

- b) A lossless compression would certainly not be able to grant the high levels of compression that can be obtained by using a lossy compression and that are necessary for applications in storage and/or distribution of images.
- c) Compressing an image is meaningfully different than compressing raw binary data.

Lossy compression is most commonly used to compress multimedia data especially in applications such as streaming media and internet telephony. By difference, lossless compression is essential for text and data files. In many cases it is advantageous to make a master lossless file that can then be used to produce compressed files for different purposes. There are two basic lossy compression systems:

- a) In lossy transform codecs, examples of image or sound are taken, sliced into small sections, transformed into a new root space, and quantized. The resultant quantized standards are then entropy coded.
- b) In lossy predictive codecs, previous and/or subsequent decoded data is used to predict the current sound sample or image frame. The fault between the projected data and the real data, organized with any additional information wanted to reproduce the prediction, is next quantized and coded.

2.3.2 Lossless Image Compression

Image compression may be lossy or lossless. Lossless compression is chosen for archival commitments and frequently for technical drawings, clip art, medical imaging or comics. This is because lossy compression methods, particularly when used at low bit rates, present compression artifacts. Lossless and lossy compression are terms that describe whether or not, in the decompression of a file, all original data can be recovered when the file is uncompressed. With lossless compression, each and every single bit of data that was originally in the file remains after the file is uncompressed. All the original information is completely restored. This is usually the technique of choice used for text or spreadsheet files, where misplacing words or financial data could pose a problem.

a) Lossless Data Compression

It is a class of data compression algorithms that allows the exact original data to be reconstructed from the compressed data. The term lossless is in difference to lossy data compression, which only permits an estimate of the original data to be reassembled, in exchange for better compression rates. For example, it is used in the ZIP file format and in the UNIX tool gzip. It is also frequently used as a component within lossy data compression technologies.

b) Lossless JPEG

It is a 1993 addition for lossless compression to JPEG standard by the Joint Photographic Experts Group. However, it might be used as an umbrella term to refer to all lossless compression techniques developed by the Joint Photographic Expert group. They include JPEG 2000 and JPEG-LS (Lossless JPEG). Lossless JPEG scheme was developed as a late addition to JPEG in 1993, using a completely different technique from the lossy JPEG standard. It uses a predictive scheme based on the three nearest (causal) neighbors (upper,

left, and upper-left), and entropy coding is used on the prediction error. It is not supported by the standard Independent JPEG Group libraries, although Ken Murchison of Oceana Matrix Ltd. wrote a patch that extends the IJG library to support Lossless JPEG. Lossless JPEG has some popularity in medical imaging, and is used in some digital cameras to compress raw images, but otherwise was never widely adopted [12].

2.4 Existing Steganography System

2.4.1 Least Significant Bit (LSB) Method

The most common algorithm belonging to spatial domain class of techniques is the Least Significant Bit (LSB) Replacement technique in which the least significant bit of the binary representation of the pixel gray levels is used to represent the message bit. In terms of embedding and extraction complexity make simple spatial domain techniques are used. These techniques use the pixel gray levels and their color values directly for encoding the message bits. This kind of data embedding leads to an addition of a noise of 0:5p on average in the pixels of the image where p is the embedding rate in bits or pixel.

Zhang and Wang has presented multiple base notational systems that has been employed for embedding data bits which is based on the Human Vision Sensitivity (HVS) . To compute the number base to be used for data embedding the variance value for a block of pixels is used. A similar class of algorithm based on HVS has been proposed by Wu and Tsai named as Pixel Value Differencing.

Shailender Gupta, Ankur Goyal and Bharat Bhushan proposed method . The Least Significant Bit (LSB) steganography is one such technique in which least Significant bit of the image is replaced with data bit. As this method is vulnerable to steganalysis so as to make it more secure the raw data is encrypted before embedding it in the image.

