International Journal of Electrical and Computer Engineering (IJECE)
Vol. 9, No. 1, February 2019, pp. 460~467
ISSN: 2088-8708, DOI: 10.11591/ijece.v9i1.pp460-467

A secure image steganography based on burrows wheeler transform and dynamic bit embedding

Ahmed Toman Thahab
Department of Electrical and Electronic Engineering, University of Kerbala, Iraq

Article Info

Article history:

Received Jan 11, 2018
Revised Oct 10, 2018
Accepted Oct 27, 2018

Keywords:

Burrows wheeler transform
dynamic embedding
Data security
Payload
XOR

ABSTRACT

In modern public communication networks, digital data is massively transmitted through the internet with a high risk of data piracy. Steganography is a technique used to transmit data without arousing suspicion of secret data existence. In this paper, a color image steganography technique is proposed in spatial domain. The cover image is segmented into non-overlapping blocks which are scattered among image size window using Burrows Wheeler transform before embedding. Secret data is embedded in each block according to its sequence in the Burrows Wheeler transform output. The hiding method is an operation of an exclusive-or between a virtual bit which is generated from the most significant bit and the least significant bits of the cover pixel. Results of the algorithm are analyzed according to its degradation of the output image and embedding capacity. The results are also compared with other existing methods.

Corresponding Author:

Ahmed T.Thahab,
Department of Electrical and Electronic Engineering,
University of Kerbala.
104 Al-Mwathafeen Road, Kerbala Province, Iraq.
Email: toeahmed@gmail.com

1. INTRODUCTION

A technique that conceals secret information inside another variety of information data is termed as steganography [1]. It is broadly used nowadays in information security strategies over the last decade since this kind of invisible communication prevents imperrmissible access to secret information [1]. Although other data security strategies such as cryptography and encryption techniques exist for secure data transmission, steganography uses files such as audio, image, video formats to transmit data oblivious to the fact that data exits in the message [1]. Watermarking is also considered as a hiding technique, but it is used to protect copyright of content such as small tags or company logos. Steganography conceals greater payload in cover media [2].

e cover file where information is hidden; the secret file and the stego-file which is considered as the output resultant from the system. When embedding secret message data in a cover media, it ought to be visually undetectable and the output stego-file should not possess any noticeable adaptation compared to the original cover image [3]. Imperceptibility is crucial factor when it comes to steganography techniques; it reflects the level of correlation between the stego and the cover file. Moreover; other factors such as security and capacity are also considered decisive factors for steganography algorithm.

Steganography has been under the scope of researcher for some time since; it is a vital field for covert communication. Secret information can be embedded in the spatial domain specifically in the pixel intensity value [4]. A pioneering embedding method used in steganography is the least significant bit technique (LSB). This technique utilizes the least significant bits of a code-word and replaces them with the most significant bits of the secret binary data. In [5], a new LSB technique is proposed using differencing.

The method takes the difference between bit No.5 and bit No.6 to hide secret message. If the result of difference is not equal to the secret information bit the bit no. 5 is transverse. Results of this method show a peak signal to noise ratio (PSNR) of 51.1803 with payload of 262144 bits. The main disadvantage of method in [5] and generally LSB substitution is the techniques' tolerance in terms of security and capacity [6], [7], therefore; it has been combined with other techniques such as encryption and source coding etc.

Authors in [8] encoded the secret data using Huffman encoding method and embed the encoded secret data in the cover image using LSB method. Although the technique achieved a PSNR of +57.43dB, the embedding capacity of the technique used is 25% of the cover image size. Huffman encoding requires a Huffman table to be known at the receiver side to decode the secret data. Processing time is also a critical factor in the performance of this algorithm. Other authors used encryption to decrease the chances of data security compromise. In [9], an improved version of least significant bit (LSB) method is proposed using encryption. The secret data embedding positions are encrypted using a secret encryption key. The maximum stego file quality (PSNR) is 53.7869dB without mentioning the embedding capacity of the algorithm.

For LSB to have more security features, authors in [10] utilized a Ron Code (RC4) to randomize the concealment of secret data bits in the cover image instead of LSB conventional sequent embedding manner. The complexity of the algorithm and its time processing consumption are disadvantageous to the proposed method.

