International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:2, No:5, 2008

Increased Capacity of Information Hiding in LSB's Method for Text and Image

H.B.Kekre, Archana Athawale, and Pallavi N.Halarnkar

***Abstract*—**Steganography, derived from Greek, literally means "covered writing". It includes a vast array of secret communications methods that conceal the message's very existence. These methods include invisible inks, microdots, character arrangement, digital signatures, covert channels, and spread spectrum communications. This paper proposes a new improved version of Least Significant Bit (LSB) method. The approach proposed is simple for implementation when compared to Pixel value Differencing (PVD) method and yet achieves a High embedding capacity and imperceptibility. The proposed method can also be applied to 24 bit color images and achieve embedding capacity much higher than PVD.

***Keywords*—**Information Hiding, LSB Matching, PVD Steganography.

I. Introduction

INFORMATION hiding techniques have been receiving much attention today. The main motivation for this is largely due to fear of encryption services getting outlawed [14], and copyright owners who want to track confidential and intellectual property copyright against unauthorized access and use in digital materials such as music, film, book and software through the use of digital watermarks. Encryption and Decryption algorithms are widely used to encrypt secret (confidential) data so that it is not directly accessible to the otherwise illegitimate person and whenever the owner or genuine person requires the data, it can be decrypted with the help of a key or with the help of a retrieving algorithm/function. Steganography has a different approach to deal with this problem. Steganography [15] is an application of information hiding. 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 [3] 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". Generally, a good steganographic technique should have good visual/statistical imperceptibility and a sufficient payload [16].

The terminology LSB replacement/ LSB matching was discussed by T.Sharp [17]. LSB substitution algorithm is the simplest scheme to hide message in a host image. It replaces the least significant bit (LSB) of each pixel with the encrypted message bit stream. Authenticated receivers can extract the message by deciphering the LSB of every pixel of the host image with a pre-shared key. Since only the least significant bit of pixels are altered, it is visually imperceptible by human. The capacity of the algorithm is 1 bit per pixel. Although this algorithm is visually imperceptible, it can be statistically analyzed by other entity without processing the pre-shared key. Fridrich and Goljan [4] have surveyed some methods to detect LSB substitution. One of those is known as histogram analysis [5][6]. Research found out that, if an image is processed with LSB substitution, the histogram of the image will be showed in a "pair-wise" manner. These pair-wise blocks are known as Pairs of Values (PoV) [6] which can be identified by $\chi 2$ -test [7]. A similar idea called LSB matching [8] has been proposed to improve LSB substitution. Yet it is also vulnerable to other designated detection algorithms [8]. This is owing to the histogram of the host image is changed. A revised version of LSB matching is proposed by Mielikainen [9] in 2006. This method greatly improves the above two methods by lowering the expected number of modifications per pixel, from 0.5 to 0.375. Therefore, the histogram affected by the scheme is less significant. Only a few detection methods for LSB matching have been proposed.

A novel approach of image embedding was introduced in [22]. The method consists of three main steps. First, the edge of the image is detected using Sobel mask filters. Second, the least significant bit LSB of each pixel is used. Finally, a gray level connectivity is applied using a fuzzy approach and the ASCII code is used for information hiding

The best-known detector for LSB matching is based on the center of mass (COM) of the histogram characteristic function (HCF) [18], [19]. [23] proposes steganalysis methods for extensions of least-significant bit (LSB) overwriting to both of the two lowest bit planes in digital images. In [24], various statistical measures are analyzed and PMF based method of detection is proposed. It uses the frequency count of the pixel

Dr. H.B Kekre is a Senior Professor with MPSTME, NMIMS University, Vile Parle(w) Mumabi 56.(hbkekre@yahoo.com).
Ms. Arachana Athawale is a Ph.D. Research Scholar with MPSTME, NMIMS University, Vile Parle (W) Mumbai .(athawalearchana@gmail.com).
Ms.Pallavi N.Halarnkar is a lecturer with MPSTME, NMIMS University, Vile Parle (W) Mumbai 56(pallavi_halarnkar@rediffmail.com).

intensities in the image to test for the detection of stego image or not. Here LSB embedding technique is used. Experimental result shows that the PMF based methods are successful for detection of hidden data along with successful estimation hidden message. Later Sun *et al.* [10] proposed a data hiding scheme based on arranging the position of pixels. This method perfectly retains the host image histogram profile. Thus histogram analysis cannot distinguish whether an image is processed with this method.

