RS RESILIENT STEGANOGRAPHY TECHNIQUE WITH GA BASED OPTIMIZATION

[1]Jigyashu Keshri & [2]Sharmila. K. P

Dept. of E&C, CMRIT
E-mail : [1]jigyashurikhil@gmail.com, [2]sharmila.kp@cmrit.ac.in

***Abstract*- With the extensive application of steganography, it is challenged by steganalysis. The most notable steganalysis algorithm is the RS attack which detects the stego-message by the statistic analysis of pixel values. To ensure the security against the RS analysis, we present a new steganography based on genetic algorithm in this project. After embedding the secret message in LSB (least significant bit) of the cover image, the pixel values of the stego-image are modified by the genetic algorithm to keep their statistic characters. Thus, the existence of the secret message is hard to be detected by the RS analysis. Meanwhile, better visual quality can be achieved by the proposed algorithm. The experimental results demonstrate the proposed algorithm's effectiveness in resistance to steganalysis with better visual quality.**

***Keywords* – cryptography, OPAP, Genetic Algorithm, Optimal Pixel Adjustment Process, GA, IWT Integer Wavelet Transform, RS Resilient Steganography, LSB (least significant bit)**

I. INTRODUCTION

The process of sending messages between two parties through a public channel in such a way that it deceives the adversary from realizing the existence of the communication is known as steganography. The ongoing development of computer and network technologies provides an excellent new channel for steganography. Most digital documents contain redundancy. This means that there are parts of documents that can be modified without an impact on their quality. The redundant parts of a document can be identified in many distinct ways. Consider an image. Typically, margins of the image do not convey any significant information and they can be used to hide a secret message. Also, some pixels of the image can be modified to carry a small number of secret bits as small modification (e.g., least significant bit of pixels) will not be noticeable to an unsuspecting user. As the redundant parts of a digital document can be determined in a variety of ways, many steganographic methods can be developed. Mainly, steganography considers methods and techniques that can create covert communication channels for unobtrusive transmission for military purposes.

Steganography is the art of hiding information imperceptibly in a cover medium. The word "Steganography" is of Greek origin and means "covered or hidden writing". The main aim in steganography is to hide the very existence of the message in the cover medium. Steganography and cryptography are counter parts in digital security the obvious advantage of steganography over cryptography is that messages do not attract attention to themselves, to messengers, or to recipients. Also, the last decade has seen an exponential growth in the use of multimedia data over the Internet. These include Digital Images, Audio and Video files. This rise of digital content on the internet has further accelerated the research effort devoted to steganography. The various applications of steganography include secure military communications, multimedia watermarking and fingerprinting applications for authentication purposed to curb the problem of digital piracy. Although these are not perfect applications of steganography, many steganographic algorithms can be employed for these purposes as well.

A. Proposed Technique

The main purpose of the project work is to establish a highly RS-resistant secure model with novel stegano algorithm along with implementation of Genetic algorithm and Integer Wavelet Transform to ensure image security and maintain image quality. The proposed method embeds the message in Discrete Wavelet Transform coefficients based on GA and OPAP algorithm and then applied on the obtained embedded image. The one of the vital goal in this project development is that the data security with the high data hiding capacity and imperceptivity. The both purpose have been considered as dominating and the GA has been implemented for better and more optimized output achievement.

In general, the steganalysis techniques can be categorized into six levels depending on how much information about the hidden messages require. These levels (ordered according to the increased amount of information acquired) are as follows:

- Differentiation between cover and stego documents—this is the first step in steganalys is and the purpose of this technique is to determine if a given document carries a hidden message.
- Identification of steganographic method—this technique identifies the type of steganographic method used and it is the so-called multi-class steganalysis.
- Estimation of the length of a hidden message—this technique reveals the amount of embedded message as the acquired information.
- Identification of stego-bearing pixels—this technique uncovers the exact locations where the pixels are used to carry the message bits.
- Retrieval of stego-key— once the transmitted data which has been already staged reaches to the receiver terminal, and then in order to access the received data a security key is required. This facilitates the authenticity of the data communication. The key is required to access the data. This technique provides access to the stego-bearing pixels as well as the embedding sequence.
- Message extraction— once the data has been embedded then it becomes available for further transmission or communication. When the transmitted data approaches to the receiver terminal then it is required to be extracted so that the text data being transmitted can be retrieved. The process of extracting the text data from the embedded or staged image data is called as message extraction. This technique normally involves extracting and deciphering the hidden message to obtain a meaningful message.