Signal transform tools such as discrete wavelet transform (DWT) is utilized in steganography to increase embedding capacity without degenerating visual quality. Lately, numerous researches have been founded on discrete wavelet transform (DWT), since the transform provides a vital frequency band localization suitable for data embedding. In [11], authors proposed a steganography method using encryption technique and DWT. The DWT is applied on the cover image; Redundancy in the cover image is determined using threshold calculation based on the image statistics. The secret message is partitioned and encrypted using RC4 algorithm. A DWT is also applied to the encrypted secret message. The embedding process implies a simple replacement of DWT coefficients of the encrypted secret message in the previously specified DWT of the cover image. Although this method is secure, it requires hard ware resources such as processor speed and memory. Although frequency domain embedding is considered a secure embedding domain, the extracted secret data will suffer some degradation from the forward and reverse transform operation.

LSB spatial embeding is considered as fragile hiding technique in terms of security, therefore it requires a combination with encryption techniques as in [10] but, these techniques require additional processing and computational expensive. Manipulating the embedding procedure will provide a massive security feature to the steganography algorithm. In this paper a new embedding procedure is developed for hiding secret data inside a cover media utilizing the Burrows Wheeler transform (BWT) for manipulating the sequence of data embedding inside the cover media according to the output of BWT. The embedding procedure is based on an exclusive–or (XOR) operation between a virtual bit and the LSB of the cover pixel coefficient.

The rest of the paper organized as follows: Section two explains the conventional LSB method. Section three presents the steps of Burrows wheeler transform. Section four illustrates the embedding and extraction method Section five gives the results of some tests for the proposed steganography algorithm and comparison with other algorithms. Section six states the conclusions and future developments.

2. CONVENTIONAL LEAST SIGNIFICANT BIT

The least significant bit technique is considered as a simple method used in the steganography field. If a secret message is considered as a binary bit stream it can be embedded into a cover media, (i.e.….Audio, image, and video) by replacing the LSB bits of the cover image with the secret bits [12]. The number of replaced bits depends on the method that is utilized in the algorithm. Suppose an image of an 8 bit grayscale have the following two pixel binary values p1=[10111000] and p2=[10110011] and the secret message that is required to be embedded is S=[10] in the pixels, the stego output pixels are S1=[10111010] and S2=[10110010]. The error between [p1, p2] and the [S1, S2] is imperceptible to the human visual system (HVS). Increasing embedding capacity requires embedding more bits into the right side of the cover stream pixels. However, increasing the embedding capacity results in a quality degrade of stego image. Therefore; a trade-off between embedding payload and quality of stego image ought to be considerable. Although this method is considered simple and fast, its main drawbacks are limited capacity and vulnerable security.

3. BURROWS WHEELER TRANSFORM

Lossless data compression is a technique used to compress data by manipulating data in a form that can be reversed to retrieve its original value without losing any vital data [13]. The Burrows wheeler transform (BWT) is a technique based on block sorting which blocks of data are manipulated to be transformed to an easier form for processing. Since the technique is used in many practical applications especially in lossless compression techniques, there have been many extensions of the original BWT [14]. Data can be transformed in the BWT domain and vice versa.

3.1. The forward burrows wheeler transform

The BWT is reversible transformation, therefore; original data is reconstructed to its original form. In order to transform the data using BWT, the output of the transformation is limited to two factors; the output (B), and the index (inx) where the location of the original input is located [13], [14]. Consider a one dimension pixel sequence P=[3 2 5 3 1]. The sequence (p) is copied to a two column table (Index and sequence) as shown in Table 1. The index (inx) exceeds (0……n-1) where n: is number of values in pixel. The sequence is then left-cyclic permutated into each consequent index row. The rows are then sorted lexicographically as in Table 2.

Table 1. Shows the Left Cyclic Permutation

Inx No.	Left Cyclic Permutation Sequences				
0	3	2	5	3	1
1	2	5	3	1	3
2	5	3	1	3	2
3	3	1	3	2	5
4	1	3	2	5	3

Table 2. Shows the Lexicographically Sorted Rows

Inx No.	Left Cyclic Permutation Sequences				
0	1	3	2	5	3
1	2	5	3	1	3
2	3	1	3	2	5
3	3	2	5	3	1
4	5	3	1	3	2

The search for the input one dimension row is executed in Table 2 and which the original pixel is indexed. In this example the original row is indexed at 3. The BWT (B) for the input pixel p=[3 2 5 3 1] is the last column B=[3 3 5 1 2] and inx=3. The proposed algorithm utilizes the transform to randomize the blocks which will be embedded. Since the lexicographical sorting depends on the input vector to the transform, unique outputs will be produced. This technique will blossom the security grade of the method.

