A new look depending on the cover image data for hiding and retrieval the secret information by using LSB technique

Atheer Hussein Zyara
Assistant Lecturer, Al-forat Al-Awsat technical University , College of Health and Medical Technique/Kufa ,Iraq

Abstract— the hiding operation of important ways to keep the Confidential data for Security institutions and the private property of individuals When transported through the internal network or Internet networks,and the objective of the hiding operation It is to remove any doubts about the existence of hidden data reverse the encryption process which raises doubts , The most important characteristic of concealment techniques that they keep pace with new technologies With the availability of different types of center-carrier or the so-called cover .In this research was to improve the hiding way Depending on the LSB Technique,and that of through hiding then jump , including a bits in a specific location of the picture And move to the next site based on the value of previous site After addition . In this way we are sending the picture with two keys They are the start and finish site.where the site value be represented by the total of ones number (1) or the total of zeros number (0)Per byte plus the Driving force (Impetus),where the impetus value be selected by user.the Features this method , Flexibility where applied to all kinds of images , Ease of hiding and retrieval, and the Durability from where of the large number of possibilities that be from the impossibility to Predict the existence of a secret message , This is what distinguishes this proposed algorithm, especially if conducted some encryption technologies on confidential data ,In order to clarify the idea of the proposed algorithm is applied to the image of the grayscaletype , the English text , and Arabic texts, The use of Efficiency standards (PSNR) and (MSE) For the purpose of measuring the level of efficiency of the proposed algorithm , Matlab language was used for the application of the proposed algorithm .

Index Terms— proposed algorithm, Steganography , least significant bit , data hiding , data retrieving , Impetus , start site , finish site.

1 INTRODUCTION

Since ancient times and there is still the efforts to find new ways and means to maintain the confidentiality of data transmitted and especially during the wars. Where many different methods have emerged in the field of maintaining the security of data , With the rapid development taking place in the field of computers and electrons and communication networks and the Internet and component remittances and inflected and electronic signature and etc. , So start the search for ways to protect the data that has led to the emergence of cryptographyand Which it was a good and convenient way but it again and with the rapid development in the field of computer become from easy detection and change the content or destructionfrom By intruders This led to search for other ways of concealment appeared aware that its principle that the message sent be hidden inside the center and another is striking to look like the pictures and texts and audio and video , and Many of the techniques of concealment is designedand The researches is still underway to get the best techniques to hide these techniques can consider them as systems of replacementthe concealment process is that we replace some of the cover data for the secret data that We are trying to send it in confidential . [1][2]

It can also be a combination of concealment techniques and encryption makes security of data sent more durability , If doubt the existence of a hidden data file inside the cover They will be sorted irregularly and is suspicious.

2 Theoretical Background

There are three types of steganography techniques, namely, as follows:

2.1 Hiding of information without a secret key (Pure Steganography)

In this way there is no secret key between the sender and the recipient, and be included the confidential information as follows :

E : M*C >> S

M : Represent a secret message the requested to send .

C : It represents a cover that will carry inside it the secret message .

While the extraction of confidential information as follows

D : S >> M

2.2 Hiding of information by a secret key(Secret Key Steganography)

In this way there is a mutual secret key between the sender and the receiver.Where the secret key is used to hide secret information inside the cover , And to extract confidential information when the receiver The process is reversed with the use of the secret key. And as follows

Hide secret data

E : M*C*K >> S

Extract secret data

D : S*K >>M

2.3 Hiding of information by public key (Public Key Steganography)

In this way there are two keys. The public key is used in hid-

ing the secret message and the private key is used in the process of extracting the secret message.

Either in this research was to use a second method , the hiding by the secret key.

3 HIDING BY DIGITAL IMAGES

There are many files that can be used in the process of hiding confidential data and from These data digital images And computer deals with digital images as a two-dimensional system . And each site where a point or what is known as (Pixel) It is the smallest unit that represents a specific location on the screen . [3]
The digital images are divided according to the composition of its colors into four types, And as follows

3.1 Binary image

The simplest types of images, take only two values,(0) means black color ,(1) means white color, It can be expressed for binary images worth of one bit per component 1Bit/Pixel .

