IMAGE STEGANOGRAPHY USING MODIFIED KEKRE ALGORITHM

Shyam Shukla[1], Aparna Dixit[2]
[1]*Information Technology, M.Tech, MBU, (India)*
[2]*Computer Science, B.Tech, GGSIPU, (India)*

ABSTRACT

The main goal of steganography is to hide information in the other cover media so that other person will not notice the presence of the information. This is a major distinction between this method and the other methods of covert exchange of information because, for example, in cryptography, the individuals notice the information by seeing the coded information but they will not be able to comprehend the information. This paper presents a secured steganography method capable of embedding high volume of information in digital cover image without incurring and perceptual distortion. This method is based compression and encryption. In this method the compression is done by Lempel-Ziv-Welch technique. Encryption and hiding of data is done by kekre algorithm.

Keywords: LSB Embedding Technique, LZW Compression, MSE, PSNR, RMSE, SSIM, Steganography.

I. INTRODUCTION

The word steganography is of Greek origin and means "covered or hidden writing". It is the science of hiding information. The basic structure of Steganography is made up of three components: the “carrier”, the message, and the key1. The carrier can be a painting, a digital image, an mp3, even a TCP/IP packet among other things. It is the object that will ‘carry’ the hidden message. A key is used to decode/decipher/discover the hidden message. This can be anything from a password, a pattern, a black-light, or even lemon juice. Basic steganography diagram is shown in figure 1. In the basic steganographic process, the secret message is hidden into a cover object. The cover object can be any of text, image, audio, video etc. A secret key is also used and the secret message is embedded into the cover object using the secret key. This new message obtained is called stego message. The stego message is transmitted over the public channel. The receiver gets the message and retrieves the message using the stego key which is same as used by the sender. In this way security is achieved by hiding the existence of the message.

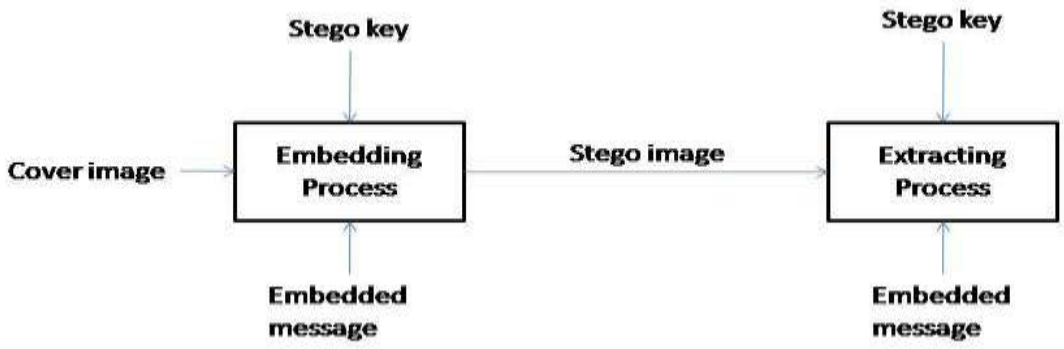

Figure 1- Basic Steganography Process

There have been a large embedding techniques proposed number of steganography in the literature. These techniques modified the cover image with different approaches, Image steganography technique can be divided into two groups:

a) Image domain also called spatial domain

b) Transform domain also called frequency domain *LSB Technique*

Most steganography jobs have been carried out on images, video clips, texts, music and sounds .Nowadays, using a combination of steganography and the other methods, information security has improved considerably. In addition to being used in the covert exchange of information, steganography is used in other grounds such as copyright, preventing e-document forging.

II. MODIFIED KEKRE ALGORITHM

Modified Kekre's Algorithm (MKA) [1] is based on Least Significant Bit (LSB) [2] method. It can be applied on 8 bit gray scale images or 24 bit Red Green Blue (RGB) color image. It uses up to five LSB's of a pixel to embed the data. The number of secret data bits that can be embedded in the pixels depends upon the pixel intensity of the pixels of the cover image. To achieve more security MKA uses 8 bit secret key to perform XOR operation to all the bytes of the secret message. While extracting the message XOR operation is also performed using the same key. The embedding algorithm maintains a matrix of pixels where up to 5 bits of message are used to embed, and this matrix is required while extracting the secret hidden message from stego-image. In Table-1 'x' shows don't care bit whose value can be either '0' or '1'. "Pixel intensity" shows the value of pixel. "Data Bit to Embed" shows current message bit used to embed into the cover-image. Suppose pixel intensity is 245 which exist in the row number 1 and 2 of the Table-1.

