International Journal of Electrical and Computer Engineering (IJECE)
Vol. 9, No. 6, December 2019, pp. 5286~5294
ISSN: 2088-8708, DOI: 10.11591/ijece.v9i6.pp5286-5294

A new approach for enhancing LSB steganography using bidirectional coding scheme

Mohammed Al-Momin, Issa Ahmed Abed, Hussein A. Leftah
Basrah Engineering Technical College (BETC), Southern Technical University (STU), Iraq

Article Info

Article history:

Received Dec 10, 2019
Revised Jul 19, 2019
Accepted Jul 27, 2019

Keywords:

Data embedding
Image processing
Information hiding
Steganography

ABSTRACT

This paper proposes a new algorithm for embedding private information within a cover image. Unlike all other already existing algorithms, this one tends to employ the data of the carrier image more efficiently such that the image looks less distorted. As a consequence, the private data is maintained unperceived and the sent information stays unsuspicious. This task is achieved by dividing the least significant bit plane of the cover image into fixed size blocks, and then embedding the required top-secret message within each block using one of two opposite ways depending on the extent of similarity of each block with the private information needed to be hidden. This technique will contribute to lessen the number of bits needed to be changed in the cover image to accommodate the private data, and hence will substantially reduce the amount of distortion in the stego-image when compared to the classic LSB image steganography algorithms.

Corresponding Author:

Mohammed Al-Momin,
Basrah Engineering Technical College (BETC),
Southern Technical University,
Basrah, Iraq.
Email: mohammed.al-momin@stu.edu.iq

1. INTRODUCTION

The dramatically increasing number of applications that are having place on the Internet, together with the massive trends of people towards social media services has put the information privacy on bottleneck [1]. Many research efforts have been dedicated for developing new algorithms of hiding private data behind elusive non-important information. The target information which needs to be transmitted in top security may be concealed beyond textual cover by including the information bits inclusively among the cover's data. The secure information may also be covered with image, audio, or video files. Only the party this information is intended to be sent to has authority to retrieve the private information from the cover data. On the other hand, the message itself can be text characters, sound tracks, or video clips.

The word steganography came from the combination of two Greek words steganos which implies hidden, and the word graphy which means writing, therefore steganography is defined as hidden writing [2, 3]. Many papers have been conducted in the field of steganography and water marking, these papers generally concentrated on two aspects, capacity and security. In capacity it is meant the maximum available space in which private data can be stored, or in other words the maximum allowable size of the hidden data [4]. In contrast, other kind of papers gives no or fewer expense to capacity, while they concentrate on the robustness of system's security [5]. Embedding messages into a cover image is very familiar in case of military applications, since such information needs to be dealt with in extreme confidentiality [6]. In image steganography, the pure visible image that does not have the embedded private message within it is called the cover or carrier image, whereas the name stego is termed to the cover when it carrier the hidden target information.

One important concern that must be taken into account when embedding messages into a cover is the imperceptibility, where as much as the image looks undistorted, it remains beyond all suspicions. Consequently, a good steganographic algorithm is necessarily characterized by its extreme difficulty to perceive any abnormality in the stego-image by the human's eye. Therefore, the visual quality of the stego strongly affects the performance of the steganographic system. The better the quality obtained, the greater the performance achieved. In the visual quality it is not meant the absolute quality of the stego-image, but instead, the quality related to the original cover image, or in other words, the difficulty to differentiate among the two images (stego, and cover).

In this paper, a new technique aims to minimize the number of modifications in the carrier image data when embedding confidential information has been suggested. Cover image is first partitioned into smaller sized blocks, and the similarity between each individual block bits and their corresponding bits of the message wanted to be concealed is taken into consideration when coding the secret message into a stego-image. Message bits are encoded in one direction when there is a dominant similarity, and encoded in the reverse direction when the dissimilarity is dominant between the message and the corresponding cover bits. The LSB of the cover image is to be amended according to this criterion, in order to accommodate the secret message. This way the number of modifications in the bits the carrier's LSB will be reduced resulting in an enhanced visible quality of the stego-image, and hence enhanced imperceptibility. In this paper, section II discusses the different strategies used for image steganography. In addition, section III explains the conventional LSB steganography system. Section IV on the other hand, shows the experimental results of our proposed algorithm, and compare with the other existent algorithms in terms of the Mean Square Error, MSE, and the Structural Similarity, SSIM, index. Finally, section IV concludes this paper.

