

A New Steganography Technique Based on Difference Scheme of RGB Channels and Text Using Histogram Analysis

Manoj Kumar Sharma, Noor Mohd, Ratika Sharma

M. Tech Scholar, Department of Computer Science and Engineering, Graphic Era University Dehradun, India
Assistant Professor, Department of Computer Science and Engineering, Graphic Era University Dehradun, India

B. Tech Scholar, Department of Comp.Sci.and Engg., Accurate Institute Of Management and Technology, Greater Noida, India

Abstract

Steganography is the art of hiding the fact that communication is taking place, by hiding information in other information. The motivation for this work includes provision of protection of information during transmission without any detection of information. We also propose a new technique namely 'A New Steganography Technique Based on Difference Scheme of RGB Channels and Text Using Histogram Analysis'. The proposed technique can encode any colored image files in order to protect confidential text data from unauthorized access. It can be applied to a very small images (24×24) as well as large images (1024×1024). We use Image quality parameters Peak Signal to Noise Ratio and Mean Square Error. The proposed method has higher PSNR value and lower MSE values. The PSNR value is greater than 50 and MSE value is in fractions.

Keywords— Steganography, Cryptography, Data Hiding, PSNR, MSE, LSB, Histogram.

I. INTRODUCTION

Information security means protecting information and information systems from unauthorized access, use, disclosure, disruption, modification or destruction. The terms information security, computer security and information assurance are being used frequently interchangeably. These fields are interrelated and share the common goals of protecting the confidentiality, integrity and availability of information. However, there are some subtle differences between them. These differences lie primarily in the approach to the subject, the methodologies used and the areas of concentration. Information security is concerned with the confidentiality, integrity and availability of data regardless of the form the data may take electronic, print or other forms. Computer security can focus on ensuring the availability and correct operation of a computer system without concern for the information stored or processed by the computer.

The word steganography comes from Greek word "steganos" meaning "covered" and the "graphy" means "writing". Thus, steganography literally means "covered writing". Moreover, it has taken some meaning from the dinosaur called the Stegosaurus. Stegosaurus was an ancient creature that had the unique feature of being covered by a series of vertical plates along its spine, hence the "Stego" part in its name.

Fig.1.1 shows a simple representation of the generic embedding and decoding process in steganography.

'In this process, a secret data is being embedded inside an original image to produce the stego image.

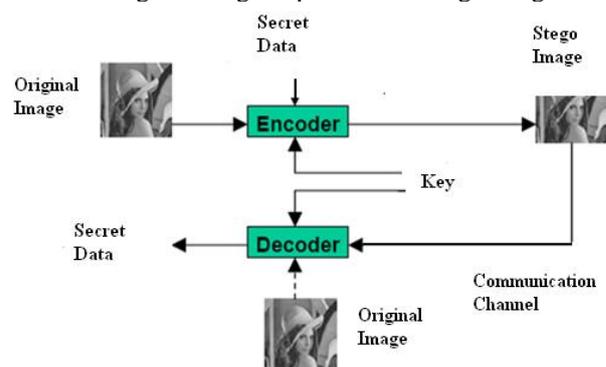


Fig. 1.1 Generic Processes of Encoding and Decoding

A. Requirements of Hiding Data Digitally

There are many different protocols and embedding techniques that enable us to hide data in a given object. However, all of the protocols and techniques must satisfy a number of requirements so that steganography can be applied correctly [15].

The following is a list of main requirements that steganography techniques must satisfy:

- The integrity of the hidden information after it has been embedded inside the stego object must be correct. The secret message must not change in any way, such as additional information being added, loss of information or changes to the secret information after it has been hidden. If secret information is changed during

steganography, it would defeat the whole point of the process.

- The stego object must remain unchanged or almost unchanged to the naked eye. If the stego object changes significantly and can be noticed, a third party may see that information is being hidden and therefore could attempt to extract or to destroy it.
- In steganography, changes in the stego object must have no effect on the hidden data. Imagine if you had an illegal copy of an image that you would like to manipulate in various ways. These manipulations can be simple processes such as resizing, trimming or rotating the image. The hidden data inside the image must survive these manipulations, otherwise the attackers can very easily destroy the data and the point of steganography will be broken.
- Finally, we always assume that the attacker knows that there is hidden information inside the stego object.

B. Objectives of This Proposed Technique

In this paper, various Steganographic techniques have been studied and analyzed. A New Steganography Technique based on Difference Scheme of RGB Channel and Text using Histogram Analysis is proposed.