Wu and Tsai [2] utilized the difference between the two consecutive pixels in the cover image to determine what size the secret message is to be hidden. In a paper by Zhang [11] Pixel Value Differencing (PVD) was successfully attacked. This was done by analysis of the histogram of the stego image. In [12] a method based on PVD is proposed which tries to increase the embedding capacity of PVD. They use LSB embedding for smooth regions and PVD embedding for edged areas. Another method [20] based on PVD was proposed instead of pixel values their remainder was computed using the modulus operation, and then secret data is embedded into the two pixels by modifying their remainder. An Adaptive data hiding method was proposed in [21].In this method Pixels located in the edge areas are embedded by a K-bit LSB substitution method with a larger value of than those of the pixels located in smooth areas. When compared to the Wu *et al.*'s PVD and LSB replacement method, their method provides both larger embedding capacity and higher image quality. Vajiheh Sabeti, et.al. [13] have presented a steganalysis method for PVD.

II. PIXEL VALUE DIFFERENCING

A. Embedding Stage

The cover images used in the PVD method are supposed to be 256 gray-valued ones. In the embedding phase a difference value d is computed from every non-overlapping block of two consecutive pixels, say p_i and p_{i+1} of a given cover image. The way of partitioning the cover image into two-pixel blocks runs through all the rows of each image in a zigzag manner. Assume that the gray values of p_i and p_{i+1} are g_i and g_{i+1}, then d is computed as $g_{i+1} - g_i$ which may be in the range from -255 to 255. A block with d close to 0 is considered to be an extremely smooth block, whereas a block with d close to -255 or 255 is considered as a sharply edged block. The method only considers the absolute values of d (0 through 255) and classifies them into a number of contiguous ranges, such as R_k where $k=1,2,...,q$. These ranges are assigned indices 1 though n. The lower and upper bound values of R_k are denoted by l_k and u_k, respectively. The width of R_k is u_k-l_k+1.In PVD method, the width of each range is taken to be a power of 2

Every bit in the bit stream should be embedded into the two-pixel blocks of the cover image. Given a two-pixel block B with gray value difference d belonging to k^{th} range, then the number of bits, say n, which can be embedded in this block, is calculated by n=log (u_k-l_k+1) which is an integer. A sub-stream S with n bits is selected from the secret message for embedding in B. A new difference d' then is computed with equation 1.

$$d' = \begin{cases} l_k + b & d \geq 0 \\ -(l_k + b) & d < 0 \end{cases} \quad (1)$$

where b is the value of the sub-stream S . Because the value b is in the range [0, u_k-l_k], the value of d' is in the range from l_k to u_k . If we replace d with d', the resulting changes are presumably unnoticeable to the observer. Then b can be embedded by performing an inverse calculation from d' to yield the new gray values (g_i', g_{i+1}') for the pixels in the corresponding two-pixel block (p_i, p_{i+1}) of the stego-image. The inverse calculation for computing (g_i', g_{i+1}') from the original gray values (g_i, g_{i+1}) of the pixel pair is based on a function given in equation 2.

$$(g_i', g_{i+1}') = \begin{cases} (g_i - \lceil m/2 \rceil, g_{i+1} + \lfloor m/2 \rfloor \; if \; d \; is \; even \\ (g_i - \lfloor m/2 \rfloor, g_{i+1} + \lceil m/2 \rceil \; if \; d \; is \; odd \end{cases} \quad (2)$$

where m is d'-d .the embedding is only done for pixels which their new values would fall in the range of [0,255].

B. Retrieving Stage

In the extracting phase, the original range table is necessary. It is used to partition the stego-image by the same method used for the cover image. Calculate the difference value d*(p_i, p_{i+1}) for each block of two consecutive pixels Then, find the optimum R_i of the d* same as in the hiding phase. Subtract l_i from d*(p_i, p_{i+1}) and b_0 is obtained. The b_0 value represents the secret data in decimal number. Transform b_0 into binary with t bits, where t = [$\log_2 w_i$]. The t bits can stand for the original secret data of hiding.