B. Optimized Steganography

The major scopes of this project work are listed below.

- Blind steganalysis: The proposed system has developed a framework in order to distinguish a stego image from a cover image. Mainly, it has been broken several steganographic methods from the literature. This technique uses an image processing technique that extracts sensitive statistical data, which employs a better technique to determine the existence of a secret message. In addition, this technique can be refined and used to detect a different type of steganographic method. This property is important when dealing with an unknown and new steganographic method.
- Use of IWT and GA: The proposed system is extended to determine the best fitness function along with RS analysis to produce the stego image. This is important information that allows an adversary to mount a more specific attack. From the literature review, it can be said that the proposed system is better resilient to statistical attack
- Message length estimation. It has been designed a simple yet effective technique based on first-order statistic to estimate the length of an embedded message. This estimation is crucial and normally is required if it has been intend to extract a hidden message. It has been have identified that the notches and protrusions can be utilized to approximate the degree of image distortion caused by embedding operation. In particular, this technique attacks the steganographic method developed in past.
- Steganographic payload locations identification. It has been presented a technique to identify the locations where hidden message bits are embedded. This technique is one of the very few researches in the literature that is able to extract additional secret information. Eventually, this information is very important for an adversary who wishes to remove a hidden message or deceive communication.
- Enhancement of existing steganalysis techniques. It has been proposed improvement to existing image steganalysis. Specifically, it has been selected and combined several types of features from several existing steganalysis techniques by using a feature selection technique to form a more powerful blind steganalysis. It has been shown that the technique has improved the detection accuracy and also reduced the computational resources. It has been also shown that by minimizing the influence of image content, the detection accuracy can be improved.

II. HIGH LEVEL DESIGN

Design is one of the most important phases of software development. The design is a creative process in which a system organization is established that will satisfy the functional and non-functional system requirements. Large Systems are always are decomposed into sub-systems that provide some related set of services. The output of the design process is a description of the Software architecture. With continued research and improvement in algorithm design, steganography can be taken as a serious means to hide data and the present work appears that it was more efficient in hiding more

data (payload). GA is employed to obtain an optimal mapping function to reduce the error difference between the cover and the stego image and use the block mapping method to preserve local image properties and to reduce the algorithm complexity, and then applied the Optimal Pixel Adjustment Process (OPAP) to increase the hiding capacity of the algorithm in comparison to other systems.

In this high level system design the complete system design and development is to be carried out. The system development with the proper sequence and the synchronization with the all connecting modules will be covered in the process of high level designing. The Genetic algorithm implementation is also one of the important step for the high level system design. In this development process the GA has been used for the RS analysis. This has been done so as to protect the data communication from the RS attacker module. This facility and off course the enhancement will be providing a better applicability and usefulness as compared to the other module developed for the same problem.

A. Design Considerations

The proposed work presents a new steganographic technique in order to embed large amount of data in colored images while keeping the perceptual degradation to a minimum level using Integer to integer wavelet transform (IWT) and Genetic Algorithm. This technique allows hiding an data in an uncompressed color image. Our motivation to hide information in images is to provide security to images that contain confidential information. Our proposed technique is based on LSB technique which will replace more than one bit from each pixel to hide secret data. But the security of the secret message can be enhanced by combining the Least Significant Bit Technique (LSB) and Wavelet Transform (WT). The purpose of the design is to plan the solution of a problem specified by the requirements document. This phase is the first step in moving from problem to the solution domain. The design of the system is perhaps the most critical factor affecting the quality of the software and has a major impact on the later phases, particularly testing and maintenance. The project work is basically an experimental test-bed for evaluation of RS-attack using LSB as well as Genetic Algorithm. So the design to be considered in this project work should be a framework application in Matlab preferably in an integrated development environment considering all the parameters to protect the information using advance steganography.

