

**REFERENCE-FREE STEGANALYSIS VIA NOISE
ESTIMATION IN THE HISTOGRAM-CORRELATIVE
DOMAIN**

**Hind R. Almayyali¹ , Amal Fadhil Mohammed² ,Intisar Hadi
Abbas³, Hadeel Qasem Gheni⁴**

¹Faculty of Computer Science and Mathematics, University of Kufa, Najaf,
Iraq

²Faculty of Biotechnology, Al-Qasim Green University, Babylon, Al Qasim,
Iraq.

³Computer Science Department, Babylon Education Directorate, Hilla - Iraq

⁴Computer Science Department / College of Science for Women /
University of Babylon

Emails: hindalmayyali@gmail.com, amal.f@uoqasim.edu.iq,
intisarza92@gmail.com wsci.hadeel.qasem@uobabylon.edu.iq.

Abstract

It is difficult to ensure a reliable database of original images for the purpose of classification to detect image tampering or steganography. Hence reference-free analysis is necessary in care no database is available. In this paper we present a reference-free technique for detecting image steganography in the spatial domain. No need for an image database. The approach is based on estimating the noise introduced in the histogram autocorrelation of the cover image by LSB tampering while hiding a secret message in the cover image. Simulation results have shown the effectiveness of the proposed technique.

Keywords— steganography, steganalysis, image tampering, reference-free.

I. INTRODUCTION

There has been an increasing demand for sharing information between individuals and institutions. Some data should be kept confidential, for instance, some medical and military should be kept secret. For secure exchange of information, we can utilize cryptography with a secret key so that the recipient can uncover the message and verify that the message is untampered with. However, many cryptographic methods can be attacked by aggressors to decrypt the message, unless a complicated cryptography is used. However, for a simpler secure system we should adopt the methods of steganography which include an apparently-innocent cover image with the secret message embedded inside in a way so that moves the least doubt to excite aggressors. However, the approach of steganography can equally be used by criminals, hence it gets necessary to distinguish the stego data and attempt to get the information embedded

in it. The way of detecting information inside an image is called steganalysis, which is similar to cryptanalysis which is used to decrypt a suspected message. The task of steganalysis is to find whether there is a message embedded inside an image, or to further discover any extra information like the message length, or the method of data hiding utilized in the cover image [1,2].

II. BACKGROUND TO STEGANALYSIS

Steganalysis is the science of identifying the existence of possible hidden messages inside a stego medium like an image. Steganalysis can follow one of two paths: either passive or active. In the passive approach, one may try to find whether the medium (image) contains a stego message. In the active approach, the aggressor can distinguish the stego method and may tamper with the hidden message. Generally, the embedding of a message should be intangible; however, it can change the statistics of the cover image. To discover statistical changes is a main task of steganalysis [2].

Steganography (especially LSB) makes benefit of the fact that the human visual system (HVS) is not sensitive to small changes in image pixel illumination level. Communication and information systems utilize compression methods for transmitting or saving data. Compression techniques remove redundant information by transforming the image to another domain where the signal is sparse. This reduction also takes advantage of the HVS as not sensitive to small changes. Hence, compressed images (like jpeg) cannot be used efficiently for data-hiding.

There are two main approaches to analyze a stego image using steganalysis as follows [2,3]:

1. Specific steganalysis: In this approach, the analyzer attempts to decide what kind of embedding and analyze the image statistics. However, this kind of steganalysis is limited to a particular data-hiding approach and probably needs a specific image format.
2. Blind or universal steganalysis: This approach is not limited to a specific data-hiding technique and no information is needed before steganalysis.

III. RELATED WORKS IN STEGANALYSIS

There are many approaches for detecting hidden message in a cover image. Some of these approaches handle the cover image in the spatial domain, while other techniques analyze the cover image after being transformed to another domain, like the discrete-cosine domain or the wavelet domain. As most special techniques hide data in the LSB, some of the steganalysis approaches specialize only in detecting LSB hidden data. Some of these techniques specialize in analyzing color images, others work only on gray-scale images, while some of them can work on both image kinds [4,5,6,7].

The method in [6] was proved to be effective in finding sequential and non-sequential LSB embedding in digital images.

In [7] a more general approach for LSB steganalysis is proposed. This approach can also detect the message length.

