*IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)*
*e-ISSN: 2278-2834, p-ISSN: 2278-8735*
*PP 37-44*
*www.iosrjournals.org*

# GA Based Reversible Data Hiding in Encrypted Images by Reserving Room before Encryption

Patil K.U.[1]&Nandwalkar B.R.[2]
[1,2](Comp. Engg. Dept., GNS COENashik, SPP Univ., Pune(MS), India)

***Abstract*—***Information Security has become important aspect now-a-days. For hiding secret information in images, there are various steganographic techniques. Some of them are more complex than others & all of them have strong & weak points. Algorithm for information hiding is design to obtain the three important parameters capacity, robustness & invisibility. This paper propose a novel reversible data hiding technique for digital color image. This technique guarantees effective retrieval of hidden data in color image without degradation in the quality of color image. The propose method makes use of Least Significant Bit (LSB) of three color channels that is red, green, and blue for embedding the secret data. The cryptographic methods are used to encrypt & decrypt the data to increase the security. The Genetic Algorithm (GA) is used for selection of optimized regions for hiding multiple data files.*

***Keywords*-***Reversible Data Hiding, Genetic Algorithm, Steganography, Cryptography, LSB Substitution.*

## 1. INTRODUCTION

In recent years, information Security domain has attracted considerable research interest. To overcome this problem, a technique of cryptography was created for securing the secrecy of communication. In order to keep the data secret, many different methods for encryption & decryption of data have been developed. Unfortunately, keeping only the contents of a message secret is not the solution to this problem. There is requirement keeping the existence of message secret. The technique used to implement this is called steganography. Steganography is useful technique for information hiding. The aim of steganography is to incorporate secret data into digital cover media like images, video, audio, etc. file formats. Several characteristics of steganography like invisibility, storage capacity & resilience against attack make it an efficient data hiding technology. The data of any format like text, image, audio, video, etc. can be hided using steganography.

Data hiding techniques are generally divided into two groups: Spatial and Frequency Domain. The first group makes use of Least Significant Bit (LSB) of image pixel for embedding the message. The second group makes the use of frequency coefficients of image for embedding purpose. In case of reversible data hiding, the lossless recovery of original cover after the extraction of embedded data is an important aspect. The application of genetic algorithm & Cryptography in steganography can significantly increase the capacity or imperceptibility of information.

In this paper, reversible data hiding using genetic algorithm is propose. Genetic algorithm is used for selecting the embedding regions according to the number and size of secret data and the optimized threshold values of regions. Here multiple data files are hidden in the cover image. The digital images are considered as cover image. The LSB of three color channels that is RGB in the given digital image is used for embedding the secret message.

## 2. LITERATURESURVEY

W. Zhang, B. Chen, and N. Yu have proposed a system which uses a decompression algorithm for embedding the data and proved that using this construction they can achieve the rate-distortion bound as long as the compression algorithm reaches the entropy. This establishes equivalence between data compression and RDH for binary covers. By using the binary codes they have improved three RDH Schemes that used binary features sequence as covers. The advantages of the system are: 1. This System reduces the embedding distortion. 2. This System improves RDH Schemes for spatial, JPEG and binary image. The disadvantages of the system are: 1. In this system only two simple methods are used to modify histogram. 2. This system has not considered gray scale covers for designing recursive codes [2].

J. Fridrich, M. Goljan, and D. Rui have proposed a system having general framework for RDH. They are the first to use extracted compressible features of original cover. In this system by compressing proper bit-planes that have minimum redundancy, space to hide data is created. If the image is not noisy, lowest bit-planes which offers lossless compression can be used. In completely noisy image some bit-planes have strong correlation. These big planes are used to find room space to store the hash. The advantages of the system are 1. High Capacity.2.High Security Level. 3. Can be applied for authentication purpose of JPEG file, audio file, digitized hologram etc. The disadvantages of the system are 1. Noisy images forces to embed information in higher bit-plane. 2. Single bit-plane in small image does not offer enough space to hide hash. 3. Capacity is not high enough to embed large payload [3].

