Color Document Image Authentication with Data Repair Capability

Ms.Lekshmi.C.R, Mr.K.Saravanamoorthy

Abstract— A new authentication technique for color images with data repair capability.The RGB channels are first transformed into three secret embedding channels SEC using reversible integer transform.Agroup of 2×2 blocks which is the strongest points of non overlapping spacially adjacent pixels in S,E and C image is selected as the valid blocks for embedding the secret message.Then adaptive mod 4 embedding operation is further applied to all the valid blocks to embed a pair of binary bits using the shortest route modification scheme.Each secret message is alsoencrypted by RSA encryption algorithm to provide the system wih more security.Data will be embedded inside the image using the pixels.Then the pixels of the stego image can then be accessed back inorder to retrieve back the hidden data inside the image.However, a secret key is needed by the receiver inorder to retrieve back the data.This secret key is generated by RSA decryption algorithm.By using the secret key to retrieve the data,it maintains privacy,confidentiality and accuracy of the data.The proposed method was tested on different color images.From the experimental results,compared with the some well known adaptive and non adaptive authentication techniques,the proposed method provides larger embedding capacity,while being less detectable by steganalysis methods.

Index Terms— Adaptive mod 4 embedding,Data repair,Image authentication,Reversible integer transform(RIT),RSA encryption and decryption algorithm,Secret Embedding Channels(SEC).

1.INTRODUCTION

THE picture is the most common and convenient way of conveying or transmitting information. A picture is worth a thousand words. Pictures concisely convey information about positions, sizes and inter-relationships between objects. Human beings are good at deriving information from such images, because of our innate visual and mental abilities. About 75% of the information received by human is in pictorial form.

Digital image is a form for preserving important information. However, with the fast advance of digital technologies, it is easy to make visually imperceptible modifications to the contents of digital images. How to ensure the integrity and the authenticity of a digital image is thus a challenge. It is desirable to design effective methods to solve this kind of image authentication problem, particularly for images of documents whose security must be protected. It is also hoped that, if part of a document image is verified to have been illicitly altered, the destroyed content can be repaired. Such image content authentication and self-repair capabilities are useful for the security protection of digital documents in many fields, such as important certificates, signed documents, scanned checks, circuit diagrams, art drawings, design drafts, last will and testaments, and so on.

However with the fast advance of digital technology gray scale images are replaced by color images. In this paper new authentication technique has been employed for color images. It maintains privacy, confidentiality and accuracy of the hidden data and preserves the stego image quality.

Several methods for image authentication for gray scale images have been proposed in the past. Wu and Liu manipulated the so-called flippable pixels to create specific relationships to embed data for authentication and annotation of binary images. Yang and Kot proposed a two-layer binary image authentication method in which one layer is used for checking the image fidelity and the other for checking image integrity. In the method, a connectivity-preserving transition Tzeng and Tsai’s criterion for determining the flippability of a pixel is used for embedding the cryptographic signature and the block identifier. Yang and Kot proposed a pattern-based data hiding method for binary image method in which one layer is used for checking the image fidelity and the other for checking image integrity. In the method, a connectivity-preserving transition Tzeng and Tsai’s criterion for determining the flippability of a pixel is used for embedding the cryptographic signature and the block identifier. Yang and Kot proposed a pattern-based data hiding method for binary image authentication in which three transition criteria are used to determine the flippabilities of pixels in each block, and the watermark is adaptively embedded into embeddable blocks to deal with the uneven embed ability condition in the host image. In the method, a set of pseudorandom pixels in a binary or halftone image are chosen and cleared, and authentication codes are accordingly computed and inserted into selected random pixels. Lee et al. proposed a Hamming-code-based data embedding method that flips one pixel in each binary image block for embedding a watermark, yielding small distortions and low false negative rates. Lee et al. improved the method later by using an edge line similarity measure to select flippable pixels for the purpose of reducing the distortion.