4. PROPOSED IMAGE STEGANOGRAPHY

In this section, we describe a novel color image steganography. As mentioned in previous section, the main drawbacks of the LSB technique are security compromise and limited embedding capacity. In order to overcome these drawbacks, a new image steganography approach is proposed using generated virtual bit. The method increases the embedding capacity of the conventional LSB without introducing noticeable distortion to the stego-image. The method also possesses security features protecting secret data inside the cover image using Burrows wheeler transform to enhance the embedded data security. Figure 2 shows the block diagram of the method.

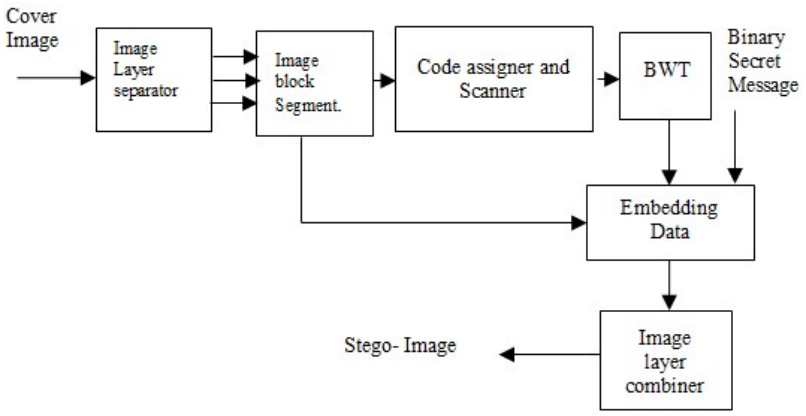


Figure 2. Block diagram of the proposed hiding algorithm

A RGB image is input to the algorithm, and separated into three layers R, G, and B. Layers of the color image are divided into k*k non-overlapping blocks. Utilizing the Burrows wheeler permits the assign of alphabat codes or decimal numbers to the blocks. In the proposed work, each block is assigned a one decimal code. For example using a 64*64 block for 128*128 image will have a four (64*64) blocks per layer, therefore; four decimal codes exist. The decimal numbers representing blocks are converted to a one dimension vector using either Hilbert scanning, raster scanning, zig-zag scanning….etc. In this paper, the Hillbert scanning path is used since this type of scanner scans every block of any square image size. The one dimension vector is input to the Burrows wheeler transform for randomizing the block's sequences to be embedded with secret data, this will increase the security of the secret data since intruders are unaware of the sequence of the blocks, size of blocks, type of the scanner used and type of block asigning. This prior embedding process is considered as first line of defense for the embedded secret data. Blocks of one layer are input to BWT and other two layers are embedded due to the same sequence of the first layer. This will reduce time and complex processing required by BWT; in addition the algorithm does not require the inverse BWT for extracting the data.

4.1. Hiding secret data

The method hides secret data in the spatial domain of a color image without utilizing any signal transform. The embedding method assumes a virtual bit at the least significant bits of each cover image pixel which is generated from a two MSB bit according to (1):

$$(\mathrm{Vir})_{\mathrm{i}} = \mathrm{MSB}_{8\text{-}\mathrm{i}+1} \oplus \mathrm{MSB}_{8\text{-}\mathrm{i}+2} \quad (1)$$

where : *Vir*: virtual bit generated.
MSB: the i^{th} most significant bit of cover bit & *i=8, 7…n*.

The XOR between the LSB bit and the virtual bit according to (2) and the result (Res) is compared to the secret data bit. If the result is equal to the data bit, the LSB bit is unchanged. Otherwise change the LSB bit to the difference bit. In this operation one bit is embedded in the ith LSB bit.

$$(Res)_i = LSB_i \oplus Vir_i \quad (2)$$

Where*: LSB:* the i^{th} least significant bit of cover bit.

Embedding operation is continued for n bits. If n=8, then only one bit is embedded in the LSB of the cover pixel. If n=7, then two bits are embedded in the cover pixel etc. For an illustration of the procedure for the embedding process. Table 3. shows the embedding procedure for the secret data in the sixth bit of the cover data. The table assumes that the virtual bit is 0. Table 4. shows the illustration for the same embedding process of Table 3. for virtual bit of 1. It is clear from Table 3. and Table 4. that each embedded codeword yields to a unique code word, therefore; a complete extraction of secret data is applicable.