J. Tian has proposed a system which uses difference expansion method for embedding data. This system uses the features which are compressed by expansion i.e. the differences between two neighboring pixels. Some differences are selected for expansion by one bit i.e. the difference is multiplied by 2. Thus, LSB's of the differences are all zero and this LSB's can be used for embedding messages. The advantages of the system are: 1. Use of compression and decompression causes no loss of data. 2. This system is also applicable to audio and video data. 3. The compressed location map and changeable bit streams of different numbers are encrypted which increases the security. The disadvantages of the system are: 1. as there is division by 2 there may be some round off errors. 2. Depends largely on the smoothness of natural image that's why can't be applied to images who's capacity is zero or very low. 3. Degradation of visual quality due to bit replacements [4].

Z. Ni, Y. Shi, N. Ansari, and S. Wei have proposed a system which uses histogram shift strategy for RDH. In this system, the space is saved for embedding the data by shifting the bins of histogram of gray values. The authors make use of zero point and a peak point of given image histogram to embed messages. In this system, the embedding capacity is the number of pixels with peak point. For embedding, the whole image is searched for peak point. The advantages of the system are: 1. Simple to implement. 2. Constant PSNR of 48.0dB is obtained. 3. Distortions are quite invisible. 4. Capacity is high. The disadvantages of the system are: 1. Capacity is limited by the frequency of peak-pixel values in histogram. 2. Time consuming as image is searched several times [5].

X. L. Li, B. Yang, and T. Y. Zeng have used a hybrid algorithm which makes the combination of three techniques of PEE (i.e. Prediction-error Expansion), adaptive embedding and pixel selection. In the proposed system, depending on the threshold values the image pixels is divided into two parts. Then the pixels are selected depending on their capacity-parameter and threshold. The smooth pixels are selected from two parts. Finally, data is embedded by modifying the histograms that are derived from selected pixels. The advantages of the system are: 1. By decreasing the modifications to pixel values, the system reduces the embedding impact. 2. More sharply distributed prediction-error histogram can be obtained. 3. The visual quality of watermarked image is greatly improved [6].

L. Luo et al. have used an interpolation technique for developing their reversible image watermarking system. This system can embed a large amount of converted data into images with imperceptible modification. The interpolation errors which are residuals of this technique have greater decorrelation ability. The highly efficient reversible watermarking scheme is developed by applying additive expansion to these interpolation-errors. The advantages of the system are: 1. High image quality. 2. Greater embedding capacity. 3. Less Computational cost [7].

G. Xuan, J. Chen, J. Zhu, Y.Q. Shi, Z. Ni, and W. Su have proposed a integer wavelet transform based lossless data hiding technique. This system hides the authentication information. For preventing gray scale overflowing during data embedding, the histogram modification or integer modulo addition techniques are used. This method uses second-generation wavelet transform IWT. The information is hided into middle bit-plane and in the high frequency sub-bands respectively. This makes the watermarked image greatly as same as the original image. Also the PSNR value is increased. The advantages of the system are: 1. High embedding capacity. 2. Security level is raised due to the use of secret key during embedding of data. The disadvantages of the system are: 1. only gray scale mapping is done. 2. Often multiple bit planes are needed to have enough space [8].

In this paper, V. Sachnev, H. J. Kim, J. Nam, S. Suresh, and Y.-Q. Shi have proposed a system which gives reversible or lossless watermarking for image without using a location map. In this system data embedding depends on the prediction errors. Prediction errors based on magnitude of its local variance can be recorded using sorting technique. Using the sorted prediction errors and reduced size location map whenever needed improves the data embedding capacity by decreasing the distortion. The histogram shift significantly reduces the size of location map. The double embedding scheme in this system allows using each pixel for hiding data. The

advantages of the system are: 1. Capacity can be significantly increased. 2. Double embedding scheme is used. 3. Less distortion [9].