In [8], the Authors propose an approach to detect hidden data in natural images. The approach utilizes the image histogram, where the approach can also estimate the ratio of the message to the cover. However, the method may not work for low ratios of message to cover.

Reference [9] explained a steganalysis algorithm that works with randomized LSB message. In [10] the Authors presented a method that depends on the estimated change of cover image statistics.

In [11] the Authors proposed an approach that is based on statistical moments of the image wavelet transform.

In [12,13] the Authors proposed a reference-free approach that detects the existence of hidden data based on analyzing the derivatives of the auto-correlation function of the cover image histogram.

In this work we extend the reference-free approach by estimating the noise power in the autocorrelation function of the cover histogram rather than using the derivatives of this function.

IV. THE PROPOSED HISTOGRAM-CORRELATIVE APPROACH

It is expected that a clear, non-compressed natural image has a smooth autocorrelation of its histogram. Hence, any tampering (even as a small quantity) in the spatial domain of the cover image can be reflected as disturbance in the autocorrelation function of the image histogram.

Assuming the message as not correlated with the cover image, the disturbance caused by embedding this message in the LSB bits of the cover image as Gaussian or semi-Gaussian noise. This modeling proved effective in the detection of tampering on a large number of trials with different images, especially natural images.

The proposed approach can be outlined as follows.

- (a) Compute the autocorrelation function $R_h(k)$ of the cover image histogram.
- (b) Estimate the noise power σ^2 in the histogram autocorrelation $\{R_h(k)\}$. This computation is based on the finite-difference filtering approach which was first proposed in [14]. Of course, this is done under the assumption of Gaussianity of noise.

The autocorrelation function is a measure of the amount of similarity between signal samples. Note that in this approach we do not consider the correlation between image pixels, but the correlation between the histogram samples. If we consider the cover image histogram $\{h(i)|i = 0, \dots, N-1\}$ as a signal with N samples (N is the number of levels in the gray-valued illumination, being $2^8 = 256$ for 8-bit pixels), the autocorrelation is calculated as [15]:

$$R_h(k) = \frac{1}{2N-1} \sum_{i=0}^{2N-1} h(i)h(i+k) \quad \dots\dots (1)$$

where $\{h(k)|k = 0, \dots, N-1\}$ is the cover image histogram given by:

$$h(k) = n_k = \text{Number of pixels with intensity } k$$

For untampered image, the histogram autocorrelation function

$$\{R_h(k) | k = 0, \dots, 2N - 1\}$$

is expected to be smooth and differentiable.

When a signal $\{s(k)\}$ is disturbed by additive noise $\{z(k)\}$, the resultant noisy signal is given by:

$$x(k) = s(k) + z(k) \quad \dots\dots (2)$$

The noise power σ^2 can be estimated from the noisy signal power σ_x^2 after applying the n^{th} difference filtering as follows [14]:

$$\hat{\sigma}^2 = (C_n^{2n})^{-1} \text{var}(\Delta^n x) = \frac{(n!)^2}{(2n)!} \text{var}(\Delta^n x) \quad \dots\dots (3)$$

where,

$$\Delta x(k) = x(k + h) - x(k),$$

$$\Delta^n x(k) = \Delta\{\Delta^{n-1}x(k)\},$$

and C_n^{2n} is the binomial coefficient. Normally we take $h = 1$.

Normally taking $n = 5$ is sufficient. It should be noted that we didn't use the 2D difference filtering as explained in [14] because we are dealing with the 1D signal $x(k) = R_h(k)$, which is not the image itself but its histogram correlation.

After estimating noise power, we put a threshold on this power, σ_o^2 , which is normally a small amount below at which the image is considered as clean, otherwise it is considered as tampered with (or have hidden data).

In this work we generally used the following threshold for decision:

$$\sigma_o^2 = 0.1 \quad \dots\dots\dots (4)$$

Hence, our decision is as follows:

$$\hat{\sigma}^2 (2N - 1)^2 \leq \sigma_o^2 \quad \dots\dots\dots (5)$$

However, the value in Equation (4) could be changed to a lower one when failure occurs.

