International Journal of Engineering Research & Management Technology ISSN: 2348-4039

January-2017 Volume 4, Issue-1

Email: editor@ijermt.org www.ijermt.org

A NOVEL STEGANOGRAPHY TECHNIQUE BASED ON DIFFERENCE SCHEME OF RGB CHANNELS AND TEXT USING HISTOGRAM ANALYSIS

Ms. Bhawana Chaudhary
Department of Computer Science Engineering
Marathwada Institute of Technology
Bulandshar

Mr. Manoj Kumar Singh
Department of Computer Science Engineering
Marathwada Institute of Technology
Bulandshar

ABSTRACT: The art and science of hiding information has gained much attention. We are also surrounded by a world of secret communication, where people of all types are transmitting information as innocent as an encrypted credit card number to an online-store than and as insidious as a terrorist plot to hijackers. Steganography derives from the Greek word steganos, meaning covered or secret, and graphy (writing or drawing) . In this paper, we propose a new image steganography scheme for colored images based on the cluster analysis. In this scheme, we analyze the secret data in order to make its clusters. The secret data can be textual, image/video or audio/speech. A cluster contains ASCII values of characters if the secret data is text, sample values for audio/speech and pixel values in case it is image/video. We then calculate the difference value between the secret data and the minimum value contained in the cluster. We do not hide actual secret data; the difference value is embedded equally into two channels of the image. The experimental results show that our proposed method has enhanced security as compared to the modified Kekre algorithm [7] and pixel intensity based high capacity data embedding method [9]. Furthermore, our scheme has good hiding capacity, high PSNR value, and very low MSE value.

KEYWORDS: Steganography, Stego-Key, Stego-Image. Channel, cluster analysis, difference value, cover image, stego image, PSNR, MSE**.**

1. INTRODUCTION

Steganography is a mechanism of embedding secret data into the cover image so that the hidden message cannot be detected. The secret message can be text, audio/speech, or video/image.

Steganography comes from the Greek word steganos that means ‘hidden’, and graphy refers to ‘writing’. ‘Hidden writing’ has been existed since ancient Greek times around 440 B.C. [3]. The security of the hidden data can be increased if it is encrypted before hiding.

However, the encrypted data needs be decrypted before reading. Stegoimage consists of cover image, secret key, and secret data, i.e. stego image = cover image + secret key + secret data. The secret key is an encryption-decryption key and the secret data can be textual, audio/speech, or image/video data. There have been discussed several techniques in literature for hiding the secret data in an image. While hiding the data in an image the visual quality of the original image should not change much; otherwise a masquerade can suspect the presence of the secret data.

There are three main principles for hiding the data in an image: a) insertion of least significant bit (LSB) in some pre-specified pixels, b) masking and filtering, and c) transform based techniques. The high PSNR and low MSE values jointly signify that the secret message cannot be detected from the Stegoimage.

The high data capacity tells that the large amount of data can be conveyed/ hidden in the host image.

