Segmenting and Hiding Data Randomly Based on Index Channel

Emad T. Khalaf[1] and Norrozila Sulaiman[2]

[1,2] Faculty of Computer Systems & Software Engineering, University Malaysia Pahang, Kuantan, 26300, Malaysia

Abstract

Information hiding is a technique of hiding secret using redundant cover data such as images, audios, movies, documents, etc. In this paper, a new technique of hiding secret data using LSB insertion is proposed, by using the RGB channels of the cover image for hiding segmented data. One of the three channels became the index to the two other channels. Firstly, the secret data are segmented into Even segment and Odd segment. Then, four bits of each segment is hidden separately inside the two channels depending on the numbers of "1"s inside the index channel. If the numbers were Even, then four bits of Even segment will be hidden. However, if they were Odd then four bits of Odd segment will be hidden. The opposite process retrieve the secret data from image by reading the bits of the index channel and check the numbers of "1"s to extract the Even segment and Odd segment. Finally, recombining the two segments to extract the secret data. Experimental results show that the proposed method can provide high data security with acceptable stego-images.

Keywords: *Steganography, Data hiding, Data segmenting, Index channel*

1. Introduction

Information security requirement became more important, especially after the spread of Internet applications [1]. However, Owners of sensitive documents and files must protect themselves from unwanted spying, copying, theft and false representation. This problem has been solved by using a technique named with the Greek word "steganography" it is mean hiding information [2]. Steganography is the art and science of hiding information. The data-hiding system design challenge is to develop a scheme that can embed as many message bits as possible while preserving three properties: imperceptibility, robustness, and security [4].In addition, proposing an effective method for image hiding is an important topic in recent years [5],[6].There have been many techniques for hiding information or messages in images in such a way that the alterations made to the image are perceptually indiscernible. Common approaches include [7]:

(i) Least significant bit insertion (LSB)
(ii) Masking and filtering
(iii) Transform techniques

Information hiding is an emerging research area, which encompasses applications such as copyright protection for digital media, watermarking, fingerprinting, and steganography [8]. All these applications of information hiding are quite diverse [8] and many encoding methods was proposed, a reversible image hiding scheme based on histogram shifting for medical images was proposed in [5]. An image-in-image hiding scheme, based on dirty-paper coding, that is robust to JPEG and additive white Gaussian noise (AWGN) attacks was proposed in [9]. Chen et al. [10] used a vector quantification method, but the method required a set of look-up tables. Moreover, the decoded images were little distorted from original images. Wang et al. [11] proposed a least significant bit technique to hide information. The technique could improve the visual quality of cover images, but the reconstruction processes were very complicated calculations. Chang et al. [12] proposed two kinds of hiding techniques and the hiding techniques secured better visual quality. However, the information capacity of these hiding techniques was low. Yang and Lin [13] used a basal-bit orientation method to hide images, and the method had large hiding capability and good visual quality of the secret image. In this paper we proposed a new method of segmenting and hiding the secret data in bmp color image by segmenting these data into two segment, i.e. Even segment and Odd segment. Then those two segments of characters will be hidden separately and randomly inside the cover image. By using random pixels to insertion secret data with modifying those data, this could avoid the detection by comparison of modified image with original image [3]. Two channels were used for hiding data in 24-bit BMP image and the third channel was used as index channel for the hidden data.