III. NEW APPROACH

In this section we describe our proposed method [1], which is then compared with the PVD method. The cover image used is a gray scale image. Before Embedding the data we use 8 bit secret key and XOR with all the bytes of the message to be embedded. Message is recovered by XOR operation by the same key. Every pixel value in this image is analyzed and the following checking process is employed

1. If the value of the pixel say g_i, is in the range $240 \leq g_i \leq 255$ then we embed 4 bits of secret data into the 4 LSB's of the pixel. This can be done by observing the first 4 Most Significant Bits (MSB's). If they are all 1's then the remaining 4 LSB's can be used for embedding data.
2. If the value of g_i (First 3 MSB's are all 1's), is in the range $224 \leq g_i \leq 239$ then we embed 3 bits of secret data into the 3 LSB's of the pixel.
3. If the value of g_i (First 2 MSB's are all 1's), is in the range $192 \leq g_i \leq 223$ then we embed 2 bits of secret data into the 2 LSB's of the pixel.
4. And in all other cases for the values in the range $0 \leq g_i \leq 192$ we embed 1 bit of secret data in to 1 LSB of the pixel.

Similarly, we can retrieve the secret data from the gray values of the stego image by again checking the first four MSB's of the pixel value and retrieve the embedded data.

IV. COMPARISON OF THE NEW APPROACH AND PVD

The proposed method is quite simple to implement as compared to PVD. The proposed method was implemented on Gray scale images as well as on 24 bit color images. The Embedding capacity obtained is much more than PVD as can be seen from the results Table No I. In PVD During the embedding phase there is a possibility that the gray values of the two- Pixel block may fall off the boundary value which needs a checking process to be employed in the embedding phase and incase of value fall off the boundary value the block is abandoned for inserting the secret data. The same checking process is again repeated in the retrieval phase. Due to this there can be certain blocks which are not utilized for embedding secret data. In the proposed method there is no such checking process required and it also utilizes every pixel of the image to embed the secret data.

V. EXPERIMENTAL RESULTS

In our Experiments , Four cover images "Lena", "Baboon", "Peppers" and "Mahalakshmi" were used, each with size 512X512. Three of the cover images used to embed text message which is Abraham Lincoln's letter to his son's teacher. They were compared with PVD and the results obtained are shown in Table No.II. In addition to this we have also introduced a new parameter in our experiment which is known as Average Fractional change in Pixel value abbreviated as (AFCPV). The results for the AFCPV for the proposed method are also included in Table No III. In Table No III results are shown where the message used is a image of an ATM card. The proposed method was also implemented on color image. Figure 1 shows the 24 bit color image used as cover image and the result is shown in Table no IV.Figure 2 shows the stego images in PVD and the proposed method.

TABLE I
EMBEDDING CAPACITY OBTAINED IN PVD AND THE PROPOSED METHOD

Cover Image	PVD Method I Embedding using the range widths of 8, 8, 16, 32, 64, and128 Capacity in Bytes	PVD Method II Embedding using the range widths of 2, 2, 4, 4, 4, 8, 8, 16, 16, 32, 32, 64, and64 Capacity in Bytes	Our Method Capacity in Bytes
Lena	50,960	25,940	35,827
Baboon	56,291	36,061	34,235
Peppers	50,685	27,269	60,317

TABLE II
VALUE OF RMSE AND PSNR'S OF STEGO IMAGES IN WHICH A FILE CONSISTING OF TEXT IS EMBEDDED

Cover Image	PVD Method Embedding using the range widths of 8, 8, 16, 32, 64, and 128		PVD Method Embedding using the range widths of 2, 2, 4, 4, 4, 8, 8, 16, 16, 32, 32, 64, and 64		Our Method	
	RMSE	PSNR	RMSE	PSNR	RMSE	PSNR
Lena	2.07	41.69	0.97	48.43	0.28	59.05
Baboon	3.25	37.90	1.59	44.10	0.27	59.36
Peppers	2.09	41.73	1.20	47.19	0.39	56.24

TABLE III
VALUES OF MSE, RMSE, PSNRS AND AFCPV OF STEGO- IMAGE IN WHICH AN ATM CARD IMAGE IS EMBEDDED.