Table 3. Shows the Embedding Scheme for Virtual Bit 0

	Embedding Bit 0		
No.	Cover bits B5 B6 B7 B8	Sec. data	Stego bits
1	0000	0	0000
		1	0100
2	0001	0	0001
		1	0101
3	0010	0	0010
		1	0110
4	0011	0	0011
		1	0111
5	0100	0	0000
		1	0100
6	0101	0	0001
		1	0101
7	0110	0	0010
		1	0110
8	0111	0	0011
		1	0111

Table 4. Shows the Embedding Scheme for Virtual Bit 1

	Embedding Bit 0		
No.	Cover bits B5 B6 B7 B8	Sec. data	Stego bits
1	0000	0	0100
		1	0000
2	0001	0	0101
		1	0001
3	0010	0	0110
		1	0010
4	0011	0	0111
		1	0011
5	0100	0	0100
		1	0000
6	0101	0	0101
		1	0001
7	0110	0	0110
		1	0010
8	0111	0	0111
		1	0011

4.2. Extracting secret data

Most of the steganography algorithms require a lot of keys in order to extract the secret data from the stego image. These keys are either embedded or sent as header information. In addition to keys, other algorithms require knowledge of the cover image to reconstructed secret data. The proposed embedding algorithm does not require the cover image to extract secret data making the algorithm as a blind steganography algorithm and keys required are block sizes of the segmentation (k) of image and the number of bits embedded per pixel (n). Figure 3 shows the secret extraction block diagram.

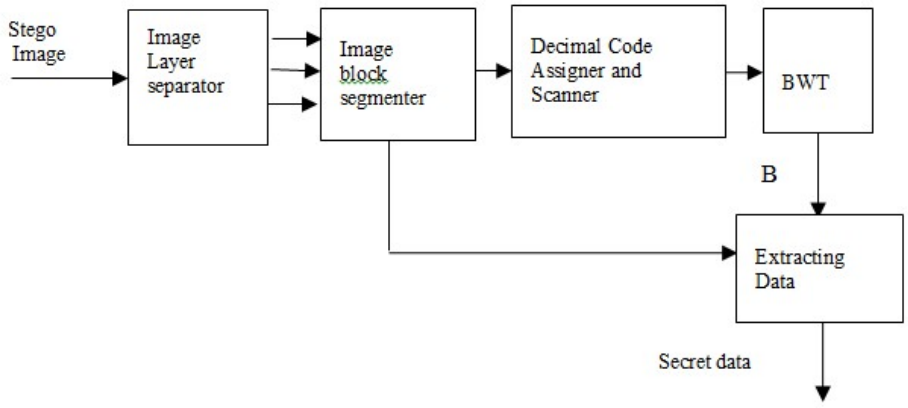


Figure 3. Block diagram secret data extraction

The stego image is separated into R, G and B layers. Each layer is divided into (k*k) blocks and have the same assignment and type of scanner. The burrows wheeler transform is applied to find the embedding sequence of the blocks. The extraction of the ith bit from the stego pixel is according to (3).

$$(\overline{Res})_i = Steg_i \oplus Vir_i \quad (3)$$

Steg: the i^{th} least significant bit of steg bit.
If the $\overline{\text{Res}}$ is equal to the i^{th} bit of the stego, the secret bit is $\overline{Res}$.
Otherwise: secret bit is the inversion of the i^{th} bit.
The blindness proposed method can extract the secret data without creating errors.

5. RESULT AND DISCUSSION

The basic performance of any stenographic algorithm is the stego-image quality and the embedding capacity that the algorithm provides to carry secret data. The stego quality is gauged using the image metric peak signal to noise ratio (PSNR) which is given by (4) [15] & (5) [15] for a square image. Some authors determine the PSNR for each layer of the color image; therefore, the quality metric (PSNR) will be averaged over the three layers:

$$PSNR = 10\, log_{10}\left(\frac{(255)^2}{MSE}\right) \quad (4)$$

$$MSE = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\widehat{im} - im)^2 \quad (5)$$

where: $\widehat{\text{im}}$ and im are the stego and cover pixels respectively.
N: is the number of pixels in row or column of a square image.