V. RESULTS AND DISCUSSION

The proposed approach for steganalysis using noise estimation has been tested on MATLAB using LSB steganography with 1 and 2 LSB-bits for message hiding in a larger cover image. In this test, only gray-scale images are considered, although extension to color images is straightforward. If color images are used, then the noise power $\hat{\sigma}^2$ in Equation (5) would be the average of three RGB estimated noise powers:

$$\frac{\hat{\sigma}_R^2 + \hat{\sigma}_G^2 + \hat{\sigma}_B^2}{3} (2N - 1)^2 \leq \sigma_o^2 \dots\dots\dots (6)$$

noting that the average in Equation (6) is related to noise power and not to the different RGB weights normally used in RGB-to-gray conversion.

We used MATLAB built-in images as shown in Figures 1 and 3 for a message with two different covers.

The ratio r of the message size to the image size is defined as:

$$r = \frac{\text{Message Size}}{\text{Cover Image Size}} \dots\dots\dots (7)$$

To test the performance under different ratios r , the message was resized.

Figures 2 and 4 show the normalized autocorrelation of the histograms of the tampered cover images using LSB steganography with 1 and 2 bits for different message-to-cover ratios.

The decision about the existence of tampering using the criterion of Equation (5) is shown as high/low dashed red line. A high line (approaching the limit of 1) means tampering, otherwise no tampering. It is evident that the original image (top sub-figure) is clear, with the dashed line kept at a low level.

It can be seen from Figures 2 and 4 that the disturbance in the histogram correlation is more evident for 2-LSBs than 1-LSB tampering. For the same k -LSB steganography (k fixed), the disturbance (noise) is more evident for larger message-to-cover ratio (r). In addition, disturbances are clearer in Figure 4 than in Figure 2, hence, the combination in Figure 1 represents a more difficult case.

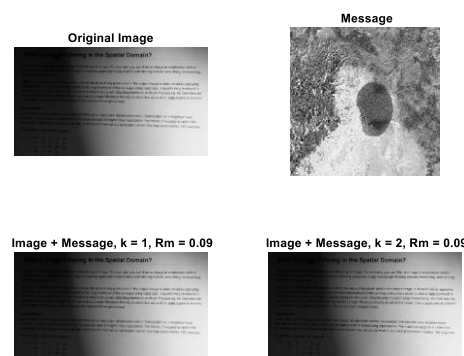


Figure 1: First cover image and message.

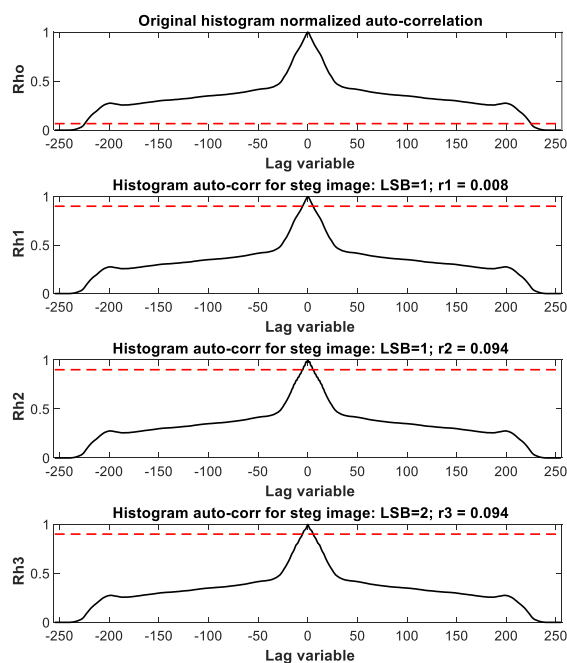


Figure 2: Normalized autocorrelation of the histograms of the first cover image (with and without a message) using LSB steganography with 1 and 2 bits.

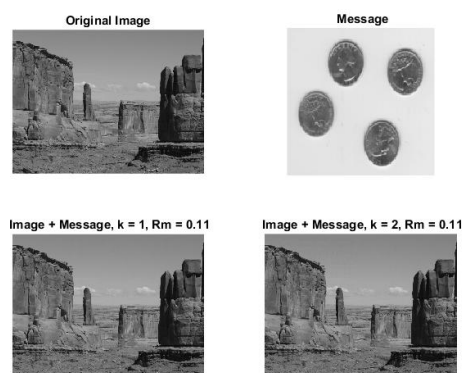


Figure 3: Second cover image and message.