Later on extension is given to prediction-based schemes, by designing a high-capacity reversible watermarking algorithm based on adaptive pixel prediction. This algorithm is designed to achieve high embedding capacity while preserving high perceptual quality It computes to be predicted pixel by linearly combining its neighboring values.

3. Proposed System

The figure 3.1 depicting the view of propose system architecture. Proposed system can be described as follows; a database sample shown in the diagram is to be used for storing the various images and also it will store Encrypted images. Then need arises to go for encryption using generating a key for encrypting the image. Then user will go for compression by applying the technique based on Einstein's algorithm and after compression can embed the additional data. Following diagram shows every aspect of the proposed system. Here is depicting the various modules, functionalities present in the system.

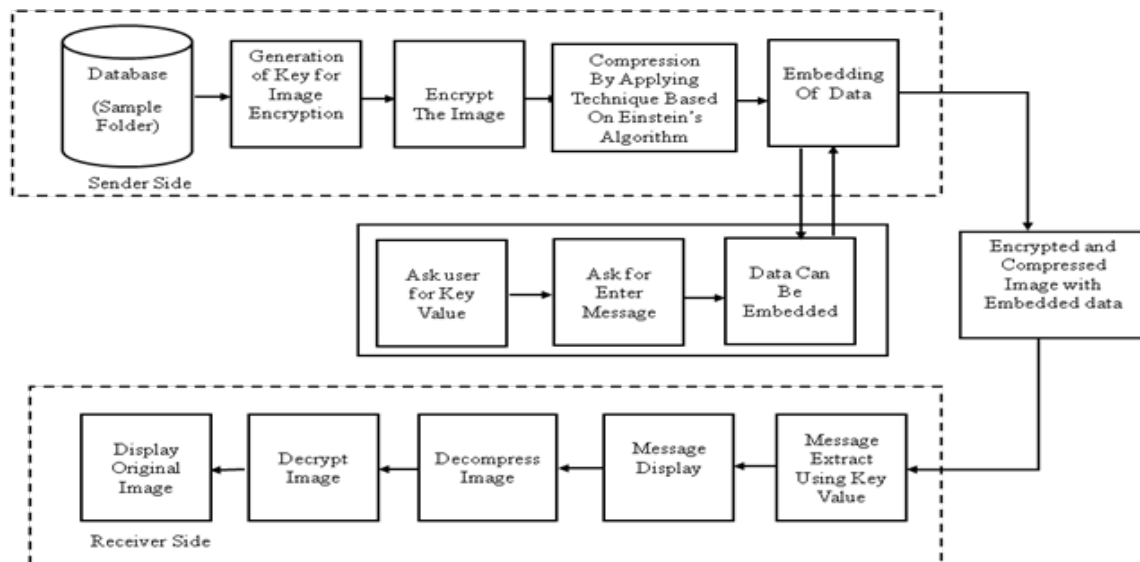


Figure 3.1: Flow of Proposed System at Sender Side

The steps for encryption and key generation algorithm are as follows:

- 1) Select the image to be encrypted.
- 2) For variable length key generation find the n no. of bit values.
- 3) Generate array to store n no. of bit value.
- 4) Generate array to store byte value.
- 5) Convert n no. of bit value into byte value.
- 6) XOR the final key stream with the image to be encrypted to get the cipher image.

3.1.1 Image Compression Based on Einstein's Algorithm

Algorithm: Compression Based on Einstein's Algorithm

Input: Encrypted Image

Output: Compressed Image

Algorithm Steps:

- 1) Select the "Encrypted Image".
- 2) Converts the encrypted image into the "Matrix" form which contains the numeric values of each pixel present in that image.
- 3) Apply the "Compression Technique Based on Einstein's Algorithm" to compresses the matrix.
 - a) Separates the R-G-B values present in that matrix of each pixel position.
 - b) Then for compressing the values of R-G-B, multiply;
 - i) $R \times 256 \times 256$
 - ii) $G \times 256$
 - iii) $B + 1$.
 - c) Perform Addition of above mentioned formulae in 3(b) (i) (ii) and (iii) to find the "Index" number.
 - d) Repeat step 3(b) and 3(c) until the last pixel in the image.
 - e) "Index number", will specify the corresponding "Pixel" which can be compressed and stored into the "Data Base".
- 4) End.