The main objectives of this proposed work are as follows:

- To study various available image steganography techniques.
- To design and develop a new image steganography technique that can hide data in the colored images.
- To study histogram analysis of cover image in (R-G-B) channel and secret message.
- To study quality metrics (MSE, RMSE, PSNR) of the cover image and stego images in (R-G-B) channel.
- To study histogram analysis of cover image and stego image in (R-G-B) channel.
- To evaluate effectiveness of proposed scheme with previous best known schemes by comparing results.

II. RELATED WORK

Due to rapid development in both computer technologies and internet, the security of information is regarded as one of the most important factors of information technology and communication. Accordingly, there is needed to take measures which protect the information securely.

Generally, secret information may be hidden in one of three ways, such as cryptography, digital watermarking and steganography. Digital watermarking is only used for authentication purpose. For confidentiality and integrity purpose,

cryptography and steganography methods are used. The methods of cryptography render the data unintelligible to outsiders by various transformations, whereas the methods of steganography conceal the existence of messages. Among the methods of steganography, the most common thing is to use images for steganography. This is called image steganography. The image used to camouflage the secret data is called the original image or cover image while the cover image with the secret data embedded in it is called the stego image.

A. Spatial Domain Image Steganography Technique

Now we discuss spatial domain image steganography techniques that are based on complexity, spread spectrum, adaptive LSB and pair wise bit.

a) Complexity Based Region Segmentation Method to Embed Image

Niimi et al. (1997) [17] have discussed a technique to embed secret data into a dummy image by using image segmentation based on a local complexity measure. The key idea to this approach is that a binary image can be categorized as “informative” and “noise like” regions, which are segmented by a “complexity measure”. If the embedding data is noise-like, one can hide it in the noise like region of the dummy image. If a part of embedding data is simple, then one can apply “image conjugate” operation to it. This operation transforms a simple pattern into a complex pattern. This conjugate operation does not lose any information. The PSNR value found is about 30-33 db which shows a low quality stego image that is open for further improvement.

b) Pair Wise Bit Based Data Hiding Approach on 24 bit Image

Ghosal (2011) [25] has given a method to hide information within the spatial domain of the 24 bit color image. This technique works by considering the three channels (viz. red, green and blue) of each pixel of the cover image one by one up to the (maximum, if desire) last pixel and calculating the number of ones and zeroes in the red channel. Then, we calculate the absolute difference value of the number of zeroes and number of ones which is again divided by the total embedding channel numbers viz. green and blue which is 2 for a 24 bit color image. The resultant number of bits of the hidden data is embedded on the LSB part (in bit range of 0-3) of the green and blue bytes (channels) of each pixel of the cover image respectively.

e.g. (R, G and B) bit pattern for two consecutive pixels of a 24-bit color image is as shown below

```

11011011    00010110
10000011

01001100    00110110
10101011
    
```

Now, if we want to embed a character A (has the binary value 01000001), we need to follow the above method. So, as per our method:

The number of 1's in Red byte is 6
 The number of 0's in Red byte is 2
 So the absolute difference value is (6-2) =4
 Dividing the above result by 2 yields=4/2=2.
 So, bit embedded on the LSB part of the green and blue byte is 2. Also, for the second R byte the bit embedded on the LSB part of the blue bite is 1.
 Now, the bit stream of the stego image will be as follow shown in below:

```

11011011    00010101
10000000

01001100    00110100
10101001
    
```

So, by replacing only 6 bits in 4 numbers of selected bytes, we can hide the binary string 01000001.

III. PROPOSED MEHOD

The proposed algorithm is based on difference scheme of RGB channels and text using histogram analysis which improves the quality parameters like Peak Signal to Noise Ratio (PSNR) and Mean Square Error (MSE). The security feature of this technique is more reliable to protect confidential text data from various other previous steganography techniques. This proposed scheme can encode any image files (JPEG, BMP, DIB) having 24 bits per pixel in order to protect confidential text data from unauthorized access. The technique can be applied to very small images (24 x 24) as well as to large images (1024 x 1024).

In this technique, the text data is not hidden directly into any of the RGB channel pixel but the difference of the text data from one of the RGB channel pixel is hidden into the first six pixels LSB of the third channel. The channel selected for taking the difference between the channel pixels and text data is determined by the Histogram Analysis called selected channel. The second channel pixel LSBs are used as the pixel indicator which indicates the corresponding selected channel pixel which is used for the difference with the first character of the text message. The results show that the quality parameters, PSNR values are much higher and MSE values are lower in fractions compared to previous existing image steganography techniques.

