International Journal of Scientific Engineering and Science

Volume 1, Issue 2, pp. 48-52, 2017.

A Study on Various Secret Data Embedding Techniques in Digital Images for Secure Communication

Sruthy Raj[1], Annmary Thomas[2]

[1]PG Scholar, Dept. of ECE, MBC College of Engineering and Technology Peermade, Kerala, India-685531

[2]Assistant Professor, Dept. of ECE, MBC College of Engineering and Technology Peermade, Kerala, India-685531

Email address: [1]sruthyraj6096@gmail.com

Abstract— *Covert communication using images is relatively very young and fast growing field combining image and signal processing with cryptography, communication theory, coding theory, signal compression, and the theory of visual perception. Image steganography aims at hiding the very presence of communication, to make the communication undetectable. This is done by adding secret data in a suitable multimedia cover, e.g., image, audio, and video files. There exists various techniques to embed secret data in images. Different image file formats have different embedding schemes, with their own advantages and disadvantages. In this paper a study on available secret data embedding techniques in an image file in various spatial domain and frequency domain techniques are done along with masking and filtering techniques. Techniques for the recovered image quality assessment is also discussed in this paper since performance of each embedding techniques is assessed by the quality of recovered image and the distortion less extraction of secret data.*

Keywords— *Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), Least Significant Bit (LSB),and filtering Masking, Steganography.*

I. Introduction

For a significant period of time, individuals tried to generate creative techniques for secret communication since security has emerged to be a prominent issue. Steganography is an area of data security, which covers data in a cover media for secret communication. While information is transmitted over network it may get exposed to eavesdropping. In many applications such as military and medical imaging and also covert police and military application confidentiality of transmitted information is of high priority. Cryptography is a technique that aids in encryption of the information to be transmitted. Then the cipher text (encrypted information) can be directed over the network to the receiver who can decrypt the data using the private key [1]. But the cipher text may lead to eavesdropping no matter how strong the algorithm is. So information may become disposed to to disturbance as the intruder may modify it to give a fake data. Here, the prerequisite of steganography arises where the secret information can be masked in some other medium that may be text, image, audio or video and then transferred to the receiver [2]. The blend of steganography and cryptography help in increasing the security level of the information being communicated.

An intensive history of steganography can be found in the script [3]. The word steganography is of Greek source. It is gotten from two Greek words "stegos" which signifies "cover" and "grafia" which signifies "expressing" [4]. It is for the most part utilized for the protected communication to veil it from attackers that make challenges for unintended client to extricate the data. Just the beneficiary of stego-media can remove the secret information. Steganalysis [5] is utilized for the identification of concealed information. It takes after the comparative computational steps as utilized by a watermarking strategy. Nevertheless, the objectives are different of both the plans.

There exists various techniques to embed secret data in images. Different image file formats have different embedding schemes, with their own advantages and disadvantages.

In this paper a study on available secret data embedding techniques in an image file in various spatial domain and frequency domain techniques are done along with masking and filtering techniques. Techniques for the recovered image quality assessment is also discussed in this paper since performance of each embedding techniques is assessed by the quality of recovered image and the distortion less extraction of secret data.

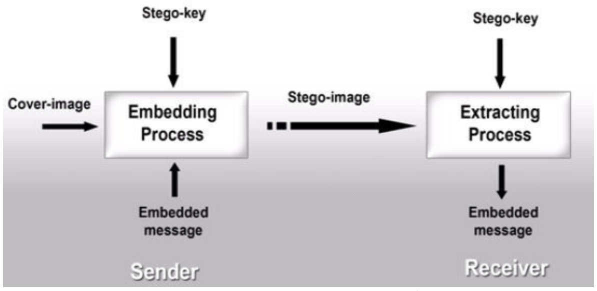
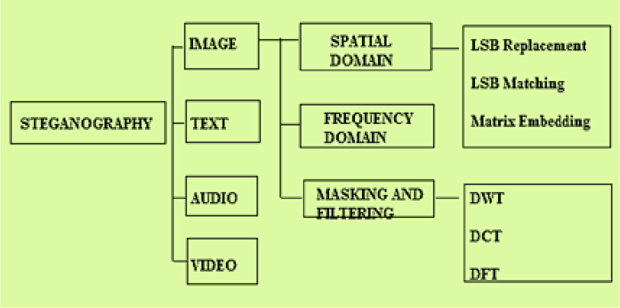


Fig. 1. A generalized data embedding framework.