2. IMAGE STEGANOGRAPHY

It is very familiar to hide information within an image as a consequence to the relatively high storage capacity as well as satisfying imperceptibility it offers due to the large number of redundant bits it contains [7]. Images are dealt with in computer as fixed size matrices of pixels. These pixels are represented in many different ways depending on the type image coding used. In gray-scaled image type, each pixel is expressed by a single 8-bit value which represents the degree of the gray scale this pixel has, graded from 0 to 255[8]. In contrast, indexed images, use the pixel's value to refer to a separate stand-alone index, which in turn defines the hue of this pixel accurately. Black and white images use the value of 0 to represent the black color, and 1 to represent white, where no blackness grading is offered [8]. On the other hand, RGB image utilizes three 8-bits numbers to define the quantity of redness, greenness, and blueness of the corresponding pixel [8].

Since the bandwidth of communication networks is limited due to the vast number of users, and the numerous bandwidth eating applications that are used over these networks, this makes uploading and downloading big sized images an inefficient way [9]. Alternatively, images are used to be compressed before being forwarded on the network's links. Different approaches with different performances exist in this area. These approaches can be roughly classified into two main types, lossy and lossless compression techniques. In lossy system, unnecessary information of image is removed to achieve a substantial difference in the image size, whereas lossless techniques used some statistical strategies to reduce the number of redundancies in image, and hence reduce the original image size [10].

One of the most common compression technologies used is the JPEG (Joint Photographic Experts Group) image coding. The latest uses a lossy compression scheme, where the image is transformed from its special coordinates into frequency domain coordinates. The major idea is standing beyond the inability of human eyes to recognize changes of images in high frequency spectrum [11]. Discrete Cosine Transform (DCT) which is a modified version of Discrete Fast Fourier Transform (DFFT) is used to represent the image in frequency domain. Only the small portion of image information which is located at low frequency space is kept survived, whereas all other data is omitted. DWT (Discrete Wavelet Transform) is also sometimes used to compress image files [12].

In image steganography, as the name implies, the medium to hide secret information behind is an image, where this technique exploits the limited capability of human's eye to percept some kinds of adjustments in the cover image. Many aspects of human eye's limitations were invested to accommodate the secrete message within the carrier image without being noticed by untargeted watchers. Note that lossy image compression techniques are not suitable for some types of steganography since the important message bits can be missed as victim of applying the compression method [13]. According to the way of hiding the private data within the cover file, image steganography can be classified into three major kinds.

a. Data Insertion Strategy
This kind depends on inserting the data of the secret message in places of the carrier image where the watcher will not perceive. Of course, this redundant data addition will be accompanied by image size enlargement which may be considered as a clear drawback of the system. However, this method of steganography possesses good degree of imperceptibility since private data are used to be embedded in locations that are usually ignored by the application that displays the image, such as in the header or trailer of the image's file [14].

b. Data Replacement Strategy
In contrast to the data insertion technique, this method does not make any change to the size of the carrier image file, since it does not insert any more bits. Instead, this technique replaces some bits of the original cover image which have minor effects on the appearance of the resulting image. The space that is available for embedding the private data is totally governed by the number of insignificant bits of the cover image. The imperceptibility measure is this method depends on the degree of significance of the replaced bits, and the degree of similarity to the original image, the more similar the stego-image obtained, the more the performance achieved [14].

c. Cover Image Generation Strategy
In contrast to the other strategies, this one uses a one-way algorithm to construct a unique cover image corresponding to each private message. This way strengthens the system against malicious attacks which utilizes a comparison between, the stego and cover images to discover the hidden message. The major problem with this method is the random behavior of this algorithm, where the generated image consists of random shapes and colors, which may be suspected by the end users [14].

3. IMAGE STEGANALYSIS

The term *steganalysis* refers to techniques used to reveal the hidden information embedded in the cover file [15, 16]. Many algorithms were introduced for this purpose. The steganography system is considered broken as soon as the hidden data is extracted from the cover file, even though such data have not been decoded or deciphered to its original secret message. In fact any change that is made to the cover file has its own effect on the characteristics of the cover image; this reality opens the door to the stegranographic efforts. The following three main types of image steganalysis techniques can be distinguished. (i) Passive steganalysis attacking technique: The aim of this technique is to discover and read the private message stored in the carrier image without any trend to modify the stego image data. Therefore this method keeps the setgo-file intact [17]. (ii) Active attacks: This attack changes the content of the stego-image by adding some noise to the image in order to prevent any probable secret information from being conveyed as a precautionary procedure even if the transmitted image was out of suspicion [18]. (iii) Malicious attacking technique: In this technique, the attacker not only has the ability to amend the stego-image, but further they plays a role of one on the authorized communicating parties by replacing the original stego-image by a fake one to convey fabricated the information they want to the end users [18].