Cover Image	Our Method			
	MSE	RMSE	PSNR	AFCPV
Lena	0.14	0.38	56.42	0.001029

Fig. 1 Mahalakshmi.bmp(512X512)

TABLE IV
VALUES OF MSE, RMSE AND PSNRS OF STEGO- IMAGES IN WHICH A FILE CONSISTING OF TEXT IS EMBEDDED AND COVER IMAGE USED IS 24 BIT COLOR.

Cover Image	Our Method				
	MSE	RMSE	PSNR	AFCPV	Embedding Capacity
Mahalakshmi	0.07	0.28	59.10	0.000231	1,12,930 bytes

Our Method
MSE: 0.08

PVD Method
MSE: 4.28

Fig. 2 Stego Images for the proposed method and PVD

VI. CONCLUSION

As seen from Table I the embedding capacity by our method is more than PVD method II but less than PVD method I. However for Baboon image the embedding capacity by our method is much higher than both PVD I and II. This shows that embedding capacity is very much image dependent.

In Table No II we have compared the RMSE and PSNR of the cover images for text embedding. It is seen that our method gives better performance in all the parameters as compared to PVD method I and II.

We have also used our method on gray scale image (Table III) and 24 bit color image of Mahalaksmi (Table IV) for embedding same text and ATM card image, the results given are very good. However they are not compared with PVD as PVD was implemented only on gray scale images [2].

We have given New Algorithm to increase the embedding capacity which is simple and performs better.