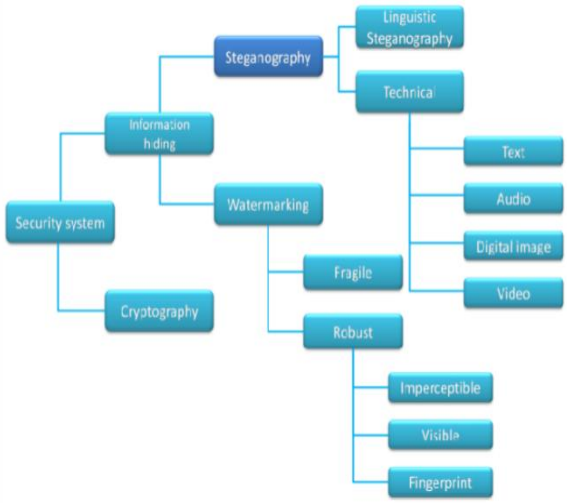


Fig 1. Security System

In our work, we use least significant bit (LSB) substitution principle for hiding the secret data into LSBs of the pixels. For example, in simple LSB substitution method, for colored images in which each pixel is represented in 24 bits (8 bits/pixel/color) can hide 3 bits of data, one bit in each color of the pixel: red (R), green (G) and blue (B) colors/channels. In our work, we first construct clusters of the secret data. The R channel maintains information about each cluster. The B and G channels are used for hiding the secret data. Some of the recently developed important methods for data hiding are modified Kekre algorithm (MKA) [7] and pixel intensity based high capacity data embedding method (PIBHCDEM) [9]. The MKA method hides the secret data based on the intensity of pixels and maintains a matrix for extracting the secret data. The PIBHCDEM works in similar way as the MKA works, but it increases the data hiding capacity. Our method has relatively better performance as compared to the MKA and PIBHCDEM in terms of PSNR, MSE and its data hiding capacity is also good. Moreover, it has enhanced security because it embeds the difference between the minimum value of each cluster and its other remaining elements rather than the actual data. The rest of the paper is organized as follows. Section II discusses the related work. Section III discusses the proposed work and, in section IV, the experimental results are discussed. Finally, in section V, the paper is concluded.

EMBEDDING PROCESS OF STEGANOGRAPHY

Figure 1.5 shows a simple representation of the generic embedding and decoding process in steganography. In this example, a secret data is being embedded inside an original image to produce the stego image.

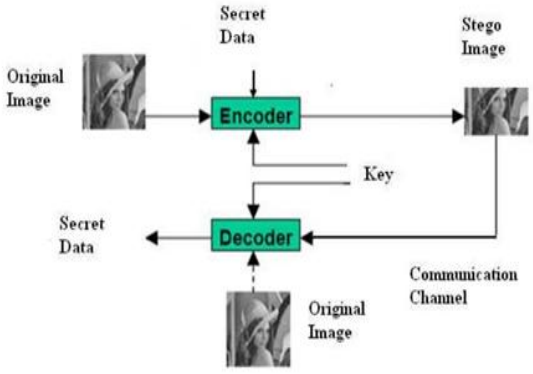


Figure 2: Generic Process of Encoding and Decoding

The first step in embedding and hiding information is to pass both the secret message and the original image into the encoder. Inside the encoder, one or several protocols will be implemented to embed the secret information into the cover message. The type of protocol will depend on what information you are trying to embed and what you are embedding it in. For example, an image protocol to embed information inside images is being used.

2. RELATED WORK

One of the important digital steganography methods is spread spectrum image steganography [2]. This method can hide a message of any form, be it image, text, audio and its capacity is of significant size. It can recover the hidden data and also does not change the size of original image. Chan and Cheng discuss a method - a simple LSB substitution method for hiding the secret data into the cover image [1]. In this method, the LSBs of pixels of the cover image are directly replaced by the bits of the secret data. The optimal LSB substitution is effective even in worst case. However, when the secret data is to be hidden in large number of LSBs of the cover image, it requires large computation for both embedding and extraction. The optimal solution in such cases is obtained by using the genetic algorithm.

Table 1. Pibhcdem Scheme [9]

Pixel Intensity	Data bit to embed	Matrix Entry	Utilize bits/Bit
240-255	1	1	5
240-255	0	-	4
224-239	0	1	5
224-239	1	-	3
192-223	0	1	3
192-223	1	-	2
32-191	0	1	2
32-191	1	-	1
16-31	0	1	3
16-31	1	-	2
0-15	0	1	5
0-15	1	-	4

However, in these schemes, no characteristics of human visual system have been explored. Wu et al. [6] have discussed an important method that exploits the characteristics of the human visual system and makes use of the least significant bit (LSB) substitution and pixel value differencing (PVD). In PVD method, the image is divided into non-overlapping blocks and then the difference value between two consecutive pixels for each block is calculated, which ranges from -255 to 255.

