Encrypted Reversible Data Hiding on Compressed Image

Akshay Jain
Research Scholar
Vishwakarma Institute of Technology, Pune

Prof. Dipak Pawar
Assistant Professor,
Vishwakarma Institute of Technology, Pune

ABSTRACT

The paper presents a new approach in Image steganography. Information security is the important research field. Steganography process hides message into cover file and forms a stego file. In image steganography there is a need of method which will increase the security, reduce the distortion in the stego file and recovers the data without any loss. In the era of multimedia and internet there is need of reducing time for transmission. The proposed approach is combination of compression, data hiding technique and encryption. To make the transmission and storage of digital data faster, lossy compression is used. On the compressed image data hiding is done. The stego image is encrypted using AES to ensure user authentication. If the receiver has encryption key and data hiding key then only he can obtain the secret message. Nodes are selected randomly in data hiding stage. On the randomly selected nodes lossless LSB steganography is used.

Keywords

Vector quantization, Reversible data hiding, Encryption

1. INTRODUCTION

Reversible data hiding means embedding data into digital image. The data should be recovered without any error and the quality of degradation of the image after data embedding should be low. The rapid progress in computer network has spread the digital data over the world. Steganography is a technique to hide data into some cover media such as image, audio, video, text and such stego cover file is transferred over the network. The steganography is based on 3 basic facts security, robustness and embedding capacity. In reversible data hiding much work is done for authentication. Increasing the embedding capacity while maintaining the quality is the major objective. The main goal of steganography is to make sure that the intruder should not be able to detect that the secret message is hidden in cover media. The stego cover file quality must remain intact.

All digital multimedia formats can be used to hide message. But the formats with high redundancy are suitable. Image contains large amount of redundant data. So the digital image steganography is a best way to protect sensitive information like authentication details. When we use image as cover file it is known as image steganography. In medical field the patient's details can be embedded within image so reducing the transmission time. This approach can be used in applications like military applications or in applications where secret communication is required. [1]

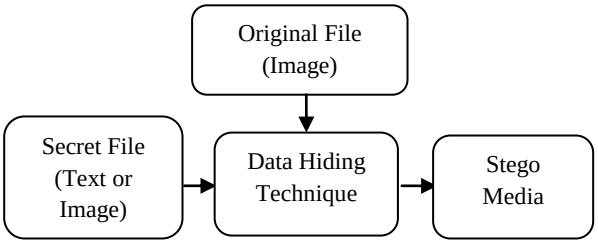

Fig 1: Overview of Steganography Process

The overview of steganography process is shown in Fig 1.In image steganography there are different domains spatial, transform and compression. In Spatial domain techniques the intensities of the original image are used directly to hide data. In transform domain the original pixels are converted to frequency coefficients and then message is embedded. In compression domain, first the image will be compressed; on compressed image data is hidden. [2]

For efficient transmission of digital data over low bandwidth channel there is need of compression. The basic goal of compression is to represent data with minimum number of bits with acceptable image quality. Compression is achieved by removing redundancy. The Image compression is classified into 2 type's lossless image compression and lossy image compression. In lossless image compression the original data can be reconstructed without any loss. In lossy compression some part of data is lost in the quantization process.

Encryption ensures strong user authentication. To provide more security to data hidden encryption is done on the compressed stego image. The sequence in which compression, encryption and steganography are done has greater impact on embedding and compression ratio. The proposed approach is combination of first compression followed by data hiding then encryption. The proposed method uses vector quantization for compression, AES-128 bit for encryption and random selection based data hiding. A novel approach combination of compression, data hiding and encryption increases security and recovers the data without any loss.

The rest of the paper is organized as follows. Section 2 describes the literature work. In section 3 describes proposed methodology and experimental results in section 4.

2. LITERATURE SURVEY

The literature survey has been done on 3 phases namely compression, encryption and steganography.

2.1 Survey on compression [3]

Compression is classified into lossless and lossy. For high compression ratio with acceptable distortion lossy compression is used.

