Modulo based Image Steganography Technique against Statistical and Histogram Analysis

V. Nagaraj
Research Scholar,
ECE Dept., Pondicherry
Engineering College.
Pondicherry, India.

Dr. V. Vijayalakshmi
Assistant Professor,
ECE Dept., Pondicherry
Engineering College.
Pondicherry, India.

Dr. G. Zayaraz
Associate Professor,
CSE Dept., Pondicherry
Engineering College.
Pondicherry, India.

ABSTRACT

Steganography is the art and science of communicating secret data by embedding it into a multimedia carrier. The ultimate goal here is to conceal the very existence of the embedded data. Although the term Steganography has been known for thousands of years, its digital version came about only lately and research was intensified after the depressing event of Twin towers (11th Sep 2001). Steganalysis, which is the official counter attack science, defeats Steganographic algorithms whether they are based on the traditional spatial domain or the transform one. Steganography's ultimate objectives, which are undetectability, robustness and capacity of the hidden data, are the main factors that separate it from other relating techniques, namely watermarking and cryptography. This paper focuses on improving the embedding capacity in steganography. For the improvement in the capacity, secret bits are encoded into cover image by three types of modulo functions. More specifically, to alleviate further color distortion and obtain a higher hiding capacity, the R, G and B component is encoded by Mod u, Mod v and Mod w functions respectively. Simulations were performed on cover images with different message sizes. The proposed technique showed that improved data hiding capacity in the cover image can be obtained without any compromise in histogram and statistical analysis. It also gives good perceptual quality in Stego image and greater security.

Keywords

Steganography, adaptive algorithm, spatial domain, frequency domain, security, embedding payload

1. INTRODUCTION

Steganography or Stego as it is often referred to in the IT community, literally means, covered writing which is derived from the Greek language. Steganography is defined by Markus Kahn as follows, "Steganography is the art and science of communicating in a way which hides the existence of the communication". In contrast to Cryptography, where the enemy is allowed to detect, intercept and modify messages without being able to violate certain security premises guaranteed by a cryptosystem, the goal of Steganography is to hide messages inside other harmless messages in a way that does not allow any enemy to even detect that there is a second message present. In a digital world, Steganography and Cryptography are both intended to protect information from unwanted parties. It is for this reason that most experts would suggest using both to add multiple layers of security. Steganography can be used in a large amount of data formats in the digital world of today. The most popular data formats used are .bmp, .doc, .gif, .jpeg, .mp3, .txt and .wav. mainly because of their popularity on the Internet and the ease of use of the steganographic tools that use these data formats. These formats are also popular because of the relative ease by which redundant or noisy data can be removed from them and replaced with a hidden message.

Steganographic technologies are very important face of the future of security and privacy on open systems such as the Internet. Steganographic research is primarily driven by the lack of strength in the cryptographic systems on their own and the desire to have complete secrecy in an open systems environment. Many governments have created laws that either limit the strength of cryptosystems or prohibit them completely. This has been done primarily for fear by law enforcement not to be able to gain intelligence by wiretaps, etc. This unfortunately leaves the majority of the Internet community either with relatively weak and most of the time with breakable encryption algorithms or none at all. Civil liberties advocates fight this with the argument that "these limitations are an assault on privacy". This is where Steganography comes in. Steganography can be used to hide important data inside another file so that only the parties intended to get the message even knows a secret message exists. To add multiple layers of security and to help subside the "crypto versus law" problems previously mentioned, it is a good practice to use Cryptography and Steganography together. As mentioned earlier, neither Cryptography nor Steganography are considered "turnkey solutions" to open systems privacy, but using both technologies together can provide a very acceptable amount of privacy for anyone connecting to and communicating over these systems.

This paper focuses on a new technique of modulo substitution of encrypted data. This technique achieves greater embedding capacity of data along with greater security. This paper is organized as introduction to steganography in section 1. Related work on steganography is presented in section 2. Encoding and decoding algorithms of the proposed modulo based image steganography technique are discussed in section 3. Section 4 gives the experimental results. Section 5 and 6 provides discussions on histogram and statistical analysis. Finally, the conclusion of the paper is given in Section 7.

