CREAT. MATH. INFORM.
Volume **26** (2017), No. 1,
Pages 61 - 67

Online version at `https://creative-mathematics.cunbm.utcluj.ro/`
Print Edition: ISSN 1584 - 286X; Online Edition: ISSN 1843 - 441X
DOI: https://doi.org/10.37193/CMI.2017.01.08

Strong image steganography based on last significant bit substitution

OVIDIU COSMA, GHEORGHE ARDELEAN and ADRIAN PETROVAN

ABSTRACT. Steganography is a method of hiding secret information into an innocent looking container. Although a large number of steganography techniques have been proposed, there is a constant need for new variants because they have a relatively short lifespan, being defeated by steganalysis techniques. This article explores the possibilities of eliminating the main drawback of LSB steganography: weakness. The problem is solved by rearranging the image pixels and by applying a compensation technique that preserves their statistical properties. The proposed method contains an encryption block that adds an extra layer of security to the hidden data.

1. INTRODUCTION

Steganography is a data securing technique complementary to cryptography, whose purpose is to hide information into a container which will unveil its true content only to the informed observer. The image steganography techniques use image files as containers. They operate either in the spatial domain or in the frequency domain. The hidden data is known as payload and the image with hidden data is called stego image. The main applications of steganography are: copyright control, image search engines, feature tagging, secret communication, video-audio synchronization, etc.

The known steganography techniques can be classified by robustness strength and capacity. Robustness is the property of the payload to withstand some of the processing operations applied to the image container (geometric transformations, contrast or brightness adjusting, histogram correction, compression, cropping, etc.). Strength indicates how well hidden is the payload inside the container, and as a consequence, how difficult it is to detect the presence of hidden data. The art of detecting steganography is called steganalysis. Capacity is the amount of hidden data that the container supports, and is usually represented in bits per pixel (bpp).

In order to increase the strength of steganography, the data embedding process needs to preserve the statistical properties of the image container, in addition to the perceptual quality.

2. SPATIAL DOMAIN IMAGE STEGANOGRAPHY

A digital image is composed of pixels, each of which is represented by an integer, usually having 32 bits in length. The first 8 bits represent the pixel opacity and the following three groups of 8 bits are the intensities of the components that represent the pixel color in the Red Green Blue (RGB) color space.

One of the best known image steganography techniques is based on the use of one or more of the least significant bits (LSB) from one or more of the pixel RGB components for placing hidden data. The number of bits per pixel that are amended, are reflected on the

Received: 23.05.2016. In revised form: 20.09.2016. Accepted: 04.10.2016
2010 *Mathematics Subject Classification.* 68U10.
Key words and phrases. *Strong image steganography LSB substitution.*
Corresponding author: Ovidiu Cosma; ovidiu.cosma@yahoo.com

final aspect of the image. The distortions induced by the changes of the pixel values are more easily visible in the image areas that are lacking in details. In consequence, if large payloads are desired, the uniform use of all image pixels for storing secret data is not the best option.

In order to increase the payload, Bit Plane Complexity Segmentation (BPCS) [10] divides the image into 8 x 8 pixel blocks, that it classifies according to their complexity (detail content), for selecting the best places to embed secret data, and for detecting the regions in which alteration of pixels is not recommended. A data embedding method using BPCS[1] proposes an improvement over BPCS in the way that the complexity of the image blocks is evaluated.

Gray level modification steganography [11] proposes a technique to embed secret data by altering the gray level of pixels. A technique based on pixel value differencing is presented in [14]. The secret data is embedded by altering the differences between pixels. The amount of data stored in each pixel depends on the complexity of the region in which it is situated. In [15] is presented a technique of steganography that can be used in conjunction with a redundancy reduction scheme based on prediction. The main advantages presented by the authors are the high embedding capacity, and the perfect reconstruction of the original image.

A LSB substitution that uses the I-component of the hue-saturation-intensity (HSI) color model is presented in [6]. The I-component is divided into four sub-images that are rotated with different angles using a secret key. The secret information is encrypted and embedded into the rotated sub-images based on a specific pattern using LSB substitution. An efficient steganography method for RGB images based on gray level modification and multi-level encryption is presented in [7]. The secret key and secret data are encrypted before mapping them to the grey-levels of the image. Then a transposition function is applied on cover image prior to data hiding. A video steganographic method is presented in [5]. The secret message is encoded and embedded in the regions of interest using a face detection algorithm and adaptive LSB substitution. A secure image steganographic framework is proposed in [8], in which a stego key is encrypted using a two-level encryption algorithm, the secret data is encrypted using a multi-level encryption algorithm and the encrypted information is embedded in the image using an adaptive LSB substitution method. In [9] is proposed an image steganography method with dual-level of security. The cover image and its planes are rotated at different angles using a secret key prior to embedding. The secret information is divided into three blocks using a specific pattern, and encrypted by a three-level encryption algorithm. The image is scrambled using a secret key, and the encrypted message is embedded using cyclic LSB substitution.

