego-image, key
Procedure:
Step I. Convert the secret message into bit stream (length L)
Step II. Generate L number of pseudo random number using seed key
Step III. Calculate the non-collide L pixel positions in the cover image
Step IV. For i=1 to L
{ Get lsb of pixel denoted by ith pixel position
Append this lsb into secret bit stream
}
Step V. Convert secret bit stream into secret message
End
Output: secret message

Advantages:
- There is less chance for degradation of the original image.
- More information can be stored in an image i.e. hiding capacity is more.
- Simple and less complex.

Disadvantages:
- Less robust, the hidden data can be lost with image manipulation.
- Hidden data can be easily detected by simple attacks.
- Requirements of high transmission rate due to large size of stego image.

3.2 Distortion:
Distortion techniques require the knowledge of the original cover in the decoding process. We apply a sequence of modifications to a cover in order to get a stego-image in such a way that it corresponds to a specific secret message for embedding. Receiver measures the difference to the original cover in order to reconstruct the sequence of modifications applied by sender, which corresponds to the secret message. [9][10][16]

Embedding Algorithm:
Input: cover image, key, secret message.
Procedure:
Step I. Convert the secret message into bit stream (Length L)
Step II. Generate L number of pseudo random number using seed key
Step III. Calculate the non-collide L pixel positions in the cover image
Step IV. While complete bit stream not embedded
{ if secret bit =1
{ if pixel value < 128
{ Increase pixel value by x }
Else
{ Decrease pixel value by x }
}
}
End

Output: stego-image

On the receiver side, first of all the differences between the pixel values of cover image and stego-image is calculated. Then pixel positions are calculated in the same way with the use of the key and pseudo random number generator. If the difference at a location is 0 then secret bits is taken as o otherwise it is taken as 1. The algorithm for extraction process is as:

Extraction Algorithm:
Input: cover image, key, stego-image
Procedure:
Step I. Convert the secret message into bit stream (Length L)
Step II. Generate L number of pseudo random number using seed key
Step III. Calculate the non-collide L pixel positions in the cover image
Step IV. Calculate the difference between Cover image and stego- image
Step V. for i=1 to L
{ If value of pixel difference =0
{ Secret bit =0
Else
Secret bit =1 }
}
Step VI. Convert the bit stream into secret message.
End
Output: Secret message

Advantage:
- Less degradation of cover image than LSB i.e. less MSE will produce.
- Embedding capacity is highest.
- Simple and less complex.

Disadvantages:
- In many applications, this technique is not useful, since the receiver must have access to the original covers.
- If someone also has access to original cover, he can easily detect the cover modifications and has evidence for a secret communications.
- Requirements of high transmission rate due to large size of stego-image.

4. PROPOSED WORK

Jpeg Steganography:
Embedding Algorithm:
Input: secret message, cover image
Procedure:
Step I. Convert the secret message into bit stream
Step II. Divide the cover image into 8*8 blocks
Step III. Calculate DCT coefficients for each block
Step IV. Quantize the coefficients
Step V. While complete message not embedded do
{ get next DCT coefficient
If DCT != 0, DCT != 1 and DCT != -1 then
{ get next bit from message
Replace DCT LSB with message bit
}
Step VI. De-quantize and take inverse DCT to obtain stego-image
End.
Output: Stego-image

Extracting Algorithms:
Input: stego image
Procedure:
Step I. Divide the stego image into 8*8 blacks
Step II. Calculate DCT coefficients for each block
Step III. Quantize the coefficients
Step IV. While secret message not completed do
{ Get next DCT coefficient
If DCT! = 0, DCT! =1 and DCT!= -1 then
{ Concatenate DCT LSB to secret message bit stream }
}
Step V. Convert secret message bit stream bit stream into the message.
End.
Output: Secret message.

5. EXPERIMENTAL RESULT

After implementation of the jpeg steganography algorithm that there is no so much change in perceptual quality of image to detect visual changes i.e. quality of embedded image is not degraded by these technique.