3.1.2 BPCS-Steganography

Bit Plane Complexity Segmentation (BPCS) was introduced in 1998 by Eiji Kawaguchi and Richard O. Eason to overcome the shortcomings of the traditional Least

Significant Bit (LSB) manipulation techniques. While the LSB manipulation technique works very well for most gray scale and RGB color images, it is severely crippled by its limitation in capacity, which is restricted to about one eighth the size of the base image. BPCS is based on the simple idea that the higher bit planes could also be used for embedding information provided they are hidden in seemingly "complex" regions.

BPCS stands for Bit plane complexity segmentation in which it splits the region image into bit planes and finds out the complex regions. The main reason behind finding the complex region is that at that place it can store additional data and that it will be invisible for the end user. BPCS-Steganography is a type of digital steganography. Digital steganography can hide confidential data (i.e. secret files) very securely by embedding them into some media data called "vessel image". The vessel image is also referred to as "carrier, cover, or dummy data". In BPCS-Steganography true color images (i.e., 24-bit color images) are mostly used for vessel image.

The embedding operation in practice is to replace the "complex areas" on the bit planes of the vessel image with the secret data. The most important aspect of BPCS-Steganography is that the embedding capacity is very large. In comparison to simple image based steganography which uses solely the least important bit of data, and thus (for a 24-bit color image) can only embed data equivalent to 1/8 of the total size, BPCS-Steganography uses multiple bit-planes, and so can embed a much higher amount of data, though this is dependent on the individual image. For a normal image, roughly 50% of the data might be replaceable with secret data before image degradation becomes apparent.

Algorithm for BPCS:

In figure 3.7 the flow of embedding the message into encrypted image is shown. First it takes "Compressed Image" as an Input, it then finds out the "Complex Regions". In the same area of image it stores the data. Also it finds out the area based on the length of the black-and-white border.

Another portion it finds is the connected areas that could be used to find the complex regions in an image.

Algorithm: BPCS

Input: Encrypted and Compressed Image

Output: Stego Image

- a) Convert the carrier image (original / cover image) from Pure Binary Code (PBC) to Canonical Gray Code (CGC).
- b) Segmentation on carrier image will be performed i.e. convert each bit plane of the carrier image into informative and noise like regions by using threshold values α_0 that means complexity of image will be calculated.
- c) Group the bytes of secret files into series of secret blocks.
- d) If the block is less complex than the threshold α_0 , then conjugate it to make it more complex block.
- e) The conjugated block must be more complex than α_0
- f) Embed each secret block into the complex region or replace all the noise like regions with the series of secret block where maximum color changes will be observed.
- g) Convert the embedded dummy image from CGC back to PBC.

4. Experiment and Analysis

4.1 Evaluation Parameters

To compare the aspects that defined in problem statement and measuring by the Evaluation parameters those are correctly recovered image and incorrectly recovered image. The correctly and incorrectly recovered image shows the percentage of data recovered in lossless manner.

4.1.1 Peak Signal to Noise Ratio (PSNR)

Peak signal-to-noise ratio, frequently abbreviated as PSNR, is an engineering term for the ratio between the supreme possible power of a signal and the power of corrupting noise that disturbs the loyalty of its illustration. Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale [28]. A series of pseudo random numbers as the secret bit stream are embedded into the cover image. The peak signal to noise ratio is utilized to evaluate the quality of the stego image. For an $M \times N$ color image the PSNR value is defined as:

$$PSNR (dB) = 10 * \log_{10} \left(\frac{255^2}{MSE} \right)$$

$$MSE = \sum_{i=1}^M \sum_{j=1}^N \left(\frac{p_{ij} - q_{ij}}{M * N} \right)^2$$

Where M and N represent the image size. In the formula p (i, j) stands for the original pixel value, and q (i, j) represents the pixel value in position (i, j) with the secret data already hidden in. the greater PSNR means the lower degree of distortion after the hiding of data.