ASSUMPTIONS AND DEPENDENCIES

- The primary assumption of the project work is that the user is taking the input of original image and not any processed image or manipulated image. The secondary assumption is that the user is expected to use the standard encryption algorithm in a most secure system and network.
- The basic dependency of the project work is to run the application, user needs the Matlab environment and to use application and evaluate its concept, user needs an original image and information in plain text format.

CONSTRAINTS

The application is based on the optimization using genetic algorithm in the current steganographic application. The major constraint here is that it has been found that whenever an image input is subjected to such types of processing, there is a minimal tendency of loss of actual quality of image. In order to resist RS analysis, the effect on the relation of pixels needs to be compensated which may not be achieved by adjusting other bit planes. The implementation may be computational infeasible in practical application. So to overcome this constraint, genetic algorithm is deployed to estimate the better adjusting mode for which the image quality will not be degraded to higher extent.

DEVELOPMENT METHODS

This project would follow Iterative Development Methodology. This would enable the product to be built in increments. Rational Unified Process (RUP) would be the approach to manage the development process of the project. The RUP is not a single concrete prescriptive process, but rather an adaptable process Framework. It encompasses a large number of different activities; it is also intended to be tailored, in the sense of selecting the development processes appropriate to a particular software project or development organization. The RUP is recognized as particularly applicable to larger software development teams working on large projects.

The advantages of the iterative process model are:-

- It is easier to accommodate the changes in requirement at alter stage.
- It is easier to control the risk. The higher risk areas are addressed in the beginning of the project.

Incorporation of the feedback of iteration into subsequent iterations will make the quality of the product better.

- The best practices of Rational Unified process are:-
- Develop software iteratively
- Manage requirements
- Use component based architecture

- Visually model software
- Verify software quality
- Control changes to software

B. System Architecture

The project work ensures the security against the RS analysis and to achieve this, the application should be designed with a strategy to overcome all the limitation considered in the previous research work. The current strategy to design the architecture of the project work depends completely on a robust process of safeguarding the input to the application. This strategy calls for implementing least significant bit for embedding the secret message of the cover image. The next issue which might be encountered is the loss of quality of the image and the planning done for safeguarding the quality of the image which is achieved by implementing Genetic Algorithm. It is a search technique used in computing to find exact or approximate solutions to optimization and search problems.

In the proposed method, the message is embedded on Integer Wavelet Transfonn coefficients based on Genetic Algorithm. Then, OPAP algorithm is applied on the obtained embedded image. We use Genetic Algorithm to fmd a mapping function for all the image blocks. In our GA method, a chromosome is encoded as an array of 64 genes containing permutations 1 to 64 that point to pixel numbers in each block. The main idea of applying OPAP is to minimize the error between the cover and the stego image. In this research work, it is adopted genetic algorithm to search for a best adjustment matrix. Genetic algorithm is a general optimization algorithm. It transforms an optimization or search problem as the process of chromosome evolution. When the best individual is selected after several generations, the optimum or sub-optimum solution is found. The three most important operations of genetic algorithm are reproduction, crossover and mutation.

The below mentioned figure represents the overall system functionalities of the developed algorithm. The overall system function can be summarized by observing the figure mentioned above. The figure represents the real operative steps of the developed design. In the processing the user interface helps so as to provide a user interface to handle the developed model and to access the developed module. At the inception, the cover image is selected where the data is to be embedded. Once the cover image has been selected then the text data or the message is to be selected and then in order to accomplish the motive of steganography the stego key is assigned so that at the other terminal the data can be retrieved by putting the key. Once the Key has been provided, the real application development for the RS analysis will be started with the help of robust GA optimization. In this technique initially the message is to be embedded. Here the GA is playing a vital role for embedding more and more data to the image. In this developed system architecture the integer to integer wavelet transform has been done. Once the data has been embedded into the image file, then after embedding the image is again recovered and then it is now ready to be transmitted over the communication channel. On the other hand at the receiver terminal or the extraction terminal with the accurate assignment of the stego key the data is retrieved accurately.