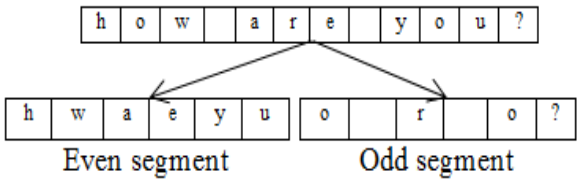
2. Steganography Techniques

Steganography is the art of embedding information in such a way that prevents the detection of hidden messages. It means hiding secret messages in graphics, pictures, movie, or sound. Steganography comes from the Greek word steganos, which means ‘covered’, and -graphy, which means ‘writing’. Covered writing has been manifested way back during the ancient Greek times around 440 B.C. Some of old steganography examples are shaving the heads of slaves and tattoo messages on them. Once the hair had grown back, the message was effectively hidden until the receiver shaved the heads once again. Another technique was to conceal messages within a wax tablet, by removing the wax and placing the message on the wood underneath [14].The most popular and frequently method of Steganography is the Least Significant Bit embedding (LSB). the level of precision in many image formats is far greater than that perceivable by average human vision. Therefore, an altered image with slight variations in its colors will be indistinguishable from the original by a human being, just by looking at it. If we using the least significant bits of the pixels' color data to store the hidden message, the image itself is seemed unaltered [15],[16] and changing the LSB’s value will have no effect on the pixel’s appearance to human eye. In our method, the 24-bit BMP image and least significant bit (LSB) insertion were used. The reason behind using BMP type is that it is more accurate in showing the image without any of compressed data and it is considered to be the most used format in hiding operation and analyzed.

3. The Proposed Method

A new steganography technique of uses the RGB images to hide the data in different channels was proposed. Two files are require to embedding a message into an image .The first is the message (the information to be hidden), a message may be plain-text, cipher-text, other images, or anything that can be embedded in a bit stream. The second file is the innocent-looking image that will hold the hidden information, called the cover image. Generally, Digital images are stored in computer systems as an array of points (pixels) where each pixel is consisting of three channels: (Red, Green, and Blue) and $0 \leq R,G,B \leq 255$ [17]. In our method the data is hidden into two of the RGB pixel channels based on the third channel. However, when using a 24 bit color image, two bits of two colors components can be used, so a total of 2 bits can be stored in each pixel

At the beginning, the secret data is split programmatically into array of characters and this array was segmented into two segments, Even segment and Odd segment
For example:
The secret data is: "how are you?"
Thus, the segmenting process will be:



Each segment is hiding separately in random two channels. One of the three channels was used as index channel to the next two channels by counting the number of "1"s in the index channel. If it is Even, then four bits of the Even segment data will be hidden inside the least significant bit of channel 1 and the least significant bit of channel 2. If it is Odd, then four bits of the Odd segment data will be hidden inside the least significant bit of channel 1 and the least significant bit of channel 2, The following example explain the idea.

Suppose that three adjacent pixels (nine bytes) with the following RGB encoding are used.

	Index Ch.	Channel1	Channel2	
Pixel(1):	10010101	00001100	11001001	*Even*
Pixel(2):	11010111	00001110	11001011	*Even*
Pixel(3):	10011011	00010000	11001010	*Odd*

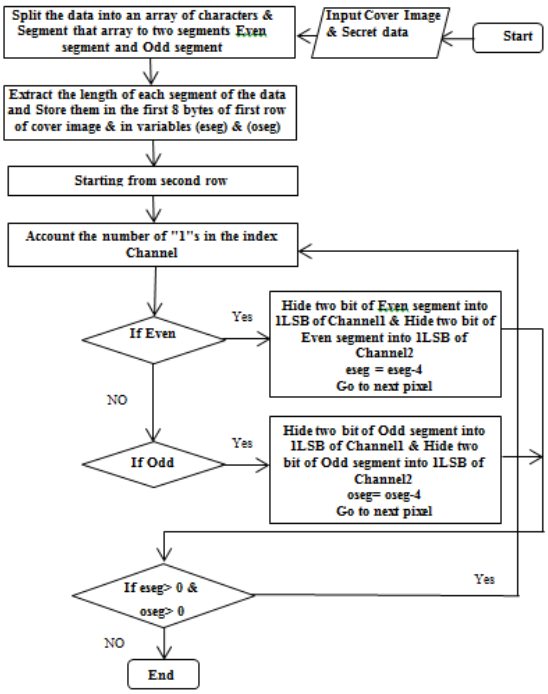
Now, in pixel1 4bits from Even segment of data will be hidden because number of "1"s in index channel is Even. In addition, 4bits from Even segment of data will be hidden in pixel2. However, in pixel3, 4bits of Odd segment will be hidden because number of "1"s in index channel is Odd. In this example, the red color was used as index channel. To improve security, the index channel is not fixed. The indexes are chosen sequentially, the first index is Red, and the subsequent indexes are Green and

Blue respectively. The index LSB bits are naturally available at random, based on image profile and its properties. Table 1 shows the relationship between the index channel and the hidden data inside the other channels.