2. RELATED WORK

In recent years steganography techniques received much attention from the research world, many researchers have spent considerable effort in designing several steganography algorithms.

F. Petioles *et al* [1] explained the information hiding techniques and also briefed the future directions in hidden communication system. An obvious method was to hide a secret message in every nth letter of every word of a text message. Text steganography using digital files is not used very often since text files have a very small amount of redundant data.

W. Bender *et al* proposed a technique for data hiding techniques for addressing the data hiding process and evaluated these techniques in light of three applications: copyright protection, tamper proofing, and augmentation data embedding [2]. H. Farid in his work [3] presents improved methods for information hiding. J. Fridrich *et al* in his work [4] describes liable and accurate method for detecting Least Significant Bit (LSB) non sequential embedding in digital images.

Practical Steganalysis of Digital images has explained [5] the problems encountered in deploying jpeg images and the universal blind detection schemes and special cases such as JPEG compatibility steganalysis. A Double Layered “Plus-Minus One” Data Embedding Scheme has been developed by W. Zhang *et al*., [6]. This work can hide a longer message than simple LSB embedding.

Attacks on Steganography Systems were analyzed [7] by A. Westfeld and A. Pfitzmann. X. Zhang and S. Wang in their paper [8] efficient steganography embedding by exploiting modification direction deals with the steganography embedding process especially on the Internet. H. Zhang and H. Tang proposed in their paper [9] a novel image steganography algorithm against statistical analysis. This is an easiest method for embedding messages in an image with high capacity. It is not detectable by statistical analysis such as RS and Chi-square analysis.

J. Kang *et al* in their paper [10] Steganography using block based adaptive threshold method used encoder and decoder design for image steganography. Attack against Statistical Steganalysis and given the proliferation of digital images on the Internet, and given the large amount of redundant bits present in the digital representation [11] of an image was dealt in depth by N. Provos.

The basic concept of a modulus operation in steganography is developed by C.C. Thien and J.C. Lin. In their method use of the modulus function improves capacity by overcoming the falling-off-boundary problem and provides good image quality by minimizing the changes in pixel values [12].

C.M. Wang, *et al* [13] proposed method, implements modulo function and an optimal approach to alter the remainder so as to greatly reduce the image distortion caused by the hiding of the secret data. It can also solve the falling-off-boundary problem by readjusting the remainder of the two pixels, staying secure against the RS detection attack.

J.C. Joo, *et al* in their paper [14] Steganalytic measures is designed using these weaknesses of the modulus (Pixel value differencing) PVD steganography method. Through experiments, they proved that the steganalytic measures successfully defeat the modulus PVD steganography method.

C.F. Lee and H.L Chen [15] proposed a technique using the simple modulus function, higher embedding capacity can be obtained for a larger divisor, while also obtaining a higher visual quality of a stego image using a smaller divisor. This scheme additionally yields lower computational costs and memory needs in its embedding and extracting procedures.

J.C. Joo, *et al* in their[16] paper made two main contributions to steganographic technique: a turnover policy and a novel adjusting process. The turnover policy yields histograms close to those of the cover image and maintains the symmetry of the PVD histogram. The novel adjusting process helps to remove fluctuations around the border of the subrange. The proposed method therefore maintains the PVD histogram similar to that of the cover and is secure against RS analysis.

S. Lyu and H. Farid [17] explains the first and higher order magnitude and phase statistics are relatively consistent across a broad range of images, but are disturbed by the presence of embedded hidden messages. The proposed method is examined on a large collection of images, and on eight different steganographic embedding algorithms.

Many data hiding schemes have been developed and proposed. In steganography, one simple and well known scheme is the least significant bit (LSB) replacement scheme, C.K. Chan and L.M. Cheng proposed simple LSB substitution with an optimal pixel adjustment process (OPAP). The LSB bits of cover pixels to embed secret data. In the LSB replacement scheme, a secret bit b is embedded into a cover pixel x by performing (x-1 + b) for the odd value x or (x + b) for the even value x. The embedded data are not easy to detect because the modification is asymmetric. Many existing LSB based schemes [18] have been used to try to improve the hiding capacity or to reduce the distortion of the stego image.