Table-1

S. No	Pixel Intensity	Data bit to Embed	Matrix Entry	Utilize Bit/ bits
1	240-255	1	1	5
2	240-255	0	-	4
3	224-239	0	1	5
4	224-239	1	-	3
5	192-223	X	-	2
6	0-191	X	-	1

For embedding the secret data bits its 5 or 4 LSBs can be utilized, depending on current data bit to embed. If current data bit is 0 then 4 LSBs otherwise 5 LSBs are used for embedding data bits. Mark a 1 bit entry into maintained matrix pixel position to identify that if this pixel contains 5 bits of data. Same procedure is applicable for other pixels of the image according to Table 1

Same procedure will be run to extract the data bits of message using maintained matrix because it keeps the track of the pixel position where 5 LSBs are utilized. At the end, 8 bit secret key with XOR operations applied on the extracted message to regenerate original message which was embedded. Hussain [3] discusses a method that is an improvement of MKA. It also applies 8 bit secret key with XOR operation on all bytes of message to change the originality of message. It maintains a matrix for those pixels which will embed 5, 3 and 2 LSBs of data. In Table- 2 "Pixel intensity" shows the value of pixel. "Data Bit to Embed" shows current message bit used to embed into the cover-image. "Matrix Entry" maintains a matrix which denotes the 5 LSB are embedded.

"Utilize Bits" shows the total number of bits embedded into a pixel. If pixel intensity is 33 which exist in row number 7 and 8 of Table-2. For data embedding, 2 or 1 bits of pixel can be utilized depending on current data bit to embed. If current message (want to embed data) bit is 0 then 2 bits otherwise 1 LSB are used for embedding data bits. If this pixel contains 2 bits of data, mark a 1 bit entry into maintained matrix pixel position for identification of extra bits. Same procedure is applicable for other pixels of the image according to Table-2.

Table-2

S. No	Pixel Intensity	Data bit to embed	Matrix Entry	Utilize Bit/ bits
1	240-255	1	1	5
2	240-255	0	-	4
3	224-239	0	1	5
4	224-239	1	-	3
5	192-223	0	1	3
6	192-223	1	-	2
7	32-191	0	1	2
8	32-191	1	-	1
9	16-31	0	1	3
10	16-31	1	-	2
11	0-15	0	1	5
12	0-15	1	-	4

III. LEAST SIGNIFICANT BIT (LSB) TECHNIQUE

Least significant bit (LSB) insertion is a common, simple approach to embedding information in a cover image. The least significant bit (in other words, the 8 bit) of some or all of the bytes inside an image is changed to a bit of the secret message. When using a 24-bit image, a bit of each of the red, green and blue colour components can be used, since they are each represented by a byte. In other words, one can store 3 bits in each pixel. An 800 × 600 pixel image, can thus store a total amount of 1,440,000 bits or 180,000 bytes of embedded data.