Table 1.Meaning of Index channel values

No of "1"s in the index Channel	Channel1	Channel2
Even	*2bit of Even segment*	*2bit of Even segment*
Odd	*2bit of Odd segment*	*2bit of Odd segment*

As shown in the table, if the index channel is Red, channel1 will be Green and channel 2 will be Blue and the sequence will be RGB. In the second pixel, the index is Green. Channel 1 and channel 2 will be Red and Blue respectively. Hence, the sequence is GRB. In the third pixel, the index is Blue. Therefore, channel 1 is Red and channel 2 is Green. The sequence is BRG. The processes method is as shown in Figure 1. First process is used to input the cover image and the secret data. Then, the data will be segmented and the length of each segment will be stored in the first 8 bytes of the beginning of the image. The hiding process starts from the second row and it depends on the numbers of "1"s in the index channel.



The recovery processes for the proposed method is shown in Figure 2. It is the exact reverse of the hiding process, starting with extracting the two segment of the data length from the first 8 bytes of the image.

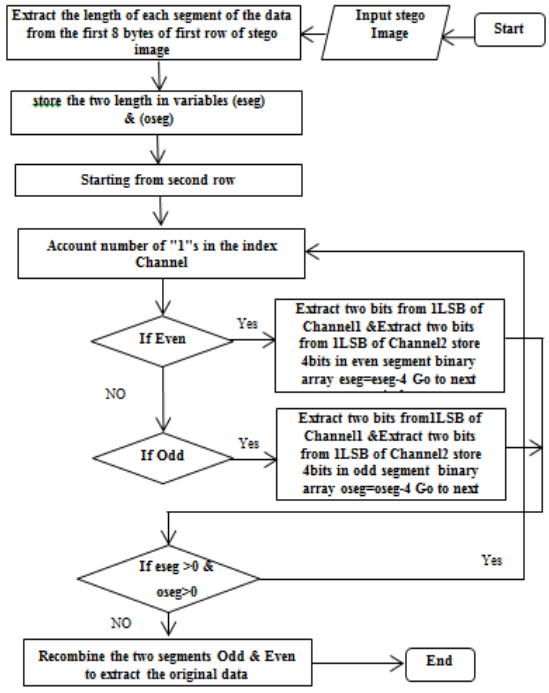


Fig. 2 Recovery and Recombine process flowchart

4. Experimental Results and Analysis

We have tested our algorithm for different sets of images as well as text messages. Histogram analysis has been implemented to the image before and after embedding data, to compare the original channels, before and after modifying channels. This can give a clear idea of the security and if change is minimal, then the method is considered secure. Figure 3a shows the original image of Mosul City and Figure 3b shows the stego image of Mosul City. Another image is an original image of a bird, as shown in Figure 4a and its stego image is as shown in Figure 4b. The Red, Green and Blue histograms for Mosul City image is as shown in Figure5, Figure6, Figure7 and Figure8. The modified images after applying the method did not show any identifiable visual difference.