In this paper a new authentication technique has been employed for color images with data repair capability. The RGB channels are first transformed into three secret independent channels, SEC using reversible integer transform A group of 2×2 blocks of non-overlapping spatially adjacent pixels, which is the strongest points in S,E and C image is selected as the valid block for embedding the secret message. The modulo 4 arithmetic operation is further applied to all the valid blocks to embed a pair of binary bits using the shortest route modification scheme. Each secret message is also encrypted by RSA encryption algorithm to provide the system with more security. Data will be embedded inside the image using the pixels. Then the pixels of stego image can then be accessed back in order to retrieve back the hidden data inside the image. However, a secret key is needed by the receiver in order to retrieve back the data. This secret key is generated using the RSA decryption algorithm. By using the secret key to retrieve the data, it maintains privacy, confidentiality and accuracy of the data.

The remainder of the paper is as follows .In Section I RSA Encryption technique, the key generation, encryption and decryption on which the proposed method is based is first reviewed. In Section II, the details of the proposed method, including,, Reversible integer transform to transform RGB to SEC and viceversa.In Section III Adaptive mod 4 embedding for embedding the secret data and tampered data repairing, are described. Experimental results and a comparison of performances of the proposed method with others are shown in Section IV, followed by conclusions in Section V

I.REVIEW OF RSA ENCRYPTION TECHNIQUE

Ron Rivest, Adi Shamir, and Leonard Adleman introduced a cryptographic algorithm in 1978 named RSA, which was essentially to replace the less secure National Bureau of Standards (NBS) algorithm. RSA involves a public key and a private key. The public key can be known to everyone and is used for encrypting messages. Messages encrypted with the public key can only be decrypted using the private key. The keys for the RSA algorithm are generated the following way:

Algorithm 1: Key generation

Step 1.Choose two distinct prime numbers ***p*** and ***q*.**

Step 2.Compute ***n* = *pq*. *n*** is used as the modulus for both the public and private keys

Step 3.Compute **φ(*n*) = (*p* – 1)(*q* – 1),** where **φ** is **Euler's totient function**.

Step 4.Choose an integer ***e*** such that **1 < *e* < φ(*n*)** and greatest common divisor of **(*e*, φ(*n*)) = 1**; i.e.**,** ***e*** and **φ(*n*)** are co prime. ***e*** is released as the public key exponent

Step 5.Determine ***d*** as:

$$d = e^{-1}(mod\ \varphi(n)\ )$$

i.e.**,** ***d*** is the multiplicative inverse of ***e*** mod **φ(*n*).** The **public key** consists of the modulus ***n*** and the public (or encryption) exponent ***e*.** The **private key** consists of the modulus ***n*** and the private (or decryption) exponent ***d*** which must be kept secret**.**

Algorithm.2:Encryption

Step 1: The public key is given by (n,e).If one wants to send a message M, we first turns **M** into an integer m, such that 0<=m<n by using an agreed-upon reversible protocol known as a padding scheme.

Step 2: Compute the cipher text Ccorresponding to

$$c = m^{e}(mod\ n)$$

Step 3: Given m one can recover the original message **M** by reversing the padding scheme. using the square-and-multiply algorithm for modular exponentiation

$$m = c^{d}(mod\ n)$$

II.REVIEW OF REVERSIBLE INTEGER TRANSFORM

It is evident that the R, G and B channels of an image are highly dependent on each other whereas for a good encryption technique this relation must be broken before encryption so that the robustness of the technique increases. Usually, the RGB channels are transformed into independent transform like HSI, YCbCr etc. But the main problem with these existing independent transforms is non-perfect reconstruction. Hence, for a perfect color space conversion, there must be a transformation which maps integers to integers. For this purpose and to enhance security, RGB channels are first transformed into three secret independent channels (SEC channel) using reversible integer transform (RIT) and then the encryption is done either in all or any of S, E and C channels.

The SEC channel is obtained from RGB channel as.

$$\begin{bmatrix} S \\ E \\ C \end{bmatrix} = \left(A \begin{bmatrix} R \\ G \\ B \end{bmatrix} \right)$$

$$\begin{bmatrix} S \\ E \\ C \end{bmatrix} = \left[\begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \right]$$

and the RGB channel is obtained from SEC channel as

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \left(A^{-1} \begin{bmatrix} S \\ E \\ C \end{bmatrix} \right)$$

$$\begin{bmatrix}R\\G\\B\end{bmatrix} = \left[\begin{pmatrix} {}^{1}/_{a_{11}} & 0 & 0 \\ -{}^{a_{21}}/_{a_{22}} & {}^{1}/_{a_{22}} & 0 \\ -{}^{a_{31}}/_{a_{33}} & -{}^{a_{32}}/_{a_{33}} & {}^{1}/_{a_{33}}\end{pmatrix}\begin{bmatrix}R\\G\\B\end{bmatrix}\right]$$