W. Luo, F. Huang and J. Huang [19] in their paper LSB matching revisited image steganography and propose an edge adaptive scheme which can select the embedding regions according to the size of secret message and the difference between two consecutive pixels in the cover image. For lower embedding rates, only sharper edge regions are used while keeping the other smoother regions as they are. When the embedding rate increases, more edge regions can be released adaptively for data hiding by adjusting just a few parameters.

H. Yang, Xingming SUN and Guang SUN [20] proposed scheme obeys the principle that edge areas cannot endure abrupt changes, and can embed a large number of secret data while achieving high quality of the stego image.

Masoud Afrakhteh, and Subariah Ibrahim [21] in their proposed method used at least four numbers of eight surrounding pixels of a target pixel. More the pixels used for estimating the capacity, higher is the image quality achieved. In this paper the main focus is to represent an improvement made in terms of imperceptibility in comparison with existing method.

3. PROPOSED MODULO BASED IMAGE STEGANOGRAPHY TECHNIQUE

The human eyes are more sensitive to the change in color images than gray scale images. A little color distortion could apparently appear on the resulting stego images. The objective of this paper is to propose a data hiding method for color image with better embedding capacity and to provide greater security. For attaining this Mod u, Mod v and Mod w functions are employed into the proposed method. The details of the proposed modulo based image steganography technique is briefly described in the following subsections.

3.1 Encoding Process

The proposed modulo based image steganography technique requires cover image, secret message and stego key as inputs for the encoding process. In the encoding process, stego key is used as seed for pseudo random index generator. Let G_r , G_g and G_b be the sets of pixels which are selected by pseudo random numbers. Similarly secret message bits are divided into S_r, S_g and S_b.

The cover image is split into three parts, namely, R-plane, G-plane, and B-plane. The top left pixels of each plane are numbered as 0, 1 and 2 and are called as flag pixels. These pixels are set as flags to identify the planes. If the flag pixel of the plane is set to 0, then encoding of secret bits S_r in R plane using Mod u function is carried out. On the other hand, if the LSB of the flag pixel is 1, then it encodes secret bit S_g in G plane by using Mod v function. If the flag pixel is 2, then the encoding of secret bit S_b in B plane is done by using Mod w function. Here for R plane, the embedding of secret bits is carried out by using mod u function where $u = 2^m$ and *m* indicates the number of host bits that are used to embed the secret data. This is same for mod v and mod w functions of G plane and B plane respectively. The encoding process of embedding encrypted secret S_r into R plane is done by using equation (1)

$$x`_i = x_i + e_i – (LSB_m(x_{i-1})+LSB_m(x_i))Mod\ u \qquad (1)$$

where x` is the new value of pixel after encoding of the secret encrypted data, The value of *x`* is closest value to *x* among all possible values satisfying above equation, Here *x* is the host value of pixel in R plane, *e* is the secret bit value and *i* is the position of the pixel in the cover image. In the similar way encoding process for mod v and mod w are carried out by varying *m* value from 1 to 3 bits per pixels. However, different value of *m* should be assigned for mod u, mod v and mod w functions respectively. Fig.1 shows the encoding process of the Proposed Modulo Based Image Steganography Technique.

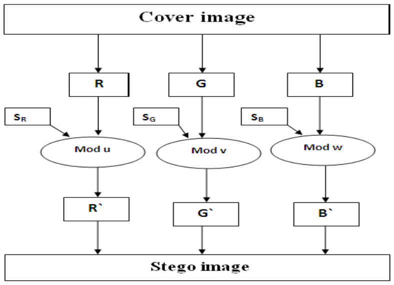


Fig.1 Encoding process of the proposed modulo based image steganography technique

Encoding of the secret bits in the RGB planes will result new R`G`B` planes will result combining these three new planes in the resultant of the stego image.