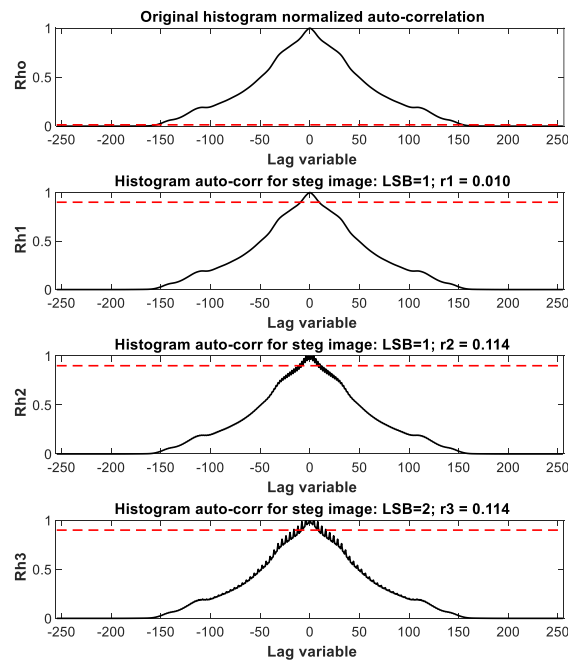


Figure 4: Normalized autocorrelation of the histograms of the second cover image (with and without a message) using LSB steganography with 1 and 2 bits.

VI. LIMITATIONS AND THRESHOLD VARIATION

The proposed approach may fail for very small message to cover ratio r . Failure can also be cover and/or message – dependent. Figures 5 and 6 show a failure to detect steganography for $r < 0.005$ when we use a threshold of $\sigma_o^2 = 0.1$ as shown in Equation (4). However, this can be treated if we lower the threshold to $1/70$, as shown in Figure 7. Hence, if failure occurs, we should attempt at lowering the threshold before confirming the failure of the approach.

VII. CONCLUSION

A reference-free approach for steganalysis or tampering detection in the spatial domain is presented. The approach is based on detecting disturbance in the autocorrelation of the cover image histogram via estimating the noise power in the correlative-histogram domain using a finite-difference filter. The approach has been tested for LSB steganalysis, but it could be applied to other spatial steganography methods.

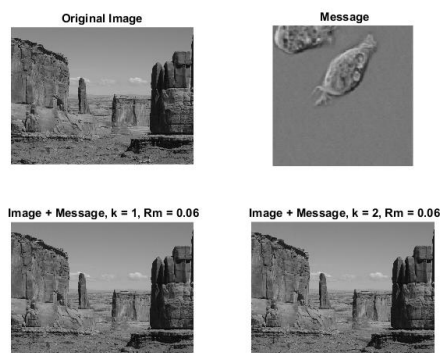


Figure 5: Third cover image and message.

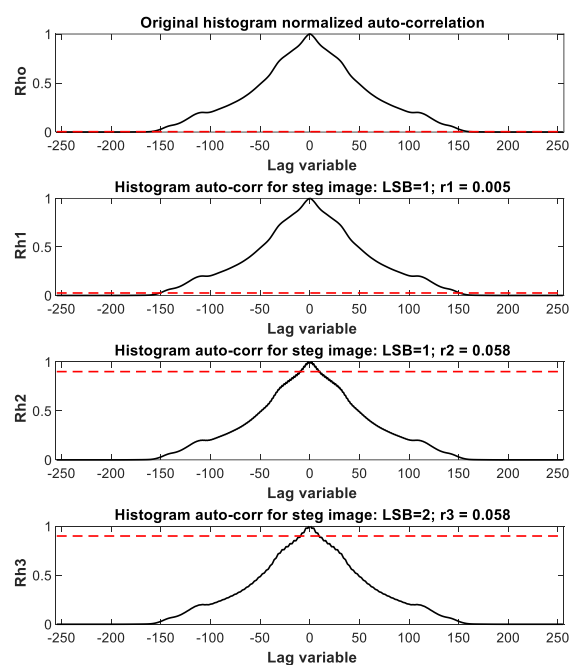


Figure 6: Normalized autocorrelation of the histograms of the third cover image (with and without a message) using LSB steganography with 1 and 2 bits. Notice the failure at $r1$. Threshold of 0.1 is used here. However, for the same value of $r1$, we can remedy this failure as shown in Figure 7.