M. Johnson, P. Ishwar, V. M. Prabhakaran, D. Schonberg, and K. Ramchandran have tried to first encrypt the data and then compressing it, so that the compressor don't know the encryption key. The encrypted data is compressed using distributed source coding principles. The key will be available only to the decoder. They have shown that encrypted data can be compressed to same rate as that of original unencrypted data could have been. The perfect secrecy and original cover recovery is obtain in this system [10].

In the above methods, the data hiding is done but the size of the data hided is limited. Also, some of the systems have considered gray-scale image as the cover image. So, We are trying to overcome this problems through our system.

## 3. IMPLEMENTATION DETAILS

In this paper, we proposea Genetic Algorithm based reversible data hiding for digital image. Here, we aregoing to propose the system for embedding multiple data files in the digital cover media.

A reversible data hiding is a technique which gives lossless or distortion-free data hiding. Using this technique, not only the security of data is obtained but also the exact data and original image after extraction is obtained. In this system the digital image used for embedding the data files is called cover image. The following diagram shows the data hiding process:

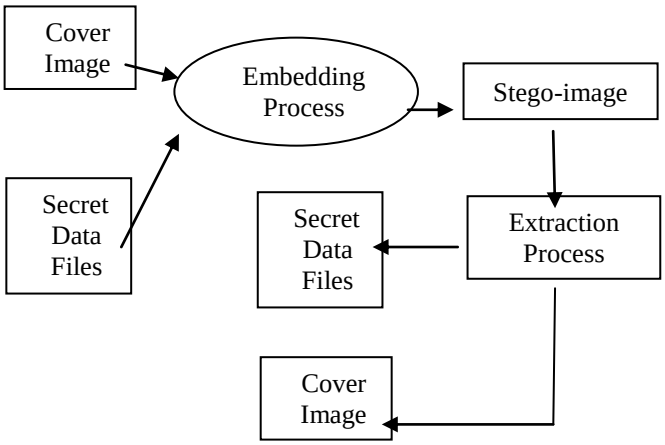


Fig.3.1 reversible data hiding process

### 3.1. PLATFORM: Microsoft Visual Studio 2010

An integrated development environment provided by Microsoft is Microsoft visual studio. Using visual studio we can develop the graphical user interface, web application, web sites etc. like other IDE, visual studio has code editor which also provides syntax highlighting.

### 3.2. Propose system architecture

The propose system architecture is the extension of the existing system architecture. In the propose system Selection of optimized regions block is added. This block is added for obtaining best fitted regions for embedding data files.

Here in the first stage, we are searching for optimized region's using Genetic Algorithm then that regions are reserved for data embedding and using the encryption key the image is converted into encrypted version. The second stage embeds the data files into the encrypted image using the data hiding key and gives marked encrypted image. The third stage has data extraction and image recovery phases. The data is extracted using data hiding key and image is recovered using encryption key.