Fig. 3a: Original image (Mosul City) length 500x330

Fig. 3b: Stego image (Mosul City) with text size 1420 characters

Fig. 4a: Original image (Bird) length 760x570

Fig. 4b: Stego image (Bird) length 760x570 with text length 2300

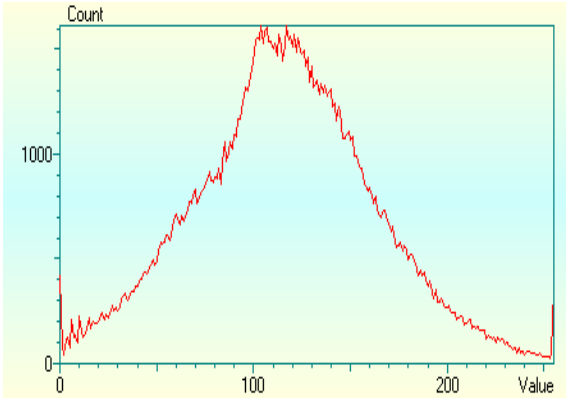


Fig. 5a: Original image (Mosul City) histogram of Red channel

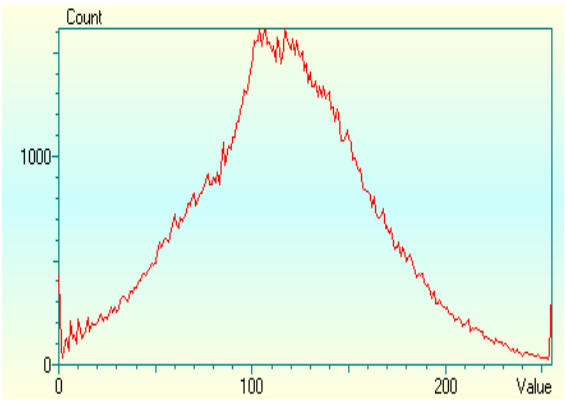


Fig. 5b: Modified image (Mosul City) histogram of Red channel

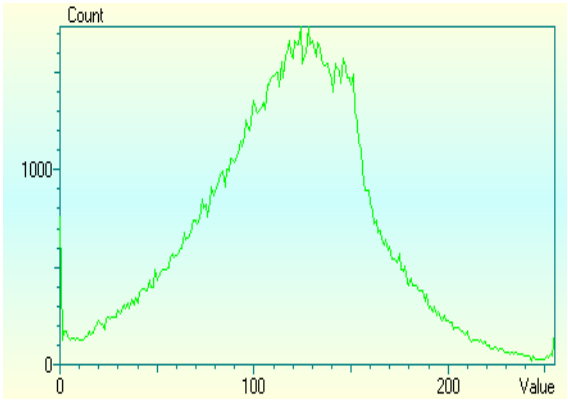


Fig. 6a: Original image (Mosul City) histogram of Green channel

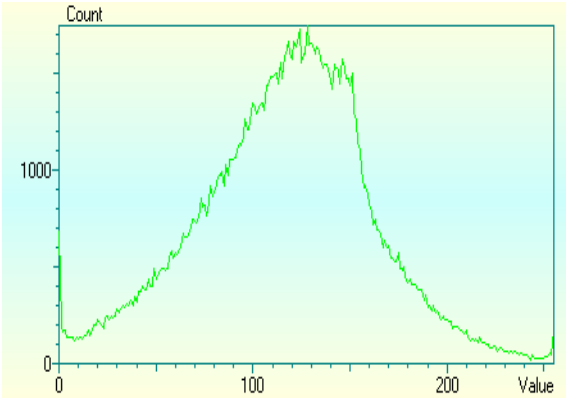


Fig. 6b: Modified image (Mosul City) histogram of Green channel

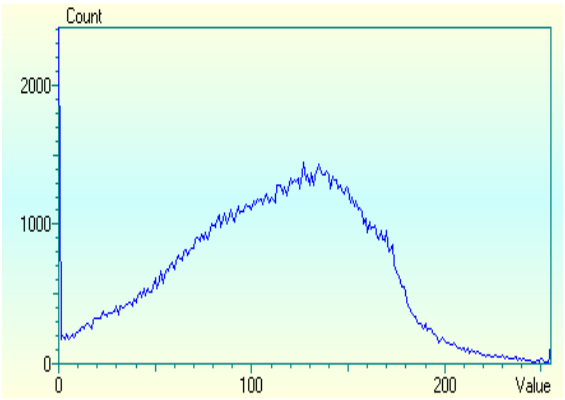


Fig. 7a: Original image (Mosul City) histogram of Blue channel

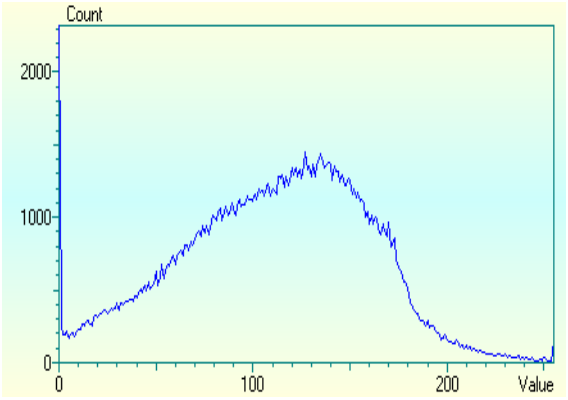


Fig. 7b: Modified image (Mosul City) histogram of Blue channel

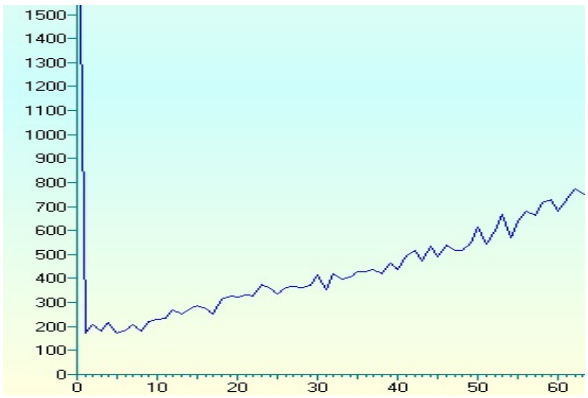


Fig. 8a: Histogram Zooming of Blue channel of original image (Mosul

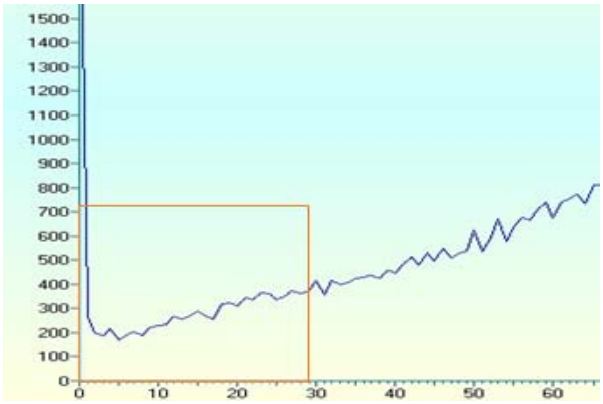


Fig. 8b: Histogram Zooming of Blue channel of Modified image (Mosul City)

By comparing the three RGB channels before and after hiding the data, the security aspect can be discussed. By observing the channels histograms before and after the modification, i.e. Figure 5a, Figure 5b, Figure 6a, Figure 6b, Figure 7a & Figure 7b, the change cannot be easily detected. However, if part of histogram is enlarged, some changes in the curve can be seen as shown in Figure 7a & Figure 7b. This creates some future work to be investigated including the reasons and implications of this issue. From many test runs, different distributions between the three channels were identified, which continued varying between the channels with no detected pattern. This undetectable pattern changing within RGB channels promise that the proposed technique may be considered random or pseudorandom based on the