4. CONVENTIONAL LSB STEGANORARHY METHOD

Each pixel in a gray scaled image represents the darkness of that point of the image where this pixel is located. The value of this pixel ranges from 0 to 255 to interpret 256 degrees of darkness. The value 0 corresponds to a pure black color, and the amount of pixel's whiteness increases with the increase this value, whereas the value 255 corresponds to a pure white color. These 256 levels of whiteness are represented by an 8-bit number. It is clear that changes to the very last bit (LSB) do not make a significant change in the value of the number, and hence in amount of darkness of the pixel corresponding to that number. For instance, if one changes the LSB of a pixel from 10100001 which reads 161 in decimal to 10100000 which reads 160, the value of darkness does not increase that much that the human eye can recognize. This fact is employed in LSB steganography, where the LSBs of the image pixels (well known as bit 0 plane) are replaced by the private message wanted to be hidden.

In order to test the significance of each bit plane of an image in the total appearance of that image, the image in Figure 1 was used. Figures 2-5 show the plot of bit 0, 4, 5 and 7 planes. It is clear that bit 0 plane does not have that much information that affect the total appearance of the image. In contrast, as we move towards the MSB, more information is extracted.

Figure 1. The original image [19]

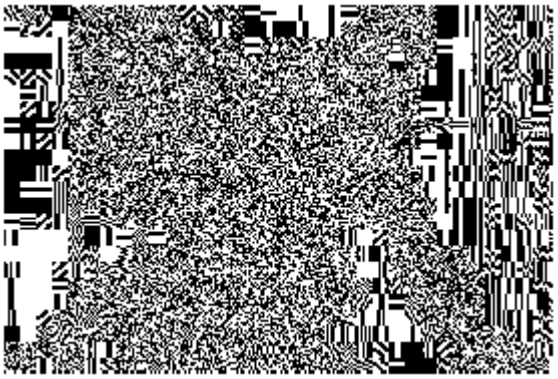

Figure 2. Bit 0 plane

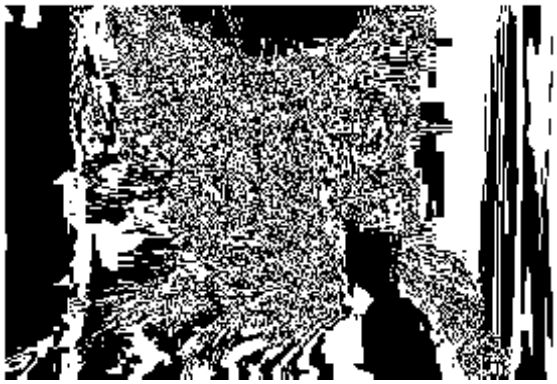

Figure 3. Bit 4 plane

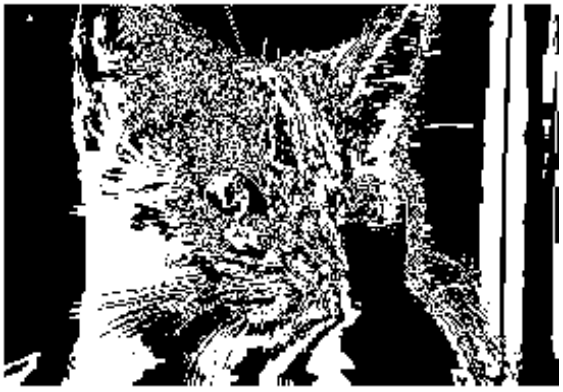

Figure 4. Bit 5 plane

Figure 5. Bit 7 plane

5. THE PROPOSED BI-DIRECTIONAL LSB STEGANOGRAPHY METHOD

In this section, the proposed steganography technique is illustrated. The LSB plane of the cover image is divided into fixed size blocks. Bits of each block are XORed with the corresponding message bits that are planned to be placed there. Regardless of whether the message is ciphered or in its plain form, the message's bits are XORed with the cover's bits. If the number of resulting zeros exceeds the number ones, this implies that there are more similarities between the message and cover bits than dissimilarities. On the other hand, if there are more ones yielded from the XOR operation than zeros, this means that the number dissimilarities between the cover and message bits is greater than the number of similarities. Similarity occurs when the bit of the cover image is similar to the corresponding bit of the message image, whereas if these bits differ in their values, dissimilarity occurs.