**Modules:**

1. Regions Selection
2. Encrypted Image Generation

3. Data Hiding In Encrypted Image
4. Data Extraction and Image Recovery

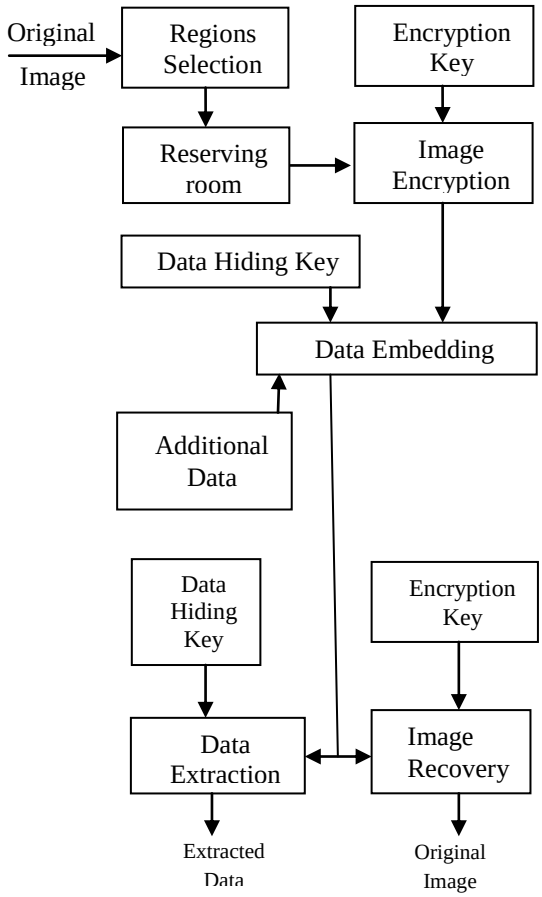


Fig.3.2 Propose system architecture

Modules Description:

1. Regions Selection:
   Input: In this we consider original digital image as the input.
   Algorithm - Genetic Algorithm
   Output: This will give optimized regions for data embedding as the output.'
2. Encrypted Image Generation:
   Input: In this the image passed by first module, encryption key are the inputs.
   Algorithm: Advanced Encryption standard Algorithm
   Output: This module outputs the marked encrypted image.
3. Data Hiding In Encrypted Image:
   Input: In this the marked encrypted image, the data hiding key, the data files to be embedded are taken as inputs.
   Algorithm- L.S.B substitution
   Output: this module gives the data containing encrypted image as the output.

4. Data Extraction and Image Recovery:
   Input: In this the data embedded image, data hiding key and encryption key is taken as input.
   Algorithm- A.E.S Decryption Algorithm

Output: this module gives the original data files and image as output.

3.3. Genetic Algorithm

Genetic Algorithm is a procedure used to find exact or approximate solutions to the optimization and search problems. The basic components of Genetic Algorithm are:

1. Representation of problem to be solved
2. Genetic operators (Selection, crossover, Mutation)
3. Fitness Function
4. Initialization Procedure

In this paper, we are using Genetic Algorithm for selecting the embedding regions in the digital image. The Genetic algorithm will find the best position for data embedding and also optimize the quality of steganographic image. In this optimization we have to manage the four conflicting goals: larger data files, higher image quality, larger data capacity and better robustness.

Genetic Algorithm Operations

The first step to model the problem as a Genetic algorithm problem is to determine the chromosome, G.A operators and fitness functions. The steps for Genetic Algorithm procedures are:

1. Initialization of population

The initial population contains the random strings which are called as chromosomes. The initial population is formed using B * B pixel blocks and binary encoding. In binary encoding the chromosome is the random string of 0, 1 bits.

2. Selection

In the selection procedure the best suitable pair among individuals is selected. The fitness values of the individuals are added to get the fitness. Then the individuals having fitness value greater than 50% are randomly selected.

3. Crossover

Crossover is nothing but a genetic parameter which combines two chromosomes to form a new chromosome. The new chromosome formed has the better qualities than both of the parent chromosomes, if it consumes the best characteristics from each of the parent. New chromosome has combination of properties from both parent chromosomes.

4. Mutation

This often takes place only after the crossover. The mutation depends on the crossover and also on binary encoding. Mutation randomly changes the newly formed chromosome. The mutation randomly changes a bit from 0 to 1 or 1 to 0.

5. Fitness function

The fitness function calculates the fitness value of each individual. The formula for fitness function is

$$F(x) = \frac{1}{m * n} \sum_{i}^{m} \sum_{j}^{n} (Iij - I'ij)^2 \quad (1)$$

Where, m*n is height and width of original cover image, Iij and I'ij are the pixel values of co-ordinates in cover image and rotated image respectively. The algorithm for Genetic Algorithm is as follows:

1. Compute image statistics.
2. Generate an initial population.
3. Compute the fitness value.
4. While not (meets the max. iteration) Do
   i) Select the individuals.
   ii) Generate new population using crossover and mutation.
   iii) Repeat step.1 using new population.
5. Optimized region is selected.
6. Repeat from step.1 for another region selection.