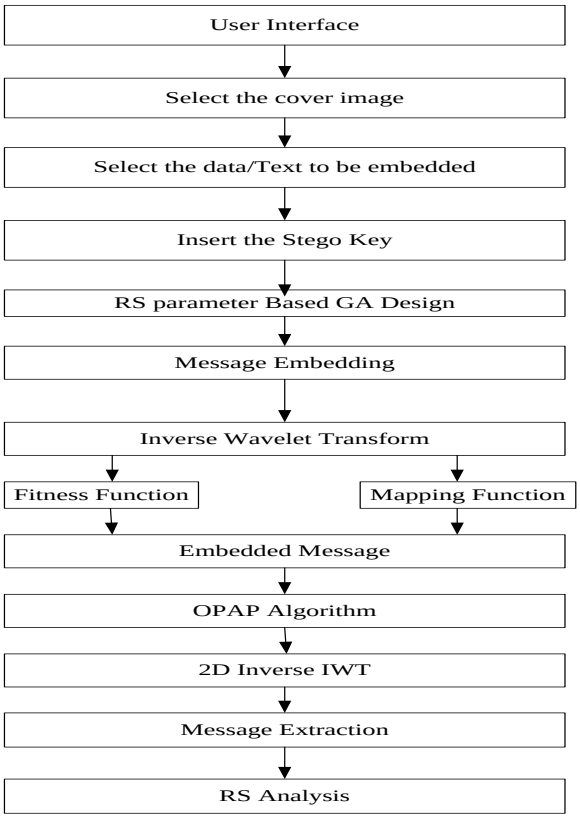

Fig 2.2 : The overall functional flow diagram

III. DETAILED DESIGN

The in-depth study, analysis and development are carried out in the detailed system design part of the presentation. Detailed design of the proposed steganographic project gives in depth picture of the most components described in the system architecture. Meanwhile here we will be discussing the detail design of the system proposed and hence developed. In this section details and flow chart of each module has been described. The control flow is shown by the structure chart, the functional description of which are presented in the flow chart diagrams.

A. Module Specification

The proposed system is designed with two fundamental modules as explained below:

- Embedding Module: The main task of this module is to embed a secret text within the cover colored image using encryption key. The complete cover image is divided into 8x8 blocks before any further processing. The frequency domain representation of the respective created blocks is estimated by two dimensional Integer wavelet transform in order to accomplish 4 sub bands LL1, HL1, LH1, and HH1. 64 genes are generated containing the pixels numbers of each 8x8 blocks as the mapping function. The message bits in 4-LSBs IWT coefficients each pixel according to mapping function are embedded. Based on fitness evaluation, Optimal Pixel Adjustment Process on the Image is applied. Finally, inverse two dimensional integer wavelet transform is computed in this module in order to generate the stego image.

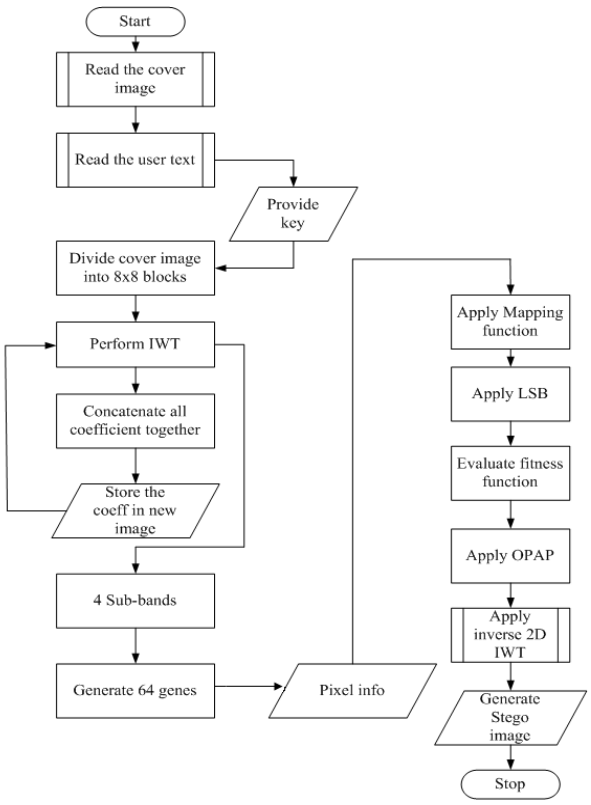

Fig.3.1.1 Flow Chart of the Message Embedding process

The embedding is carried out after segmenting and getting in depth of the pixels-world. The data hiding can be done only when the data is hided in the segmented or space created in the image data.