The spatial steganography techniques have the advantage of simplicity and high capacity, and the disadvantage of low robustness and strength.

3. Transform domain steganography

The transform domain steganography techniques are characterized by the fact that the secret data is embedded after a step in which the image is converted from the spatial domain to the frequency domain. The embedding is performed by altering some of the transform coefficients.

There have been proposed various techniques for selecting the coefficients to be used for embedding the payload. Jsteg [12] is a steganographic tool designed for JPEG compressed images. The embedding of the payload is performed by replacing the LSB of the quantized DCT coefficients whose values are different from 0, 1 and -1. JPHide [3] is different from Jsteg by the fact that it randomly selects the coefficients that will store the payload, and it can make use of their second LSB. A steganographic system based on a genetic algorithm is presented in [4]. The genetic algorithm optimizes the localizations in which the payload is embedded in the DCT coefficients of the image container. The authors claim that the method defeats almost all known steganalysis methods.

4. Steganalysis techniques

Steganalysis is the art of detecting steganography. The main disadvantage of LSB steganography is the fact that it can be easily detected. The LSB steganography is based on the assumption that the LSB of the pixels in an image contain random data, and therefore they may be replaced with bits of secret data without creating suspicion.

The authors of [13] state that actually the pixels LSBs are correlated, and those correlations can be observed with the naked eye if the bits are displayed (representing 0 and 1 as black and white). The pixels LSBs of the picture in Figure 1 are shown in Figure 2. The correlations are clearly visible. Consequently the presence of hidden data can be visually detected because it disrupts those correlations. Figure 3 shows the LSB of the pixels in the same picture, after the first half of them were replaced with random data.

In reality such strong correlations (like those in Figure 2) do not appear in any picture. The phenomenon can be rarely observed especially in the case of older images converted into a new format, or images that contain large areas of smooth background, or have undergone a filtering operation. A usual raw photo contains very weak correlations at pixel LSB level, or nothing at all. Figure 5 shows the LSBs of the pixels in an ordinary image (the one in Figure 4). The source of Figure 4 is presented in [2]. However, the visual method can be effective even if there are no correlations in the LSBs of the pixels, if the hidden data consists of printable characters, because a pattern emerges, due to the properties of their ASCII codes. Such a texture can be seen in Figure 6, which shows the LSB of the pixels in a stego image filled to 50% of the capacity with ASCII text. The original image container is the one in Figure 4. Even though this effect disappears if the image pixels are processed in a pseudo random sequence, the proposed method contains a data encryption block that adds a level of security to the hidden data.

In order to overcome the limitations of the visual steganalysis method, a statistical method has been proposed in [13], known as the chi-square attack. According to the method, the image pixels are divided into groups based on the values of all the bits excepting the LSB. Thus in each group there will be 2 pixel values (one for the LSB = 0 and the other one for LSB = 1). The replacement of the LSBs with random data tends to bring nearer the two values within each group, thus disturbing the natural distribution of the two pixel values within the pixel groups. The chi-square attack calculates a probability for the presence of hidden data depending on the length of the analyzed sample. Figures 7, 8 and 9 show the results of the chi-square attack performed on the image in Figure 4, before and after being filled to 25% and 50% of capacity with random data.

5. The proposed method

The proposed method is a variant of LSB substitution, modified to increase strength and data security. Only the least significant bit of the three RGB pixel components are used to store secret data. Thus the induced distortions can't be observed with the naked eye, even in the case of the pixels located in areas lacking details. The operations of the proposed method are shown in Figure 11. In order to avoid the vulnerability to the visual attack, the image pixels are not processed in their natural order. The position of the secret data bits within the image is determined in a pseudorandom pixel reordering step. Figure

