International Journal of Engineering Research and General Science Volume 3, Issue 2, March-April, 2015
ISSN 2091-2730

Data Security using Image Steganography and Processing

Sachin lawande, Aniket Pawar, Kishor Dhumal

Dr. D. Y. Patil collage of Engineering Pune India

sachin1lawande@gmail.com

Abstract— Steganography is an art that contained communication of secret data in an appropriate carrier. Steganography's goal is to hide the embedded data so as not to arouse an eavesdropper's suspicion. For hide secret data in digital images, large varieties of steganography techniques are available same more complex and hard than others. Steganography has various use and different application. It covers and integrates of recent research work without going in to much details of steganography, which is the art and science of defeating steganography techniques.

Keywords— Digital image steganography, data hide, cover image, Data encryption, Data stego.

INTRODUCTION

Steganography word is of Greek source and essentially means that concealed writing. Protection of the transmitted data secret from being intercepted or tampered has led to the development of various steganography techniques. Steganography has been manifested long way back during the ancient Greek Times. Greek tyrant Histiaeus in 499 BC shaved the head of Back, slave was dispatched with the hidden message. Pliny the His slave and wrote message on his scalp. After the hair grew Elder explained how the milk of the Thithymallus plant dried to transparency when applied to paper but darkened to brown when subsequently heated, thus providing the way for hide Information. Giovanni Battista Porta described how to conceal a message within a hardboiled egg by writing on the shell with a special ink. In World War II long sentences of regular letters were used to disguise secret messages. With the tremendous advancement in digital signal processing, use of internet, Computing power, steganography has gone digital. The data hide process starts by identifying a cover image's redundant.

1.1. Application of steganography

Steganography can be used for wide range of applications such as, in defense organizations for safe circulation of database, in military and secrete agencies, in smart identity's cards where personal details are embedded in the photograph Itself for copyright control of materials. In medical imaging's patient details are embedded within image providing protection of information and reducing transmission time , cost[1], in online voting system so as to make the online election secure and robust against a variety of fraudulent behaviours[2] .

1.3. Classification of Steganography Techniques

Classifications of steganography techniques based on the types of cover files as shown in Fig 2. Almost all digital file formats can be used for steganography, however only those with a degree of redundant bits are preferred. The larger size of audio and video files makes them less popular as compared to images. The term protocol steganography refers to embedding information within network protocols such as TCP

In Spatial domain, cover-image is first decomposed into bits planes and then least significant bit (LSB) of the bits planes are replaced with the secret data bits. Advantages are high embedding capacity, ease of implementation and imperceptibility of hidden data. The major drawback is its vulnerability to various simple statistical analysis methods. Frequency domain embedding techniques, which first transforms the cover-image into its frequency domain, secret data is then embedded in frequency coefficients. Advantages include higher level of robustness against simple statistical analysis. Unfortunately, it lacks high embedding. In compression domain, secret data is

Embedded into compression codes of the cover-image which is then sent to the receiver. It is of paramount importance where bandwidth requirement is a major concern.

2. SPATIAL DOMAIN-BASED STEGANOGRAPHY TECHNIQUES

The most direct way to represent pixel's color is by giving an ordered triplet of numbers: red (R), green (G), and blue (B) that comprises particular color. The other way is to use a table known as palette to store the triplet, and put a reference into the table for each pixel. The spatial domain-based steganography techniques use LSB algorithm for embedding/extraction of data as shown in Fig 3.

2.1 E Stego Data Hide

Ez Stego data hide scheme was given by Machado[4, 5]. In this method palette is first sorted by luminance to minimize the perceptual distance between consecutive colors. Ez Stego then embeds the secret data into the LSB of the indices pointing to the palette colors. This approach works quite well in gray scale images and may work well in images with related colors. The major drawback is, since luminance is a linear combination of colors R, G, and B (Luminance = 0.299 R + 0.587 G + 0.144 B), occasionally colors with similar luminance values may be relatively far from each other. Other drawbacks are the ease of extraction of hidden data, dependency of stegno-image quality on number of palette colors, and ease of detection of presence of data using simple statistical histogram analysis.

Fridrich[7] proposed a palette modification scheme for hide data. In this method, both the cost of removing an entry color in a palette and the benefit of generating a new one to replace it are calculated. If the maximal benefit exceeds the minimal cost, entry color is replaced. His method remarkably reduces the distortion of the carrier images, but suffers with the low embedding capacity as Ez Stego does. Cheng[8], *et al*. proposed high embedding capacity technique that can hide.