- Extraction Module: The main task of this module is the extraction of the actual user text information from the stego image to understand the effectiveness of message embedding process. It considers the stego image as input along with key for decrypting the hidden text from the stego image. Once the data has been transmitted over the communication channel and when the receiver receives the embedded image file, then it becomes necessary to again segment the image data and then take out the text data available at the space covered by the text data at the time of message embedding. The extraction can be summarized in a simple sentence as to take out the data that has been embedded.

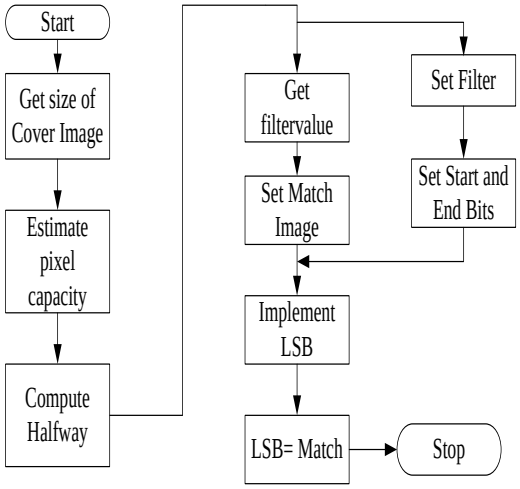

Fig.3.1.2 Flow Chart of the LSB Implementation

B. Structure Chart

A Structure Chart (SC) in software engineering and organizational theory is a chart, which shows the breakdown of the configuration system to the lowest manageable levels.

A genetic algorithm (GA) is a search technique used in computing to find exact or approximate solutions to optimization and search problems. Genetic algorithms are categorized as global search heuristics. Genetic algorithms are a particular class of evolutionary algorithms (EA) that use techniques inspired by evolutionary biology such as inheritance, mutation, selection, and crossover.

The below mentioned figure represents the structural chart representation for the proposed system development. Here it represents the overall processing and the step by step presentation of the project module.

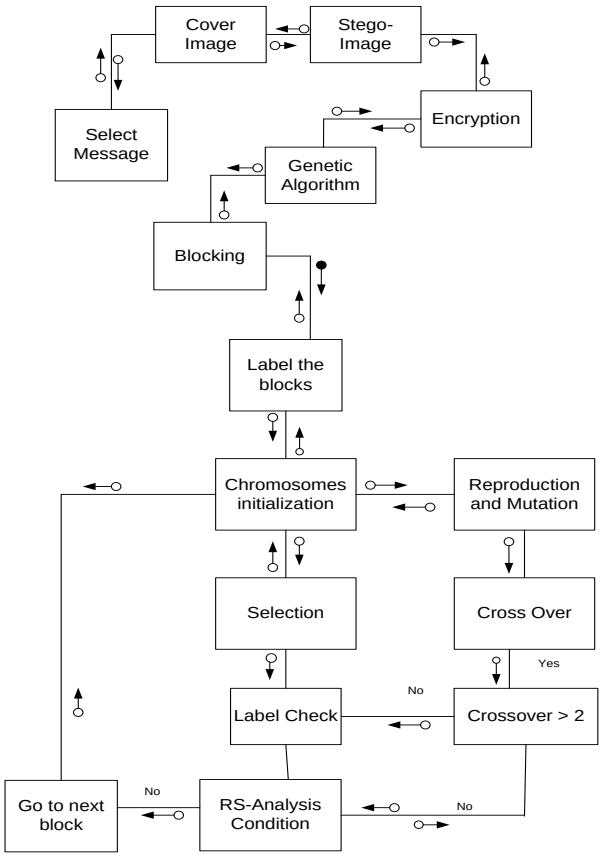

Fig.3.2. Structure Chart of the GA utilization of the proposed system

C. LSB Implementation

This process flow chart will show the section where LSB is implemented. The major operation takes place when the application starts getting the size of the cover image and then it creates a tree structure for ease in computation. It then get all the filter value of the pixels, where the application initialize the filter and configure the start and end bits, which finally set the match image.