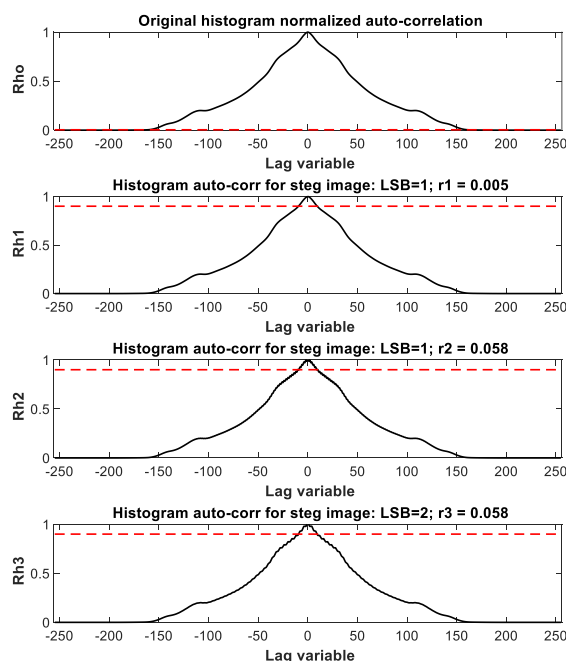


Figure 7: Normalized autocorrelation of the histograms of the third cover image (with and without a message) using LSB steganography with 1 and 2 bits. A threshold of $1/70$ is used here to remedy the failure in Figure 6.

REFERENCES

- [1] P. C. Mandal, et al, "Digital image steganography: A literature survey," Information Sciences, Elsevier, 2022.
- [2] N. Subramanian et al., "Image Steganography: A Review of the Recent Advances," IEEE Access, 2021.
- [3] S. Das, et al, "Steganography and Steganalysis: different approaches", in arXiv preprint arXiv: 1111.3758, 2011.
- [4] S. Lyu and H. Farid, "Steganalysis using color wavelet statistics and one-class support vector machines", Proc. SPIE 5306, 35–45 -2004.
- [5] A. Ker, "Improved detection of LSB steganography in grayscale images", International Workshop on Information Hiding, 2004.
- [6] J. Fridrich, M. Goljan, R. Du, "Reliable detection of LSB steganography in grayscale and color images", Proceeding of ACM, Special Session on Multimedia Security and Watermarking, Ottawa, Canada, 2001, pp. 27–30.
- [7] A. D. Ker, "A General Framework for Structural Steganalysis of LSB Replacement", International Workshop on Information Hiding, 2005.
- [8] T. Zhang, X. Ping, "A new approach to reliable detection of LSB steganography in natural images," Signal Processing, Elsevier, 2003.

- [9] X. Kong, T. Zhang, X. You, and D. Yang, "A new steganalysis approach based on both complexity estimate and statistical filter," in Proc. IEEE Pacific-Rim Conf. on Multimedia, 2002.
- [10] M. U. Celik, et al, "Universal Image Steganalysis Using Rate-Distortion Curves", Proc.SPIE: Security, Steganography, and watermarking of Multimedia Contents,2004.
- [11] G. Xuan et al, "Steganalysis Based on Multiple Features Formed by Statistical Moments of Wavelet Characteristic Functions," International Workshop on Information Hiding, 2005.
- [12] N. M Abdali, Z. M. Hussain, "Reference-free detection of LSB steganography using histogram analysis," 30th International Telecommunication Networks and Applications Conference (ITNAC), IEEE, 2020.
- [13] N. M Abdali, Z. M. Hussain, "Reference-free differential histogram-correlative detection of steganography: Performance analysis," Indonesian Journal of Electrical Engineering and Computer Science, 2022.
- [14] J. Kowalski, G. Mikolajczak, J. Peksinski, "Noise Variance Estimation in Digital Images using Finite Differences Filter," 2018 41st International Conference on Telecommunications and Signal Processing (TSP), 2018.
- [15] Jose Maria Giron-Sierra, Digital Signal Processing with MATLAB Examples, Springer, 2017.