If the cover's block in total is more similar to the corresponding message part, then the information is to be coded in forward direction, where logic 0 is used to interpret the bit value 0, and logic 1 is used to interpret the bit value 1. In contrast, if the number of dissimilarities exceeds the number of similarities, then backward coding is used. In backward coding technique, it is meant that logic 0 in the cover image is used to embed the value 1 of the message's corresponding bit, and logic 1 is used to represent the value 0. The aim of this bidirectional coding technique is to minimize the number of amendment in the cover image, in order to keep the stego image (the image with the embedded private message) as much similar to the cover image as possible. Consequently, this leads to a substantial reduction of the error which is defined as the difference between the original cover image and the image with embedded information (stego-image).

The size of blocks and some other parameters regarding the ciphering technique, if any ciphering used, are to be agreed between the two communicating users in advance. Since the receiving terminal has no prior information about the construction of the secret message, and to what extent a particular block of the cover image is similar to the corresponding part of private message, then the receiver cannot make a decision about whether the embedded data were coded in forward direction or in backward direction. Such information is cardinal to the receiver in order to be able to decode and extract the embedded data from the stego-image. This problem can easily be tackled if the sender leaves a footprint in some predefined (agreed) location in the stego-image to indicate the similarity index.

Similarity index which is ranged from 0 to 1 is defined as the amount of similarity between the bits of the confidential message and the cover image. The value 0.5 0f this index refers to an equal number of similarities and dissimilarities. If the similarity index is greater than 1, then the number of similarities is greater than the number of differences, else there are more differences than likeliness. Similarity index is not that costly to embed in the stego-image. It can be represented by a mere one bit, where logic 0 refers to an overall similarity (similarity index ≥ 0.5), and logic 0 refers to an overall dissimilarity (similarity index < 0.5), and it can be placed in the first bit of each block as we have used in our simulation, or in the same location of the block's first bit but in the second LSB bit plane (bit-1 plane). The receiving user decodes each 0 as 0 and each 1 as 1 if the similarity index bit is 1, whereas it interprets each 0 as 1 and each 1 as 0 if the similarity index bit is 0. Figure 6 below shows the sending part flow chart of the proposed steganography approach, and Figure 7 illustrates the receiving part algorithm.

The performance of this newly introduced algorithm was measured according to two major criteria: mean squared error, and the structural similarity. These two factors are commonly used to study the quality of resulting images after different kinds of manipulations. In mean squared error, the average error of a particular processed image with respect to a standard reference image is computed according to equation (1) below. Ref is the reference image, and Im is the actual image required to be compared to the reference image. M and N are the dimensions of the images.

$$\boldsymbol{MSE} = \frac{\sum_{i=1,j=1}^{i=M,j=N} (\boldsymbol{Ref_{ij} - Im_{ij}})^2}{\boldsymbol{M \times N}} \quad (1)$$

The other factor that dimensions the performance of the proposed system performance is the structural similarity between the stego-image and the original cover image. The built-in MATLAB function ssim was used to compute the structural similarity index in a window based manner. Two gray scaled images were taken to test the performance of the new system; the cat image in Figure 1, and the well-known Lena image.

Figure 8 below shows Lena's image after embedding 32KB of confidential text within it. It is clear that this image does not raise any suspicion to watchers, and hence keeps the steganography system imperceptible. Figures 9 and 10 compare the Mean Square Error (MSE) and the Structural Similarity index (SSIM) of the proposed bidirectional coding technique with the conventional one for the first test image, cat's image. It is clear from these figures that the newly proposed system outweighs the traditional technique to a substantial extent, especially for smaller blocks sizes. Figures 11 and 12 on the other hand present an overall comparison between the conventional and proposed LSB systems in term of the mean square error and the structural similarity index for the second test image, Lena's image. It is also obvious in these figures that the proposed system still more beneficent than the traditional LSB system, and this benefit increases with the decrease of the blocks sizes.