3.2 Grayscale image

It is Monochrome image ,contains lighting information only ,Each point contains 8 Bit / pixel , They allow of the 256 of lighting levels from (0 is black) for (255 white) .

3.3 Color image

It consists of three bands (3 bands) , Each band represents a single byte,So we can say that each element in the color image is represented by (3 bytes),And this gives the reason for the large size for color images compared with the previous images .[4]

3.4 Image multiple spectrum

This image contains more than three bands(3 bands) , may up to seven bands (7 bands).
These digital images can be used as a cover for secret data Which is stored in the form of a binary number is added to the least significant bit (LSB)And when changing the least significant bit(LSB) does not affect the picture significantly.[6]

4 MODEL COVERAGE SYSTEM

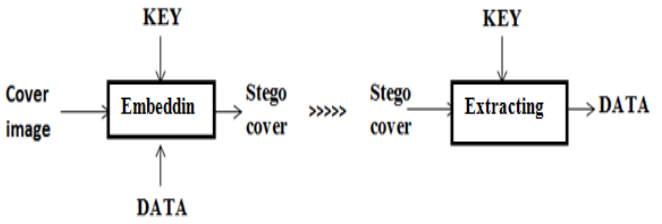

Fig. 1. Block diagram of data hiding and retrieval.

Coverage system model consists Of the following components

1. Midfield carrier or what is known as cover C(Cover image)
2. Secret data D (Data)
3. Embedded cover S(Srego-cover)
4. Secret Key K(Key)
5. the embedding process EM (Embedding)
6. the extracting process EX (Extracting)

5 THE PROPOSED METHOD

The proposed method is based on hiding then jump and Depending on the LSB Technology, Where one bit is hiding at a particular location of the image (pixel) (That address be the first secret key) Then find the sum of the number of Ones plus the Impetus or find the sum of the number of Zeros plus the Impetus (The impetus is a positive integer greater than zero where It is very necessary and be agreed between the sender and the receiver or sends as secret key.There are two reasons for the use of the first driving force in some cases be combined sum of Ones or zeros per byte is zero and this means that there is no transmission, and the second reason is to increase the strength of the algorithm in terms of the large number of possibilities) Then jump to the site the other depending on the output sum of the previous location, and so onConcealment process continues and find the output sum then jump to another location until last bit from secret data (The site address of last bit will be the second secret key),and thus send two keys , the start location address(SL), the last location address(LL) , and Embedded cover (stego_cover) . and from Through secret keys and embedded cover will can the receiver to retrieval the secret data .

6 THE RULES OF COVER AREA AND COLLECT BITS

After converting the image from decimal to binary(dec2bin) , we will deal with the following components .
C={c1,….,ci}
Where C is the cover Image
i Represents the location address,Where i ≥ 1
ci Represents the location content , where ci is 8 bits per one byte .
sumones(ci)= + impetus ……..(1)
sumzeros(ci) =) + impetus ……...(2)
k is the number of bits that will sum , where $0 < k \leq$ number of bits per one locationAnd it depends on what type of image (binary image , grayscale image,color image, Image of multiple spectrum).
j is number of bit that will begin the summation from it , where $0 < j \leq k$.
c(j) is the Bit value , where c(j) =1 | 0 .
impetus is The value of the jump, which is added to the sumones or sumzeros value , where impetus > 0 .