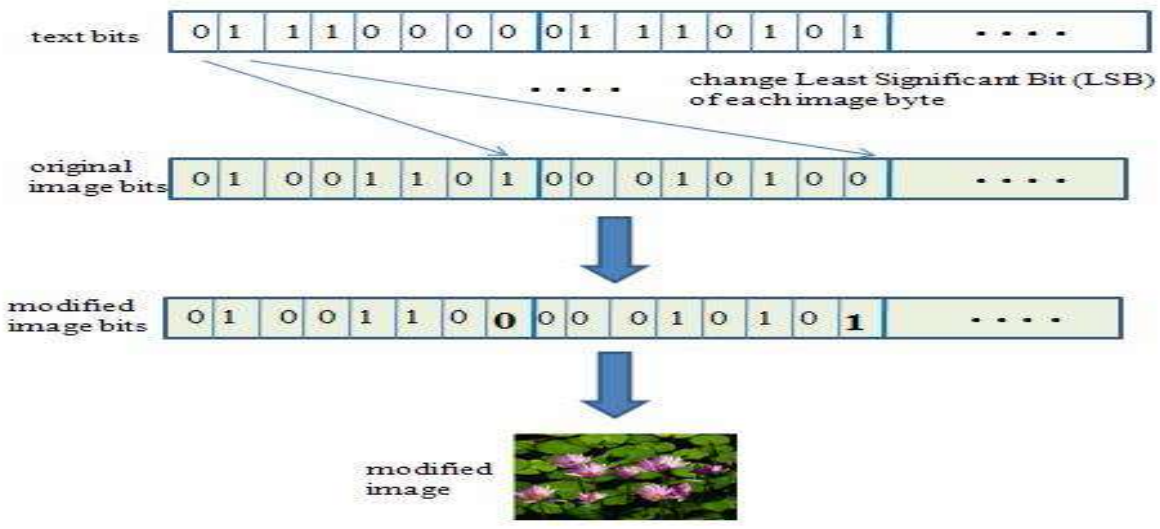

Figure 2- LSB Technique

IV. IMAGE QUALITY METRICS

The image quality metrics are figures of merit used for the evaluation purpose of the image quality. These metrics provide some measures of the closeness between two digital images by exploiting the differences in the statistical distribution of pixel values. The most commonly used quality metrics are:

☐ Mean Square Error (MSE)

☐ Root Mean Square Error (RMSE)

☐ Structural Similarity (SSIM)

☐ Peak Signal to Noise Ratio (PSNR)

4.1. Mean Square Error (MSE)

The mean square error is defined as the square of the difference between the pixel values of the original image and the stego image and then dividing it by size of the image. The mathematical formula for computing mean square error between x and y images of sizes M*N is given below

$$MSE = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} [x(m,n)-y(m,n)]^2$$

The lower value of Mean Square Error (MSE) signifies lesser error in the stego image in other words better quality.

4.2. Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) is calculated by getting the square root of the mean square error (MSE). The RMSE can be calculated as follows.

RMSE =√MSE

4.3. Structural Similarity (SSIM)

The SSIM metric was given by Wang el. This method is used to measure the similarity between two images [4]. It is designed by modeling any image distortion as a combination of three factors that are loss of correlation, luminance distortion and contrast distortion. Mathematically, the SSIM is calculated as follows:

$$\text{Luminance} : l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

$$\text{Contrast} : c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

$$\text{Structure} : s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$

$$SSIM(x,y) = [l(x,y)]^{\alpha}[c(x,y)]^{\beta}[s(x,y)]^{\gamma} = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

4.4. Peak Signal to Noise Ratio (PSNR)

The Peak Signal to Noise Ratio (PSNR) measures the estimates of the quality of stego image compared with an original image and is a very commonly used metric way to measure image reliability or conformity. The mathematical formula to calculate the PSNR value is as follows:

PSNR = 20log10 [MAXPIX / MSE]

Where MAXPIX is the maximum pixel value and MSE is the Mean Square Error.

In PSNR, 'signal' is the original image and 'noise' is the error in the stego image resulting due to encoding and decoding. PSNR is a number that reflects the quality of the stego image and is measured in decibel (dB). Mathematically, PSNR is inversely proportional to the MSE, which implies the lower the value of MSE higher is its PSNR. Thus higher the Peak Signal to Noise Ratio (PSNR) is better.

V. RELATED WORK

Steganography is an area of invisible communication. Steganography used for protecting the unauthorized access of the information but is an ancient technique which is in existence since 440 B.C. The most basic and important image Steganographic Technique is Least Significant Bit [5] embedding technique. In this technique, least significant bit of each pixel is replaced with secret message bit until message end. This way data can be hidden in the least significant bits of the cover image and the human eye would be unable to notice the hidden image in the cover file. This technique can be used for hiding images in 24-bit, 8-bit or gray scale format. But this technique has less embedding capacity and easy to detect. Marvel

[6] discusses spread spectrum image steganography technique. In spread spectrum techniques, hidden data is spread throughout the cover-image making it harder to detect. In spread spectrum image steganography the secret message is embedded in noise and then combined with the cover image which results into the stego image. Since the power of the embedded signal is much lower than the power of the cover image, the embedded image is not perceptible to the human eye or by computer analysis without access to the original image. This technique provides high security than the previous techniques. However, this method does not provide sufficient data payload. Hamid et al. in

[7] discussed a texture based image steganography technique. This technique divides the texture areas into two groups. One is simple texture area and other is complex texture area. In Simple texture area 3 LSB bits of Red channel, 3 LSB bits of Green channel and 2 LSB bits of blue channels are used for embedding the secret data. In Complex texture area data is embedded into the 4 LSB bits of the pixel. This method increases the embedding capacity of the covered image.