III.REVIEW OF ADAPTIVE MOD4 EMBEDDING METHOD

An adaptive image steganographic model is proposed here that is based on mod-4 embedding algorithm to maximize the embedding capacity while maintaining image fidelity. We use 4 pixels to represent two bit of the message.

A. *Proposed method for Data Hiding*

The Adaptive mod-4 embedding method is used for information hiding within the spatial domain of any color image. The input messages can be in any digital form, and are often treated as a bit stream.

Stage I-The embedding scheme

The block diagram of the proposed method is shown in fig below, where we consider the JPEG compressed image for presentation purpose.
By definition, a valid block is given by expression

$$valid\ blocks(Q): C = \left(\frac{1}{4}\sum_{x\in Q} |x - m_Q|\right) > T$$

where mq is the mean gray level value of the pixels in the block and T is the minimum contrast defined by the user.TakeT=10,inthiswork.

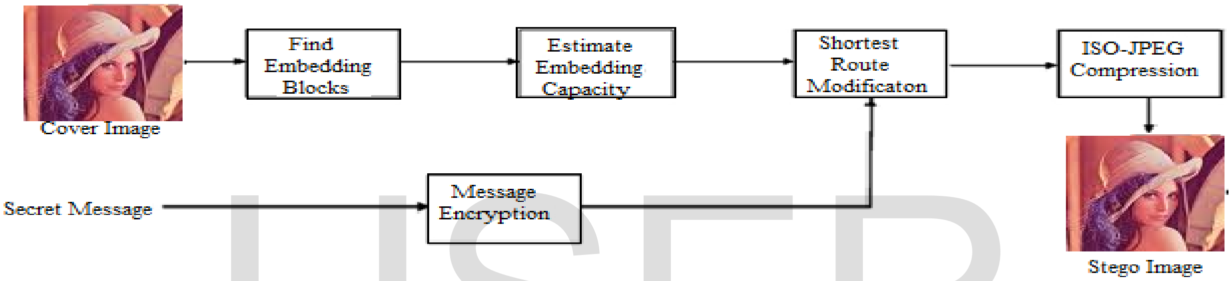


Fig.1.Mod 4 Embedding

Stage II Find the embedding blocks

The 8×8 blocks of qDCTs (each denoted by FQT) are decompressed and extracted from the JPEG data stream. Each FQT is then divided into 16 groups of 2×2 spatially adjacent quantized DCT coefficients (GQC). A GQC is characterized as a valid GQC (vGQC) if it satisfies the following conditions:

$$\left|\{x: x\varepsilon GQC, x > \varphi_1\}\right| \geq \tau_1$$

$$\left|\{x: x\varepsilon GQC, x < -\varphi_2\}\right| \geq \tau_2$$

where |X| denotes the cardinality of the set X. $\varphi_1$and $\varphi_2$are magnitudes that govern the coefficients to be modified. $\tau_1$and τ_2 are positive thresholds indicating if the candidate GQC is a validGQC.

Stage III:Estimating Embedded Capacity

The vGQCs are extracted from the FQTs, and loaded into an empty dynamic array β in the order determined by a secret key $\mathcal{K}$, This process randomizes the order of extraction of vGQCs and thus makes Mod4 satisfy the Kerckhoff's principle when all vQGCs are exhausted; the length of β determines the maximum embedding capacity Ω of the cover image. In particular, ̓Ω depends on the cover image, the JPEG quality factor, two magnitudes Φ1 and Φ2, and two positive threshold τ_1 and τ_2. Ω is twice of the number of the vGQCs because each vGQC holds exactly two bits.