TABLE 1
EXPLAIN THE METHOD TO COLLECT THE ONES PLUS THE IMPETUS OR THE ZEROS PLUS THE IMPETUS

	1100 1100	**1111 1111**	**0000 0000**	**1010 1100**	**0011 0000**
Ones	4	8	0	4	2
Ones+1(impetus=1)	5	9	1	5	3
Ones+4 (impetus=4)	8	12	4	8	6
Zeros	4	0	8	4	6
Zeros + 1 (impetus=1)	5	1	9	5	7
Zeros +4(impetus=4)	8	4	12	8	10

Algorithm (1) Embedding Process

Input: - IMAGE gray scale (IMG) , Message (S),Start location (SL), Impetus(P)
Output :- Stego IMAGE , Start location (SL) , Last location (LL)
Step 1 :-
1) Convert the IMAGE (IMG) intobinary data and put the result in (BIMG).
2) Convert the Massage (S) intobinary data and put the result in (BS).
3) Calculate the size of BIMG (Row, Column) and put the result in (Rm , CM).Where CM depends on what type of image (binary image , grayscale image,color image, Image of multiple spectrum).
4) Calculate the Length of BS and put the result in (LenS).
Step2 :-
At least If LenS*P> Rm
Then return the cover is small to hidingthe message
stop .
Step3 :-
1) i= SL, embed first bit from BS in ci (0 <i ≤ Rm) .

2) find sumones(ci)or sumzeros(ci) , $\sum_i^k ci(j)$, and put the resilt in (suml) , where k ≤ CM .
3) find sum P and suml,and put the result in (Tsum) , (where P > 0)
4) transfer to another location depending on the (Tsum) of location previos
5) continue to last bit of the (BS)
6) put last location address in LL = i (where last ci is The location last address , now become LL is second address).
Step4:-end.

Algorithm (2) Extraction process

Input :- stego IMAGE , start location(SL) , last location(LL), Impetus(P)
Output :- hidden text(HT)
Step1:-
1) Convert stegoIMAGE (StIMG) intobinary data and put the result in (StBIMG).
2) Calculate the size of StBIMG (Row Column) and put the result in (Rm CM).
Step2:-
1) i =SL
2) HT(a)=ci(b) , where a =1, and b ≤ CM .
find sumones(ci) or sumzeros(ci) , $\sum_i^k ci(j)$, and put the resilt in (stegsum) ,where k ≤ CM .
3) find sum P and (stegsum) , and put the result in (Tstegsum) , (where P > 0)
4) transfer to another location depending on the (Tstegsum) of location previos
5) continue until i ==LL (last location)
Step3:- end

Process of hiding and retrieval

The imposition of that data to be hidden is a text message , and the cover is a gray scale image .After converting the text message to the binary system allows you to hide 1 bit per pixel,Where the hiding it be based on the LSB technique.

Suppose that we have the following binary data (1101100110) , We have the following image and we assume it numbered according to the following locations .

1	1111 0001	2	1100 0001	3	1100 0011	4	0001 1100	5	1111 0000
6	1111 0000	7	1100 0011	8	1111 1001	9	1110 0001	10	0000 0000
11	0000 0000	12	0000 1110	13	1000 0011	14	1111 1000	15	1111 1111
16	1111 1111	17	0101 0011	18	0001 1100	19	0000 0011	20	0000 0111
21	0000 0111	22	1110 0011	23	1110 0001	24	1111 1000	25	1100 0001
26	0000 0111	27	1010 1011	28	1111 1000	29	1110 0011	30	1100 0011
31	1100 0001	32	0101 0011	33	0000 0011	34	1010 1010	35	0000 1110
36	1100 0011	37	0011 0011	38	1111 1000	39	0000 0111	40	0011 0011
41	0000 1110	42	0000 0011	43	1110 0011	44	1100 0001	45	0000 0011
46	0101 0011	47	0000 0001	48	1010 1010	49	1100 0011	50	0000 0001
51	1110 0011	52	0001 1100	53	0101 0011	54	0000 1110	55	0001 1100
56	1010 1010	57	1100 0011	58	0011 0011	59	0101 0011	60	0011 0011
61	1111 0000	62	0000 1110	63	0000 0011	64	1110 0011	65	0000 0011
66	0000 0000	67	0011 0011	68	0000 0001	69	1010 1010	70	1111 1001
71	1111 1111	72	0000 0011	73	0011 1001	74	0101 0011	75	1100 0001
76	0000 0111	77	1111 1001	78	0001 1101	79	0011 0011	80	1100 0011

EXAMPLE (1) :

We suppose that we will jump between locations depending on account The total number of the Ones , Impetus(P) = 1 , Start location(SL) = 1 , will be the addition to the first bit from the right , Therefore the secret message be hidden in the following locations .