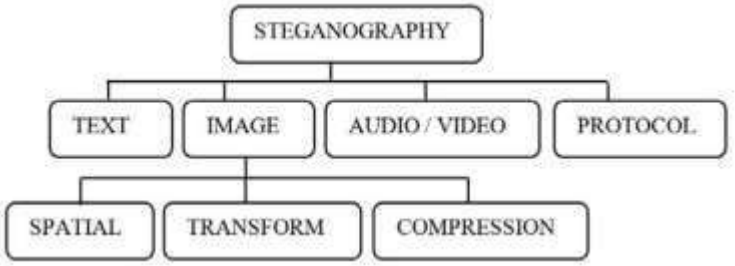

Figure 2. Classifications of steganography technique

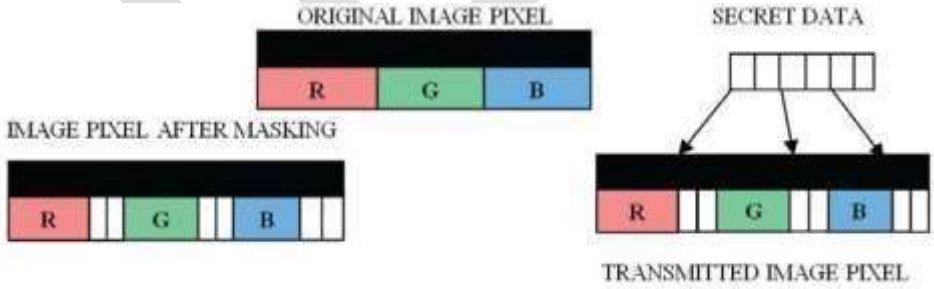

Figure 3. Basic spatial domain data hide.

2.2 S-Tools, Hide & Seek, StegoDos, White Noise Storm, and other techniques

S-Tools by Andy Brown[5,6] reduces the number of colors on. In each segmented bit-plane its complexity is A from 256 to 32 while maintaining the image quality. Instead of and based on a threshold value block is divided into simply going with adjacent colors as Ez Stego does, S-Tools 'informative region' and 'noise-like region' and the secret manipulate the palette to produce colors that have a

difference of data is hidden in noise regions without degrading image one bit. As compared to Ez Stego, non-linear insertions in S-quality. BPCS provides high embedding capacity and least Tools method make the presence and extraction of secret data degradation of the cover-image as compared to traditional more difficult and achieve better results in terms of visual LSB manipulation techniques. Maya13, *et al.* uses variance of perceptibility. Figures 4 and 5 shows cover image before and image block as a parameter for complexity measure. Prime after embedding data. Hide & Seek given by Maroney[5] uses LSB advantages achieved are high embedding capacity and of each pixel to encode characters of secret data and has robustness against noise as compared to BPCS technique. Embedding capacity which is restricted to $1/8^{th}$ of the size of the cover-image. StegoDos[5] works only with 320 X 200 pixels.

2.4 Information Theory-based Data Hide

Image and involves much effort in encoding and decoding of the Hadhoud[14], *et al.* proposed a technique based on entropy secret message. White Noise Storm includes encryption to Calculation. In this method entropy of the '4' most randomize the bits within an image and suffers with the problem significant bits (MSBs) are calculated first which contains of using large cover file. Most detail of each pixel. If the entropy is > 2 then it inserts '4'

Younes[10], *et al.* proposed a method in which data is bits into the '4' LSBs, if not then the entropy of the '5' MSBs is inserted into LSB of each byte within the cover-image in calculated. If it is > 2 then it inserts '3' bits into the '3' LSBs, if encrypted form. Mandal[11] proposed a method with minimum not then it inserts '2' bits into '2' LSBs. Flowchart for entropy deviation of image fidelity resulting high quality stegno based data hide is shown in Fig 6. This method provides high image with better embedding capacity. Embedding and high level of image transparency.