The flowchart of Genetic algorithm is as follows:

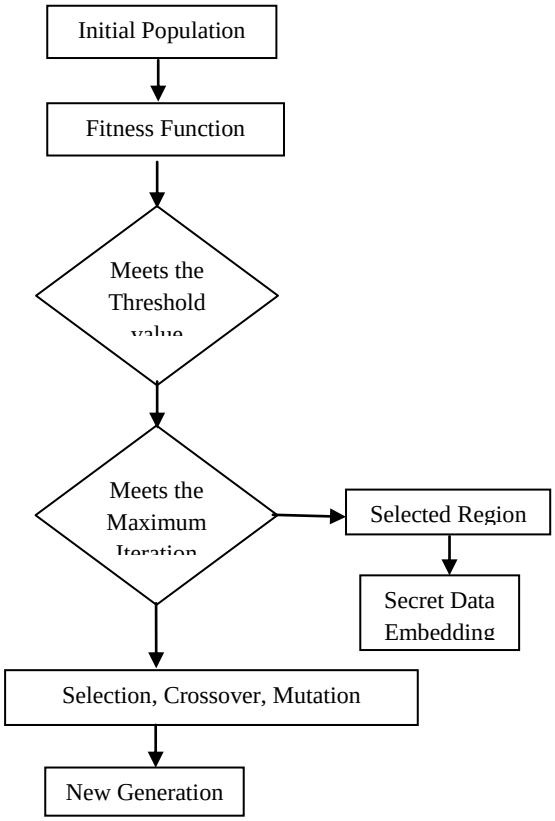


Fig.3.3 Flowchart of Genetic Algorithm

## 4. RESULTS

4.1. Data Used

In this paper digital color image of JPEG format and other formats like bmp, png are used as cover image. The data files used for embedding are of txt, docx, pdf etc. format. The size of image depends on the data size.

4.2. Results

The original image and stego-image are as below:

The figure of data hiding and data recovery are given below:

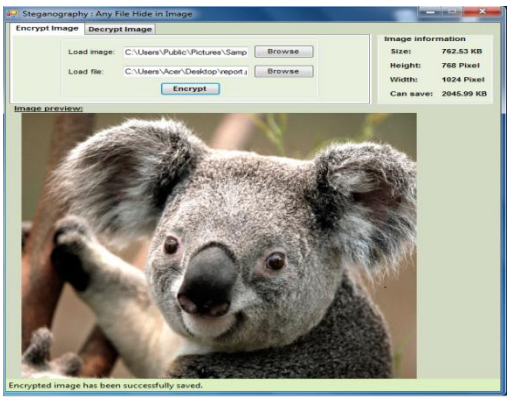


Fig.4.1. Data hiding and image recovery

## 5. CONCLUSION AND FUTURE SCOPE

In this paper, we arepropose a reversible data hiding technique based on Genetic Algorithm (GA) to embed the secret data in color image with high security, imperceptibility and robustness. Due to the selection of optimized region, a good balance between security and image quality is achieved. The data embedding capacity will be significantly increased in this system. Multiple data files will be embedded in multiple optimized regions selected using Genetic Algorithm. The security of the data is much more increased due to the use of two keys: encryption key and data hiding key. The detection of embedded data regions is not easy as regions are not fixed.

In future, we will extend this system considering audio, or video files as the cover. In this paper only digital image is considered as cover.

## ACKNOWLEDMENT

First and foremost, I would like to thank my guide, Prof. B. R. Nandwalkar, for his guidance and support. I will forever remain grateful for the constant support and guidance extended by guide, in making this report. Through our many discussions, he helped me to form and solidify ideas. The invaluable discussions I had with him, the penetrating questions he has put to me and the constant motivation, has all led to the development of this project.