1	1111 0001	2	1100 0001	3	1100 0011	4	0001 1100	5	1111 0000
6	1111 0000	7	1100 0011	8	1111 1001	9	1110 0001	10	0000 0000
11	0000 0000	12	0000 1110	13	1000 0011	14	1111 1000	15	1111 1111
16	1111 1111	17	0101 0011	18	0001 1100	19	0000 0011	20	0000 0111
21	0000 0111	22	1110 0011	23	1110 0001	24	1111 1000	25	1100 0001
26	0000 0111	27	1010 1011	28	1111 1000	29	1110 0011	30	1100 0011
31	1100 0001	32	0101 0011	33	0000 0011	34	1010 1010	35	0000 1110
36	1100 0011	37	0011 0011	38	1111 1000	39	0000 0111	40	0011 0011
41	0000 1110	42	0000 0011	43	1110 0011	44	1100 0001	45	0000 0011
46	0101 0011	47	0000 0001	48	1010 1010	49	1100 0011	50	0000 0001
51	1110 0011	52	0001 1100	53	0101 0011	54	0000 1110	55	0001 1100
56	1010 1010	57	1100 0011	58	0011 0011	59	0101 0011	60	0011 0011
61	1111 0000	62	0000 1110	63	0000 0011	64	1110 0011	65	0000 0011
66	0000 0000	67	0011 0011	68	0000 0001	69	1010 1010	70	1111 1001
71	1111 1111	72	0000 0011	73	0011 1001	74	0101 0011	75	1100 0001
76	0000 0111	77	1111 1001	78	0001 1101	79	0011 0011	80	1100 0011