3.2 Decoding Process

In decoding process, initially Stego image is separated into RGB planes. The same stego key is used to generate the pseudo random number at the receiver side. Fig.2 shows the decoding process of the proposed modulo based image steganography technique.

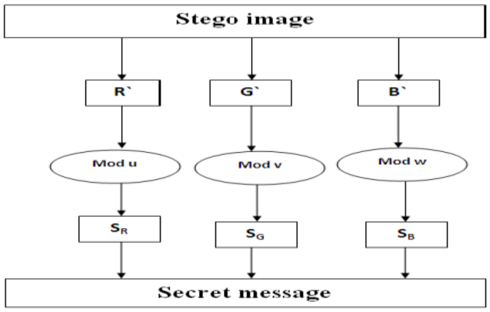


Fig.2 Decoding Process of the Proposed Modulo Based Image Steganography Technique

The pseudo random numbers select the set of pixels G_r, G_g and G_b in RGB planes. The selected pixels can be decoded with following equation:

$$e_i = (LSB_m(x_{i-1})+LSB_m(x_i))Mod\ u \qquad (2)$$

The extracted values *e* which is the secret data as encrypted is stored S_r, S_g and $S_{b.}$ It is obvious that the original data can be obtained by combining the three parts of secret pixels S_r, S_g and S_b from the stego image.

4. EXPERIMENTAL RESULTS

The proposed modulo based image steganography technique is simulated with Lena1997 image as cover image with size of 9067 bytes. Fig.3 shows cover (Lena1997) image.

Fig.3 Cover (Lena1997) image

For encoding process, cover image is divided into three planes named as the R, G and B planes. The encoding of secret bits in each of RGB plane is carried out by using mod u, mod v and mod w functions respectively.

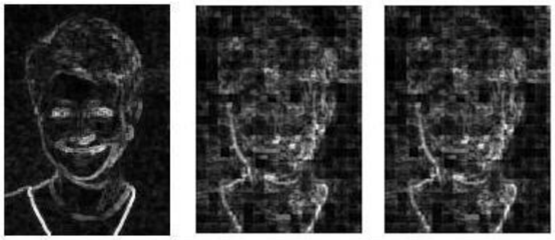
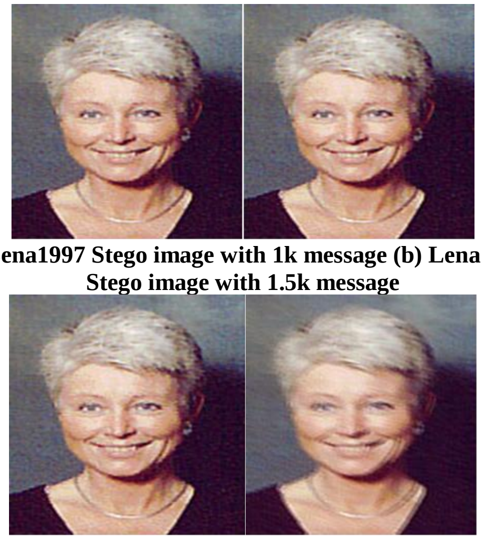


Fig.4 Image Separated By RGB Planes

Fig.4 shows the separated cover image by RGB planes with the secret bits encoded. The proposed modulo based image steganography technique was simulated with Lena1997 image with different message sizes. The combining these three RGB planes gives the resultant of stego image. The proposed technique was analyzed by encoding with varying message size from 1 k bytes to 5 k bytes on Lena1997 cover image.

The stego image of secret message size of 1k, 1.5k, 3k and 5k are shown in the Fig.5. From Fig.5 it is inferred that the proposed modulo based image steganography technique can encode secret message less up to 5k bytes having a good perceptual image quality. The resultant stego images are also analyzed by both histogram and statistical analysis.