The following three Phase of Proposed algorithm are::

- **Histogram Analysis.**
- **Embedding of secret message in cover image.**
- **Extraction of secret message from stego image.**

a. Histogram Analysis

In a computer, images are represented as arrays of values. These values represent the intensities of the three colors R (ed), G (reen) and B (lue), where a value for each of the three colors describes a pixel. Figure3.1 shows an image containing group of pixels. According to RGB model, a pixel contains three color planes that are red, green and blue planes.

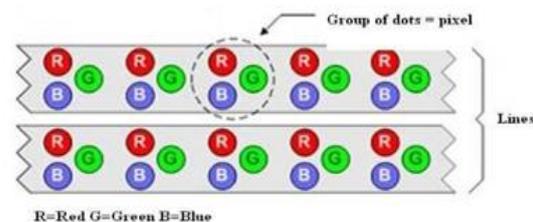


Fig. 3.1 Image containing Group of Pixels

b. Embedding of Secret Message in Cover Image

The proposed new steganography technique embeds the data in a image. Firstly the second channel, used as the pixel indicator it's every pixel LSBs are set to zero. Now we start traversing each row sequentially of selected channel matrix and note the difference of the first character of the secret message with the first pixel of this channel till the difference falls in the range of -63 to 63 then this difference is noted and the corresponding second channel pixel LSB is set to one. Similarly, we move forward sequentially to note difference between the selected channel pixels and the characters of the text to be embedded.

The difference noted first time is embedded into the first six pixels LSBs of the third channel, the second difference is noted into the next six pixels LSBs of the third channel and so on. Thus, the message is embedded into the cover image by embedding the difference between the pixel of a selected channel and the character of the secret text message into this channel. This text embedded image is called the stego image.

Embedding Phase: We will describe the embedding process of this paper proposed scheme. We summarize the process step-by-step as follows.

Algorithm 1. (The embedding process)

Input: A cover-image of $w \times h$ and secret data.

Output: A stego-image of $w \times h$.

- **Step 1:** Select a cover-image of $w \times h$ and secret data for hiding.

- **Step 2:** Compute the histogram corresponding to the secret data with R, G and B plane of cover image.
- **Step 3:** Initialize the LSBs of planes as zero say channel c_i ; c_j , except the one say channel c_k which corresponds to the maximum probability of matching with secret data.
- **Step 4:** Compare stego key and secret data with channel c_k sequentially in the span of -63 to 63.
- **Step 5:** At matched point embed the difference with LSBs of channel c_i and a '1' on the LSB of channel c_j .

c. Extraction of Secret Message from Stego Image

Extraction of the secret message from the images can be done in the reverse way of the same technique, but the key must be the same as used in embedding the secret message into the images i.e. sender and receiver should know common stego key. If the key is not same, it is not possible for the receiver to decode the secret message from the stego image. Thus, this feature of matching key at the receiver side makes this technique a bit more secure in terms of attacks. The information of the histogram analysis is also transmitted from the sender to the receiver i.e. the receiver must know which channel are used for embedding difference, and pixel indicator.

Here, firstly we separate the three R(ed) G (reen) B (lue) channel from the stego image. After the stego key is matched, the second channel which is used as the pixel indicator channel is traversed sequentially and each pixels LSBs is observed, if the LSB of the pixel observed found to be 1, it's corresponding first channel pixel value is noted down and at the same time the difference value stored in the first six pixels LSBs of the third channel is recorded, now the difference of these two values gives the secret text character of the message. Similarly the process is repeated to extract the whole secret message.

Extracting Phase: The extracting process is just reverse of the embedding process. We summarize the process step-by-step as follows.

Algorithm 2. (The extracting process)

Input: A stego-image of $w \times h$.

Output: A secret data.

- **Step 1:** Separate the c_i ; c_j ; c_k channel from the stego image.
- **Step 2:** Scan for '1' in LSB of channel c_j .
- **Step 3:** At matched point put corresponding values of channel c_k in array A[i] and difference from LSB of channel c_i in array B[i].
- **Step 4:** Retrieve the secret message from the data array A[i] and difference array B[i].