Characteristics of compression

1. Coding redundancy
2. Interpixel redundancy
3. Psychovisual redundancy

The compression techniques are classified as follows

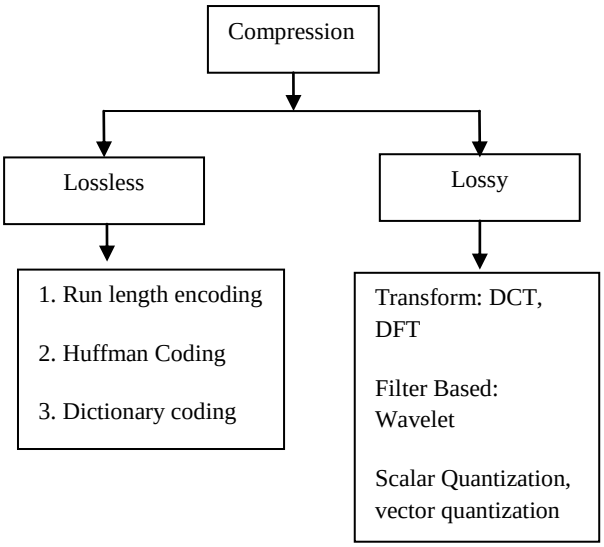

Fig 2: Compression techniques

In lossless compression percentage of compression is less compared to lossy method

2.2 Survey on steganography

The data hiding schemes are classified as follows. The data hiding schemes in spatial domain are generally classified as LSB Substitution [4, 5], pixel Value Differencing [6], Kekre [7], LSB Matching [8], Adaptive LSB [9], Edge based embedding method, and pixel intensity based method. In all spatial domain methods least significant bit (LSB) is the most popular data hiding technique.

2.3 Survey on encryption

Encryption is categorized into public key and private key encryption. Public key encryption uses a private key and public key. Private Key encryption uses secret keys that need to be shared between two entities. Public key encryption methods like RSA, ECC takes longer time for image encryption. [10] Private Key encryption methods like AES, DES, and triple DES are widely used. AES is most secure and fast algorithm. AES is proved to be best algorithm by National Institute of standards and Technology (NIST). [11]

3. PROPOSED METHODOLOGY

The proposed methods consist of 3 phases. First compression is applied to reduce transmission time. In second phase data hiding is accomplished and then encryption algorithm is applied to increase security. The sequence in which compression and encryption are applied has greater impact on the compression ratio. The compression, steganography and then encryption have best results. The main objectives of this work are to increase the security, increasing embedding capacity and lossless recovery of data. The algorithms in each phase are discussed as follows. The Fig 3 shows the different phases in the proposed methodology.

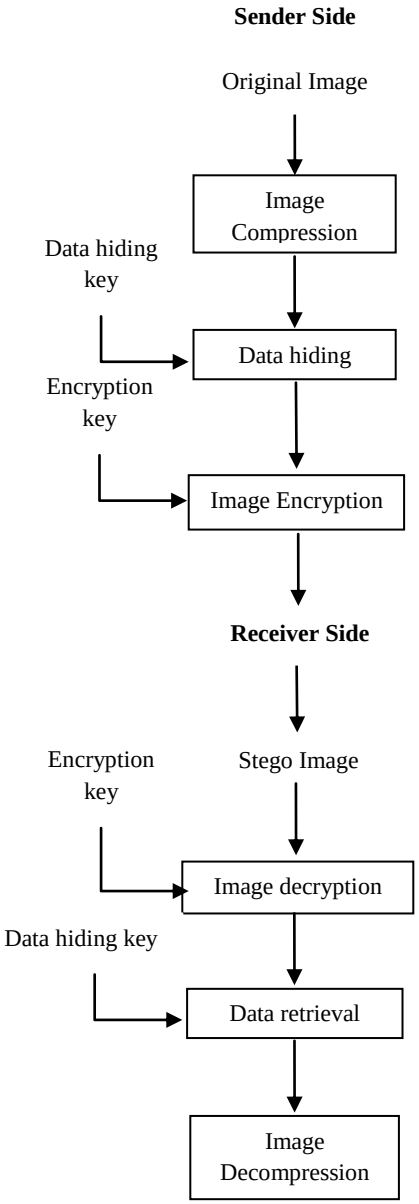

Fig 3: Proposed Approach

3.1 Data Compression

In lossy compression vector quantization is widely used quantization technique. In vector quantization codebook is essential part of image quality. In vector quantization there are 2 parts namely encoding method and decoding method. At the encoder, an input image is divided into blocks called as input vectors. The lowest distorted code vector in the codebook is found for each input vector. The closest code vector is found