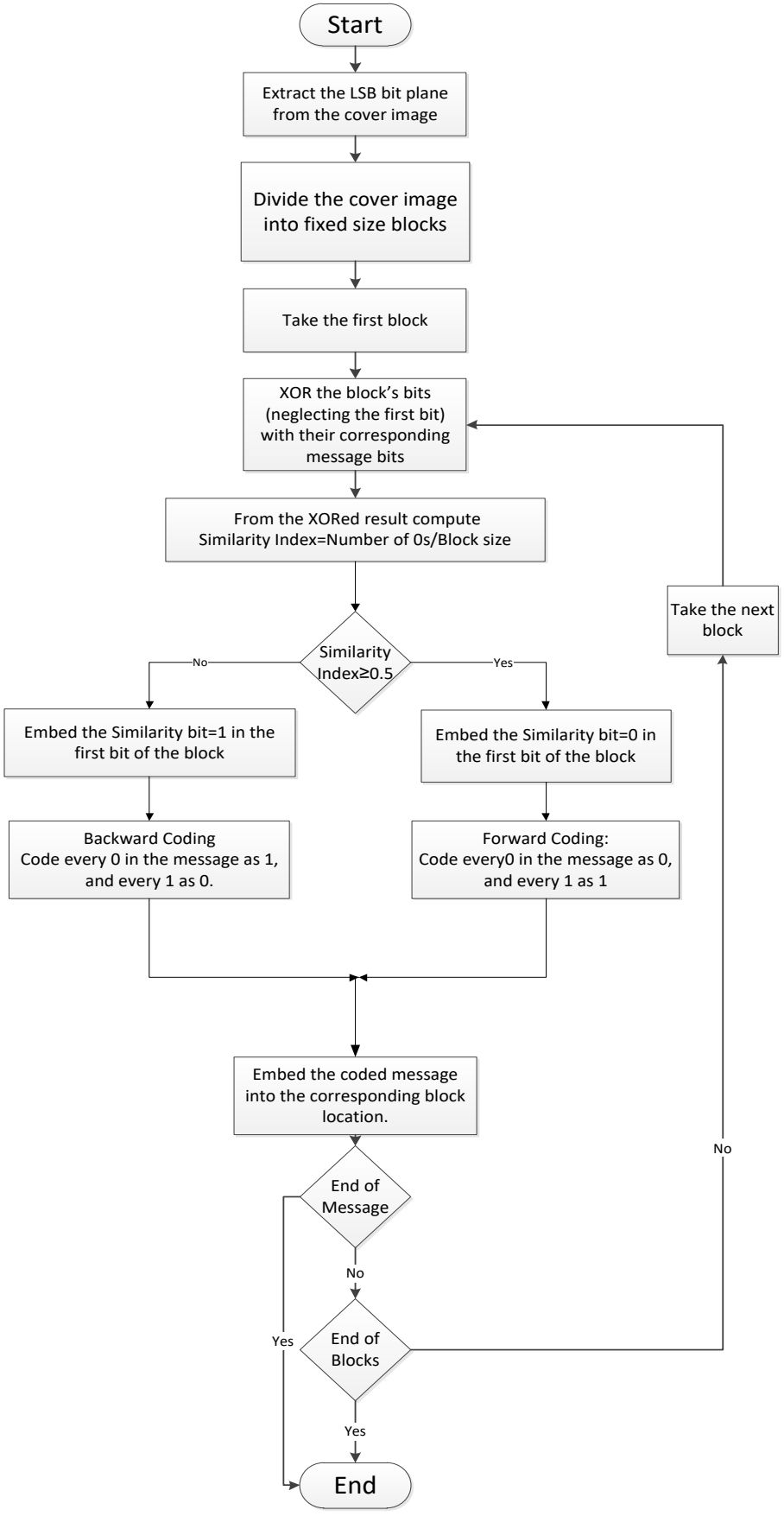

Figure 6. The sending part flow chart of the proposed steganography