Stage IV:Shortest Route Modification

Mod4 embeds pairs of message bits into the ordered vGQCs in β. Denote the ith vGQC from β by Bi, and the ith pair of message bits from the secret message by *xyi* *ε*{00,01,10,11}. Unlike the ordinary direct modification method, SRM leads to lower expected number of modification on the coefficients. The embedding scheme constrained by SRM is illustrated in Table 1, show a representative example to embed *xyi* = 00. The first column mod (bi, 4) indicates the remainder of the sum of the qDCTs in Bi divided by 4

Table1.Shortest route modification

mod($b_{i,}$4)	**$\oplus$**	**$\ominus$**	**Possible Routes**	**Shortest Route**
00	0	0	No change	N/A
01	3	1	-1 or +3	-1
10	2	2	+2 or-2	Conditional
11	1	3	+1 or -3	+1

.

The second column $\oplus$ records the number to be added to the positive coefficients to achieve Mod (bi', 4) =00 where bi' denotes the new sum after modification(s). Similarly, the third column $\ominus$ indicates the number to be subtracted from the negative coefficients to achieve mod (bi', 4) =00

Stage V-ISO– JPEG compression

The modified Bi's and GQCs (with location association) form blocks F'QT. Each F'QT then undergoes the same zig zag scan, and differential and run length coding as in the JPEG compression scheme. The resulting steganogram is transmitted to the intended receiver as a JPEG image.

B.The Extraction Scheme.

The extraction process can be carried out by reversing the embedding procedure and the block diagram is shown in fig.2.

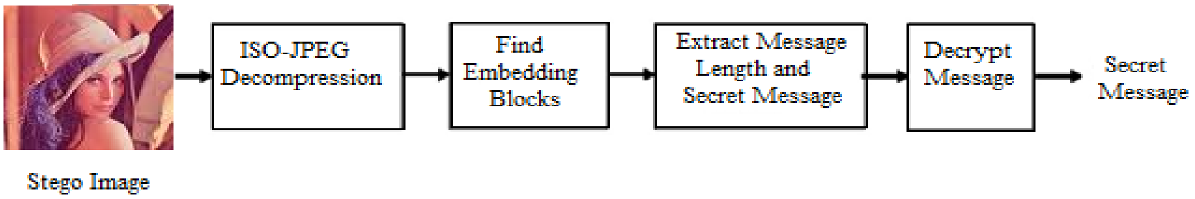


Fig.2.The Extraction

IV.EXPERIMENTAL RESULTS AND COMPARISON WITH EXISTING METHODS

A.Data Embedding and Data Retrieval

The first results that we show here come from our experiments using a input RGB color image (256×256) in jpeg format. shown in Fig.3(a)

The result of applying reversible integer transform into Fig.3(a) is shown in Fig.3(b). This is the SEC image.

The S,E and C region is separated. The S image is shown in Fig.3(c). Then intensity points are found in Fig.3(c) and shown in Fig.3(d).Then Adaptive mod 4 embedding technique is applied to embed the secret data in Fig.3(c).The cover image produced by data embedding in Fig.3(c) is shown in Fig.3(e).Similar process is repeated on E and C region to embed the data. The Stego SEC image produced as a result of the above processes is shown in Fig.3(f).