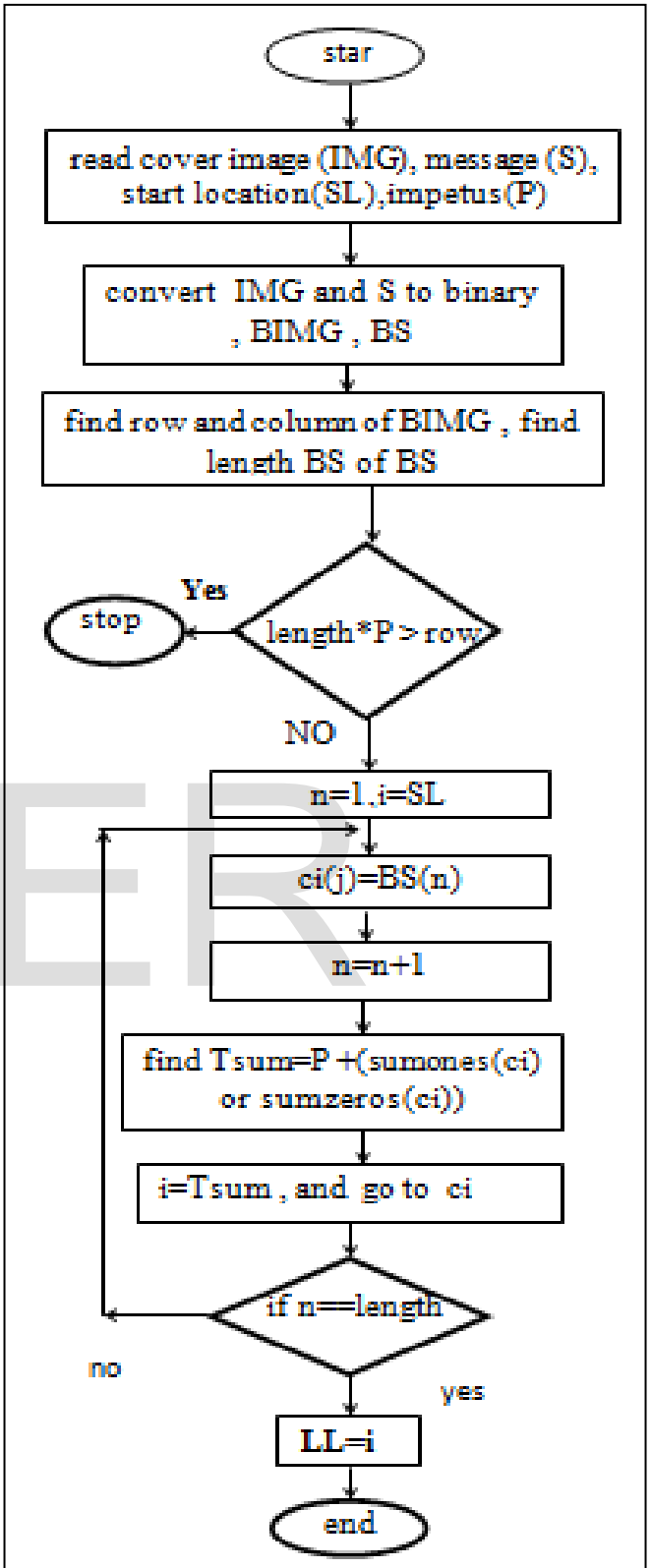

Fig. 2. The flowchart for proposed algorithm to embedding a secret data

Retrieval of secret data from the Embedded cover

(stego_cover)

Assuming the example above will we receive the following values (1,49) From the sender with the embedding cover , and thus find the sum of the number of Ones plus one of the first siteTo jump through it to the other locationWhich will be the site number seven(7) . And so on until the access to the site 49and This proves the validity of the data .When taking the first bit of each site within the chain is made up our secret message .

EXAMPLE (2) :

We suppose that we will jump between locations depending on account The total number of the Ones , Impetus(P) =1, Start location(SL) =80, will be the addition to the first bit from the right ,Therefore the secret message be hidden in the following locations .

80	1100 0011	5	1111 0001	11	0000 0000	12	0000 1111	17	0101 0011
22	1110 0010	27	1010 1010	32	0101 0011	37	0011 0011	42	0000 0010

EXAMPLE (3) :

We suppose that we will jump between locations depending on account The total number of the Zeros , Impetus(P) =1, Start location(SL) =1, will be the addition to the first bit from the right ,Therefore the secret message be hidden in the following locations .

1	1111 0001	5	1100 0001	9	1100 0011	15	0001 1100	16	1111 0000
17	1111 0000	23	1100 0011	29	1111 1001	33	1110 0001	40	0000 0000