I wish to express my sincere thanks to the Head of department, Prof. N. R. Wankhede, also grateful thanks to the M.E Co-coordinator Prof. Ms. J. V. Shinde and the departmental staff members for their support.

## REFERENCES

[1] Kede Ma, Weiming Zhang, Xianfeng Zhao, Nenghai Yu, and Fenghua Li, “Reversible data hiding in encrypted images by reserving room before encryption,” IEEE Transactions on Information Forensics And Security, vol. 8, no. 3, pp. 553-562, March 2013

[2] W. Zhang, B. Chen, and N. Yu, “Improving various reversible data hiding schemes via optimal codes for binary covers,” IEEE Trans.Image Process., vol. 21, no. 6, pp. 2991–3003, June. 2012.

[3] J. Fridrich, M. Goljan, and D. Rui, “Invertible Authentication”, In Proc. of SPIE Photonics West, Security and Watermarking of Multimedia Contents III, San Jose, California, USA, Vol. 3971, pp. 197-208, January 2001.

[4] J. Tian, “Reversible data embedding using a difference expansion,”IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 8, pp. 890–896,August. 2003.

[5] . Ni, Y. Shi, N. Ansari, and S. Wei, “Reversible data hiding,” IEEE Trans. Circuits Syst. Video Technol., vol. 16, no. 3, pp. 354–362, March.2006.

[6] X. L. Li, B. Yang, and T. Y. Zeng, "Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection," IEEE Trans. Image Process., vol. 20, no. 12, pp. 3524–3533, December.2011.
[7] L. Luo et al., "Reversible image watermarking using interpolation technique," IEEE Trans. Inf. Forensics Security, vol. 5, no. 1, pp. 187–193,March. 2010.
[8] G. Xuan, J. Chen, J. Zhu, Y.Q. Shi, Z. Ni, and W. Su, "Lossless Data hiding based on Integer Wavelet Transform", In Proc. of IEEE International Workshop on Multimedia Signal Processing. Marriott Beach Resort St. Thomas, US Virgin Islands, 9-11 December 2002.
[9] V. Sachnev, H. J. Kim, J. Nam, S. Suresh, and Y.-Q. Shi, "Reversible watermarking algorithm using sorting and prediction," IEEE Trans. Circuits Syst. Video Technol., vol. 19, no. 7, pp. 989–999, July. 2009.
[10] M. Johnson, P. Ishwar, V. M. Prabhakaran, D. Schonberg, and K. Ramchandran, "On compressing encrypted data," IEEE Trans. Signal Process., vol. 52, no. 10, pp. 2992–3006, October. 2004.
[11] K. Hwang and D. Li, "Trusted cloud computing with secure resources and data coloring," IEEE Internet Comput., vol. 14, no.5, pp.14–22, Sep./October.2010.
[12] X. Zhang, "Separable reversible data hiding in encrypted image," IEEE Trans. Inf. Forensics Security, vol. 7, no. 2, pp. 826–832, April. 2012.
[13] Y. Chakrapani and K. SoundaraRajan," Genetic algorithm applied to fractal image compression," in ARPN Journal of Engineering and Applied Sciences , vol. 4, no. 1, February.2009
[14] MedisettyNagendra Kumar and S. Srividya,"Genetic Algorithm Based Color Image Steganogaphy Using Integer Wavelet Transform And Optimal Pixel Adjustment Process," in International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-3, Issue-5, October 2013.
[15] X. Zhang, "Reversible data hiding in encrypted images," IEEE Signal Process. Lett., vol. 18, no. 4, pp. 255–258, April. 2011.
[16] W. Hong, T. Chen, and H.Wu, "An improved reversible data hiding in encrypted images using side match," IEEE Signal Process. Lett., vol. 19, no. 4, pp. 199–202, April. 2012.