A superior image steganography algorithm performance is remarked when the algorithm achieves a high PSNR and embedding capacity. Moreover, the vital issue on message recovery side is the errorless message retrieval from the cover media. The proposed algorithm is assessed with four images with various number of embedding bits. The size of the image is 256*256*3 and 512*512*3 color images. Figure 4 shows the original images of Baboon and Lena. Other image metric such as normalized cross correlation (NCC) is

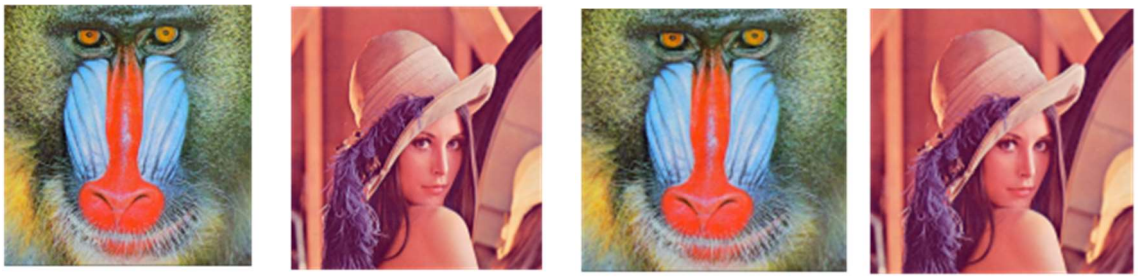
utilized to illustrate the high correlation presented between the stego-image. The calculation of NCC for a square image is given in (6).

$$NCC = \frac{1}{N^2}\sum_{I=1}^{N}\sum_{J=1}^{N}\left[\hat{im} - im\right]^2 \quad (6)$$

As stated in section 4, the image is segmented into k*k blocks since the input images are a size of 512*512 then k=64, therefore; there will be 64 blocks of 64*64. All blocks will be assigned decimal codes and converted to a one dimension block using Hilbert scanning. The one dimension code vector is randomized according to Burrows wheeler transform. The algorithm will embed one bit (n=8) and two bits (n=7) in the cover image, Table 5 illustrates the PSNR for various bit embedding payloads. The secret data stream is generated by a psedu-random generator.

Table 5. Stego Image Quality versus Payload Covers Images

Image Name and size	Number of bits	Payload in bits	PSNR in dB	NCC
Babbon 512*512*3	n=8	786432	51.1412	0.999
	n=7	1572864	44.1509	0.989
Lena 512*512*3	n=8	786432	51.1462	0.999
	n=7	1572864	44.1400	0.988
Viptraffic 128*128*3	n=8	491520	51.1120	0.999
	n=7	983040	44.1383	0.979
Tajmahal 256*256*3	n=8	196608	51.1418	0.999
	n=7	393216	44.1490	0.989
Kola 256*256*3	n=8	196608	51.1487	0.999
	n=7	393216	44.1322	0.979



a. original Baboon and lena b. Stego image Baboon and lena

Figure 4. Shows the original and stego color 512*512 Baboon and Lena images

With many experiments made on other various images the range of PSNR for present technique is 51.5-51.12 dB. When three bits are embedded in a cover pixel, the PSNR is in the range of 40-35dB with payload 2359296 bits. Embedding more bits in the cover image will degrade the quality of the image with respect to the original image. Stego images shown in Figure 4(b) are highly correlated to the original images compared with the originals. Therefore, the algorithm preserves image quality for high payloads and is highly secure for data embedding. The proposed algorithm compared with other proposed algorithms in [15] [16] is less complex, in addition to the required processing time. However, the method presents better PSNR and high capacity embedding. Table 6 and Table 7 show the comparison between these techniques respectively.

Table 6. Comparison of Stego Image Quality and Payload for 256*256 Image

Image	Reference [15]		Proposed Method n=8	
	APSNR in dB	Payload in bits	PSNR in dB	Payload in bits
Lena	50.6208	147456	51.1353	196608
Baboon	50.6167	147456	51.1412	196608
Tagmahal	50.5972	147456	51.1418	196608
Kola	50.6279	147456	51.1487	196608

Table 7. Comparison of Stego Image Quality and Payload for 512*512 Image

Image	Reference [16]		Proposed Method n=8	
	PSNR in dB	Payload in bits	PSNR in dB	Payloadin bits
Lena	44.0886	786436	51.1462	786432
Baboon	44.0307	785110	51.1412	786432
Sailboat	43.9728	786832	51.1372	786432
Paper	44.0092	785392	51.1419	786432