REFERENCES

[1] Dr. H. B. Kekre, Ms. Archana A. Athawale, "Information Hiding using LSB Technique with Increased Capacity", International Journal of Cryptography and Security, Special issue on Steganography, 2008.(Accepted for publication).
[2] D. C. Wu and W. H. Tsai, A steganographic method for images by pixel-value differencing," *Pattern Recognition Letters*, vol. 24, no. 9-10, pp. 1613-1626, 2003.
[3] Johnson,Neil F,"Steganography", 2000: www.jjtc.com/stegdoc/index2.htm
[4] J. Fridrich and M. Goljan, "Practical steganalysis of digital images—State of the art," in Proc. SPIE Security Watermarking Multimedia Contents, vol. 4675, E. J. Delp III and P. W. Wong, Eds., 2002, pp. 1–13.
[5] N. Provos, "Defending Against Statistical Steganalysis", 10th USENIX Security Symposium, Washington, DC, 2001.
[6] A. Westfeld and A. Pfitzmann, "Attacks on Steganographic Systems," Lecture Notes in Computer Science, vol.1768, Springer-Verlag, Berlin, 2000, pp. 61–75.
[7] W. Dixon, F. Massey: Introduction to Statistical Analysis. McGraw-Hill Book Company, Inc., New York 1957.
[8] A. Ker, "Steganalysis of LSB Matching in Grayscale Images." IEEE Signal Processing Letters, vol. 12(6), pp. 441–444, 2005.
[9] J. Mielikainen, "LSB Matching Revisited," IEEE Signal Processing Letters, vol. 13, NO. 5., MAY 2006
[10] Hung-Min Sun, Yao-Hsin Chen, and King-Hang Wang, "An Image Data Hiding Scheme being Perfectly Imperceptible to Histogram Attacks," Image and Vision Computing New Zealand IVCNZ 2006, November 27-29, 2006
[11] X. Zhang and S. Wang, "Vulnerability of pixel-value differencing steganography to histogram analysis and modification for enhanced security", Pattern Recognition Letters, 2004.
[12] H.C. Wu, N.I Wu, C.S. Tsai and M.S. Hwang, "Image Steganographic scheme based on pixel-value differencing and LSB replacement methods",VISP(152), No. 5, October 2005
[13] Vajiheh Sabeti, Shadrokh Samavi, Mojtaba Mahdavi, Shahram Shirani, "Steganalysis of Pixel-Value Differencing steganographic Method", 2007 IEEE. pp. 292-295
[14] Petitcolas F, Anderson R, Kuhn M: 'Information Hiding – A Survey' Proceedings of the IEEE, Vol. 87. July 1999.
[15] N. Johnson and S. Jajodia, "Exploring steganography: seeing the unseen," IEEE Computer, pp. 26-34, February 1998.
[16] X. Zhang and S. Wang, "Steganography using multiple-base notational system and human vision sensitivity," *IEEE Signal Process. Lett.*, vol. 12, no. 1, pp. 67–70, Jan. 2005.
[17] T. Sharp, "An implementation of key-based digital signal steganography," in *Proc. Information Hiding Workshop*, vol. 2137, Springer LNCS, 2001, pp. 13–26.
[18] J. Harmsen andW. Pearlman, "Steganalysis of additive-noise modelable information hiding," in *Proc. SPIE Security Watermarking Multimedia Contents*, vol. 5020, 2003, pp. 131–142
[19] A. Ker, "Steganalysis of LSB matching in greyscale images," *IEEE Signal Process Letter*, vol. 12, no.6, pp. 441– 444, Jun. 2005.
[20] Chung-Ming Wang , Nan-I Wu, Chwei-Shyong Tsai, Min-Shiang Hwang, "A high quality steganographic method with pixel-value differencing and modulus function" , Science Direct,The Journal of Systems and Software 81 (2008) pp. 150–158
[21] Cheng-Hsing Yang, Chi-Yao Weng, Shiuh-Jeng Wang, *Member, IEEE*, and Hung-Min Sun, "Adaptive Data Hiding in Edge Areas of Images with Spatial LSB Domain Systems", IEEE Transactions on Information Forensics and Security, VOL. 3, NO. 3, September 2008 pp. 488-497.
[22] Rehab H. Alwan, Fadhil J. Kadhim, and Ahmad T. Al-Taani, "Data Embedding Based on Better Use of Bits in Image Pixels", International Journal of Signal Processing 2;2 © www.waset.org Spring 2006.
[23] Andrew D. Ker, *Member, IEEE,* "Steganalysis of Embedding in Two Least-Significant Bits", IEEE Transactions on Information Forensics and Security, Vol. 2, NO. 1, March 2007
[24] Shila P. Hivrale, S. D. Sawarkar, Vijay Bhosale, and Seema Koregaonka, "Statistical Method for Hiding Detection in LSB of Digital Images: An Overview", Proceedings of World Academy of Science, Engineering and Technology Volume 32 August 2008 ISSN 2070-3740.

Dr. H. B. Kekre has received B.E. (Hons.) in Telecomm. Engineering. from Jabalpur University in 1958, M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S.Engg. (Electrical Engg.) from University of Ottawa in 1965 and Ph.D. (System Identification) from IIT Bombay in 1970. He has worked Over 35 years as Faculty of Electrical Engg. and then HOD Computer Science and Engg. at IIT Bombay. For last 13 years worked as a Professor in Department of Computer Engg. at TSEC, Mumbai. He is currently Senior Professor working with Mukesh Patel School of Tech. Mgmt. and Engg., NMIMS University, Mumbai, INDIA. His areas of interest are Digital Signal processing and Image Processing. He has more than 200 papers in National / International Conferences / Journals to his credit. Recently five students working under his guidance have received best paper awards.

Ms. Archana A. Athawale has Received M.E.(Computer Engineering) degree from V.J.T.I., Mumbai University in 1999, currently pursuing Ph.D. from NMIMS University, Mumbai. She has more than 9 years of experience in teaching. Currently working as - Assistant Professor in Department of Computer Engineering at Thadomal Shahani Engineering College, Mumbai. Her area of interest is Image Processing, Signal Processing and Computer Graphics. She has 8 papers in National /International Conferences/journal to her credit.

Ms. Pallavi.N.Halarnkar has received BE Computer degree from P.C.C.E, Goa University in 2003, currently persuing M.E. Computer from TSEC, Mumbai University. She has teaching experience of 4 years. Currently working as a lecturer in Department of Computer Engineering at MPSTME, NMIMS University, Vile Parle, Mumbai.