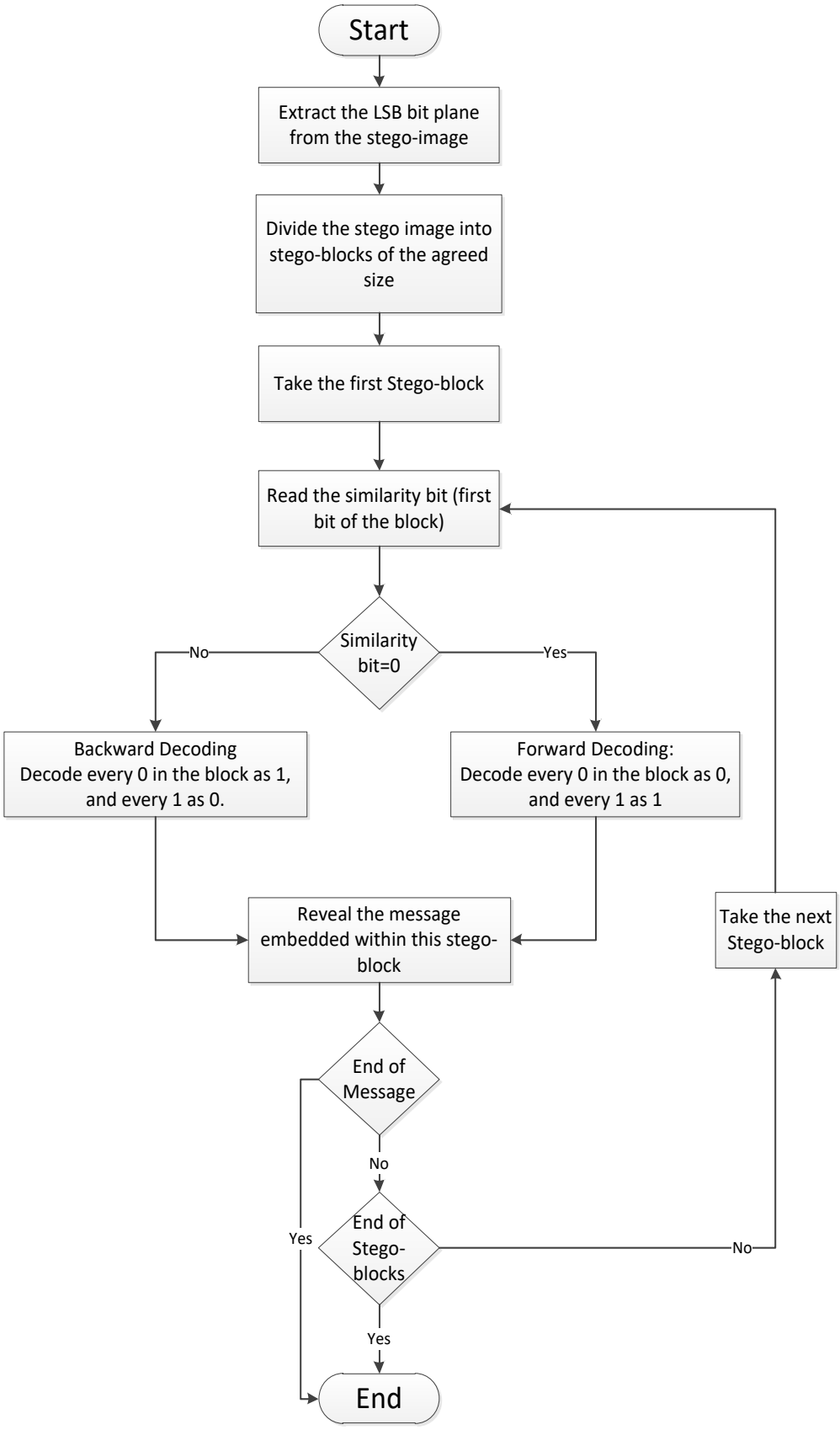

Figure 7. The receiving part flow chart of the proposed steganography approach

Figure 8. Lena Image [20]