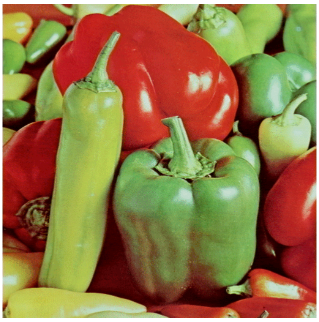

FIGURE 1

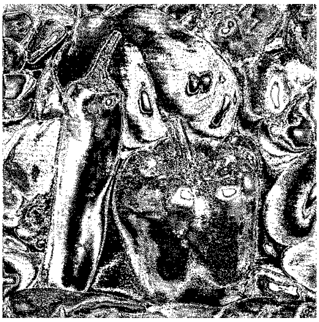

FIGURE 2

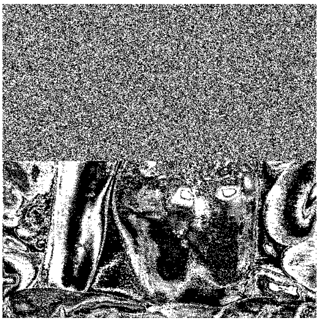

FIGURE 3

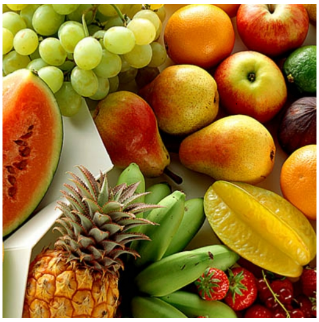

FIGURE 4

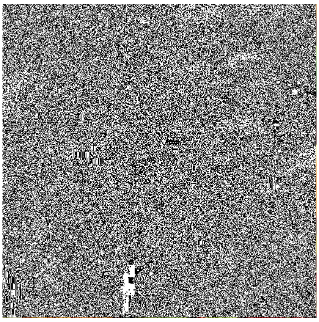

FIGURE 5

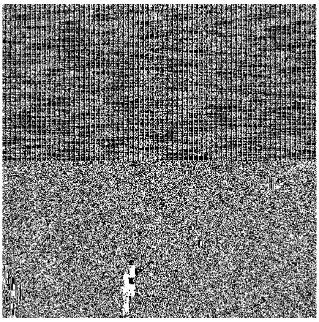

FIGURE 6

13 shows the position of the pixels that keep secret data, when 20% of the image container capacity is used. The original image is the one in Figure 4.

In order to properly combat the chi-square attack, the image pixels are divided into groups based on their value of which the LSB is ignored. So the pixels in each group will have two possible values: one for the LSB = 0 and the other one for the LSB = 1. The secret data embedding algorithm will use two types of pixels: main pixels and auxiliary pixels. The main pixels are used for storing the secret data bits in their LSB, and the auxiliary pixels are used in a compensation step. In the initialization stage all the image pixels are marked as main pixels. Then for each main pixel a pair is searched in the remaining of the image, using the following criteria: the pair must belong to the same pixel group as the main pixel, and it is not pair of another main pixel. If such a pixel is found, it will be marked as auxiliary. The first pixel found to satisfy these conditions will be chosen, because it is as close as possible to the main pixel in the pair, and it will be marked as auxiliary. If a pair can be found for each of the main pixels, then at the end of the operation 50% of the image pixels will be main pixels and the other 50% will be auxiliary pixels.

In order to avoid the vulnerability to the visual attack, the pairing step is followed by a pseudorandom rearrangement of the image pixels. The final result determines the position of each data bit in the image, and the positions of the pixels used in the compensation operations.

The compensation step has the purpose to preserve the distribution of the two pixel values in each group. For the implementation of this operation all the changes to the pixels in each group will be stored in a separate variable. When the LSB of a pixel changes from 0 to 1 the variable of the group is incremented, and if the change is in the opposite sense, the variable is decremented. In the compensation step, if the last modified pixel