5.1 Embedding Capacity

It is the sizes of the secret data that can be embed in cover image without destroying the integrity of the cover image. It can be represented in bytes or bit per pixel. It depends upon the characteristics of cover image and the embedding algorithm used for steganography[6][12][13].

The table shows the embedding capacity of the 3 techniques for different cover image:

Table 1. Embedding Capacity

Image Size(N)	LSB Steganography	Distortion Steganography	Jsteg Steganography
N=256	23437	23437	61
N=500	93750	93750	136
N=1024	393216	393216	5798
N=2048	1572864	1572864	7843

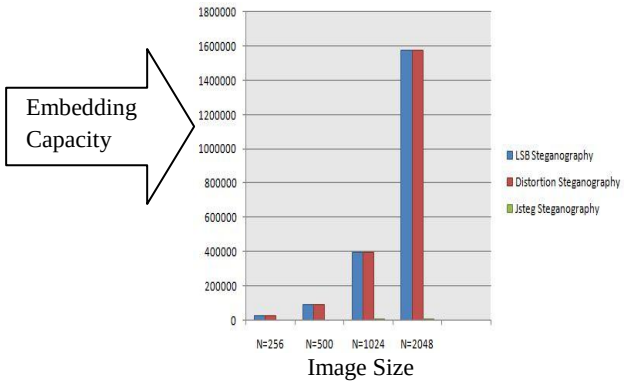

Image Size

The table show that embedding capacity of spatial domain capacity is fix large quantity for a cover image size but capacity of transform domain techniques is very less and it is not fixed for a given size of cover image, it depends upon characteristics of cover image.[7][16][14]

5.2 Peak Signal to Noise Ratio (PSNR)

It is the measure of quality of the image by comparing the cover image with the stego-image. High PSNR indicates good perceptual quality of stego-image. The result of PSNR[8] for all the techniques is in the following table:

Table 2. Peak Signal to Noise Ratio

Image Size (N)	LSB Steganography	Distortion Steganography	Jsteg Steganography
N=256	110.4659	117.6791	36.9118
N=500	122.5315	129.7203	39.1027
N=1024	142.169	142.7986	39.3242
N=2048	153.9334	154.8398	39.2944

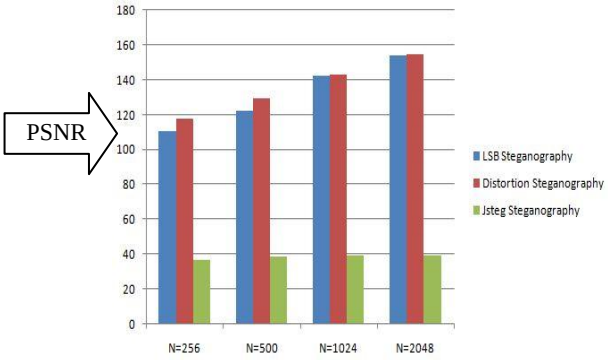

Image size

5.3 Mean square Error (MSE)

The results of the all techniques for the given setup parameters are in the following table:

Table 3. Mean Square Error

Image Size (N)	LSB Steganography	Distortion Steganography	Jsteg Steganography
N=256	5.84	1.11	13.24
N=500	3.63	6.94	7.99
N=1024	3.95	3.41	7.59
N=2048	2.63	2.13	7.64