II. Existing Secret Data Embedding Schemes in Digital Images

A. Spatial Domain Techniques

This technique utilizes encoding in Least Significant Bits. LSB addition strategy is a simple approach for embedding information into the cover image. There are various adaptations of this technique; every one of these strategies dependably change a portion of the bits in the estimations of picture pixels for concealing secret information. LSB dependent steganography is one of the significant strategies that hide secret messages in the LSBs of some pixel values with no detectable distortions [6] since for our human eye, varieties in the LSB are unnoticeable. Embedding of additional bits of information can be done either randomly or simply. LSB techniques and also Matrix embedding are some spatial domain techniques.



1. LSB replacement

In this data hiding technique, the cover pixel LSBs are substituted with a bit of the message that must be embedded. Before embedding, the message is changed into a sequence of bits which are then embedded successively where the LSBs are located [7]. This data hiding technique is perceptible regardless of the possibility that there is low embedding rate.

Advantages:

- Degradation of the cover image is difficult
- Higher embedding rate

Disadvantages:

- Low robustness
- Destruction of embedded data even with small attack

2. LSB matching

This method is quite enhanced over LSB substitution technique. In this procedure 1 is either arbitrarily summed up or subtracted from the estimation of the cover pixel on the off chance that the bit of the classified message is not proportional to the LSB that come from the cover pixel [8]. When contrasted with LSB Replacement technique it is difficult to recognize LSB matching.

3. Matrix embedding

This strategy encodes the cover image and additionally the message by a error correction code. It changes the cover image with respect tocoding. In this procedure the conceivable message bits are inserted randomly per an embedding change subsequently it helps in expanding embedding efficiency.

B. Frequency Domain Steganography

This method is commonly used in JPEG file format. This file format is quite common because of its small size and wide usage in digital photography. The lossless compression of JPEG includes various steps like transformation of RGB to YUV where Y refers to brightness, U refers to chrominance and V for colour [9]. Then Discrete Cosine Transformation is applied and later quantization and Huffman encoding is done. It is recommendable to insert information before applying Huffman encoding for lossless recovery. Transform domain hide information in the particular zones of the original image giving much more embedding capacity and robustness. Transform domain techniques are superior to LSB strategies because of the way that they embed data in especially those regions of original image which are very little exposed to processing of images.

Frequency domain techniques include:

1. Discrete wavelet transformation

This transformation is widely used in multi-resolution demonstration. Wavelets are described as the functions obtained over a fixed interval with zero as average value. In this technique a signal is disintegrated into a number of constituents in frequency domain. 1-D DWT fragments a cover image into two components known as approximate component and detailed component [10]. A 2-D DWT is utilized to segment a cover picture into four sub components: one approximate component (LL) and the other three comprise detailed components represented as (LH, HL, HH).

LL	HL
LH	HH

Fig. 2. DWT components.

2. Discrete cosine transformation

This transformation techniques is helpful for separating an image into various parts of varying significance (which is related with the picture's quality). It changes an image from its spatial domain into frequency domain. In this method, for each color constituent, the JPEG configuration of picture makes utilization of cosine transform to convert pixel blocks of size 8 x 8 into a check of 64 cosine coefficients each. For each 8x8 pixel block having pixel value f(x,y), the coefficients f(u,v) are given as [11]

$$F(u,v) = \frac{1}{4} C(u)C(v) \left[\sum_{x=0}^{7} \sum_{y=0}^{7} f(x,y) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2x+1)v\pi}{16} \right]$$

Where $C(u) = \begin{cases} \frac{1}{\sqrt{2}}, if\ u \le 0 \\ 1, if\ u > 0 \end{cases}$

3. Discrete fourier transformation

This technique is mainly is critical as it isolates a picture into the sine and cosine values. It changes over space and time subordinate data into the recurrence based data. It is helpful for various applications including picture separating and remaking and in addition picture pressure. It doesn't incorporate all frequencies that outcome to frame a picture yet constitutes of just the arrangement of those specimens which are adequate to portray the first picture. The DFT for the vector x having length n is some other vector y having length n [12]:

$$y_{p+1} = \sum_{j=0}^{n-1} w^{jp} x_{j+1}$$

Where w signifies root nth for unity

$$w = e^{-2\pi i/n}$$

C. Masking and Filtering

This technique is commonly used in grey scale images. Covering up is like setting watermarks on a printed picture.