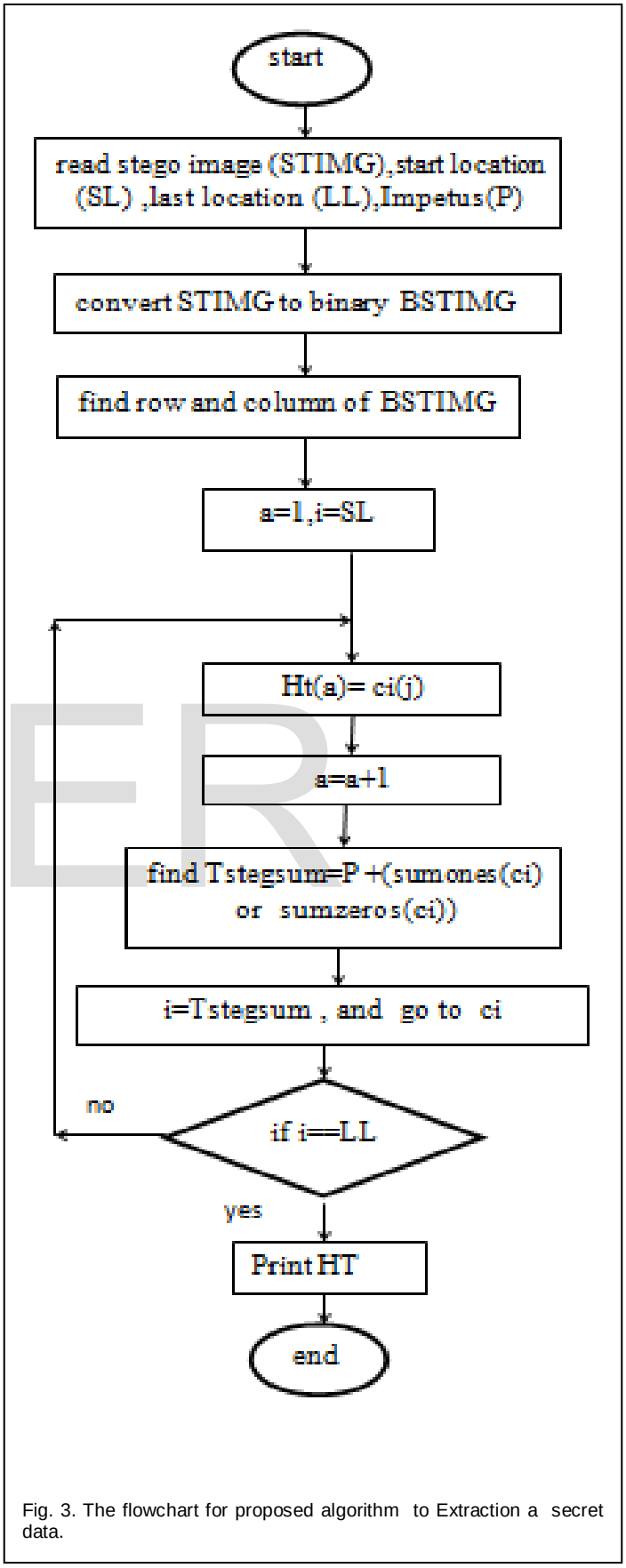

Fig. 3. The flowchart for proposed algorithm to Extraction a secret data.

7 RESULTS

The proposed algorithm has been tested to hide on a grayscale image with different sizes (Note : that the proposed algorithm can be applied to all the different types of images Such as Binary image , color image , Image multiple spectrum) .

Use have been the following standards (PSNR),(MSE)to measure the strength of the proposed algorithm and non-discrimination the hidden data by the human eye , Which measures the error square between the original image (cover image) and embedded image (the image that contain hidden data) by applying the following equations . [5]

$$MSE = \frac{1}{RM * CM} \sum_{i=1}^{RM} \sum_{i=1}^{CM} (cij - scij)$$

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

Where :
RM : represents the row for cover image .
CM : represents the column for cover image .
cij :represents an image unit before the hiding .
scij :represents an image unit after the hiding .
L : represents The level of signal strength , where L = 255 For the image of 8 bits per image unit.

TABLE 1
THE FOLLOWING TABLE SHOWS THE RESULTS OF THE APPLICATION OF THE PROPOSED ALGORITHM IN DIFFERENT WAYS ON THE IMAGES AND TEXTS OF DIFFERENT SIZES

	Image size	Text length by bits	PSNR	MSE	Start location (SL)	Last location (LL)	Impetus	Sumones or sumzeros	location LSB per byte
Ima1	250*250	121	78.880	0.00085	1	469	1	1	1
Ima2	250*250	121	78.270	0.00098	1000	7465	50	1	1
Ima3	250*250	2596	65.116	0.02018	3000	28156	20	1	1
Ima4	250*250	2044	60.164	0.06310	3000	53365	20	0	2
Ima5	250*250	574	71.196	0.00498	30000	37635	10	1	1
Ima6	250*250	2044	54.039	0.25856	3000	53553	20	0	3
Ima7	250*250	2044	54.052	0.25779	3000	50381	20	1	3
Ima 8	250*250	2044	48.048	1.02707	3000	50554	20	1	4
Ima 9	500*500	2044	72.036	0.00410	3000	214451	100	1	1
Ima 10	500*500	2044	71.895	0.00424	3000	216611	100	0	1
Img 11	500*500	2044	71.994	0.00414	250000	213621	100	0	1
Img12	500*500	4746	69.030	0.00819	1	62003	10	1	1
Img 13	500*500	8512	66.496	0.01468	1	111123	10	1	1
Img 14	500*500	5334	68.466	0.00933	2777	206339	35	1	1