by using squared Euclidean distance. The code vector which has minimum squared Euclidean distance with input vector is chosen. The corresponding index associated with the searched code vector is transmitted to the decoder. Compression is achieved as we are replacing image blocks with the index of the closest code vector. [12]

The important task in vector quantization is to generate a good codebook. To generate codebook, generalized Lloyd's algorithm is proposed by Linde, Buzo, Gray referred as LBG algorithm. Advantage of compression is that it makes the original data random and reduces the transmission time.

The vector quantization works as shown in Fig 4

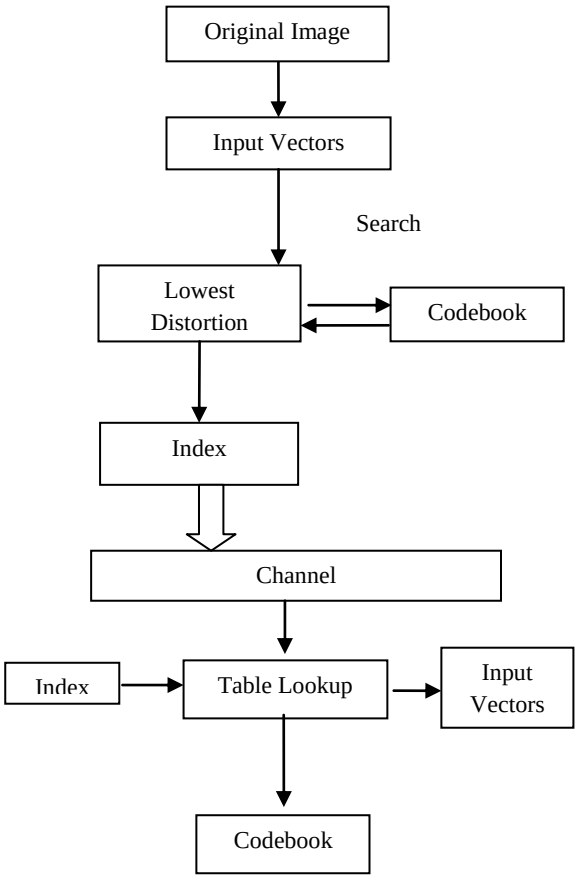

Fig 4: Vector Quantization Method

The LBG algorithm works as follows.

Linde Buzo Gray (LBG) algorithm [13, 14, 15]

Consider Training set Consist of *M k*-dimensional vectors as input to design codebook

$$X_M = x_1, x_2, \dots\dots, x_M$$

Codebook of *N* Code vectors

$$CodeBook_N = CV_1, CV_2, \dots\dots, CV_N$$

$0 < \epsilon < 1$: precision of the optimization process

The LBG algorithm is an iterative method for codebook generation. For codebook initialization there are 3 ways. Randomly selecting code vectors from image, by splitting method and by pair wise nearest neighbor method. Codebook initialization with splitting method is described here.

The steps in LBG algorithm are described as follows

1. Initial code vector in codebook is set as the average of the entire training sequence.
2. For each training vector calculate the closest code vector according to squared Euclidean distance. Add that training vector into the corresponding cluster of the closest code vector found.
3. Split the code vector into 2
$$CV1 = (1 + \epsilon)CV$$
$$CV2 = (1 + \epsilon)CV$$
4. Calculate distortion between training vector and code vector of the associated cluster by using following equation
$$Distortion = \frac{1}{K}\sum_{i=1}^{K} ED_i$$
Where ED_i is the square of the Euclidean distance between the code vector and the i[th] training vector and K is the number of training vectors in the associated cluster.
5. Split Training vectors into two clusters depending on the difference between the training sequence and code vectors.
6. Obtain centroid for new cluster and replace old centroid by new centroid.
7. Compute the difference between training sequence and new centroid and let that difference be $Distortion'$.
8. if $(Distortion^1 - Distortion) / Distortion^1 < \epsilon$ then go to next step otherwise go to step 3
9. Calculate steps 3 to 8 until desired number of code vectors are obtained.

