A novel Digital watermarking technique based on STD (standard division)

Aymen Mudheher Badr, Mohammed Layth Talal, Ghassan Sabeeh Mahmood

Abstract— the data hiding, it's one of the most important means used by the security institutions with critical communications in all countries of the world, they provided the technology of high security, especially in the communication networks and the Internet. In this paper, we modified LSB algorithm to become more secure and robust, Cryptography and Steganography works separately from the other to ensure a secure content. We design scheme to increase the safety and reliability of the copyrights to images, videos, books …etc., that's published on the internet by providing a watermark image hiding inside the original file that we want protect it. Firstly, reading first image (binary scale) and encrypted it by XOR algorithm and using a key agreed upon by the two parties (sender and recipient), followed by the division of the cover image to be hide data where clips blocks size (8*8) and account values standard deviation (Standard Deviation (STD)) for each section are then finding less and the largest value of a standard deviation in addition to the median value, then isolate sections where the value of the standard deviation less or equal of median value to be key concealment (ie be adopted as locations to hide) and that by including all bit of message into the (LSB) for each section of the selected sections. The effective of the proposed scheme has been estimated by Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR), BER. This paper also illustrates how security has been enhanced using this algorithm.

Index Terms— Cryptography, XOR, gray code, LSB, steganography, watermark, STD, PSNR, MSE, BER.

1 INTRODUCTION

Along with several developments information technolo gies leads to the problem of illegal copying and redistri bution of digital media. A digital watermark was one from a lot of solution. It is a pattern or digital signal embedded into the host media (image, text, video or audio) to be protected it. That contains useful certifiable information for the owner of the host media, such as his name, company logo, etc. the watermark can be detected or extracted later to make an assertion about the host media [1],[2]. Watermarks can either be invisible or visible.

There are two important properties of a watermark: firstly, the watermark embedding inside any host media should not alter the quality and visually of the host image and it should be perceptually invisible. Secondly, the robustness with respect to image distortions. This means that the watermark is difficult for an attacker to remove or modify it and it should be also robust to common image processing and geometric operations, such as filtering, resizing, cropping and image compression. [1].

The scientific data encryption and steganography that represented on hidden writing a watermark and ways to provide security and confidentiality of the data transmitted, where they work together to encode the message transmitted and hide the existence of a connection between the two parties [6].

Watermarking technique can be classified into: Spatial domain and Transform domain.

1.1 Spatial Domain

In this technique, the watermark embedding is achieved by directly modifying on the pixel values to the host image. The easy and most commonly used method in the spatial domain technique is the least significant bit (LSB). In the least significant bit (LSB) of each pixel in the host image was modified to embed the secret message, however, it is not very robust against attacks [3],[4].

1.2 Transform domain

This technique is also called transform domain. Values of certain frequencies are altered from their original. There are more methods like: discrete cosine transform (DCT), discrete Fourier transform (DFT), discreet wavelet transform (DWT), etc. [4]. Transform domain coefficients are altered by the watermark. Embedding the watermark in the low frequency increases the robustness with respect to image distortions. The high frequency band of an image is more prone to dropping due to quantization and it will be lost by compression or scaling attacks. The middle frequencies embedding of the watermark avoid the most visual important parts of the image and it is robust to compression and noise attacks [3],[5].

In our schema, a spatial domain was based on it, LSB technique, encrypted watermarking technique by using XOR technique with gray code and mixed them with standard division STD to become more secure and robust.

The watermark is a binary image, which is permuted using a

- *Aymen mudheher badr is a master's degree student in college of computer science in Chongqing University, china - and lecturer in engineering college in Diyala University, Iraq.*
- *Mohammed layth talal is a master's degree student in 3Yrmouk University College, Cankaya University, Turkey - and lecturer in Administration and Economic in Diyala University, Iraq.*
- *Ghassan Sabeeh Mahmood is a master's degree student in school of information science and engineering in central south university, china - and lecturer in computer science in Diyala University, Iraq.*

secret key.

A) The Least Significant Bit (LSB)

One of the earliest method to surface were those referred to as Least Significant Bit Substitution techniques, so called because of how the message data is embedded within a cover image. The term Least Significant Bit (LSB) refers to the smallest (right-most) bit of a binary sequence [6].