After performing this operation, LSB algorithm will be implemented in the cover image, where the pixels values of the steg-image are modified by the genetic algorithm to keep their static characters.

Table 4.1.1 Comparison of PSNR of Images for variant value of K

Cover Image	PSNR			
	K=3	K=4	K=5	K=6
Lena	46.83	39.94	32.04	24.69
Jet	51.88	45.20	37.45	29.31
Boat	48.41	40.44	31.17	23.60
Baboon	47.32	40.34	32.79	24.80

IV. RESULT ANALYSIS

A. Experimental Scenario

The proposed method is applied on 512x512 8-bit grayscale images “Jet”, “Boat”, “Baboon” and “Lena”. The messages are generated randomly with the same length as the maximum hiding capacity. Table I shows the stego image quality by PSNR. Human visual system is unable to distinguish the grayscale images with PSNR more than 36 dB. This project embedded the messages in the k-LSBs, from k=3 to k=6 and received a reasonable PSNR. Table I shows PSNR for variant value of k. Table I presents the results and we can see that for k equal to 4 or 5, we obtain the highest hiding capacity and reasonable visual quality. Therefore, we take k equal to 4 as the number of bits per pixel.

Fig.4.1.1 to 4.1.4 show the original cover image and stego image along with their histogram analyze which is used later to compare to test for imperceptibility.

Fig.4.1.1 Cover image

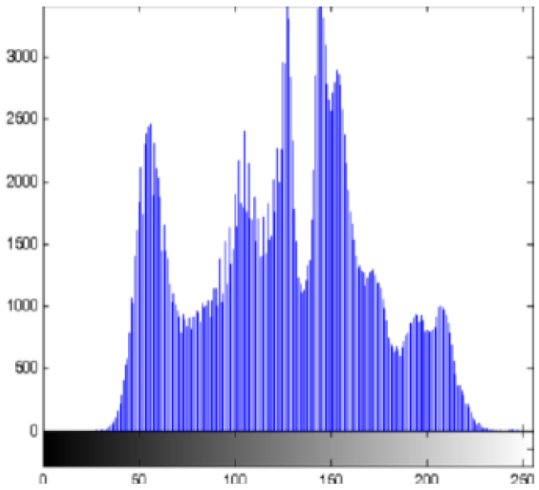

Fig.4.1.4 Stego image Histogram

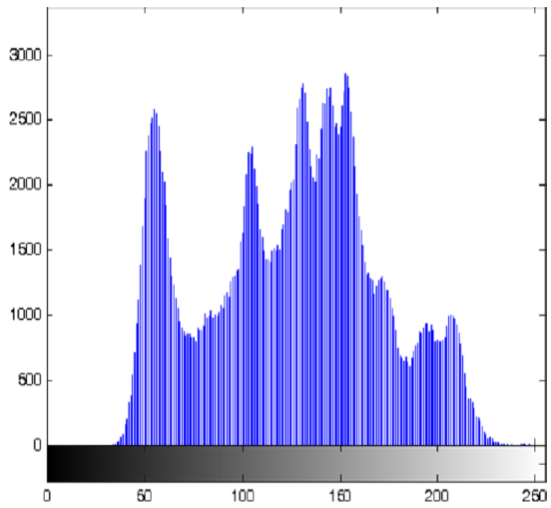

Fig.4.1.2 Cover Image Histogram

Fig.4.1.3 Stego image

B. Result Scenario

Utilization factor is a measure of percentage of LSB of the cover image used to hide secret message. As the Utilization increases, the randomness increases and hence RS ratio.