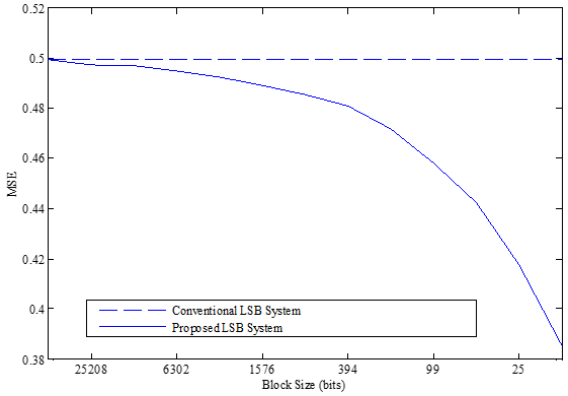

Figure 9. Mean Square Error of the resulting cat's stego-image

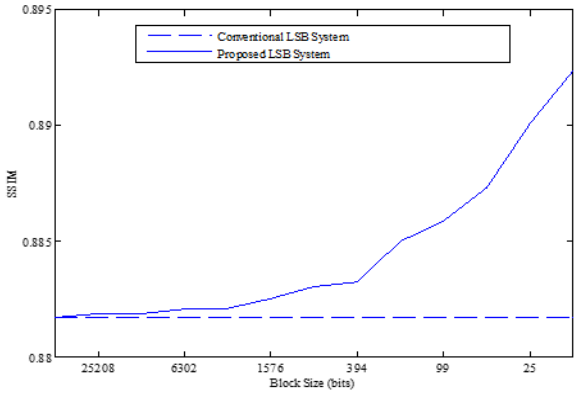

Figure 10. Structural Similarity Index of the resulting cat's stego-image

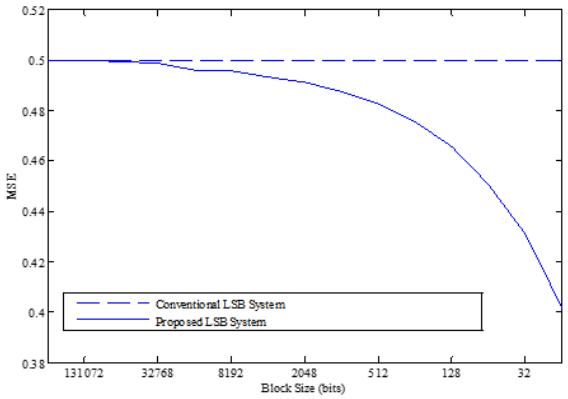

Figure 11. Mean Square Error of the resulting Lena's stego-image

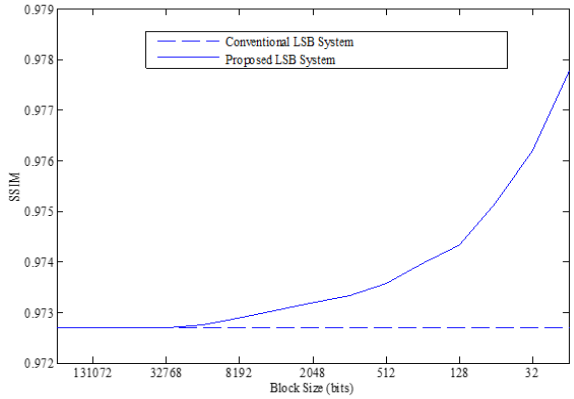

Figure 12. Structural Similarity Index of the resulting Lena's stego-image

6. CONCLUSION

In this paper, a novel idea for improving the conventional LSB steganography was introduced. The bidirectional coding technique was suggested here to minimize the number of amendments made to the original cover image bits when embedding the target confidential message. The original cover image was divided into a similar size blocks, and each block was encoded individually to the stego-image in one of the either directions. In bidirectional coding, it is meant that logic 0 of the private message bits is either encoded as 0 and logic 1 as 1, or logic 0 is encoded 1, and 1 is encoded as 0 in each stego-block depending on the extent of similarity among the cover's block bits from one side and target message bits from the other side.

Two main factors were considered here to examine the efficiency of the proposed system, namely the mean square error and the structural similarity index. The minor distortion resulted as a consequence of embedding the secret message in the cover image to produce a stego-image was precisely computed and compared with that of conventional LSB steganography systems. This new approach was examined on two test cover images. All the resulting data was supportive, where was found that the new bidirectional coded LSB steganography approach achieves a substantial reduction in the mean square error, as well as a better similarity index, especially when the cover image is partitioned into more smaller sized blocks. Future work may include improving the algorithm by expanding it to employ more than one LSB bit plane, where the embedding capacity can be extended to include the three or even four LSB bit planes. Another suggestion is to develop the system to work on colored images rather than gay scale ones.

REFERENCES

[1] Ran-Zan Wang, Chi-Fang Lin, Ja-Chen Lin, "Image hiding by optimal LSB substitution and genetic algorithm," *In Pattern Recognition*, Volume 34, Issue 3, pp. 671-683, 2001.