The structure of binary is such that each integer may only be either a 0 or a 1, often thought of as off and on respectively. Starting from the right, the value (if on) denotes a 1. The value to its left (if on) denotes a 2, and so on where the values double each time. Now let us consider the following 8-bit binary sequence:

1 0 1 1 0 0 1 **1**

The right-most value it is the LSB of this sequence. This value essentially determines whether the total sum is odd or even. If the LSB is a 1, then the total will be an odd number, and if 0, it will be an even number. However, changing the LSB value from a 0 to a 1 does not have a huge impact on the final figure; it will only ever change by +1 at most [6], [7].

B) Standard Division (STD)

This metric to Measure the dispersion of values and its deviation from the arithmetic mean, it is the square root of the variance, symbolized by a Greek letter (δ). On the conversely, equal variance standard deviation and usually refer to it by symbol (δ^2).

In the next example we explain how we can calculate STD to set of (n) values (n=100) and its mean is (μp) [8]:

$$\delta = [(1/100)\ (d12+d22+ \dots + d1002)]\ 1/2 \quad (1)$$

$$\delta = \{(1/100)[\ (x1 - \mu p)2+(x2 - \mu p)^2 + \dots + (x100 - \mu p)^2\ ]\}1/2 \quad (2)$$

C) The Gray code:

Gray Code is a form of binary that uses a different method of incrementing from one number to the next. Gray Code is the most popular Absolute encoder output type because its use prevents certain data errors which can occur with Natural Binary during state changes [9].

1.3 Related work

The researcher (Sunny Dagar, Vinay Kumar and Yogendra Bagoriya, 2013), an image steganography algorithm is proposed which uses secret key and gray codes to hide the secret file inside the cover image. This algorithm takes image of any format like .jpeg, .gif, .bmp etc. as a carrier and converts it into .bmp format. As .bmp image uses lossless compression techniques so compression of .bmp image doesn't lose any information. Although this paper will not emphasis on image compression [10].

And researcher (Saurabh and Gaurav, 2010) the proposal watermark technology that distinguish characters used to ensure the integrity of the security hidden data in the files [11].

Jiang Li and Liu Chen (2013), proposed a digital image watermarking scheme in interpolator orthogonal multi wavelets transforming domain. At first, the logistic chaotic mapping and Arnold transformation are employed to scramble the original watermarking image, then the watermark information is embedded into the middle frequency interpolator orthogonal multi wavelets transforming coefficients of the image [12].

2. PROPOSED METHOD

In our scheme, we supposed three algorithms to encryption, steganography and extraction.

2.1 Hiding information

Firstly, we encrypted the message (M) (Text file, digital image or any media) by using XOR technique with secret key (SK) (which the owner and receiver was agree about it before, but after input it to Gray code method) and generate the encrypted watermark ($\bar{M}$).

Secondly, embed the encrypted message ($\bar{M}$) in to the color image (C) data by using Least Significant Bit (LSB) and generate final digital ($\bar{C}$). That explain as following:

2.1.1 Encryption Algorithm:

Input: the message (M), here M is digital image (gray scale) with size 23*23 as example. The Cipher key (K), from the owner.

Output: encrypted watermark (M').

1- Reading the message (M), and convert it to binary.
2- Input the Cipher key (K), generate a random array by using (K) in the same size to message and convert it to binary to become (K').
3- By using Gray code method, covert (K') to the secret key (SK).
4- By using XOR method to encryption the message (M) by the secret key (SK) to generate the encrypted watermark (M'):

$$M' = M \oplus SK$$

Where $\oplus$ denotes XOR operation.

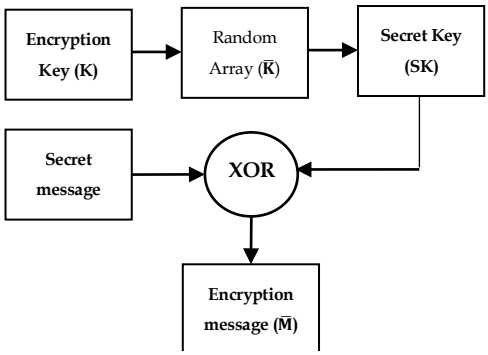


Figure 1: Encryption watermark

2.1.2 Hiding Algorithm

Input: encryption watermark (M') and color image (cover color

(C)).
Output: final color image (C')

1- Reading digital color image with JPEG, BMP, GIF or TIFF and let be (C).
2- Dividind color image to blocks with size (8*8).
3- Calculate the STD for each part from and order it increasing.
4- Find the maximum value to STD and let it be (MAX) and the minimum value to be (MIN) and the medium value to be (MED).
5- Determination the blocks have value minimum or equal to medium value to STD, to hiding the message data in it.
6- Embedding message data (M') in the least significant bit LSB to each block that have (STD>= MED), embedding two bits in the second and third bit from the odd byte to that blocks.
7- Re drawing color image and save it as final watermark image (C').
8- End.