3.2 Data hiding using random node selection

The proposed data hiding approach randomly selects the nodes for steganography and then apply lossless LSB steganography on that random node. The random nodes are selected dynamically by using mathematical function. So this method helps to hide data randomly and also to reduce distortion. Because of randomness in data hiding it increases security. So this method achieves high imperceptibility.

LSB steganography is of 2 types lossy and lossless. In lossy LSB steganography only MSB bits of secret data are embedded into original cover media file, LSB bits of secret message are not embedded so it is lossy method. In lossless

LSB steganography LSB as well as MSB bits of secret message are embedded into original cover media.

Data hiding method works as follows

Sender Side

1. Select the random node generated from the mathematical function.
2. Add the secret message into the LSB of random node

Receiver side

1. Select the stego image
2. Select the random node generated from the mathematical function.
3. Recover the secret message, image.

The data hiding scheme is shown in Fig 5.

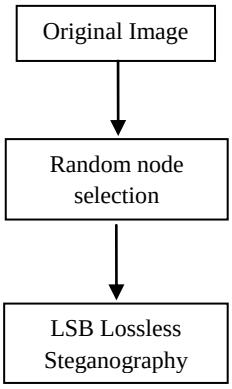

Fig 5: Data Hiding Scheme

3.3 Image Encryption

The AES algorithm is a symmetric block cipher.AES has a fixed block size of 128 bits and a key size used is 128 bits. AES operates on 4*4 array of bytes termed as state, on which basic operations of AES are performed. 10 cycles of repetitions are required for transforming plain text into cipher text. The steps of AES algorithm are as follows

1. Key Scheduling: Expand the 16 byte key to get the actual key block.
2. 16 byte fixed block called as state
3. XOR the state with fixed block

For each round the following operations are performed

1. Substitute bytes: Apply S-box operation to each of the fixed block.
2. Shift rows : Rotate state by k bytes
3. Perform Mix columns operation
4. Add round key

These transformations are applied for 9 rounds. In the 10th round mix column operation is not applied**.** Reverse rounds are applied to convert cipher text into fixed block using same encryption key.

Security of AES against attacks
AES is proved to be best algorithm by National Institute of standards and Technology (NIST). It takes 1 billion billion years to crack 128 bit AES using brute force attack.

4. EXPERIMENTAL RESULTS

Matlab tool is used for experimentation. For conducting the experiment 100 color bmp images of size 256*256 and 512*512 are used. The data that can be embedded is of type text, image. In compression code vectors and input vectors are of dimension 2*2. Maximum data can be embedded is half the size of original cover file.

Table 1: Performance calculation of proposed approach using combination of compression, encryption and steganography

Performance parameter	Lena (256*256*3 bmp)	Peppers (256*256*3 bmp)	Baboon (256*256*3 bmp)
PSNR (dB)	28.0520	25.2534	26.6476
MSE	101.8319	193.9748	140.7098
Compression Percentage (%)	75	75	75

Table 2: Performance calculation using combination of encryption and steganography

Image (256*256*3 bmp)	PSNR (dB)	Payload Size (bytes)
Lena	76.3393	70
Bird	63.2875	102
Airplane	58.3892	246
Penguin	52.2589	822
Tiger	45.5176	3126
Baboon	45.9413	3126
Peppers	45.9913	3126
Jet	39.4885	12342
Aero plane	38.6252	12342
House	34.4817	49152

Image (512*512*3 bmp)	PSNR (dB)	Payload Size (bytes)
Lena	49.1548	3126
Baboon	36.6567	49206
Peppers	30.4705	196,662

Table 3: Performance of previous approach [16]

Image (512*512)	PSNR (dB)	Payload Size (bytes)
Lena	37.9	32 bytes

The comparison of table 2 and table 3 shows that by using combination of steganography and encryption, the embedding capacity of proposed approach is much better compared to previous approach [16]