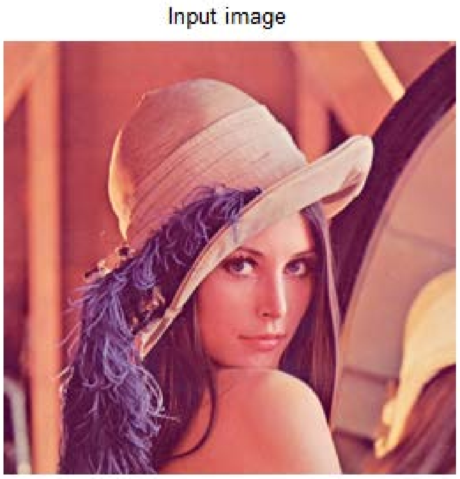


Fig.3(a):Input image

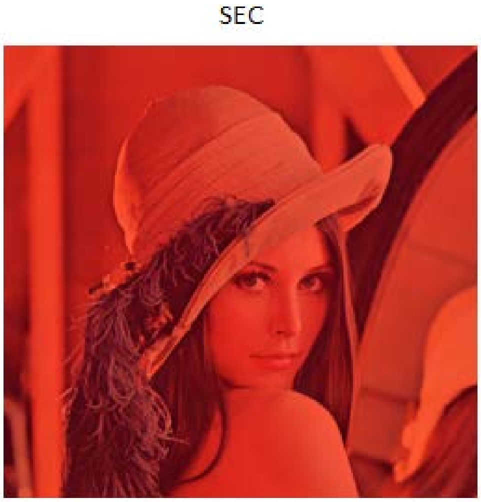
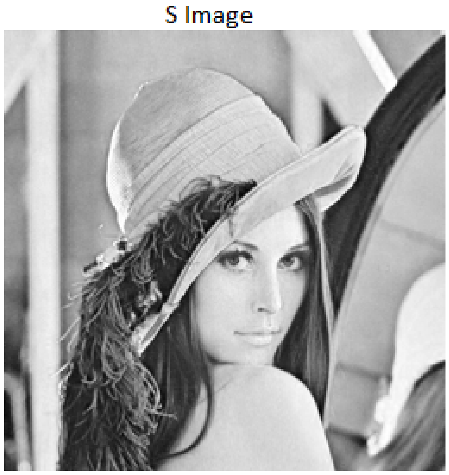


Fig.3(b):SEC Image



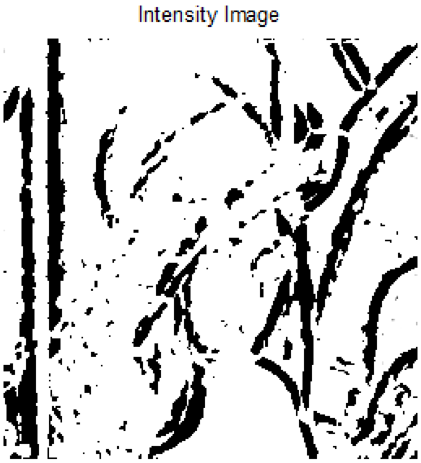


Fig3(c):The S image and Fig.3(d):Intensity image

The Stego SEC image is converted to RGB image by applying RIT to the Fig.3(f). As shown, the stego-image shown in the latter is visually almost identical to the cover image shown in the former, although the latter image includes the embedded data. Fig.3(g) is the transmitted image. At the receiver the reverse operation is performed to obtain the secret data. The received RGB image is converted to SEC using RIT. Then block division is performed on it to recover the secret data. Using the secret key the secret data is authenticated. The SEC image with block division is shown in Fig.3(h).The reversible image produced as a result of applying RIT to Fig.3(h) is shown in Fig.3(i).