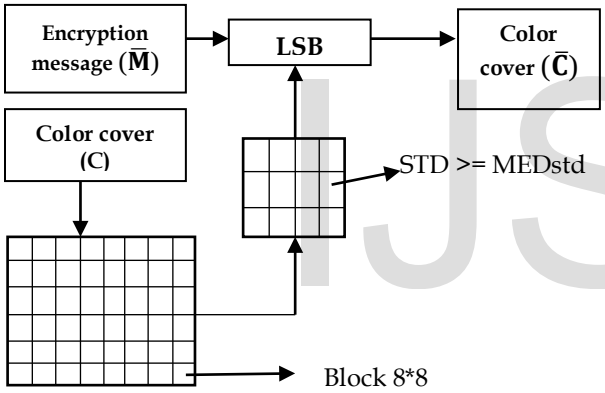


Figure 2: Embedding watermark

2.2 Extraction the watermark

After the owner sent the digital image via internet (as example), the receiver when he received it and he know the encryption key, he will extracting the watermark from the color cover to Verification if this color image is original or not, and check if it Was attacked or no.

2.2.1 Extracting Algorithm:

Input: the color image received it from internet let be (Cr).
Output: the original message (M).

1- Reading the color image and let be (Cr).
2- Dividing (Cr) to blocks with size (8*8).
3- Calculate the STD for each part from and order it increasing.
4- Find the maximum value to STD and let it be (MAXr) and the minimum value to be (MINr) and the medium value to be (MEDr).
5- Determination the blocks have value minimum or equal to medium value to STD, to hiding the message data in it.
6- Extracting the message data (M'r) from the least significant bit LSB to each block that have (STD>= MED), extracting two bits from the second and third bit from the odd byte to that blocks.
7- By using the Cipher key to generate random array (K'r) and by using Gray code method to generate the secret key (SK).
8- By input the secret key and secret message (M'r) to XOR method to de encryption to (M) and convert it to be the original message (Mr).
9- Convert it from binary to gray scale and show it.
10- End.

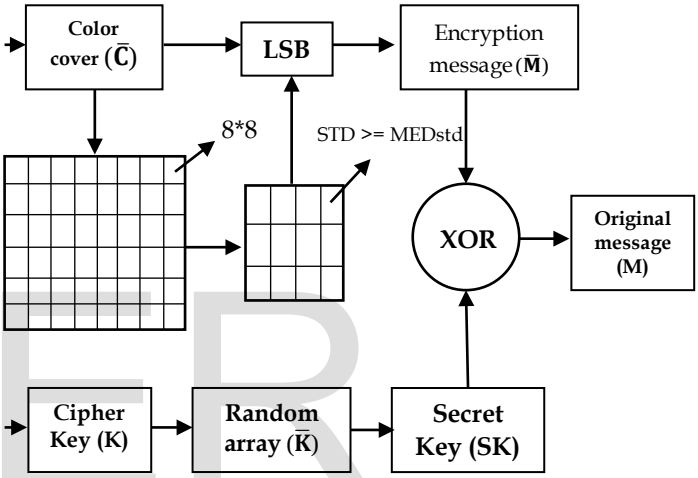


Figure 3: Extracting watermark

3. RESULTS AND DISCUSSION

3.1 Results:

The standard deviation (δ) is more accurate and better measures of dispersion (Numerical measures of the amount of the discrepancy between the data to measure), for ease of dealing with him in the analysis. It was calculated by the following equation [8]:

$$\delta = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(xi - \mu)^2} \quad (3)$$

$$\mu = \frac{1}{N}\sum_{i=1}^{N} xi \quad (4)$$

Where:

i: 1 → n , xi: image pixels, N: the number of pixel in image and μ: the mean.

And calculate the PSNR and MSE to extracted image [6][13][14]:

- Peak Signal to Noise Ratio (PSNR):

Image Name	PSNR	MSE	BER
Baghdad.jpg	59.198	0.021	0
Lena.jpg	52.158	0.036	0
Baghdad.bmp	60.785	0.017	0
Lena.bmp	52.154	0.026	0
Baghdad.gif	55.099	0.044	0
Lena.gif	49.447	0.029	0

$$PSNR = 10 \log_{10} [C_{max}^2 / MSE] \quad (5)$$

- Mean Squared Error (MSE):

$$MSE = 1/N*M*(S-C)^2 \quad (6)$$

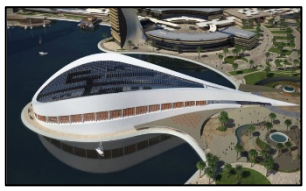
- Bit error rate (BER):

Bit error rate, BER is a key parameter that is used in assessing systems that transmit digital data from one location to another.

$$BER = (\text{no. of errors} / \text{total no. of bits sent})*100 \quad (7)$$

Where:
M, N: Image dimensions, S, C: the cover image and image after hiding Straight, Cmax: the maximum value to color pixel in image. SW: Represent the values of array that contains the watermark.