(a)Lena1997 Stego image with 1k message (b) Lena1997 Stego image with 1.5k message

(c)Lena1997 Stego image with 3k message (d) Lena1997 Stego image with 5k message

Fig.5 Image Qualities for Lena1997

5. HISTOGRAM ANALYSIS

The histogram analysis plays a vital role in image steganography. The histogram analysis is performed on both cover image and stego image. The stego image shows minimum changes in the histogram compared to the cover image histogram. From these minimum changes in the histogram of the stego image, it is difficult to infer that secret data is hidden. Fig. 6 shows the histogram of the Lena1997cover image.

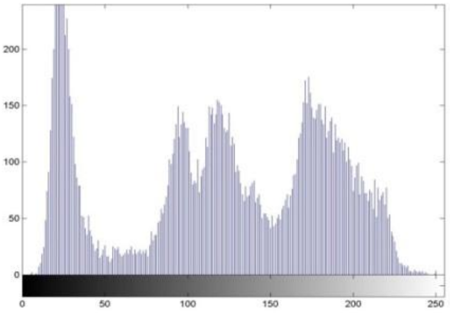
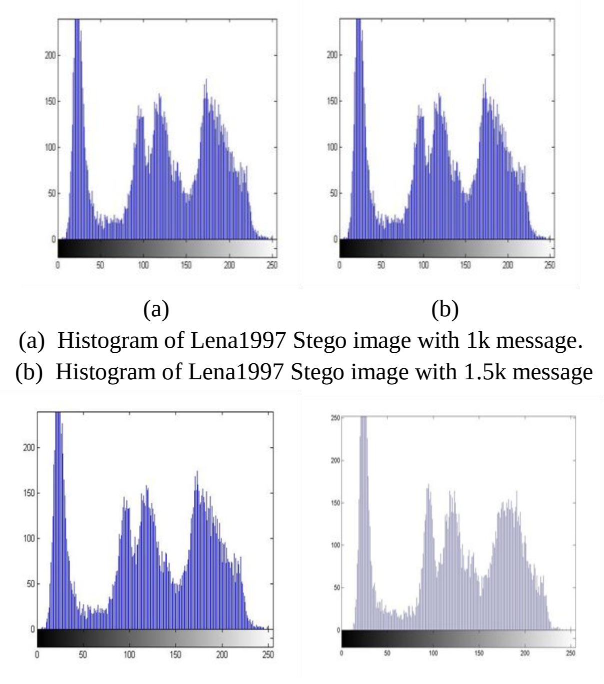


Fig.6 Histogram of Lena 1997 cover image

The histograms of stego image with different data size of 1k, 1.5k, 3k and 5k bytes are shown in the Fig.7.



(a) (b)

(a) Histogram of Lena1997 Stego image with 1k message.

(b) Histogram of Lena1997 Stego image with 1.5k message

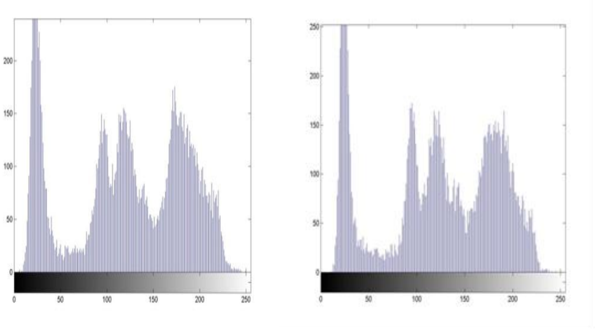
(c) (d)

(c) Histogram of Lena1997 Stego image with 3k message

(d) Histogram of Lena1997 Stego image with 5k message

Fig.7 Histogram of Lena 1997 Image

From Fig.7 it is inferred that the proposed modulo based image steganography technique can encode secret message up to 3k bytes with very minimal level of changes in the histogram. Histogram comparison of cover image and stego image with message size of 5 k bytes are shown in Fig.8



(a) (b)

Fig.8 Comparison of histogram with (a) Cover image and (b) Histogram of Lena1997Stego image with 5k byte message

Fig.8 shows the changes in histogram are noticeable. So the embedding capacity of secret message is limited to less than 5k bytes for Lena1997 stego image. From the above histogram analysis it is infered that the proposed modulo based image steganography technique can efficitvely resist Lena1997 image upto 5k byte of secret message.