IV. RESULT AND ANALYSIS

A. Image Quality Parameters

Image quality parameters are figures of merit used for the evaluation of imaging system or processes. The image quality parameters can be broadly classified into two categories [6], subjective image quality and objective image quality. Subjective image quality is a method of evaluation of images by the viewers and it emphatically examines fidelity and at the same time considers image intelligibility. In the objective measures of the image quality metrics, some statistical indices are calculated to indicate the reconstructed image quality. The image quality parameters 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 parameters for comparing stego image and original image are:

- Mean Square Error (MSE),
- Peak Signal to Noise Ratio (PSNR),
- Root Mean Square Error (RMSE)

a) Mean Square Error (MSE)

The mean of pixel values of the image and by averaging the sum of squares of the error between two images.

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

Where $x(m, n)$ and $y(m, n)$ are the two images of size $M \times N$. In this case x is the original image and y is the stego image.

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

b) 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 standard (benchmark) way to measure image reliability or conformity.

$$PSNR = 20 \log_{10} \left[\frac{MAXPIX}{MSE} \right]$$

Where MAXPIX is the maximum pixel value and RMSE is the Root Mean Square Error of the image (it quantifies the average sum of distortion in each pixel of the stego image i.e. average change in pixel caused by encoding algorithm)

$$RMSE = \sqrt{MSE}$$

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).

C. Results of Proposed Algorithm

The simulation part of this paper, works with two cover images Lena and Peppers. The results of the simulation for these two images are tabulated and shown here. The original cover image and its stego image for Lena and peppers are shown in a figures 4.1 and 4.2 and results based on the quality metrics (i.e. values of MSE, RMSE, and PSNR) are tabulated in Table 4.1 and 4.2.



Fig. 4.1(a) Cover Image Fig. 4.1(b) Stego Image

TABLE 4.1 MSE, RMSE and PSNR value of Lena image

Cover Image	MSE	RMSE	PSNR(dB)	Capacity (bytes)
Lena	0.1116	0.3341	52.8918	25,000
Lena	0.1115	0.3339	52.8950	43,690



Fig. 4.2(a) Cover Image Fig. 4.2(b) Stego Image

TABLE 4.2 MSE, RMSE and PSNR value of Peppers image

Cover Image	MSE	RMSE	PSNR(dB)	Capacity (bytes)
Peppers	0.1093	0.3306	52.9374	25,000
Peppers	0.1094	0.3308	52.9348	43,690

The histogram analysis of the cover image and stego image for Lena and peppers is done shown in the figure 4.3 and 4.4 respectively which shows that after embedding the data the distortion in the stego image is not noticeable as the histogram of the cover image and stego image are very similar which accounts for better stego image quality.

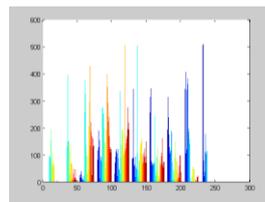


Fig. 4.3(a) Cover Image

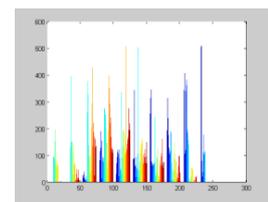


Fig. 4.3(b) Stego Image

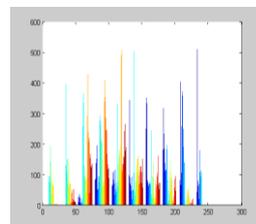


Fig. 4.4(a) Cover Image

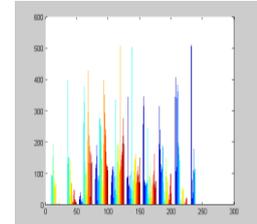


Fig. 4.4(b) Stego Image

Also, from the results analyzed above we see if the data embedded into the image is increased then also there is a minor change in the values of MSE, RMSE and PSNR it can either increase or decrease too, which specifies if the embedded data stored in the cover image is increased, then also the quality of the retrieved stego image is better.

D. Comparative Performance

The best known steganography techniques discussed in the literature survey are compared with the results of the proposed work on the basis of quality metrics MSE, RMSE and PSNR values for the image lena (based on the availability of the results). Proposed scheme provides better results than all the previous schemes.

The results compared for the MSE values of Lena and Peppers image are shown in table 4.3 and 4.4 respectively.