Figure 4. S-Tools: Before embedding

Figure 5. S-Tools: After embedding

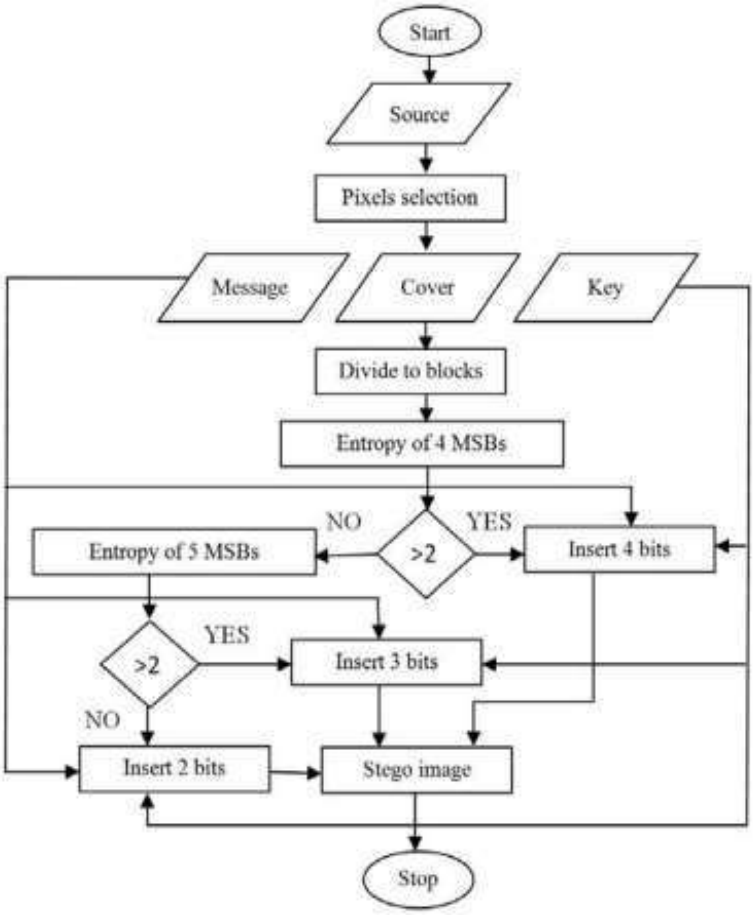

Figure 6. Entropy-based data hide

3. DATA HIDE TECHNIQUES IN FREQUENCY DOMAIN

Frequency domain methods hide messages in significant Areas of the cover-image which makes them more robust to attacks such as compression, cropping or image processing methods than LSB approach and moreover they remain imperceptible to the human sensory system as well. Many transform domain variations exist, one of which is discrete cosine transform (DCT). Some of the important frequency domain-based steganography data hide methods are:

3.1 J Steg, J Steg-Shell, JP Hide, and Out Guess

JSteg developed by Derek Upham[25,26] sequentially replaces the LSB of the DCT coefficients with the message's data. This technique does not require a shared secret; as a result, anyone who knows the steganography system can retrieve the message easily, thus not so secure. JSteg-Shell is a windows user interface to JSteg developed by Korejwa[26]. It supports encryption and compression of the content before embedding the data with JSteg. Both methods can be easily detected using χ^2 -test given by A. Westfield in 1999. JP Hide steganography system was given by Allan Latham[26]. Two versions 0.3 and 0.5 are available. Version 0.5 supports additional compression of the secret message. As the DCT coefficients are not selected sequentially from the beginning of the image, JP Hide is not vulnerable to χ^2 test; however detected using its extended version[25].

Outguess was proposed by Provos[25, 26] as a response to the statistical tests given by Andreas Westfield. It improves embedding by selecting DCT coefficients randomly. Two versions are available: Outguess 0.13b which is vulnerable to extended version of χ^2 -test and Outguess 0.2 which has the ability to preserve frequency counts statistics and hence remain undetected. Provos observed that while embedding not all the redundant bits were used and thus it is possible to use the remaining bits to correct statistical deviations that embedding created. Outguess 0.2 uses this phenomenon to avoid class of 2 -tests. χ

3.2 Data Hide Techniques: F3, F4 and F5

F3 decrements the non-zero coefficient's absolute value only if the LSB does not match with the secret bit. Zero coefficients are skipped completely. Advantage is its 2 -test). Major resistance to statistical attack (χ shortcomings are its less capacity, surplus of even coefficients caused by shrinking and repetitive embedding required since receiver cannot differentiate between skipped 0 and the 0 generated due to shrinkage. The F5 algorithm was introduced by German researchers

Pfitzmann and Westfeld[27]. F5 embeds message bits into randomly-chosen DCT coefficients and employs matrix embedding that minimizes the necessary number of changes to embed a message of certain length. F5 comes after a series of F3 and F4. F5 is similar to F4 except of the fact that F4 does not use matrix encoding in embedding process. The major strengths of F5 are its high embedding capacity without sacrificing security and its resistance to statistical and visual attacks.

3.3 Genetic Algorithm-based Data Hide

Chang[28], *et al.* proposed a JPEG and quantization table Modification (JQTM) method that improves the standard JPEG quantization table for better quality of the stage-image. In this method only 26 middle frequency components of the quantized DCT coefficients for each block are used to hide the secret message. JQTM suffers with its low embedding capacity and low security level. Li and Wang[29] modified the quantization table used in JQTM and uses Particle swarm optimization [30] to approach optimal LSB substitution, which guarantees a higher security level and better quality for the cover images.