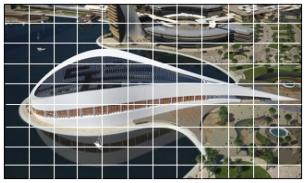
The scheme was applied on more than one image type (JPEG, BMP and GIF) was measured standard deviation (STD) to image clips (according to the table 1) was the conclusion: the less the value of the standard deviation leads to reduce the value of the dispersion of the image data, and this gives the best results.



And was the value of (BER) = 0, that mean the image extraction success and without any error.

The PSNR in (table 2) refer to all value was high and that mean no any change can be determined by the human eye in the image after hiding when we compared it with same image before hiding, show (figure 4a, 4b, 5).

Figure 4: Splitting image to blocks 8*8



(a) (b)

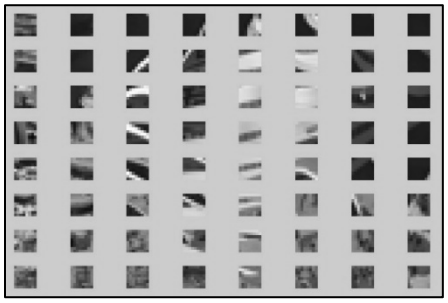


Figure 5: Dividing image to blocks 8*8 by MATLAB

Table 2: Result values

Block No.	STD Value	Block No.	STD Value	Block No.	STD Value	Block No.	STD Value	Block No.	STD Value
1	1.89	16	30.35	31	28.61	46	36.48	61	7.65
2	1.59	17	36.29	32	53.57	47	54.28	62	10.05
3	25.91	18	38.81	33	34.22	48	25.23	63	41.73
4	20.41	19	25.84	34	35.44	49	9.33	64	36.21
5	6.91	20	35.12	35	27.22	50	35.35		
6	0.35	21	26.45	36	42.21	51	33.51		
7	1.36	22	14.21	37	26.10	52	10.25		
8	1.99	23	26.39	38	31.67	53	44.48		
9	40.87	24	54.26	39	48.89	54	38.34		
10	38.62	25	14.25	40	38.55	55	21.98		
11	30.32	26	23.86	41	23.75	56	32.48		
12	27.54	27	31.60	42	26.21	57	29.06		
13	34.66	28	33.21	43	31.32	58	28.56		
14	34.40	29	28.42	44	36.21	59	24.32		
15	40.93	30	25.91	45	35.45	60	45.84		