Table 4: Embedding capacity comparison of proposed approach with previous approach [17]

	Proposed Method			Previous approach [18]		
Image Dimension	256*256	512*512	1024*1024	256*256	512*512	1024*1024
Maximum Secret data size (bytes)	16384	65536	262144	11625	54540	234058

The embedding capacity of proposed approach is more than existing approach

Table 5: File size difference in proposed approach

Sender Side

Original image size (bytes)	Compressed image size (bytes)	Secret data (bytes)	Stego file size (bytes)	Encrypted file size (bytes)
196,662	49,206	4096	49,206	49,206

Receiver Side

Decrypted File (bytes)	Stego file (bytes)	Recovered file size (bytes)	Decompressed file size (bytes)
49,206	4096	49,206	196,662

5. CONCLUSION

This work in image steganography solves the problem of less embedding capacity. The results show that embedding capacity compared with existing approaches is more while maintaining the quality. This work has increased the embedding capacity, security with acceptable distortion. Compression has reduced the transmission time and AES Encryption reduces the possibility of extracting data.

This technique can be applied on video, audio also. The compression techniques having high compression ratio can be used. Combination of transform and quantization can be used to improve compression percentage.

6. REFERENCES

[1] Abbas cheddad, "Digital image steganography : Survey analysis and current methods", signal processing 90 (2010) 727-752

[2] Babloo Saha and Shuchi Sharma, "Steganographic Techniques of Data Hiding using Digital Images", Defence Science Journal, Vol. 62, No. 1, January 2012,pp. 11-18

[3] Suryendra Kumar, "Image compression techniques for medical images: a review", IJREAS, Volume 2,Issue 2,February 2012

[4] Shaveta chutani, Himani Goyal, "LSB embedding in spatial domain –A review of improved techniques", International Journal of Computers & Technology Volume 3, No. 1, Aug 2012

[5] Chi-Kwong Chan, L.M. Cheng, "Hiding Data in images by LSB substitution", Department of Computer Engineering and Information Technology, City University of Hong kong, 11 august 2003

[6] Da-Chun Wu, "A Steganographic method for images by pixel-value differencing", Pattern Recognition Letters, December 2002

[7] H. B. Kekre, Archana Athawale, Pallavi N. Halarnkar, "Performance Evaluation of Pixel Value Differencing and Kekre's Modified Algorithm for Information Hiding in Images", International Conference on Advances in Computing, Communication and Control, pp 342-346, 2009

[8] Ling Xi, Xijian Ping, Tao Zhang, "Improved LSB Matching Steganography Resisting Histogram Attacks",IEEE,2010

[9] Hengfu Yang, Et Al., "A High-Capacity Image Data hiding Scheme Using Adaptive LSB Substitution", Radio Engineering, Vol.18, No 4, December 2009

[10] Abhinav Srivastava, "A Survey Report On Different Techniques Of Image Encryption", International Journal Of Emerging Technology And Advanced Engineering, 2012

[11] G. H. Karimian, B. Rashidi, and A. farmani, "A High Speed and Low Power Image Encryption with 128-Bit AES Algorithm", International Journal of Computer and Electrical Engineering, Vol. 4, No. 3, June 2012

[12] Linde, Y., Buzo, A., Gray R.M., "An algorithm for vector quantizer design ",IEEE Transactions on Communications COM-28, 84–95 (1980)

[13] N. M. Nasrabadi and R. A. King, "Image coding using vector quantization: A review", IEEE Transactions on Communications, vol. 36, pp.84-95, 1980.

[14] Gray, R.M. ,"Vector quantization", IEEE ASSP Magazine, pp. 4–29, (1984)

[15] Bang Huang," An Improved LBG Algorithm for Image Vector Quantization", IEEE, 2010

[16] Xinpeng Zhang, "Reversible data hiding in encrypted image", IEEE Signal processing letters, Vol. 18, No. 4, April 2011

[17] Morteza Bashardoost, "Enhanced LSB image Steganography method by using knight Tour algorithm, Vigenere Encryption and LZW compression", IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 2, No 1, March 2013