4.1.2 Mean Square Error (MSE)

Mean Square Error is a signal fidelity measure where the goal is to compare two signals by providing a quantitative score that describes the degree of similarity/ fidelity or, conversely, the level of error/distortion between them.. Let x

= { $y_i | i = 1, 2, \dots, N$ } and $y = \{ y_i | i = 1, 2, \dots, N \}$ are two finite-length, discrete signals and the number of pixels. In experiment some different color images with size 512x512 are used as the cover images and 3 of them are shown in figure 5.1. This method adopts the peak signal to noise ratio to evaluate the qualities of the recovered images.



Figure 4.1: Images Used for Experimental Results

Image (a) Baboon original cover image (b) Airplane original cover image and (c) Lena original cover image are used for finding the experimental results.

The comparison between the proposed data hiding method for color image and data hiding scheme for grayscale image in [5] with respect to PSNR values of retrieved images obtained from experiments.

Table 4.1: Comparison between Proposed Method and Scheme in [5]

Cover Image	Proposed Method	Scheme in[5]
Baboon	35.21	39.16
Airplane	35.15	39.54
Lena	33.08	39.31
Pepper	36.03	39.06

The above table 5.1 shows that, the Image quality of recovered color image is in between the range of 30 to 50 which is good as per the range of image quality. The resulting PSNR values of proposed method are taken after embedding the maximum data in single plane of the color cover image and recovered at receiver side successfully.

Here, by using proposed method the cover image of size 512x512 pixels which contains the secret data of 3000 bytes is showing the good image quality after recovering at the receiver side. Likewise, in the above table the possible results analysis is shown on the basis of data embedded in the cover image and recovered successfully. So maximum data embedded in the single plane of cover image is 32768 bytes.

Table 4.2: PSNR Report and MSE of Proposed Method

Cover Image	Data Embedded (ch)	PSNR Of Recovered Image (dB)	MSE
Baboon	3000	46.3869	0.0059
Baboon	6000	44.8817	0.0083
Baboon	12000	44.5470	0.0090
Baboon	24000	45.2776	0.0076
Baboon	30000	35.2893	0.0754
Baboon	32000	35.2177	0.0767
Baboon	32768	35.2163	0.0767
Airplane	3000	41.4856	0.0181
Airplane	6000	40.5884	0.0223
Airplane	12000	38.9853	0.0322

Airplane	24000	36.3415	0.0592
Airplane	30000	35.4013	0.0735
Airplane	32000	35.1538	0.0778
Airplane	32768	35.0641	0.0795
Lena	3000	46.3869	0.0059
Lena	6000	36.2422	0.0606
Lena	12000	34.9466	0.0816
Lena	24000	43.1478	0.0124
Lena	30000	33.0784	0.1255
Lena	32000	33.0805	0.1255
Lena	32768	33.0899	0.1225

Here the above table 4.2 depicts the report of PSNR values for the recovered image, Mean square error and the data embedded in characters in cover image are shown in figure 4.1 above, so it will be very useful to go for comparative analysis by observing the original image and recovered image quality.

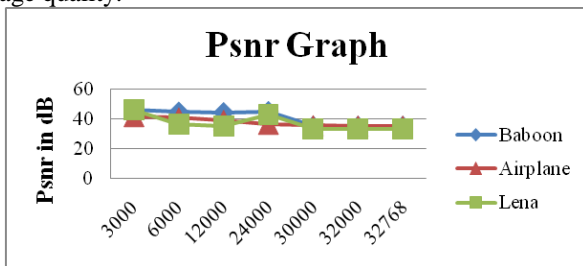


Figure 4.2: PSNR Graph

The above figure 4.3 PSNR graph has drawn on the basis table 4.2 depicted above. The PSNR graph which contains the report of successfully embedded data at sender side and data extracted also recovered image at receiver side.