8 PRACTICAL RESULTS

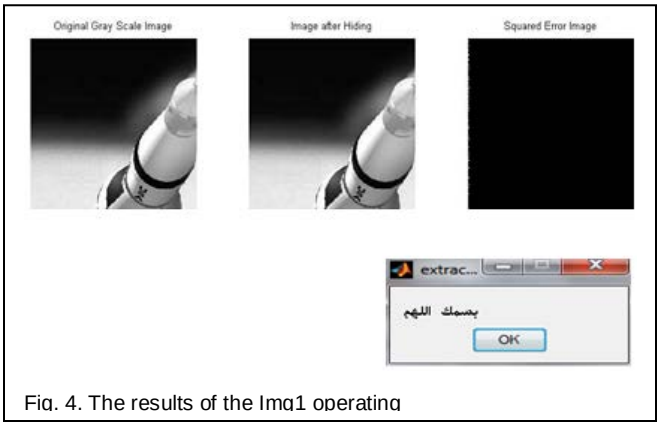

Fig. 4. The results of the Img1 operating

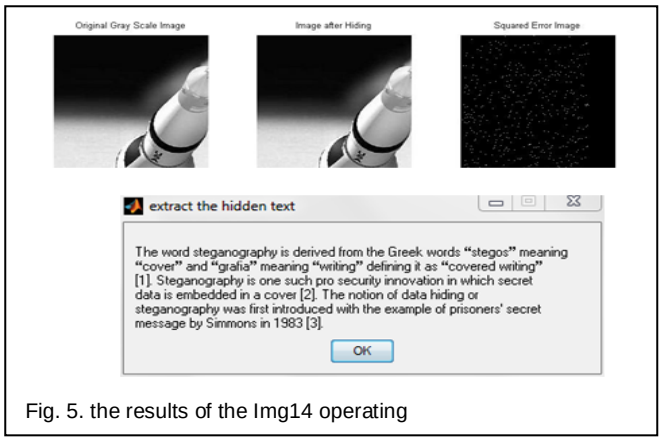

Fig. 5. the results of the Img14 operating

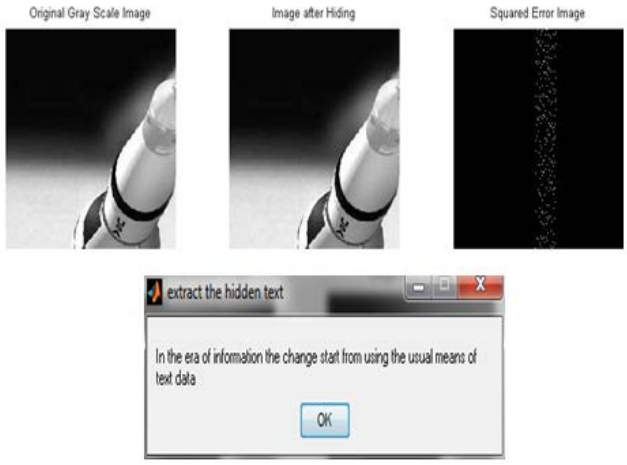

Fig. 6. the results of the Img5 operating

9 Discuss the results

Through studying the table and forms above show us
1- The value of PSNR & MSE

- PSNR Inversely proportional with size of the secret data
- PSNR Directly proportional with size of the cover image
- PSNR Inversely proportional with location LSB per byte
- MSE Directly proportional with size of the secret data
- MSE Inversely proportional with size of the cover image
- MSE Directly proportional with location LSB per byte

2- The possibilities
The number of possibilities that could arise in which the proposed algorithm , and There are seven factors affect the number of possibilities .