[2] Chi-Kwong Chan∗, L.M. Cheng, "Hiding data in images by simple LSB substitution" , *Department of Computer Engineering and Information Technology, City University of Hong Kong*, Hong Kong , August 2003.

[3] Dr. M. Umamaheswari Prof. S. Sivasubramanian S. Pandiarajan, "Analysis of Different Steganographic Algorithms for Secured Data Hiding," *IJCSNS International Journal of Computer Science and Network Security*, VOL.10 No.8, August 2010.

[4] Lee, Yeuan-Kuen, and Ling-Hwei Chen. "High capacity image steganographic model," *IEE Proceedings-Vision, Image and Signal Processing* 147.3, pp. 288-294, 2000.

[5] Wang, Shen, Bian Yang, and Xiamu Niu. "A secure steganography method based on genetic algorithm." *Journal of Information Hiding and Multimedia Signal Processing* 1, no. 1, 28-35, 2010.

[6] Kumar, Arvind, and Km Pooja. "Steganography-A data hiding technique," *International Journal of Computer Applications* 9, no. 7, 19-23, 2010.

[7] Morkel, Tayana, Jan HP Eloff, and Martin S. Olivier. "An overview of image steganography," In *ISSA*, pp. 1-11. 2005.

[8] McAndrew, Alasdair. "An introduction to digital image processing with matlab notes for scm2511 image processing." *School of Computer Science and Mathematics, Victoria University of Technology* 264 (2004).

[9] Rabbani, Majid, and Paul W. Jones. *Digital image compression techniques*. Vol. 7. SPIE Press, 1991.

[10] Singh, Manjari, Sushil Kumar, and Siddharth Singh."Various Image Compression Techniques: Lossy and Lossless,*" International Journal of Computer Applications (0975 – 8887)* Volume 142 – No.6, May 2016.

[11] Raid, A. M., W. M. Khedr, M. A. El-Dosuky, and Wesam Ahmed. "JPEG image compression using discrete cosine transform-A Survey," *International Journal of Computer Science & Engineering Survey (IJCSES)* Vol.5, No.2, pp. 39-47, April 2014.

[12] Chowdhury, M. Mozammel Hoque, and Amina Khatun. "Image compression using discrete wavelet transform," *IJCSI International Journal of Computer Science Issues* 9, no. 4, pp. 327-330, 2012.

[13] Amin, Muhalim Mohamed, Mazleena Salleh, Subariah Ibrahim, Mohd Rozi Katmin, and M. Z. I. Shamsuddin. *"Information hiding using steganography*," In Telecommunication Technology, 2003. NCTT 2003 Proceedings. 4th National Conference on, pp. 21-25. IEEE, 2003.

[14] Al-Mohammad, Adel. "Steganography-based secret and reliable communications: Improving steganographic capacity and imperceptibility." *PhD diss., Brunel University, School of Information Systems, Computing and Mathematics Theses*, 2010.

[15] Karampidis, K., Ergina, K., & Papadourakis, G., "A review of image steganalysis techniques for digital forensics," *Journal of Information Security and Applications* 40, pp.217–235, 2018.

[16] Hui Tian, Jun Sun, Yongfeng Huang, Tian Wang, Yonghong Chen, and Yiqiao Cai, "Detecting Steganography of Adaptive Multirate Speech with Unknown Embedding Rate*," Mobile Information Systems,* vol. 2017, Article ID 5418978, 18 pages, 2017. https://doi.org/10.1155/2017/5418978.

[17] Shmatok, A. S., A. B. Petrenko, A. B. Yelizarov, V. A. Tytov, and E. A. Borysenko. "Steganalysis of graphic container," *Наукоємні технології* 4, pp.426-429, 2013.

[18] Cox, Ingemar, Matthew Miller, Jeffrey Bloom, Jessica Fridrich, and Ton Kalker. *Digital watermarking and steganography*. Morgan Kaufmann, 2007.

[19] Nkondog Anselme Venceslas, Feb 15th, 2019, *Machine learning is all about math*, accessed 5 July 2019, < https://medium.com/@nkanven/machine-learning-is-all-about-math-75848cb0c93d >.

[20] Katie Sharkey and Naresh Cuntoor, March 20th , 2012, *A Major Breakthrough in Image Processing*, accessed 5 July 2019, < https://blog.kitware.com/a-major-breakthrough-in-image-processing/>.