To further increase the embedding capacity of the JQTM, Fazli[31], ***et al.*** modified quantization table proposed by Li and Wang. Fazli, *et al.* first transformed secret message using optimal substitution matrix calculated using PSO algorithm and then embed transformed results into the quantized coefficients. This technique differs from Li and Wang's method in the sense that in this substitution matrix is calculated for each 8 x 8 block of the cover-image instead of a single matrix for the whole cover-image. The great achievement of this method is a high security level, high embedding capacity, and high image quality as compared to the JQTM and Li and Wang's method.

In vector quantization (VQ) [32], a block image is imported; the VQ encoder seeks the most similar code word from the codebook to substitute for the block and the index value is then exported as the compressed code for the block. Example of
VQ encoding is shown in the Fig 9. Side-match VQ (SMVQ), improving VQ compressing performance was proposed by Kim[33]. In SMVQ instead of using the original pixels to encode the X block, Kim uses the upper U block and the left L block to encode the X block. SMVQ encoder is shown in Fig 10. Yang[34], *et al.* presented a reversible data hide scheme based on SMVQ for VQ compressed images. This method makes the corresponding code words in the current state codebook and the next state codebook close. Results show that Yang's scheme has higher capacity, better visual quality, and lower running time as compared to Chang's method[35].
Chang[36], *et al.* provided a VQ-based embedding method with high embedding capacity. In their method, a codebook is partitioned into clusters. Data are embedded into the VQ index table by transferring index values from one cluster into another cluster. Data hide schemes for VQ-compressed images are based on index modifications. These schemes may cause distortions and hence are not suitable for authentication of VQ compressed images. To overcome this limitation Jiafu[37], *et al.* proposed an image authentication scheme for VQ-compressed images. This scheme utilized an information.

4. Methodology (Design and implementation)

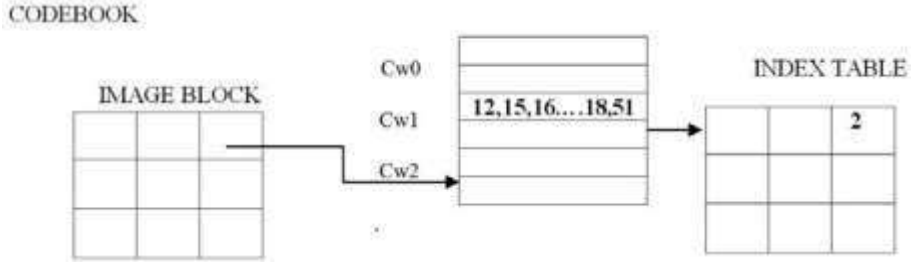

Figure 9. Example of VQ encoder

						U	
				U12	U13	U14	U15
			L3	X0	X1	X2	X3
	L		L7	X4			
			L11	X8		X	
			L15	X12			

Figure 10. Example of SMVQ encoder.

Hide method based on covering codes[38]. It only modifies few indices slightly to hide authentication information. Yang[39], *et al.* further increases the embedding capacity of VQ-based data hide scheme. Under the same sorted VQ codebook, the experimental results demonstrate that this data hide algorithm has higher capacities and better compression rates. For the VQ-based algorithms discussed above only limited amount of information can be hidden, to overcome this, Kekre[40], *et al.* proposed a method based which can achieve hide capacity of 100 per cent or more, that means secret message can be of same or more size than the cover image.

4.2 Data Hide in Block Truncation Coding

Xiaotian Wu[41], *et al.* presented a technique of data hide Method by modifying the bitmaps generated from the block truncation coding (BTC) method given by Delp and Mitchell[42]. In the encoding phase of BTC, the original image is firstly divided into non-overlapping blocks with *n* x *n* pixels. For each block, the mean value is calculated. All the pixels in the block are separated into two groups, greater and smaller than or equal to the mean value. A bitmap with the same size of the block is used to record the output of the BTC compression. The bit in bitmap is set to 1 and classified to G1, when the corresponding pixel in the block is greater than the mean value; otherwise, it is set to 0 and classified to G0. Two mean values XH and XL are calculated, representing mean of pixel values in G1 and G0. Using this each block of the original image is compressed into a bitmap and two quantization levels, XH and XL. Wu uses BTC compression where each bit of the secret message is sequentially embedded into the bitmap of the corresponding compressed non-overlapping block. It results in higher imperceptivity.