Hence, the block with large value is considered as an edge area and for small value, it is considered as smooth area. The human eyes are more sensitive to the area containing edges as compared to the smooth area. The small or large values are taken based on some pre-specified threshold value. This

method embeds more bits in edge area in contrast to the smooth area. Its main focus was on increasing the capacity of data to be embedded.

The hiding capacity increases if the image consists of more high intensity pixels otherwise the capacity remains low. Hussain [9] discusses a method that is an improvement of the MKA [7]. In that method, all bytes of the cover image are utilized and the lower intensity pixels are also significantly used. The matrix entry is maintained when 5, 3 or 2 bit of the secret data is embedded. The complete data embedding scheme of [9] is illustrated in Table-I. In this table, the pixel intensity field refers to the pixel intensity of the cover image, data bit to embed field refers to the bit of the secret data to be embedded and utilize bits/Bit refers to the number of bits of the resultant secret data embedded in the cover image. In matrix entry field, '1' is an indication of when matrix entry is maintained and '-' is used when we don't need matrix entry.

Again the MSE value is calculated. If the inverted bits of the secret data has lower MSE, then the secret data in the inverted form is embedded and the indicator bit is set as 1 for that block; otherwise the data is embedded unaltered and the indicator bit is set as 0. This process is repeated for each block of the cover image as well as for each piece of secret data. Zhang & Wang discuss a new embedding scheme called exploiting modification direction EMD embedding scheme for hiding the secret data into the cover image. In this method, each (2n + 1)-ary notational secret digit is hidden into n cover pixels, and only one pixel value increases or decreases by 1 at most. This method achieves good quality Stegoimage however results in low data hiding capacity.

3. PROPOSED METHOD

Our proposed method consists of two parts: embedding process and extraction process. These processes are discussed below.

3.1. Embedding Algorithm

- Load cover image and extract its channels (R, G and B) into arrays: red array, green array and blue array, respectively.
- Extract the secret data from the file into an array call it secret array.
- If the secret data is text, the characters are manipulated in terms of their ASCII values, which are integer values. If it is an image/video, then pixel values are manipulated, and for audio/speech data, the sample values are manipulated.
- Divide the entire resultant secret data into clusters with the following conditions:
  - A cluster should have number of elements in the range of 1 to 255.
  - Difference between the minimum and maximum values of a cluster should not be more than 63.
- Maintain the minimum value and the size of each cluster. Maximum number of clusters can be at most 16 bit long. This number is embedded in LSBs of first 16 pixels of R (red) channel.
- Represent the minimum value of each cluster in 8 bits and embed them into LSBs of pixels after 16 pixels of R channel as the first 16 pixel's LSBs contain number of clusters.
- Represent the cluster size in 8 bits and embed them into LSBs of the pixels after 8*(number of clusters +2) pixels of R channel.
- Find the difference of each element of a cluster with its minimum value for all clusters.
- Represent the difference for each cluster obtained in step (viii) into 6 binary bits.
- First embed 6 bits of the difference value into first two LSBs of three pixels of B (blue) channel, then 6 bits of next difference value into first two LSBs of three pixels of G (green) channel. Do it for each cluster.
- Concatenate all channels to reconstruct the original image, which is our stego image.

B. Extraction Algorithm

- Extract first LSB of first 16 pixels of R channel. Their decimal value gives the number of clusters.
- Extract first LSB of next pixels in group of 8 pixels. The decimal value of each group represents the minimum value of a cluster. The number of groups will be equal to the number of clusters.
- After extracting the minimum values of all clusters, we extract first LSB of the next pixels in group of 8 pixels. The decimal value of bits in a group represents the cluster size; thus such groups will be

equal to the number of clusters.