This strategy embeds secret data especially in most significant areas rather than concealing it just into the noisy area [9]. Watermarking strategies can be useful without image destruction because of the lossy technique for compression since it is more included into a image. Some different systems depend on concealing information in corners, concealing information in light of surface resemblance and so forth.

III. Major Challenges of Effective Data Embedding in Images

A. Security of Concealed Communication:

In order to avoid raising the suspicions of eavesdropper the hidden subjects must be undistinguishable both perceptually and statistically.

B. Size of Payload:

Requires adequate embedding capacity depending on the specific application set-ups, a trade-off has to be sought.

C. Steganalytic Attacks

1. Statistical attacks:

These methods use statistics of the image to reveal tiny changes in the statistical information caused by steganographic embedding and hence can effectively identify even small amounts of embedding with very high accuracy. Statistical attacks are further categorised as 'Targeted Attacks' or 'Blind Attacks'.

2. Visual attacks:

These methods try to identify the presence of secret data by visual assessment either by a computer or by bare eyes. The attack is created on predicting the embedding layer of an image (say a bit plane) and then visually examining that layer to look for any unfamiliar alterations in that layer.

IV. Recovered Image Quality Assessment in Data Hiding

For the purpose of the image quality estimation, certain quality metrics are used. These metrics convey some measures of the acquaintance between two digital images by exploiting the changes in the statistical distribution of pixel values. The most commonly used quality metrics are given below.

A. Capacity

The capacity is defined as the maximum amount of additional information that can be embedded in the cover media. Distinctive information concealing applications have diverse capacity requirements [13-16]. Capacity estimation is an essential issue of steganography, where the question is what amount of information can securely be covered up without being distinguished. For the most part, capacity is calculate dated in bits per pixel images, bits per sample in sound and bits per frame in video. Information concealing calculations with high limit and low vigor are called steganography systems, while the general term of watermarking more often than not alludes to a low-limit vigorous information concealing plan. The limit of an information concealing plan characterized as [17]:

$$Capacity = \frac{Maximum\ embedded\ data\ size}{Cover\ media\ size}$$

B. Robustness

Robustness is the ability to resist against modifications enforced on a secret data embedded image. It is the most important property for the accurate detection of embedded data. The synchronization between the data hider and, the receiver is destroyed by the attackers. So the strength of a data hiding framework is assessed by utilizing distinctive techniques, for example, bit error rate (BER) [18] and normalised correlation (NC) [23, 24]. The normalised correlation between original image and its reconstructed image is given as:

$$NC(X,Y) = \frac{\sum_{i=1}^{n}\sum_{i=1}^{n} X(i,j) \times Y(i,j)}{\sqrt{\sum_{i=1}^{n}\sum_{i=1}^{n} X(i,j)^2}\sqrt{\sum_{i=1}^{n}\sum_{i=1}^{n} Y(i,j)^2}}$$

C. Transparency or Imperceptibility

Because of the inclusion of secret data into cover image distortion is normal in the cover image. Perceptual similarity of the original image with the secret data embedded image is referred as transparency. The aim of data hiding is to include secret information which causes invisible disfigurements in cover media to keep up its commercial value. For the assessment of perceptual likeness no universal effective measure exists [19]. However, Peak Signal to Noise Ratio (PSNR) [5] and Structural Similarity Index Measure (SSIM) [19] are widely used for the purpose of perceptual similarity measure.