File name	RS Values before Optimazition			RS values after optimization			PSNR before optimization	PSNR after optimization	Utilization Factor (%)
	Rm-R_m	Sm-S_m	R/S	Rm-R_m	Sm-S_m	R/S			
Lamp	0.00024	0.00071	0.34	0.00955	0.00801	1.19	9.4344	9.5177	0.34485
Flag	0.00094	0.00100	0.94	0.00935	0.00497	1.88	4.453	4.6078	0.71586
Nature	0.00079	0.00330	0.24	0.00245	0.00205	1.20	12.6109	12.7461	0.25482
Road	0.00064	0.00184	0.35	0.00343	0.00272	1.26	7.9458	8.4372	0.22278

The R/S ratio is seen to be lowered and it is approximately lesser or equal to 1.8, which is the accepted value of R/S for an image.

The PSNR value gets lower upon embedding the message. This indicates presence of a secret communication. After optimization, the PSNR value is seen to be increased to some extent.

V. CONCLUSION

This method optimizes localization in which the message or the user specified data is to be embedded on the cover image. This overall system has been designed for steganography that facilitates the data hiding in the image file. Here the text data has been embedded into the image file. The proposed method embeds the message in Discrete Wavelet Transform coefficients based on Genetic algorithm and OPAP algorithm and then applied on the obtained embedded image. Wavelet transform has the capability to offer some information on frequency-time domain simultaneously. This process is repeated for several times and each time a section of the signal is drawn out. HAAR wavelet operates on data by calculating the sums and differences of adjacent elements. This wavelet operates first on adjacent horizontal elements and then on adjacent vertical elements. One nice feature of the HAAR wavelet transform is that the transform is equal to its inverse. Each transform computes the data energy in relocated to the top left hand corner.

In this research, we introduced a novel steganography technique to increase the capacity and the imperceptibility of the image after embedding. GA employed to obtain an optimal mapping function to lessen the error difference between the cover and the stego image and use the block mapping method to preserve the local image properties. Also we applied the OPAP to increase the hiding capacity. However by this method, the computational complexity is high, our results show that capacity and imperceptibility of image have increase simultaneity. Also, we can select the best block size to reduce the computation cost and to increase the PSNR using optimization algorithms such as genetic algorithm. The experimental results show that this method works properly and is considered to give almost the optimum solution.

A. Applications

There are many applications for digital steganography of images, including copyright protection, feature tagging, and secret communications [1, 4].

- **Copyright Protection :** A secret copyright notice or watermark can be embedded inside an image to identify it as intellectual property [5, 6]. This is the

watermarking scenario where the message is the watermark [5, 6]. The "watermark" can be a relatively complicated structure. In addition, when an image is sold or distributed an identification of the recipient and time stamp can be embedded to identify potential pirates.

A watermark can also serve to detect whether the image has been subsequently modified [7]. Detection of an embedded watermark is performed by a statistical, correlation, or similarity test, or by measuring other quantity characteristic to the watermark in a stego-image. The insertion and analysis of watermarks to protect copyrighted material is responsible for the recent surge of interest in digital steganography and data embedding.

- **Feature Tagging :** Captions, annotations, time stamps, and other descriptive elements can be embedded inside an image, such as the names of individuals in a photo or locations in a map. Copying the stego-image also copies all of the embedded features and only parties who possess the decoding stego-key will be able to extract and view the features. In an image database, keywords can be embedded to facilitate search engines. If the image is a frame of a video sequence, timing markers can be embedded in the image for synchronization with audio. The number of times an image has been viewed can be embedded for "pay-per-view" applications.
- **Secret Communications:** In many situations, transmitting a cryptographic message draws unwanted attention. The use of cryptographic technology may be restricted or forbidden by law. However, the use steganography does not advertise covert communication and therefore avoids scrutiny of the sender, message, and recipient. A trade secret, blueprint, or other sensitive information can be transmitted without alerting potential attackers or eavesdroppers.

B. Future Enhancement

Due to limitation of time and constraint of resource, the current project work is restricted to specific functionality only. But in case, such obstruction can be overcomes, the current project work could be extended to following future enhancement.