6. STATISTICAL ANALYSIS

Statistical analysis is also performed on proposed stego image. Statistical parameters like mean,variance and Root Mean Square (RMS) values are calculated from the image before and after encoding of different sizes of secret messages. The values were tabulated. The simulated results shows the good insertion capacity for the cover images.

Table 1 Statistical analysis results for Lena1997 cover image

Lena1997	**Mean**	**Variance**	**RMS**	**Message**
Cover image	121.3906747	62.7394649	136.6452940	0
Proposed Technique	121.3908498	62.7603458	136.6550381	1k
Proposed Technique	121.4054347	62.7783474	136.6762616	1.5k
Proposed Technique	121.4548096	62.7776588	136.7198056	3k
Proposed Technique	121.1322442	61.6483492	135.9173998	4k (worst case)

Statistical analysis is the one of the most accurate method for calculating pixel values in image. The values of stego image and the original cover image are compared. The change in the values reveals that something is hidden in the image. From Table.1, it is inferred that mean, variance and RMS of proposed techniques values are nearly equal to the cover image values up to 3k bytes of secret message. The variation is clearly visible when the size of secret bits in stego image increases more than 3k bytes. Embedding capacity for Lena1997 cover image is calculated with different secret message bytes sizes and the results are tabulated. Table.2 shows the embedding capacity of Lena1997 image.

Table.2 Embedding capacity for Lena 1997 image

Lena1997	**Cover image size in byte**	**Message size in byte**	**Embedding capacity in %**
Proposed Technique	9067	1000	11.02
Proposed Technique	9067	1500	16.54
Proposed Technique	9067	3000	33.08

The RGB color percentage present in the cover image before encoding of secret message is noted. In the same way after encoding secret message is also noted, then differences between the RGB values of cover image and stego image are tabulated. Table.3 shows the results of color differences. It is inferred that the color differences in the cover image and stego image was minimum.

Table.3 Color differences of Lena1997 image

Lena1997	**Red**	**Green**	**Blue**
Cover image	31.9098	9.3055	10.6332
Stego image	31.9059	9.3535	10.6622

The proposed modulo based image steganography technique was tested with various cover images. Table.4 shows the results of various cover images.

Table.4 Results of various cover images

Cover Image used	**Cover size in bytes**	**Message size in bytes**	**Embedding capacity in Proposed Technique**
Lena1997	9,067	3000	33.08
Pondicherry	21,688	4713	21.73
Beach	8,134	2178	26.77
Nature41	494,255	16803	3.39
Lena	39,718	5135	12.92

From the Table.4, it is inferred that the proposed modulo based image steganography technique provides exactly 33 % of good insertion capacity for any of the cover images.

7. CONCLUSION

In this paper, modulo based image steganography technique is implemented. The proposed technique is simulated with secret bits using Lena1997 image as the cover image. The resultant stego image obtained after encoding secret message does not show any change when compared to original cover image. Histogram and Statistical analysis performed on the stego image proved that the proposed technique can effectively resist steganalysis. Comparison of the statistical value like mean, variance and RMS for proposed technique with cover image and stego image was done. The results showed that changes in value were obtained only in the decimal places. The proposed modulo based image steganography technique provides exactly 33 % of good insertion capacity for any of the cover images. Thus, the proposed modulo based image steganography technique provides greater security and good embedding capacity of the cover image.

8. REFERENCES

[1] F. Petioles, J. Anderson, and G. Kuhn, "Information hiding- A survey", Proceeding of IEEE, special issues on multimedia content, vol. 87, no. 7, pp. 1062-1078, July 1999.

[2] W. Bender, D. Gruel, N. Morimoto, and A. Lu, "Techniques for data hiding", IBM Systems Journal, vol. 35, no.3, pp.313-336, 1996.

[3] H. Farid, "Detecting hidden messages using higher-order statistical models", Proceeding of IEEE, International conference on image processing, vol. 15, no.6, pp.68-72, Dec. 2002.