1. PSNR

It is generally utilized for the performance valuation of information concealing frameworks. In case of digital images, it is the ratio between the maximum intensity of pixel to the intensity of noise. Numerically PSNR is characterized as [20]:

$$PSNR(X,Y) = 10log_{10}\left(\frac{(255)^2}{\frac{1}{n \times n}\sum_{t-1}^{n}\sum_{j-1}^{n}(X(i,j) - Y(i,j))^2}\right)$$

where X and Y remain for the cover and the recovered images; subscripts i and j indicate the area of the pixel esteem in the particular image; and n is the dimension

2. SSIM

Structural Similarity Index Measure (SSIM) is an approach that is used to measure the similarity between the cover image (X) and embedded image (Y). It is produced by Wang et al. [21], and is considered to be connected with the quality view of the human visual framework (HVS). It is outlined by presenting an image distortion as a mix of three components that are correlation, luminance distortion and contrast distortion. The SSIM is characterized as [19]:

$$SSIM(X,Y) = \frac{(2\mu_X\mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)}$$

with $\begin{Bmatrix} C_1 = K_1L \\ C_2 = K_2L \end{Bmatrix}, \begin{Bmatrix} K_1 = 0.01 \\ K_2 = 0.03 \end{Bmatrix}$ $and\ L = 255$

where µ, 2, and X Y are the mean, difference, and covariance of the images X and Y . The constants C1, and C2 are the stabilising constants that are utilized to maintain a strategic distance from a null denominator. The estimation of SSIM list runs over the interval [0, 1]. An estimation of 0 implies no relationship amongst's pictures, and 1 implies that both pictures are same.

where X and Y remain for the first and the handled pictures; subscripts i and j signify the area of the pixel esteem in the individual pictures; and n is the stature then again width of the square picture. The bit error rate (BER) is characterized independent of the have picture measure as [22]:

$$BER = \frac{Number\ of\ incorrect\ bits}{Number\ of\ total\ bits}$$

V. CONCLUTION

Some recent secret data embedding schemes are studied in this paper. The aim was to provide the complete detail of data embedding schemes that may assist the innovative scholars to get the maximum knowledge of the area. The precise classification of embedding schemes is not possible, as many researchers have combined different approaches to develop hybrid schemes. Based on this review, the following recommendations may help interested users in data hiding for different purposes. Secret data embedding schemes are rather beneficial where the quality of cover media is of great importance such as healthcare, military application, and law enforcement. Transformed domain schemes have enhanced performance in comparison to spatial domain but computationally are expensive. Artificial intelligence based data embedding schemes are easy to implement, as these performs without user interference, and effective in comparison to conventional data embedding schemes.

VI. FUTURE SCOPE

In future, research could be extended towards designing algorithm which generate stego images which are much similar to the cover image as possible. It may be possible to design some encoding functions that can modify the message stream to make it more suitable for embedding than the original bit stream. Even if the invader knows the embedding algorithm, the exact message sequence cannot be rebuilt unless the invader has the knowledge of the encoding function. Additional potential course of research can be framed as a problem of “Image Retrieval”. It is based on probing for a suitable cover image given a message sequence and the embedding algorithm. This may be possible through preserving a huge image database and given any message sequence and a pseudo-random key for generating embedding positions. Later cover image from the database that will generate a stego image with minimum amount of changes could be done. The modification condition considered for searching can be reliant on on the features used by the matching steganalytic attacks. Roughly additional likely concepts can be rented from the area of “Visual Cryptography” which encrypts a secret information by distributing the decoding key into various images such that the message can be retrived only by a appropriate grouping of these images.

REFERENCES

[1] F. A. P. Peticolas, “Information hiding-a survey,” *Proceedings of the IEEE*, vol. 87, pp. 1062-1078, 1999.

[2] N. Provos, and P. Honeyman, “Hide and seek: An introduction to steganography,” *IEEE Security Privacy*, vol. 1, issue 3, pp. 32-44, 2003.

[3] N. Provos and P. Honeyman, “Hide and seek: an introduction to steganography,” *IEEE Security & Privacy*, vol. 1, issue 3, pp. 32–44, 2003.

[4] M. M. Sadek, A. S. Khalifa, and M. G. M. Mostafa, “Video steganography: a comprehensive review,” *Multimedia Tools and Applications*, 2014. in press.

[5] D. C. Lou, C. L. Chou, H. K. Tso, and C. C. Chiu, “Active steganalysis for histogram-shifting based reversible data hiding,” *Optics Communications*, vol. 285(10-11), pp. 2510–2518, 2012.