- A typical method for Steganalysis of the LSB substitution is the histogram attack that attempts to diagnose anomalies in the cover image's histogram. The future enhancement work would be in the direction to work on a new method for image steganography which improves over the LSB image steganography by decreasing the amount of changes made to the perceptual and statistical attributes of the cover image. Some sensitive pixels affecting the signal characteristics can be identified, and then lock and keep them from the extra bit embedding process of the LSB method, by introducing a new embedding key. Evaluation results will be expected to show that, without reducing the embedding capacity, this future idea can decrease potentially detectable changes caused by the embedding process.
- Although the developed algorithm has facilitated a better steganography technique and the outputs of the developed module has also illustrated the better performance as compared to the other existing stego techniques, then while the further development and modification cannot be ignored. In order to enhance the quality and quantity the more optimized computing like IEC (Interactive Evolutionary Computing) computing can be employed. It will provide the better transform and segmentation and this will be performing higher data embedding with higher imperceptibility.
- Here an LSB matching steganographic algorithm based on the principles of genetic algorithms, that aims to reuse the binary image color values in a controlled way so that instead of focusing to change the least significant portion of the color representation (LSBs) can be implemented. The remap of secret data will be done in a manner that reduces the color information loss up to a negligible level. The algorithm can improves the statistical analysis immunity of the steganographic image and at the same time can offer higher PSNR (an average gain of 2,4 dB) than most of the LSB matching algorithms used in our experiments.

VI. REFERENCE

[1] N. Johnson and S. Jajodia, "Exploring steganography: seeing the unseen," IEEE Computer, pp. 26-34, February 1998.

[2] D. Kahn, The Codebreakers, Macmillian, New York, 1967.

[3] B. Norman, Secret Warfare, Acropolis Books, Washington D.C., 1973.

[4] W Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," IBM Systems Journal, Vol. 35, No. 3 and 4, pp. 313-336, 1996.

[5] R. Wolfgang, C. Podilchuk and E. Delp, "Perceptual watermarks for images and video," to appear in the Proceedings of the IEEE, May, 1999. (A copy of this paper is availavle at: http://www.ece.purdue.edu/~ace).

[6] UnZign software: http://altern.org/watermark, 1997.

[7] Stirmark software: http://www.cl.cam.ac.uk/ ~fapp2/ watermarking/stirmark, 1997.

[8] N. Johnson and S. Jajodia, "Steganalysis of images created using current steganography software," Lecture Notes in Computer Science, Vol. 1525, pp. 273-289, 1998.

[9] Steganos Software: http://www.demcom.com/ english/steganos/index.htm

[10] I. Cox and M. Miller, "A review of watermarking and the importance of perceptual modeling," Proceedings of the SPIE/IST&T Conference on Human Vision and Electronic Imaging II , SPIE Vol. 3016, San Jose, CA, pp. 92-99, February 1997.

[11] M. Swanson, B. Zhu, and A. Tewfik, "Robust data hiding for mages," Proceedings of the IEEE DSP Workshop, Leon, Norway, pp. 37-40, Loen, Norway, September 1996.

[12] R. Anderson and F. Petitcolas, "On the limits of steganography," IEEE Journal on Special Areas in Communications, Vol. 16, No. 4, pp. 463-473, May 1998.

[13] J. Smith and B. Comiskey, "Modulation and information hiding in images," Lecture Notes in Computer Science, Vol. 1174, pp. 207-226, 1996.

[14] E Safy, R.O, Zayed. H. H, E Dessouki. A, "An adaptive steganography technique based on integer wavelet transform," ICNM International Conference on Networking and Media Convergence, pp 111-117, 2009.

[15] P. Chen, H. Lin, "A DWT Based Approach for Image Steganography." International Journal of Applied Science and Engineering, Vol. 4, No. 3, pp. 275-290, 2006.

[16] B. Lai and L.Chang, "Adaptive Data Hiding for Images Based on HAAR Discrete Wavelet transform," Lecture Notes in Computer Science, Vol 4319, 2006.

◈◈◈