Assuming that the image size is 500 * 500 from type of the grayscale image , and the size of secret data is 2044

1. Combining on the basis of Ones or Zeros , Where there are only two possibilities: either 0 or 1
2. Specific bits :the Specific bits to collect from each pixel, According to the type of image we have 8 bits per pixel , Where we can count on one bit or two or three, ... until ,eight .According to the following rule We can find the possibilities to the number of bits Which will combine per pixel (The number of possibilities for the supposed image is 255 possibilities) .

This rule applies to all kinds of images in order to find
The number of possibilities for the bits that will combine

$$\text{Number\_of\_eventualities\_of\_Pixel(NEP)}= \sum_{i=1}^{L} \frac{L!}{i!(L!-i!)}$$

Where L is the number of bits for every pixel

3. Start location :Depending on the assumed size of the image we have the possibility 250000 (When increasing the image size increases the number of possibilities).
4. The impetus : It depends on the size of the image and the size of the secret data , Where directly proportional with the size of the image and inversely with the size of secret data , According to the following rule can find number of possibilities for the impetus .

$$\textbf{Impetus} = \frac{imagesize}{datasize} - \textbf{sizepixel} \quad >>> \quad \textbf{Impetus} = \frac{250000}{2044} - 8 \approx 115$$

5. Least important bits : From the above example shows us that the least important bits are the first three bits from the right . Thus, we have the three possibilities .
6. The image size : Whenever increase the image size increases the possibilities .
7. the data size : Whenever increasing size of the image decreases the possibilities .

Through the seven factors that affect the number of possibilities and on the assumption that the image size is 500 * 500, and the size of secret data 2044 bits , the number of possibilities is calculated from the following rule

ALL_eventualities= 2 * NEP * DATASIZE * Impetus * LSB
ALL_eventualities= 43987500000 possibility

10 Conclusions

1. Through the impetus of roughly we could choose sites that will hide the data in it.
2. Hide large-size data without affecting the shape of the cover image .
3. Data retrieval completely without loss or errors .
4. The large number of possibilities that could reach billions of possibilities with a small gray scale image (But when you use a color image, there are trillions of possibilities) .
5. Selection of the hiding sites , have an important role to reduce the deformation of the cover image and PSNR increase and decrease MSE .
6. The proposed algorithm has high flexibility where information can be hidden in different locations .
7. We have been getting good results for the value of PSNR and MSE even with large-size data .

11 Recommendations

from through proposed algorithm, we can suggest the following recommendations

1. The text can be encrypted before embedding process ,Which increases the number of proposed algorithm possibilities and thus more difficult to detect the hidden data .

2. The proposed algorithm can be applied to other types of images .
3. You can integrate the proposed hiding algorithm with other hiding algorithms .

12 REFERENCES

[1] Rabinovich, Vlad, 1999; "Steganography–a Cryptography Layer" http://www.rit.edu/~vxr8205/crypto2/cryptopaper.html. Accessed: Jan 2004.

[2] Arampatzis, Avi T., 1999; "Data Hiding", Report Katholieke Universiteit Nijmegen, School voor Informatica, Bedrijfsgerichte Informatica.

[3] الحمامي،علاءحسين؛الحمامي،محمدعلاء، 2008
"اخفاءالمعلومات،الكتابةالمخفيةوالعلامةالمائية"؛اثراءللنشروالتوزيع؛الشارقة.

[4] Brown, C. W., Shepherd, B. J., (1994); "Graphics File Formats Reference and Guide", Manning Grennwich.

[5] Qi, Hairong; Snyder, Wesley E. & Sander, William A., 2002; "Blind Consistency-Based Steganography for Information Hiding inDigital Media". Multimedia and Expo, 2002. ICME '02.Proceedings. 2002 IEEE International Conference on Vol. 1, p.: 585- 588.

[6] Cristobal, Patricia, 2003; "Steganography: A Privacy Protector or Just a Computer Security Trick? ", SANS Institute FIRE 2003 As part of GIAC practical repository. Washington D. C.