randomness of the index channel. Imperceptibility takes advantage of human psycho visual redundancy, which is very difficult to quantify for image steganography, existing metrics to measure imperceptibility include mean-square-error (MSE) and peak-signal-to-noise ratio (PSNR) [18] which is defined as:

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(f_{ij} - g_{ij})^2$$

$$PSNR = 10\log_{10}\frac{L^2}{MSE}$$

where
M. N are the row and column numbers of the cover image,
f_{ij} is the pixel value from the cover image,
g_{ij} is the pixel value from the stego-image, and

L is the peak signal level of the cover image (for &bit gray-scale images, L = 255).

Table 2 shows the values of PSNR and MSE with different sizes of images. Referring to Table 2, the column labeled SHDRIC is our proposed

Table 2: (PSNR and MSE) of four sample images

Image name	MSE	SHDRIC PSNR	Simple LSB PSNR
Image 1	26.919	33.830	31.760
Image 2	8.801	38.686	35.554
Image 3	1.241	47.192	43.959
Image 4	0.499	51.146	51.083

From the table, it was noted that the increase in the text caused an increase in the MSE and decrease in the PSNR. However, we can see the improvement in PSNR values than in the simple LSB. So, it becomes difficult to discover the hidden text within the image.

5. Conclusion

The suitability of steganography as a tool to conceal highly sensitive data has been discussed using a new method of randomizing the secret data. The method is based on two level of security where the data will be segmented into even and odd segments, before hiding the two segments separately and randomly inside image. This suggests that an image containing encrypted data can be transmitted anywhere across the world, in a complete secured form. This method can use in any other application such as image watermarking. It can be concluded that randomizing and hiding the secret data can provide a double layer of protection.