4.3 Data Hide in Compressed Images Using Histogram Analysis

Keissarian[43] proposed a method that decomposes the host image into blocks of variable sizes according to histogram analysis of the block residuals. Variable block sizes are then encoded at different rates based on their visual activity levels. The key point is to embed majority of secrete data into smooth area of the image. Results confirmed that the proposed scheme can embed a large amount of data while maintaining satisfactory image quality. Keissarian[44] proposed further improvement in which the computation of the gray values, are carried out through analysis of the block residuals' histogram.

Experimental Results

Encryption Process

IMAGE FILE

INFORMATION FILE

BMP FILE

Decryption Process

INFO FILE

IMAGE FILE

5. CONCLUSION AND SUMMARY

This paper presented the recent research work in the Field of steganography deployed in spatial, transform, and compression domains of digital images. Transform domain techniques make changes in the frequency coefficients instead of manipulating the image pixels directly, thus distortion is kept at minimum level and that's why they are preferred over spatial domain techniques. But when it comes to embedding capacity, spatial domain techniques give better results. However, there exists a trade-off between the image quality and

the embedding capacity. Hide more data results directly into more distortion of the image. So the steganography technique deployed is dependent on the type of application it is designed for. In recent years, some researchers have concentrated on embedding secret data into the compression codes of images. Such need arises keeping in mind the bandwidth requirements.

Steganography can also be used misused like other technologies. For instance terrorists may use this technique for their secret secure communication or anti-virus systems can be fooled if viruses are transmitted in this way. However, it is evident that steganography has numerous useful applications and will remain the point of attraction for researchers.

REFERENCES:

1. Nirinjan, U.C. & Anand, D. Watermarking medical images with patient information. *In* the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Hong Kong, China, 1998, pp. 703-06.
2. Katiyar, S.; Meka, K.R.; Barbhuiya, F.A. & Nandi, S. Online voting system powered by biometric security using steganography. *In* the 2nd International Conference on Emerging Applications of Information Technology (EAIT), Kolkata, India, 2011, pp. 288-291.
3. Shirali-Shahreza, M. Improving mobile banking security using steganography. *In* the 4th International Conference on Information Technology, ITNG, Las Vegas, 2007, pp. 885-887.
4. Machado, R. EzStego, Stego Online. http://www.stego.com (Accessed on 15 April 2011).
5. Johnson, N.F. & Jajodia, S. Exploring steganography: Seeing the unseen. *IEEE Computer*, 1998, **31**(2), 26-34.
6. Westfeld, A.; Pfitzmann A. Attacks on steganography systems breaking the steganography utilities EzStego, Jsteg, Steganos, and S-tools—and some lessons learned. *In* the 3rd International workshop on Information hide, Dresden, Germany, 1999, Springer, pp. 61-76.
7. Fridrich, J. A new steganography method for palettebased images. IS&T PICS, Savannah, Georgia, 1999, pp. 285-89.
8. Cheng, Z.; Kim, Se-Min &Yoo, Kee-Young. A new steganography scheme based on an index-colour image. *In* the 6th International Conference on Information Technology: New Generations, Las Vegas, Nevada, 2009, pp. 376-81.
9. Ren, Honge; Chang, Chunwu & Zhang, Jian. Reversible image hide algorithm based on pixels difference. *In* the IEEE International Conference on Automation & Logistics, ICAL '09, Shenyang, 2009, pp. 847-850.
10. Younes, Mohammad Ali Bani & Jantan, A. A new steganography approach for image encryption exchange by using the least significant bit insertion. *Inter. J. Comp. Sci. Network Security*, 2008, **8**(6), 247-254.
11. Mandal, J.K. & Sengupta, M. Steganography technique based on minimum deviation of fidelity (STMDF). *In* the 2nd International Conference on Emerging Applications of Information Technology (EAIT), Kolkata, 2011, pp. 298-301.
12. Kawaguchi, E. & Eason, R.O. Principle and applications of BPCS-Steganography. *In* the SPIE Conference on Multimedia Systems and Applications, Boston, 1998, **3524**, pp. 464-73.
13. Maya, S.T.; Miyatake, M.N. & Medina, R.V. Robust steganography using bit plane complexity segmentation. *In* the 1st International Conference on Electrical and Electronics Engineering, 2004. Mexico, pp. 40-43.
14. Hadhoud, M.M.; Ismail, N.A.; Shawkey, W. & Mohammed, A.Z. Secure perceptual data hide technique using information theory. *In* the International Conference on Electrical, Electronic and Computer Engineering (ICEEC), Egypt, 2004, pp. 249-253.
15. Mielkiainen, J. LSB Matching revisited. *IEEE Signal Proc. Letters*, 2006, **13**(5), 285-87