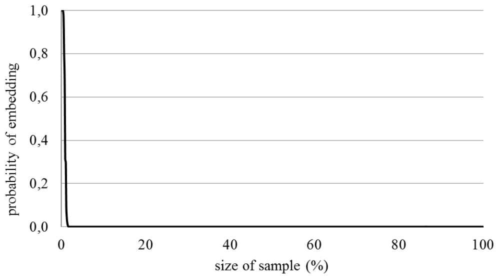

FIGURE 7

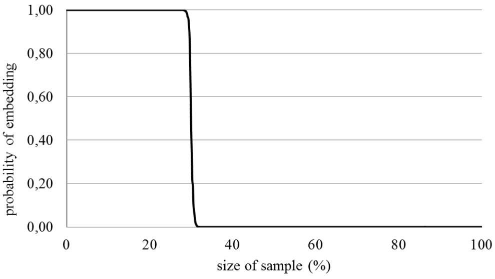

FIGURE 8

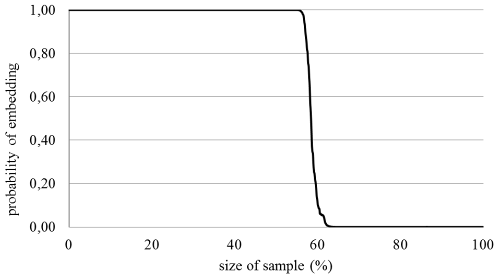

FIGURE 9

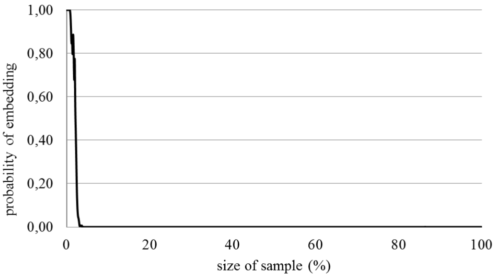

FIGURE 10

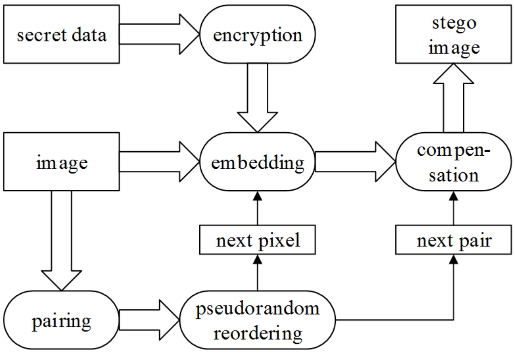

FIGURE 11

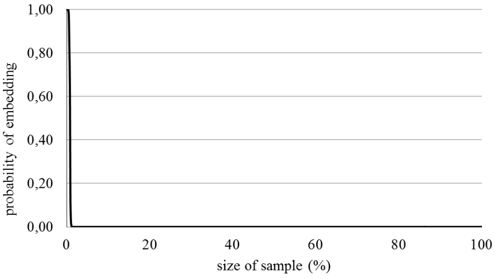

FIGURE 12

has a pair and the variable associated to the group is non-zero, then the LSB of the pair will be changed in such a way that the variable approaches 0.

6. Experimental results and discussion

Figure 14 shows the least significant bits of the image in Figure 1, after 50% of the capacity was used for random data embedding. Because the initial correlations shown in Figure 2 are still visible, we may conclude that the visual attack can be fought by a simple rearrangement of the image pixels. The random rearrangement of image pixels has proved benefic even in the case of using the chi-square attack. Figure 10 presents the detection results of the chi-square attack to the stego image, in the case of filling 50% of

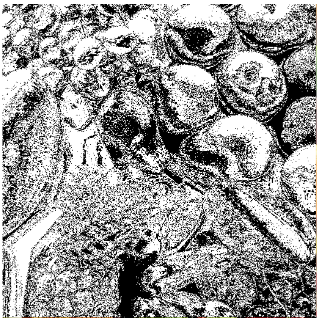

FIGURE 13

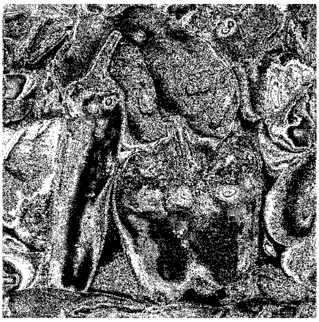

FIGURE 14

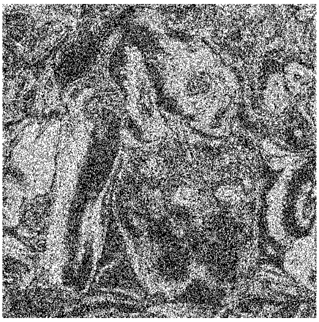

FIGURE 15

the image container capacity with random data, after a pseudorandom rearrangement of the pixels.

Figure 12 presents the results of the chi-square attack in the case that the image container in figure 4 is filled to full capacity with random data. The compensating step helps to preserve to some extent the original correlations at the LSB level. Thus the visual attack does not work, even if the full capacity of the container is used. Figure 15 shows the pixel LSB image of the picture in Figure 1, after being filled to full capacity with random data.

7. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

The proposed method contains two techniques for increasing the strength of LSB substitution steganography: pseudorandom reordering of the image pixels and chi-square compensation. By applying the two techniques, the hidden data becomes virtually invisible for both the visual and statistical steganalysis methods. The proposed method also contains an encryption block that adds an additional level of security to the hidden data. The main drawback of the proposed method lies in the fact that the compensation step reduces the capacity of the image container by approximately 50%. The capacity of an image container that is represented with 24 bits / pixel is at least:

image width * image height * 3 / 8 / 2 [bytes].

In comparison with the existing steganography techniques based on LSB substitution the proposed method has the advantage of resisting the chi square attack, even if the container is filled to full capacity with secret data. This advantage was obtained with the compensation step that aims to preserve the statistical properties of the image pixels.

A further research direction is the study of the possibility to reduce the number of pixels used in the compensation step and the impact on the method strength.

REFERENCES

[1] Hioki, H., *A data embedding method using BPCS principle with new complexity measures*, Proceedings of Pacific Rim Workshop on Digital Steganography, 2002

[2] http://1wallpaper.net/ro/food-fruits-hami-melon-watermelon-pineapple-banana-grapes-pear-compote-wallpaper.html#.V4egcriLSUl (accessed: 2016-03-05)

[3] Latham, A., *JPHIDE*, http://linux01.gwdg.de/ alatham/stego.html (accessed: 2016-03-05)

[4] Milani, A., Mohammad, A., and Varasteh, A., *A New Genetic Algorithm Approach for Secure JPEG Steganography*, IEEE International Conference on Engineering of Intelligent Systems, 2006

[5] Mstafa, R. J. and Elleithy, K. M., *A video steganography algorithm based on Kanade-Lucas-Tomasi tracking algorithm and error correcting codes*, Multimedia Tools and Applications, November 2015, 1–23

[6] Muhammad, K., Sajjad, M., Mehmood, I., Rho, S. and Baik, S. W., *A novel magic LSB substitution method (M-LSB-SM) using multi-level encryption and achromatic component of an image*, Multimedia Tools and Applications, May 2015, 1–27

[7] Muhammad, K., Ahmad, J., Farman, H., Jan, Z., Sajjad, M. and Baik, S. W., *A Secure Method for Color Image Steganography using Gray-Level Modification and Multi-level Encryption*, KSII Transactions on Internet and Information Systems, **9** (2015), No. 5, 1938–1962

[8] Muhammad, K., Ahmad, J., Rehman, N. U., Jan, Z. and Sajjad, M., *CISSKA-LSB: color image steganography using stego key-directed adaptive LSB substitution method*, Multimedia Tools and Applications, April 2016, 1–30

[9] Muhammad, K., Sajjad, M. and Baik, S. W., *Dual-Level Security based Cyclic18 Steganographic Method and its Application for Secure Transmission of Keyframes during Wireless Capsule Endoscopy*, Journal of Medical Systems, May 2016

[10] Niimi, M., Noda, H. and Kawaguchi, E., *A Steganography Based on Region Segmentation by Using Complexity Measure*, Trans. of IEICE, Vol. J81-D-II, No. 6, 1998

[11] Potdar, V. M . and Chang, E., *Gray level modification steganography for secret communication*, Proc. of 2nd IEEE International Conference on Industrial Informatics, 223–228

[12] Upham, D., *Jsteg*, http://zooid.org/ paul/crypto/jsteg/ (accessed: 2016-03-05)

[13] Westfeld, A. and Pfitzmann, A., *Attacks on Steganographic Systems Breaking the Steganographic Utilities EzStego, Jsteg, Steganos, and S-Tools—and Some Lessons Learned*, https://users.ece.cmu.edu/ adrian/487-s06/westfeld-pfitzmann-ihw99.pdf (accessed: 2016-04-10)

[14] Wu, D. and Tsai, W. H., *A steganographic method for images by pixel value differencing*, Pattern Recognit. Lett., **24** (2003), 1613–1626

[15] Wu, H. C., Wang, H. C., Tsai, C. S. and Wang, C. M., *Reversible image steganographic scheme via predictive coding*, Displays, No. 31, 2010, 31–43

Department of Mathematics and Informatics
North University Center Baia Mare
Victoriei 76, 430122 Baia Mare, Romania
Email address: ovidiu.cosma@yahoo.com
Email address: ardelean_g@yahoo.com

Department of Electrical Engineering
Electronics and Computer Science
North University Center Baia Mare
Dr. Victor Babes 62A, 430083 Baia Mare, Romania
Email address: adrian.petrovan@cunbm.utcluj.ro