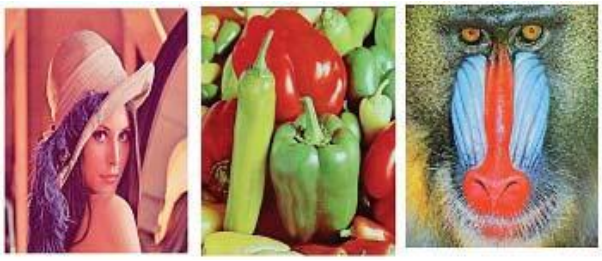
- Extract first two LSBs of pixels of B and G channels and form groups of 6 bits for each channel.
- The decimal value of this 6 bit-value from each channel (B and G) gives two difference values. It may be noted that we have already determined the size of each cluster in step (iii).
- Add these difference values obtained in step (v) of each element of a cluster to their minimum values. Perform this process for each cluster. The resultant values thus obtained give the ASCII value of two characters if secret data is text, for audio/speech it is two sample values and for video/image it is two pixel values.

Fig.3 Original color images

(a) Lena (b) Pepper (c) Baboon

Fig.4 Stego images

(a) Lena (b) Pepper (c) Baboon



4. EXPERIMENT RESULTS

Proposed model is stronger Steganography technique because without knowing the secrete keys, S-box mapping function, the extraction of secrete image from the stego image is impossible. Moreover quality of cover image is also not degrading due to variation in two LSB of each pixel which reflects only 0 – 3 difference pixel value. Additionally the proposed scheme is capable of not just scrambling data but it also changes the intensity of the pixels which contributes to the safety of the encryption

TABEL 2: CAPACITY & PSNR

Name of Image	Size (pixel)	Capacity	PSNR In DB
Baboon	64*64	25%	54.58
Cameraman	64*64	25%	55.01
Lena	64*64	25%	59.28
Pirate	64*64	25%	51.63
Living room	64*64	25%	50.19
Women-dark hair	64*64	25%	52.85

In this section, we discuss the performance results of our proposed steganography method. We have taken three commonly used color images in literature for hiding the secret data.
They include Lena, Pepper, and Baboon, each of size 512x512 as shown in Figs. 1(a)-(c). We have compared our results with that of MKA [7] and PIBHCDEM [9]. The reason for making comparison with these methods is that they are more recently developed and have good performance. The secret data taken in our experiments is same as used in MKA [7] and PIBHCDEM [9], which is Abraham Lincoln's letter to his son's teacher. The secret data which is of size 1785 bytes has been embedded into each of these images. Taking same secret data helps making meaningful comparisons of the performance of our method with these methods.

The resultant stegoimages (image with hidden secret message) by employing our proposed method are shown in Figs. 2(a)-(c). We have not shown stegoimages by employing MKA and PIBHCDEM because their visual difference is hardly discernable. The performance results in terms of quality metrics (i.e. PSNR, MSE) for hiding capacities in bytes for different steganography techniques are shown in Table II.

It is evident from Table II that our proposed method has much higher PSNR value and lower MSE value than the MKA and comparable to that of the PIBHCDEM for all three images. As far as embedding capacity is concerned, our proposed method performs better than MKA method but it is no better than the PIBHCDEM method. The reason of having higher capacity in the PIBHCDEM method is that it embeds more bits (three bits on average) of the secret data in low intensity pixels, whereas our method embeds less bits (two bits on average) per pixel Embedding more bits in low intensity pixels deteriorates the visual quality of the stego image.

Table 3. PSNR, MSE and CAP (Maximum embedding capacity) in bytes of different approaches on different cover images of size 512×512 pixels each

Cover image		Modified Kekre Algorithm	PIBHC DEM Method	Proposed Algorithm
Lena	PSNR	64.0778	65.7180	68.4471
	CAP	127402	186556	174762
	MSE	0.0254	0.0174	0.0093
Baboon	PSNR	64.0778	69.6609	98.3927
	CAP	117631	172538	174762
	MSE	0.0107	0.0070	0.0094
Pepper	PSNR	66.2061	59.5402	68.7835
	CAP	114718	196081	174762
	MSE	0.0156	0.0723	0.0086

5. CONCLUSION

In this paper, we have proposed a new steganography technique based on the similarity among the basic elements of the secret data. The basic elements in the text, image/video and audio/speech are characters, pixels and samples, respectively. This technique performs comparably better than the MKA and PIBHCDEM in terms of PSNR, MSE. Its embedding capacity is also good.