TABLE 4.3 Comparison of MSE values for Lena Image

Steganography Technique	MSE
LSB3 Technique	5.29
SSIS Technique	25.7049
HCSM Technique	2.34
OLS Technique	2.34
OLSGA Technique	2.34
OLSDPS Technique	2.34
Proposed Work	0.1116

TABLE 4.4 Comparison of MSE values for Peppers image

Steganography Technique	MSE
OLS Technique	2.35
OLSGA Technique	2.35
OLSDPS Technique	2.35
Proposed Work	0.1113

The results compared for the PSNR values of Lena and pepper image are shown in table 4.5 and 4.6.

TABLE 4.5 Comparison of PSNR values for Lena Image

Steganography Technique	PSNR
LSB3	40.89
Adaptive No. Of LSBs	40.36
PVD	41.58
2k edge correction	42.09
Proposed Work	52.93

TABLE 4.6 Comparison of PSNR values for Peppers Image

Steganography Technique	PSNR
LSB3	40.86
Adaptive No. Of LSBs	40.27
PVD	41.66
2k edge correction	42.47
Proposed Work	52.93

V. CONCLUSION AND SCOPE OF THE FUTURE WORK

In this work, we have proposed a new steganography technique based on difference scheme of RGB channels and text using histogram analysis, which is a spatial domain technique. This approach has various advantages as compared to the other techniques available in literature. Thus these advantages are:

This approach supports the image size of $m \times m$ dimension. The security of this technique is enhanced from the previous steganography techniques. Since we embed the difference of the data with one channel into the other channel thus the attacker does not have direct access to data. Since we do not need to send the cover image to receiver side, it is a blind approach so which makes it more secure.

The image quality metrics i.e. higher Peak Signal to Noise Ratio (PSNR) and lower Mean Square errors itself prove that the proposed technique has good quality of the stego-images that is highly acceptable by the human eyes. The values of PSNR obtained are greater than 52 dB and MSE value is less than 1.

The future work can be to provide an alternative traversing rather than sequential scan of pixels so that the hiding capacity can be enhanced. It can also be made rotational and transformation invariant.

REFERENCES

[1] <http://www.appliedtrust.com/resources/security/every-company-needs-to-have-a-security-program>

[2] http://www.vtc.com/products/Fundamentals_ofComputerSecurity/Basic_Security_Principles_Terms_I/72584

[3] <http://en.wikipedia.org/wiki/Cryptography>

[4] <http://knol.google.com/k/cryptography-and-data-security#>

[5] W. Stallings, "Cryptography and Network Security: Principles and Practice," Prentice-Hall, New Jersey, 1999.

[6] C. Shoemaker, "Hidden bits: A survey of techniques for digital watermarking", Independent study, EER 290, spring 2002.

[7] http://www.isoc.org/inet2000/cdproceedings/8g/8g_4.htm

[8] Herodotus, "The Histories, chap. 5 - The Fifth Book Entitled Terpsichore, 7 - The Seventh Book Entitled Polymnia", J. M. Dent & Sons, Ltd, 1992.

[9] J. Caldwell, "Steganography", United States Air Force, <http://www.stsc.hill.af.mil/crosstalk/2003/06/caldwell.pdf>, June 2003.

[10] <http://dictionary.reference.com/browse/steganography>

[11] C. Cachin, "An Information-Theoretic Model for Steganography", In Proceedings and of 2 Workshops on Information Hiding, MIT Laboratory for Computer Science, May 1998.

[12] <http://www.cs.bham.ac.uk/~mdr/teaching/modules/security/lectures/Steganography.htm>

[13] J. Cummins, P. Diskin, S. Lau and R. Parlett, "Steganography and Digital Watermarking," www.cs.bham.ac.uk/~mdr/teaching/modules03/security/students/ss5/Steganography.pdf, 2004.

[14] Shashikala Channalli et al International Journal on Computer Science and Engineering Vol.1(3), 2009.

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

[16] J M. Niimi, H. Noda and E. Kawaguch, "An image embedding in image by a complexity based region segmentation method," in Proceedings of the 1997 International Conference on Image Processing (ICIP '97), pp. 74, 1997.

[17] J W.N. Lie and L.C. Chang, "Data hiding in images with adaptive numbers of least significant bits based on the human visual system," In Proceedings of IEEE International Conference on Image Processing., vol. 1, pp. 286-290, 1999.

[18] J Y.K. Lee and L.H. Chen, "High capacity image steganographic model", IEEE Proc. Vision Image Signal Process, vol. 147, pp. 288-294, June 2000.