References

[1] J. Anitha and S. Immanuel Alex Pandian" A Color Image Digital Watermarking Scheme Based on SOFM" International Journal of Computer Science Issues, Vol 7, Issue 5, Sept 2010, Pages 302-309.
[2] Mayank Srivastava, Mohd. Qasim Rafiq and Rajesh Kumar Tiwari "A Robust and Secure Methodology for Network Communications" International Journal of Computer Science Issues, Vol. 7, Issue 5, September 2010, Pages 135-141
[3] Geeta S. Navale ; Swati S. Joshi ; Aarad_ hana A Deshmukh "M-Banking Security – a futuristic improved security approach", International Journal of Computer Science Issues, Vol 7, Issues 1, Jan 2010,Pag. 68-71
[4] I. Cox, M. Miller, and J. Bloom, Digital Watermarking, Academic Press, 2002.
[5] P. Tsai, etal. Reversible image hiding scheme using predictive coding and histogram shifting, Signal Processing, vol. 89. pp. 1129-1143, 2009.
[6] H. Sajedi, M. Jamzad, Cover Selection Steganography method Based on Similarity of Image Blocks, in Proc. of Int. IEEE 8th Conference on Computer and Information Technology, 2008.
[7] N.F. Johnson and S. Jajodia, Exploring Steganography: Seeing the Unseen, IEEE, pp. 26-34, 1998.
[8] R A Isbell, Steganography: Hidden Menace or Hidden Savior, Steganography White Paper, 10 May 2002.
[9] K. Solanki, N. Jacobsen, U. Madhow, B.S. Manjunath, and S. Chandrasekaran, Robust image-adaptive data hiding using erasure and error correction, IEEE Trans. Image Processing, vol. 13,pp.1627–1639, Dec. 2004.
[10] T. S. Chen, C. C. Chang, and M. S. Hwang, A virtual image cryptosystem based on vector quantization, IEEE Trans. on Image Process, 7 (1998) 1485.
[11] R. Z. Wang, C. F. Lin, and J. C. Lin, Image hiding by LSB substitution and genetic algorithm, Pattern Recogn., 34 (2001) 671.
[12] C. C. Chang, J. C. Chung, and Y. P. Lau, Hiding data in multitone images for data communications, IEE Proc. of Vision Image Signal Process, 151 (2004) 137.
[13] C. Y. Yang and J. C. Lin, Image hiding by base-oriented algorithm,Optical Eng-ineering, 45 (2006) Paper No. 117001
[14] Peter Wayner,Disappearing Cryptography –Information Hiding: Steganography & Watermarking–Second Edition. San Fransisco, California, U.S.A.: Elsevier Science, 2002, ISBN 1558607692.
[15] Neil F. Johnson and Sushil Jajodia, Exploring Steganography: Seeing the unseen IEEE transaction on Computer Practices. 1998.
[16] Ross Anderson, Roger Needham, Adi Shamir, The Steganographic File System, 2nd Information Hiding Workshop, 1998.

[17] Dung Dang,Wenbin Luo" Color image noise removal algorithm utilizing hybrid vector filtering" AEU-International Journal of Electronics and Communications, Vol 62, Issue 1, 2 Jan 2008, Pages 63-67
[18] Qi, Hairong; Snyder, Wesley E. & Sander, William A., 2002; Blind Consistency-Based Steganography for Information Hiding in Digital Media. Multimedia and Expo, 2002. ICME '02.Proceedings. 2002 IEEE Inter-national Conference on Vol. 1, p.: 585- 588.

Emad T. Khalaf
Graduated in Computer Information Systems and Informatics Engineering and he worked as a Technical in Internet Services Company for more than nine years. He had experience as a trainer for various computer courses. His research interests include network technology and security. He is currently studying MSc degree in the area of computer networks security.

Norrozila Sulaiman

Graduated from Sheffield Hallam University with a BSc (Hons) in Computer Studies in 1994.She worked with Employment Service in UK as a network support assistant and she involved on a research on Novell Netware. After graduated, she worked as a research officer at Artificial Intelligence System and Development Laboratory and involved in joint collaboration projects between the government of Malaysia and Japan for about 5 years. She completed her MSc degree in Information Technology and involved in a research on Wireless Application Protocol (WAP). She obtained her PhD degree in mobile communication and networks from Newcastle University in UK. Currently, she is a senior lecturer at Faculty of Computer System and Software Engineering, University Malaysia Pahang. Her main research interests include heterogeneous networks, mobile communication networks and information security.