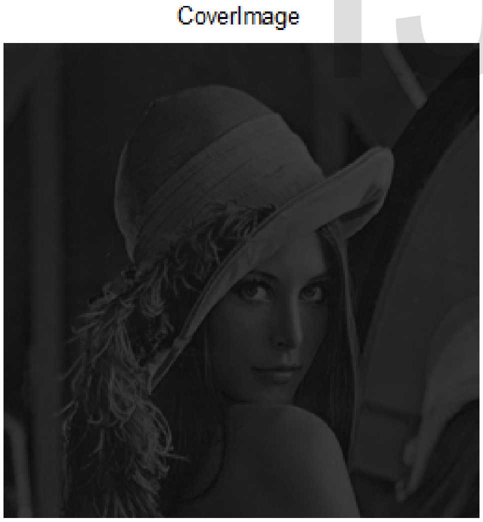


Fig.3(e):Cover image

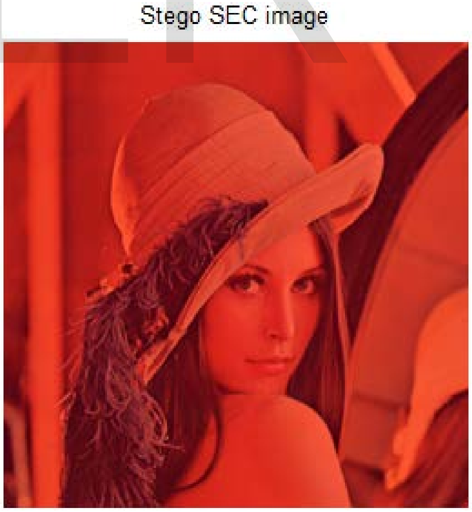


Fig3(f):Stego SEC Image

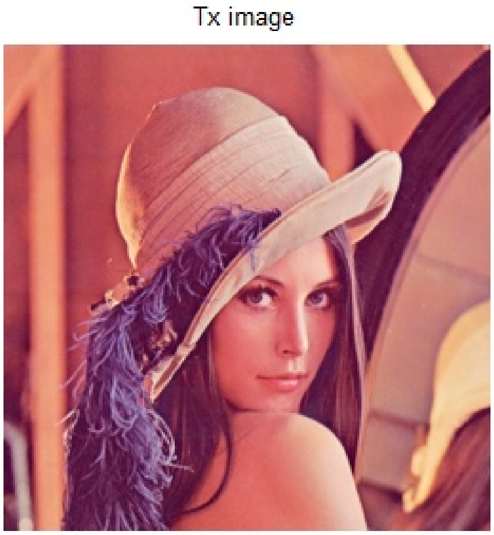


Fig3(g):The Tx image

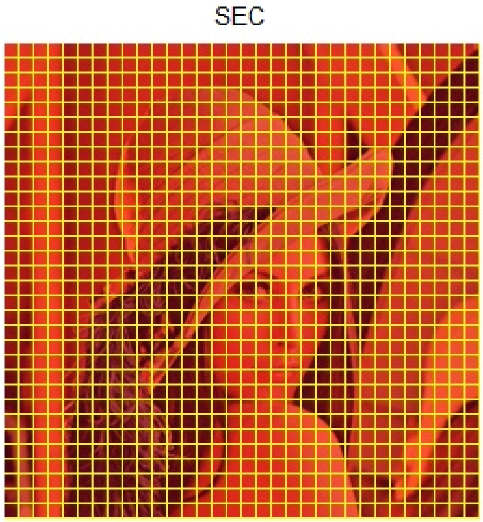


Fig3(h):SEC with block division

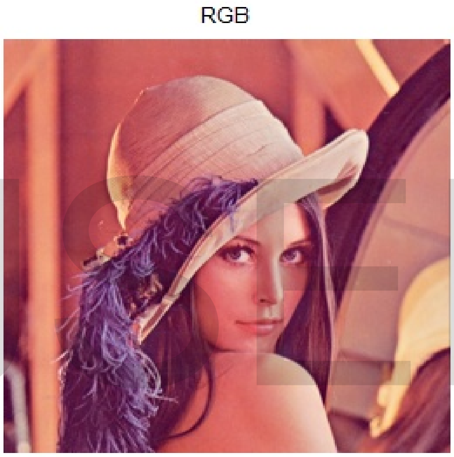


Fig3(i):RGB Image

B.Calculation of Peak Signal To Noise Ratio

PSNR is used to measure the difference between two images. It is defined as

$$PSNR = 20 * \log10 (b/rms),$$

where b is the largest possible value of the signal (typically 255 or 1), and rms is the root mean square ratio of the peak signal and the difference between two images. The PSNR is given in decibel units (dB), which measure
An increase of 20 dB corresponds to a ten-fold decrease in the rms difference between two images.

There are many versions of signal-to-noise ratios, but the PSNR is very common in image processing, probably because it gives better-sounding numbers than other measures..The PSNR values of three stego images are shown in the table below. They are all greater than 35dB,which is the empirical value for the distortion invisibility requirement. A higher PSNR would normally indicate that the reconstruction is of higher quality.

Table 2 PSNR of various images

Images	**PSNR**
Lena	+77.77dB
Baboon	+70.21dB
Pepper	+78.40dB

C.Histogram Analysis.