[8] *Yang et al*. proposed an adaptive LSB substitution based data hiding method for image. To achieve better visual quality of stego-image it takes care of noise sensitive area for embedding. *Cheddad et al*.'s[9] have proposed a region of interest (ROI) in image based adaptive steganography method. It selects required ROI in the image where it carefully hides the data bits. The selection of these regions is based on human skin tone color detection. Generally adaptive steganography methods are hard to target for attacks especially when the hidden message capacity is too small.

VI. PROPOSED WORK

In our proposed technique we preprocess the secret data before embedding it into the image. The pre-processing reduces the size of the data by a significantly large amount which permits embedding the large amount of data into the same size cover image. Our proposed technique is based on the Lempel–Ziv–Welch data compression technique. This method generates the stego image which is of very good quality. Using this method large amount of data can be embedded in cover image.

6.1 Algorithm of Proposed Work

1. Read the data from the text file and convert it from decimal to binary.
2. Pre-process the data by Lempel–Ziv–Welch technique.
3. Read selected image and represent it in the matrix form.
4. Convert the image matrix from decimal to binary.
5. Encode data and hide it in image using Kekre's algorithm.
6. At receiver side decode the data using kekre's algorithm.

6.2 Encoding Algorithm

1. At the start, the dictionary contains all possible roots and P is empty.
2. C = next character in the charstream.
3. If P+Cpresent in the dictionary then P=P+C.
4. If not then
   (i) Output the code word which denotes P to the code stream.
   (ii) Add the string P+C to the dictionary
   (iii) P=C(P now contains only the character C)
5. If there are more characters in the stream then go back to step 2.
6. If not then output the code word which denotes P to the code stream.

6.3 Decoding Algorithm:-

1. At the start the dictionary contains all possible roots.
2. cW= the first code word in the code stream.
3. Output the string cW to the charstream.
4. pW=cW.
5. cW= next code word in the code stream.
6. If string cW already present in the dictionary then
   (i) Output the string cW to the charstream.
   (ii) P=string pW.
   (iii) C= the first character of the string cW.
   (iv) Add the string P+C to the dictionary.
7. If not then
   (i) P=string pW.
   (ii) C=the first character of the string pW.
   (iii) Output the string P+C to the charstream and add it to the dictionary
8. If there are more code word in the codestream then go back to step 4.

VII. EXPERIMENTAL RESULT

The technique is implemented using java programming language. The experimental results are observed. Five sample observations are shown below. Image quality metrics are used to compare two algorithms. High PSNR value and low MSE value signifies good quality image.

PSNR values of Different Approaches on different Images

Cover Image	Modified Kekre Algorithm	Proposed Algorithm
Image1.jpg	62.0411	63.6134
Image2.jpg	75.4865	77.1353
Image3.jpg	64.5648	73.0916
Image4.jpg	70.9661	74.5371
Image5.jpg	75.2836	76.9418

MSE values of Different Approaches on different Images

Cover Image	Modified Kekre Algorithm	Proposed Algorithm
Image1.jpg	0.0525	0.0277
Image2.jpg	0.0032	0.0016
Image3.jpg	0.0058	0.0022
Image4.jpg	0.0025	0.0015
Image5.jpg	0.0020	0.0011

RMSE values of Different Approaches on different Images

Cover Image	Modified Kekre Algorithm	Proposed Algorithm
Image1.jpg	0.2389	1.1678
Image2.jpg	0.0536	0.0423
Image3.jpg	0.0977	0.0715
Image4.jpg	0.0539	0.0404
Image5.jpg	0.0604	0.0414

CAP (Maximum Embedding Capacity) values in bytes of Different Approaches on different Images.