Table 1: STD values to blocks 8*8 to image

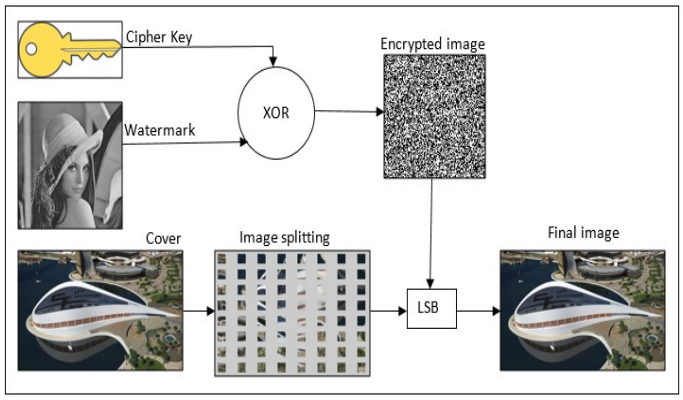


Figure 6: Hiding schema

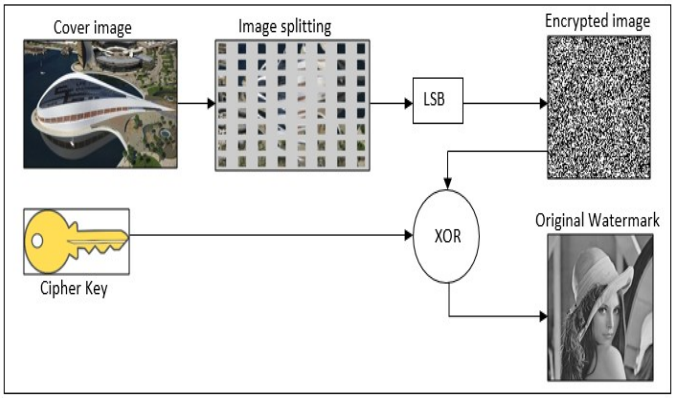


Figure 6: Extraction schema

3.2 Discussion:

The results proved the efficiency of the algorithm, as the hidden information did not occur any distortion on the cover files used. The division's cover has led to a range of sections to increase the strength of the improved algorithm. Hidden text encryption led to a secret increase improved algorithm. Hidden text encryption led to a secret algorithm to increase, especially when you do not use the secret key directly. Data retrieval has been fully and this is evident through measure values (BER). The hide of value in blocks have value less than or equal the median value to STD, better than hide in blocks have value less than the median value to STD, because that blocks have a few dispersion to data.
And finally, when the STD (standard division) decreasing that lead to increasing in PSNR values.

4. Future works

1- Algorithm development proposed addition of neural networks and genetic algorithm.
2- Use chaotic functions and merge with each other to increase the confidential way.
3- Use other Encryption algorithms to increase the security.

REFERENCES

[1] W. Bender, D.Gruhl,N.Mormoto and A.Lu,"Techniques for data hiding", IBM Systems Journal, vol. 35, no 3 pp 313-336, 1996.
[2] M. Kutter and F. A. P. Petitcolas. "A fair benchmark for image watermarking systems. In Proc. SPIE Security and Watermarking of Multimedia Contents", volume 3657, pages 226–239, San Jose, CA, USA, Jan. 1999.
[3] B. Verma, S. Jain, D. P. Agarwal, and A.Phadikar, "A New color image watermarking scheme," Info comp, Journal of computer science , vol. 5,No.2, pp. 37-42, 2006.
[4] Prabhishek Singh and R S Chadha," Review to Digital Watermarking and a Novel Approach to Position the Watermark in the", International Journal of Engineering and Advanced Technology (IJEAT), ISSN: 2249 – 8958, Vol.2, no.4, 2013.
[5] M. Wu, E. Tang, and B. Liu, "Data hiding in digital binary image," Electrical Engineering Dept., Princeton Univ., Princeton, NJ 08544,Electrical & Computer Engineering Dept., John Hopkins Univ., Baltimore, MD 21218 in Proc. Of IEEE Int. Conf. on Multimedia and Expo, New York City, pp. 393-396, 2000.
[6] Aymen Mudheher Badr, Xiao Di and Husham T. Ibrahim, "Double Hiding Information in Color Image File Based on Classical LSB Method", Research Journal of Applied Sciences, Engineering and Technology 8(3): 387-393, 2014.
[7] S. Dumitrescu, X. Wu, and Z. Wang. "Detection of LSB Steganography via Sample Pair Analysis", Lecture Notes in Computer Science, vol. 2578, pp. 355-372, 2003.
[8] Hogg, R.V., McKean, J.W., Craig, A.T., (2005), "Introduction to Mathematical Statistics", Pearson Education International, Sixth Edition.
[9] R. Varalakshmi, Dr. V. Rhymend Uthariaraj, "A New Secure Multicast Group Key Management Using Gray Code", IEEE-International Conference on Recent Trends in Information Technology, ICRTIT 2011 MIT, Anna University, Chennai. June 3-5, 2011.
[10] Sunny Dagar, Vinay Kumar, Yogendra Bagoriya, "Image Steganography using Secret Key & Gray Codes ",International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-2, Issue-5, April 2013 .
[11] Saurabh, S. and A. Gaurav, 2010. Use of image to secure text message with the help of LSB replacement. Int. J. Appl. Eng. Res., 1(1):201.
[12] Jiang Li and Liu Chen, "A Digital Watermarking Algorithm Based on Chaos in InterpolatoryOrthogonal Multiwavelets Domain", Computer and Information Science; Vol. 6, No. 2; 2013.
[13] Huajian, L., "Digital Watermarking for Image Content". Geboren, Shandong, China, pp: 47, 2008.
[14] M. A. Masud, M. Samsuzzaman, M. A.Rahman, "Bit Error Rate Performance Analysis on Modulation Techniques of Wideband Code Division Multiple Access", journal of telecommunications, volume 1, issue 2, March 2010.