Furthermore, it has enhanced security as compared to other methods. The reason being the difference between the minimum value of each cluster and its other remaining elements is embedded

rather than the actual data. Even if an attacker is able to access the hidden data, he will get only difference values not the actual secret data.

6. REFERENCES

1. Marvel, L.M., 1999. Spread Spectrum Image Steganography. IEEE transactions on image processing 8(8), 1075-1083.
2. Khalaf, E.T., Sulaiman, N., 2011. Segmenting and Hiding Data Randomly Based on Index Channel. International Journal of Computer Science. 8(3), 522-529.
3. Wang, R.Z. Lin, C.F., Lin, J.C., 2001. Image hiding by optimal LSB Substitution and genetic algorithm. Pattern Recognition. 34, 671-683.
4. Marvel, L.M., 1999. Spread Spectrum Image Steganography. IEEE transactions on image processing 8(8), 1075-1083.
5. C.K. Chan, L.M. Cheng, Hiding data in images by simple LSB substitution, Pattern Recognition 37 (3) (2004) 469–474.
6. Marvel, L.M., 1999. Spread Spectrum Image Steganography. IEEE transactions on image processing 8(8), 1075-1083.
7. Khalaf, E.T., Sulaiman, N., 2011. Segmenting and Hiding Data Randomly Based on Index Channel. International Journal of Computer Science. 8(3), 522-529.
8. Wang, R.Z. Lin, C.F., Lin, J.C., 2001. Image hiding by optimal LSB Substitution and genetic algorithm. Pattern Recognition. 34, 671-683
9. Hussain, M., Hussain. M., 2010. Pixel Intensity Based High Capacity Data Embedding Method. International conference on Information and Emerging Technologies, 1-5.
10. C. Cachin, "An Information-Theoretic Model for Steganography", In Pro-ceedings nd of 2 Workshops on Information Hiding, MIT Laboratory for Computer Science, May 1998.
11. Wu. H.C., Wu. N.I., Tsai, C.S., Hwang, M.S., 2005. Image Steganographic scheme based on pixel value differencing and LSB replacement methods. IEE Proc. Vision Image Signal Process. 152, 611-615.
12. M. Niimi, H. Noda and E. Kawaguch, "An image embedding in image by a complexity 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.
13. C.K. Chan, L.M. Cheng, Hiding data in images by simple LSB substitution, Pattern Recognition 37 (3) (2004) 469–474.
14. Marvel, L.M., 1999. Spread Spectrum Image Steganography. IEEE transactions on image processing 8(8), 1075-1083.
15. C. Cachin, "An Information-Theoretic Model for Steganography", In Pro-ceedings nd of 2 Workshops on Information Hiding, MIT Laboratory for Computer Science, May 1998.
16. www.cs.bham.ac.uk/~mdr/teaching/modules03/security/students/SS5/Steganography.pdf, 2004.
17. Shashikala Channalli et al International Journal on Computer Science and Engineering Vol.1 (3), 2009.
18. L.M. Marvel, "Spread Spectrum Image Steganography," IEEE transactions on image processing, vol. 8, no. 8, pp. 1075-1083, August 1999.
19. 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 474.
20. Wu., H.C., Wu., N.I., Tsai, C.S., Hwang, M.S., 2005. Image Steganographic scheme based on pixel value differencing and LSB replacement methods. IEE Proc. Vision Image Signal Process. 152, 611-615.
21. Amirtharajan, R., Nathella. K., Harish, J., 2010. Info Hide – A Cluster Cover Approach. International Journal of Computer Applications. 3(5), 11-18.
22. Hussain, M., Hussain., M., 2010. Pixel Intensity Based High Capacity Data Embedding Method. International conference on Information and Emerging Technologies, 1-5.