Cover Image	Modified Kekre Algorithm	Proposed Algorithm
Image1.jpg	7247	13205
Image2.jpg	123421	208400
Image3.jpg	39450	70779
Image4.jpg	118818	290675
Image5.jpg	130502	240026

Image1 (a)

Image1 (b)

Image2(a)

Image2 (b)

Image3 (a)

Image3 (b)

Image4 (a)

Image4 (b)

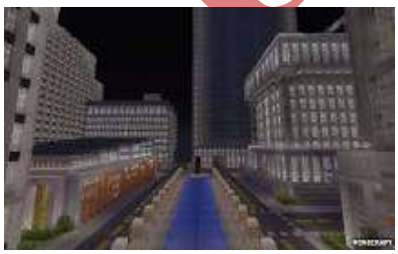

Image5 (a)

Image5 (b)

Figure 3- (A) Cover Image, (B) Stego Image

Proposed algorithm is highly secure due to preprocessing of data. For every image the value of PSNR, MSE and CAP i.e. maximum embedding capacity of our proposed technique is more than the MKA technique. Thus the proposed algorithm increases efficiency and embed large amount of data as compared to MKA.

VIII. FUTURE SCOPE

Our future work focuses upon the improvement in embedding capacity and further improvement in the efficiency of this method. Some of the ways in which we believe that Image Steganography can be made more secure is by using various ways in which the blocks in which secret data is chosen. Eg Zig-Zag scanning used in Run Length coding, Huffman coding, etc.

Areas where its scope lies:-

- Confidential communication and secret data storing
- Protection of data alteration
- Access control system for digital content distribution
- Media Database systems

IX. CONCLUSION

This work presents a scheme that can transmit large quantities of secret information and provide secure communication between two communication parties. Any kind of text data can be employed as secret msg. The secret message employing the concept of steganography is sent over the network .In addition, the proposed procedure is simple and easy to implement. Also, the developed system has many practical, personal and militaristic applications for both point-to-point and point-to multi- point communications. As steganography becomes more widely used in computing, there are issues that need to be resolved. There are a wide variety of different techniques with their own advantages and disadvantages. Many currently used techniques are not robust enough to prevent detection and removal of embedded data. The use of benchmarking to evaluate techniques should become more common and a more standard definition of robustness is required to help overcome this. The experimental results show that this method works properly and is considered to give almost the optimum solution.

REFERENCES

[1] H. B. Kekre, A. Athawale, P. 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, 2009 pp 342-346

[2] Deshpande Neeta, Kamalapur Snehal,"*Implementation of LSB Steganography and Its Evaluation for Various Bits*" K.K.Wagh Institute of Engineering Education & Research, Nashik India

[3] M. Hussain, M. Hussain., "*Pixel Intensity Based High Capacity Data Embedding Method*", Information and Emerging Technologies, International conference 978-1-4244-8003 June 2010

[4] Johnson, N.F. & Jajodia, S. (1998), "*Exploring Steganography: Seeing the Unseen*", Computer Journal

[5] Cheng-Hsing Yang, Chi-Yao Weng, Shiuh-Jeng Wang, Member, IEEE, and Hung-Min Sun, "*Adaptive Data Hiding in Edge Areas of Images with Spatial LSB Domain Systems*", IEEE Transactions on Information Forensics and Security, vol. 3, no. pp. 488-497. 3rd September 2008.

[6] L.M. Marvel "*Spread Spectrum Image Steganography*," IEEE transactions on image processing, vol. 8, no. 8, pp. 1075-1083, August 1999.

[7] Hamid, A. M., M. L. M. Kiah, et al. (2009). "*Novel Approach for High Secure and High Rate Data Hidden in the Image Using Image Texture Analysis*" Internationa Journal of Engineering and Technology (IJET).

[8] H. Yang, X. Sun, G. Sun. "*A High-Capacity Image Data Hiding Scheme Using Adaptive LSB Substitution*". Journal: Radio engineering Year: vol. 18, 4 Pages/record No.: 509-516, 2009.

[9] A. Cheddad, J. Condell, K. Curran and P. McKevitt, "*Enhancing Steganography in digital images*", IEEE - 2008 Canadian conference on computer and Robot vision, pp.326-332, 2008.

[10] H. B. Kekre, Archana Athawale, Pallavi N. Halarnkar, "*Performance Evaluation of Pixel Value Differencing and Kekre's Modified Algorithm for Information Hiding in 88 Images*", International Conference on Advances in Computing, Communication and Control, pp 342-346, 2009.