[4] J. Fridrich, M. Goljan, and R. Du, "Detecting LSB Steganography in Color and Gray-Scale Images", Magazine of IEEE Multimedia Special Issue on Security, vol no.8, issue 4.pp. 22-28, Nov. 2001.

[5] J. Fridrich and M. Goljan, "Practical Steganalysis of Digital Images – State of the Art", Proceeding of SPIE, Photonics West, Vol. 4675, Electronic Imaging 2002, Security and Watermarking of Multimedia Contents, San Jose, California, pp. 1-13, Jan. 2002.

[6] W. Zhang, X. Zhang, and S. Wang, "A Double Layered "Plus-Minus One" Data Embedding Scheme", Proceeding of IEEE signal processing letters, vol. 14, no. 11, Nov. 2007.

[7] A. Westfeld and A. Pfitzmann, "Attacks on Steganographic Systems", Proceedings of the third international workshop on Information Hiding, Springer Verlag, Sept. 2005.

[8] X. Zhang and S. Wang, "Efficient steganographic embedding by exploiting modification direction", IEEE Communication Lett., vol. 10, no. 11,pp. 781–783, Nov. 2006.

[9] H. Zhang and H. Tang, "A Novel image steganography algorithm against statistical analysis", Machine Learning and Cybernetics, 2007 International Conference, vol. 19, no.22, pp.3884-3888, Aug.2007.

[10] J. Kang, Y. You, and M. Young Sung, "Steganography using Block-based Adaptive Threshold", Proceeding of the Computer and information sciences, 2007, iscis 2007, 22nd international symposium, vol. 11, pp.234-241, Nov. 2007.

[11] N. Provos, "Defending against Statistical Steganalysis", In 10th USENIX Security Symposium, Washington, Aug.2005.

[12] C.C. Thien and J.C. Lin, "A simple and high-hiding capacity method for hiding digit-by-digit data in images based on modulus function", Pattern Recognition, vol. 36, no. 12, pp.2875–2881, June 2003.

[13] C.M. Wang, N.I. Wu, C.S. Tsai, and M.S. Hwang, "A high quality steganographic method with pixel-value differencing and modulus function," Journal of Systems and Software , vol.81, no. 1, pp. 150–158, Jan. 2007.

[14] J.C. Joo, H.Y. Lee, C. N. Bui, W.Y. Yoo, and H.K. Lee, "Steganalytic measures for the steganography using pixel-value differencing and modulus function," in Proceeding s of the 9th Pacific Rim Conference on Multimedia, vol. 5353 of Lecture Notes in Computer Science , pp. 476–485, 2008.

[15] C.F. Lee and H.L Chen, "A novel data hiding scheme based on modulus function," Journal of Systems and Software, vol. 83, no. 1, pp. 832–843, Dec. 2009.

[16] J.C. Joo, H.Y. Lee and H.Y. Lee, "Improved Steganographic Method Preserving Pixel-Value Differencing Histogram with Modulus Function", EURASIP Journal on Advances in Signal Processing, Volume 2010, March 2010.

[17] S. Lyu and H. Farid, "Steganalysis using higher-order image statistics," IEEE Transactions on Information Forensics and Security, vol. 1, no. 1, pp. 111–119, March 2006.

[18] C.K. Chan and L.M. Cheng, "Hiding data in images by simple LSB substitution ", Pattern Recognition Society, Published by Elsevier Ltd, vol. 37, pp 469 – 474, Aug. 2003.

[19] W. Luo, F. Huang and J. Huang, "Edge Adaptive Image Steganography Based on LSB Matching Revisited", IEEE Transactions on Information Forensics and Security, vol. 5, no. 2, June 2010.

[20] H. Yang, Xingming SUN and Guang SUN, "A High-Capacity Image Data Hiding Scheme Using Adaptive LSB Substitution", Radio Engineering, vol. 18, no. 4, Dec. 2009.

[21] Masoud Afrakhteh, Subariah Ibrahim," Steganography Using More Surrounding Pixels", Proceedings of the World Congress on Engineering 2010, Vol I , June 30 - July 2, 2010, London, U.K.