The histograms of orginal cover image and stego color image is shown below.The histograms of red,green and blue plane is shown separately.From the statistical analysis it is clear that the histograms of the corresponding three planes are identical.It clearly shows that there is no obvious visual distortion in the stego image

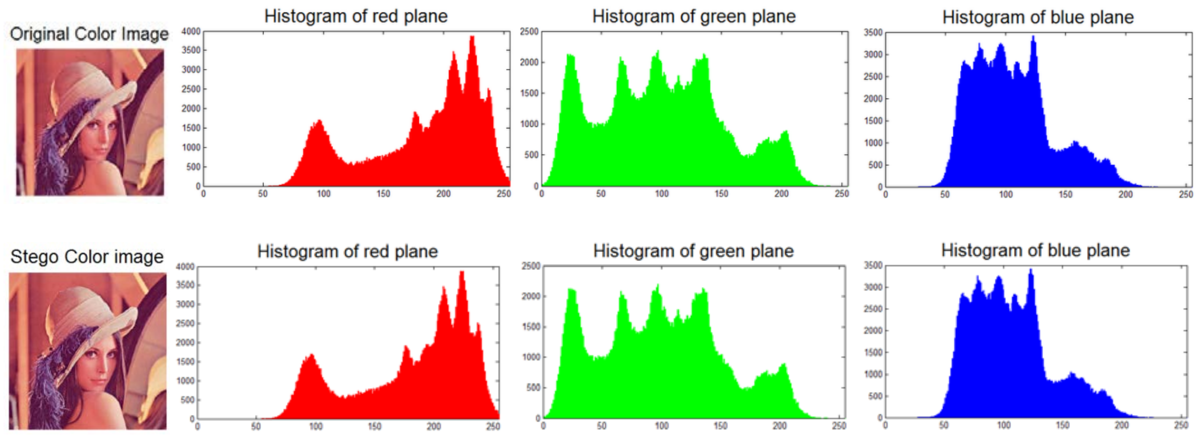


Fig 3(j):Histogram analysis

VI CONCLUSION

A new steganographic technique has been employed for color images. The RGB channels are first transformed into three secret independent channels SEC using reversible integer transform.A group of 2×2 blocks of non-overlapping spatially adjacent pixels in S,E and C image is selected as the valid block for embedding the secret message. The modulo 4 arithmetic operation is further applied to all the valid blocks to embed a pair of binary bits using the shortest route modification scheme. Each secret message is also encrypted by RSA encryption algorithm to provide the system with more security. However, a secret key is needed by the receiver in order to retrieve back the data. This secret key is generated using the RSA decryption algorithm. By using the secret key to retrieve the data, it maintains privacy, confidentiality and accuracy of the data. The proposed method was tested on different color images. From the experimental results, compared with the some well-known adaptive and non-adaptive steganography algorithms, the proposed method provides larger embedding capacity, while being less detectable by steganalysis methods. Experimental results have been shown to prove the effectiveness of the proposed method. Future works involve calculation of various ratios and employment of new methods to improve the accuracy and PSNR.

REFERENCES

[1] Che-Wei Lee *Student Member, IEEE*, and Wen-Hsiang Tsai, *Senior Member, IEEE* "A Secret-Sharing-Based Method for Authentication of Grayscale Document Images via the Use of the PNG Image With a Data Repair Capability", *IEEE Transactions on Image Processing, Vol. 21,no. 1 Jan 2012*.

[2] H. Yang and A. C. Kot, "Binary image authentication with tampering localization by embedding cryptographic signature and block identifier,"*IEEE Signal Process. Lett*., vol. 13, no. 12, pp. 741–744, Dec.2006.

[3]Gyankamal J.Chhajed "Review on Binary Image Steganography and Watermarking "*et al International Journal on Computer Science and Engineering* (IJCSE)

[4] KokSheik Wonga, Xiaojun Qib, Kiyoshi Tanaka "A DCT-based Mod4 steganographic method" .*A Signal Processing 87* (2007) 1251–1263

[5] Mohammad Tanvir Parvez and Adnan Abdul-Aziz Gutub "RGB Intensity Based Variable-Bits Image Steganography" *College of Computer Sciences & Engineering. King Fahd University of Petroleum & Minerals, Dhahra*

[6]Pramitha, K., Padma Suresh, L., and 2Shunmuganathan,K.L" Image Steganography using MOD-4 Embedding Algorithm based on Image Contrast". *International Journal of Current Research* Vol. 33, Issue, 6, pp.134-138, June, 2011.

[7]Rafael C.Gonzalez Digital Image Processing Using MATLAB Second Edition University of Tennessee Richard E. Woods, Med Data Interactive .

[8] Raphael.C.Gonzalez, Richard E Woods Digital image processing,2ndedition