6. CONCLUSION AND FUTURE WORK

The essential objective of a steganography algorithm is to address the security of information which is achieved by concealing secret communication data inside the cover media. It is crucial for the sender to conceal high amount of secret information inside a cover media and send the information over a public communication channel. However, intruders can pirate the information inside the cover media unless the stego image is highly correlated to the original media. A novel image steganography method is proposed in this paper utilizing the XOR function with generated virtual bits and a pre-embedding method based on Burrows wheeler transform. The proposed steganography method highly complies with steganography objectives. Embedding the secret data using the proposed embedding strategy produces a stego image highly correlated to the original cover, since the strategy saves the deference of the result of the XOR function. The virtual bit generated from the MSB binary cover pixel is neither affected by the noise added on the stego-image nor other compression process, therefore; the method permits a complete reconstruction of the secret data. Burrows Wheeler transform sorts either decimal number coded blocks or alphabat chericteristecs codes that are sorted lexicographically. Therefore; the assigned coded blocks of the cover image is redistributed in a random manner before embedding. The procedure enhances security of embedding data. The proposed algorithm can be enhanced using the output of digital scramblers for assigning the codes of cover image blocks since the output of the digital scramblers alters with the number of shift registers, the initial digital bit value memorized inside the shift registers and their arrangments. Nevertheless, the algorithm can be utilized in hiding images inside a cover video.

REFERENCES

[1] G. S. Swain et al, "A Novel Approach to RGB Channel Based Image Steganography Technique," *Int. Arab J. e-Technology,* vol. 2, pp. 181–186, 2012.

[2] M.Hussain et al, "Recursive Information Hiding Scheme through LSB, PVD Shifit and MPE", *IETE Technical Review,* vol. 35, pp.53-63, 2017.

[3] Kumar R et al, "A New Image Steganography Technique Based on Similarity in Secret Message*", The Next Generation Information Technology Summit (4th International Conference),* pp. 376-37, 2013.

[4] M. S Arya et al, "Improved Capacity Image Steganography Algorithm using 16- Pixel Differencing with n-bit LSB Substitution for RGB Images:, *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 6, 2016, pp 2735-2741.

[5] Ammad UI Islam et al, "Improved Image Steganography Technique based on MSB using Bit Differencing*" 6th International conference on Innovative Computing Technology (INTECH),* pp. 265-269, 2016.

[6] Weiqi Luo et al, " Edge Adaptive Image Steganography Based on LSB Matching Revisited*", IEEE Transactions on Information Forensics and Security,* pp. 201-214, 2010.

[7] Liu J et al, "Least-Significant-Digit Steganography in Low Bitrate Speech", *IEEE International Conference on Communications (ICC),* pp. 1133–1137, 2012.

[8] Das. R et al, "A Novel Steganography Method for Image Based on Huffman Encoding", *Proceedings of 3rd National Conference on Emerging Trends and Applications in Computer Science, (NCETACS),* pp.14- 18, 2012.

[9] Masud Karim S. M et al, "A new approach for LSB based image steganography using secret Key*", 14th International Conference on Computer Information Technology (ICCIT) IEEE; 2011* p. 286–291, 2011.

[10] Akhtar N et al, "Enhancing the security and Quality of LSB based Image Steganography. *5th International Conference on Computational Intelligence and Communication Networks, IEEE*; pp. 385–90, 2013.

[11] A. Al-ataby et al, "Modified High Capacity Image Steganography Technique Based on Wavelet Transform*", International Arab Journal Inf. Technol.,* pp. 358–64, 2010.

[12] C. K. Chan et al, "Hiding data in image by simple LSB substitution", *Pattern Recognition Journal, Vol. 37,* pp. 469-474, 2004.

[13] M. Burrows et al, "A Block-sorting Lossless Data Compression Algorithm", *Research Report 124, Digital System Research Center,* 1994.

[14] Ziya. Arnavut and Spyros S. Megluveras, "Block Sorting and Compression. Data Compression Conference", pp.181-190, 1997.

[15] Thanikaiselvan. V et al, "High Security Image Steganograohy Using IWT and Graph Theory", *IEEE International Conference on Signal and Image Processing Applications (ICSIPA),* pp. 337-342, 2013.

[16] Bhowmik. S et al, "A New Approach in Color Image Steganography with high level of Perceptibility and Security", *International Conference on Intelligent Control Power and Instrumentation (ICICPI),* pp. 283-286, 2016.

BIOGRAPHY OF AUTHOR

Mr. Ahmed Toman Thahab received his B.Sc in Electrical Enginnering from the University of Babylon in 2006. He was enrolled in the Graduate school of the same universty. He received his M.sc degree in the Communication and Electronic Engineering in 2011. His main study was in video compression based on wavelet Transform.