To compare the aspects that defined in problem statement and measuring by the Evaluation parameters those are correctly recovered image and incorrectly recovered image. The correctly and incorrectly recovered image shows the percentage of data recovered in lossless manner.

The table 4.2 above clears the idea, the data embedded in cover image and image recovered after applying the method for getting the results. As the results shown in table 4.2 and graph

5. Conclusion

In this work, a novel scheme for separable reversible data based Einstein's algorithm for color image is proposed, which consists of image encryption, image compression, data embedding and data-extraction/image-recovery phases. In the first phase, the content owner encrypts the original uncompressed image using an encryption key. Although a data-hider does not know the original content, he can compress the encrypted image using a novel compression method for color images, and by using data-hiding key can embed the additional data. Another beauty of this system is that this can process three dimensional gray scale images also.

To improve the embedding capacity and provide an imperceptible visual quality and security of data, a BPCS steganographic method is combined with newel raster compression algorithm for color image is presented in this work. Secret data is hidden into the bit planes of the image

as informative areas and noise like areas are categorized by the complexity threshold. The original image is encrypted with the help of variable key stream and pixel values of image are XORed and then compressed by using technique based on Einstein algorithm which reduces the complexity of algorithm used for color image compression. The simplicity of matrix addition is the major advantage of this color image compression algorithm as the compressed image requires less space in the database. The compressed matrix of color image is been created by finding the index number of each compressed pixel. The compression ratio and the quality of reconstructed image vary with different values of compression parameters as it depends upon the R-G-B values. In general, the higher the compression ratio and the smoother the original image, the better the quality of the reconstructed Image retrieved.

So conclusion of this work is that, this steganographic technique is a very strong information security technique, as it is combined with encrypted, compressed and data embedding domain. The system can process the color as well as 3D grayscale images. The data which is embedded and the encrypted image which is compressed can be extracted and original contents are recovered from the stego image in lossless manner.

6. Acknowledgements

The author would like to thank Xinpeng Zang and its members for contributing.

References

- [1] Xinpeng Zhang , "Separable Reversible Data Hiding in Encrypted Image" , IEEE Trans. on Information Forensics and Security, vol. 7, no. 2, April 2012
- [2] X.Zhang, "Lossy compression and iterative reconstruction for encrypted image," IEEE Trans. Inform. Forensics Security, vol. 6, no. 1, pp. 53–58, Feb. 2011.
- [3] X. Zhang, "Reversible data hiding in encrypted Image," IEEE Signal Process. Lett., vol. 18, no. 4, pp. 255–258, Apr. 2011.
- [4] 1, 2AlbertusJokoSantoso, 3Dr. Lukito Edi Nugroho, 4Dr. Gede Bayu Suparta, 3Dr. Risanuri Hidayat, Compression Ratio and Peak Signal toNoise Ratio in Grayscale Image Compressionusing Wavelet, IJCST Vol. 2, Iss ue 2, June 2011.
- [5] W. Liu, W. Zeng, L. Dong, and Q. Yao, "Efficient Compression of encrypted grayscale images," IEEE Trans. Image Process. vol. 19, no. 4, pp. 1097–1102, Apr. 2010.
- [6] http://en.wikipedia.org/wiki/Lossy_compression
- [7] Eiji Kawaguchi, Richard O. Eason, "Principle and Applications of BPCS Steganography." SPIE's International Symposium on Voice, Video and Data Communications, (1998-11).
- [8] <http://en.wikipedia.org/wiki/Encryption>.
- [9] Mohammed Mustaq, Mohammed Mothi, Yasser Arafat, "Einstein's Image Compression Algorithm", Version 1.00, TRIM 7 (2) July -Dec 2011