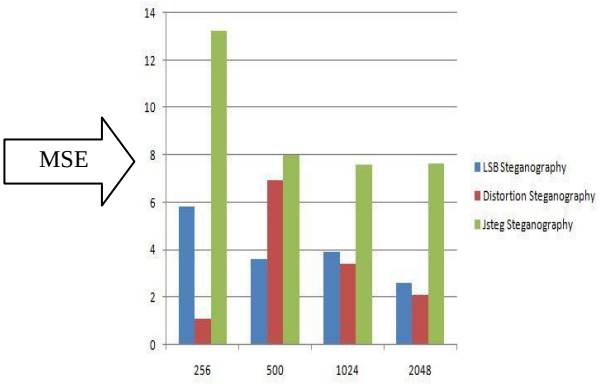

Image Size

6. CONCLUSION

Spatial domain techniques are easy ways to embed information, but they are highly harmful to even small cover modifications. Hence the size of stego-image cannot be reduced. An attacker can simply apply signal processing techniques in order to destroy the secret information entirely. In many cases even the small changes resulting out of lossy compression systems yield to total information loss. Transform domain methods hide messages in significant areas of the cover image which makes them more robust to attacks, such as compression, cropping, and some image processing. Hence lossy compression i.e Jpeg compression can be performed and size of stego-image can be reduced. But the disadvantage of Jpeg-Steganography is that only few messages can be embedded in the cover-image. The embedding capacity of Jpeg-Steganography is very less than spatial domain techniques. The spatial domain techniques provide high PSNR, high perceptual quality and high embedding capacity but these not provide robustness. On the other hand transform domain provide robustness while providing very less embedding capacity, low PSNR and low perceptual quality.

7. FUTURE SCOPE

There is tradeoff between the three properties, perceptuality, embedding capacity and robustness. The new technique should be developing to maintain the three properties at high level. The few areas which are still open in steganography are as below:

Wavelet transform can be used to increase the embedding capacity while maintain the robustness of stego-image.

Hamming coding or Matrix coding can be used to reduce the impact of steganography i.e. to increase the PSNR

Cryptography technique like RSA, ASE and hash functions can also be used with stenography to provide more security

8. REFERENCES

[1] Sanjiv Manchanda, Mayank dave and S. B. Singh, "Customized and secure Image Steganography Through random Numbers Logic". In Signal Processing: An International Journal, Volume 1: Issue (1)-2007

[2] Dipesh Agrawal, Samidha diwedi Sharma, "Analysis of Random Bit Image Steganography techniques". An International Conference on recent trends in engineering & Technology-2013.\

[3] Jayeeta Majumder, Sweta Mangal, "An Overview of Image Steganography Using LSB Technique". National Conference on advance in Computer science-2012

[4] Mamta Juneja, Parvinder Singh Sandhu, "Designing of Robust Image Steganography Technique Based on LSB Insertion and Encryption". An International Conference on Advances in Recent Technologies in Communication and Computing-2009

[5] Anil Kumar, Rohini Sharma, "A Secure Image Steganography Based on RSA Algorithm and HASH-LSB Technique". An International of Advance Research in Computer Science and Software Engineering-2013

[6] Mamta Juneja, Parvinder Singh Sandhu, "Improved LSB based Steganography Techniques for Color Images in Spatial Domain". An International Journal of Network security-2014

[7] Stuti Goel, Arun Rana, Manpreet Kaur, "A Review of Comparison Techniques of Image Steganography". Global Journal of Computer Science and technology graphic & Vision (USA)-2013

[8] Vikas Tyagi, Atul Kumar, Roshan Patel, Sachin Tyagi, Saurabh Singh Gangwar, "Image Steganography Using Least Significant Bit with Cryptography". Journal of Global Research in Computer science-2012

[9] Wallace, G. K. , "The JPEG Still Picture Compression Standard", Communications of the ACM, vol. 34, no. 4, 1991

[10] Cox, "A Secure, Robust watermark for Multimedia," in Information Hiding: First International Workshop, Proceedings,Springer-1996

[11] Xia, X, C. G. Boncelet, and G.r. Arce, "A Meltiresoltuion Watermark for Digital Images," in proceedings of the IEEE International Conference on Image Processing-1997

[12] Sandford, N. bardley and T. G. Handel, "Data Embedding Method," in Proceddings of the SPIE-1996

[13] J.J Chae and B. S. Manjunath, "A Robust embedded Data from Wavelet Coefficients", SPIE-1998

[14] J.R. Harnandez, M. amado and F. PerezGonzalez, "DCT-Domain watermarking Techniques for Still Images: Detector Performance Analysis and a New Structure", IEEE-2000

[15] J.J. eggers, R. Bauml and B. Girod, "A Communications approach to image steganography" SPIE-2002.

[16] Sangeeta Dhall, Bharat Bhushan and Shailender ,"An In-depth Analysis of various Steganography Techniques"