[6] Dr. D. Samidha and D. Agrawal, “Random image steganography in spatial domain,” *IEEE International Conference in Emerging Trends, VLSI, Embedded System, Nano Electronics and Telecommunication System*, pp. 1-3, 2013.

[7] K.A. Darabkh, I. F. Jafar, R. T. Al-Zubi, and M. Hawa, “An improved image least significant bit replacement method,” *IEEE 37th International Convention in Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 1182-1186, 2014.

[8] L. Kumar Vashishtha, T. Dutta, and A. Sur, “Least significant bit matching steganalysis based on feature analysis,” *IEEE National Conference in Communications (NCC)*, pp. 1-5, 2013.

[9] M. A. Dagadita, E.-I. Slusanschi, and R. Dobre, "Data Hiding Using Steganography,” *IEEE 12th International Symposium in Parallel and Distributed Computing*, pp. 159-166, 2013.

[10] G. Prabakaran and R. Bhavani, “A modified secure digital image steganography based on discrete wavelet transform,” *IEEE International Conference in Computing, Electronics and Electrical Technologies (ICCEET)*, pp. 1096-1100, 2012.

[11] D. R. Denslin Brabin, Dr. V. Sadasivam, “QET based steganography technique for JPEG images,” *IEEE International Conference on Control, Automation, Communication and Energy Conservation*, ISBN 978-1-4244-4789-3, 2009.

[12] Discrete Fourier Transform (DFT). Available at: http://in.mathworks.com/help/matlab/math/discrete-fourier-transform-dft.html

[13] A. Cheddad, J. Condell, K. Curran, and P. Mc Kevitt, “Digital image steganography: Survey and analysis of current methods,” *Signal Processing*, vol. 90, issue 3, pp. 727–752, 2010.

[14] J. J. Eggers, R. Baeuml, and B. Girod, “Communications approach to image steganography,” In *Security and Watermarking of Multimedia Contents IV*, vol. 4675 of *SPIE*, pp. 26–37, 2002.

[15] M. S. Subhedar and V. H. Mankar, ‘Current status and key issues in image steganography: A survey,” *Computer Science Review*, 13-14, pp. 95–113, 2014.

[16] C. H. Yang, C. Y. Weng, S. J. Wang, and H. M. Sun, “Adaptive data hiding in edge areas of images with spatial LSB domain systems,’ *IEEE Transactions on Information Forensics and Security*, vol. 3, issue 3, pp. 488–497, 2008.

[17] B. Lei, F. Zhou, E.L. Tanc, D. Ni, H. Lei, S. Chen, and T. Wang, “Optimal and secure audio watermarking scheme based on self-adaptive particle swarm optimization and quaternion wavelet transform,” *Signal Processing*, vol. 113, pp. 80–94.

[18] M. Khalil and A. Adib, ‘Audio watermarking with high embedding capacity based on multiple access techniques,” *Digital Signal Processing*, vol. 34, pp. 116–125, 2014.

[19] G. A. Papakostas, E. D. Tsougenis, and D. E. Koulouriotis, “Moment-based local image watermarking via genetic optimization,” *Applied Mathematics and Computation*, vol. 227, pp. 222–236, 2014.

[20] M. Ali, C.W. Ahn, M. Pant, and P. Siarry, “An image watermarking scheme in wavelet domain with optimized compensation of singular value decomposition via artificial bee colony,” *Information Sciences*, vol. 301, pp. 44–60, 2015.

[21] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, issue 4, pp. 600–612, 2004.

[22] M. Khalil and A. Adib, “Audio watermarking with high embedding capacity based on multiple access techniques,” *Digital Signal Processing*, vol. 34, pp. 116–125, 2014.

[23] M. Ali and C.W. Ahn, “An optimized watermarking technique based on self adaptive DE in DWT–SVD transform domain,” *Signal Processing*, vol. 94, pp. 545–556, 2014.

[24] M. Ali, C. W. Ahn, and P. Siarry, “Differential evolution algorithm for the selection of optimal scaling factors in image watermarking,” *Engineering Applications of Artificial Intelligence*, vol. 31, pp. 15–26, 2014.