

## An Algorithm for Data Hiding in Binary Images

Eman Th. Sedeek Al-obaigy

Veterinary Medicine College  
University of Mosul

Received on: 1/10/2007

Accepted on: 17/12/2007

### الملخص

تم في هذا البحث استخدام خوارزمية جديدة لأخفاء البيانات في الصور الثنائية. يتم أولاً في هذه الخوارزمية تحديد النقاط الضوئية التي يمكن أن تنقلب دون حدوث تشوهات مرئية وذلك عن طريق استخدام مجموعة من القوانين التي يتم من خلالها فحص جميع النقاط المجاورة لنقطة المركز لكل قطاع غيرمنتظم ومن ثم يتم تغيير النقطة المركزية فقط في حالة مطابقة القطاع لهذه الشروط وهذه الخاصية تسمح باكتشاف البيانات المضمنة دون الرجوع الى الصورة الاصلية. وقد اظهرت التجارب نتائج مختلفة لصور ثنائية مختلفة، تمت كتابة البرنامج بلغة . MATLAB (6.5)

### ABSTRACT

In this work a new data hiding algorithm for binary images is used. In this algorithm the pixels which can be flipped without introducing noticeable artifacts are identified at first by using some rules that examine the neighbors of the central pixels in each non-uniform block and then this central pixel only is changed in the block which matches these rules and this property allows easy detection of the embedded data without referring to the original image. Experiments show different results for different binary images. The program is written in MATLAB (6.5).

## 1. Introduction

The use of binary images become more common in our life like scanned documents, drawing and signatures and these images can have the capabilities of data hiding for authentication, annotation and copyright protection. However data hiding in binary images become more difficult than that in grayscale and color images by which the gray level or color value of a selected group of pixels is changed by a small amount without causing visually noticeable artifacts. These techniques cannot be directly applied to binary images where the pixels have either "0" or "1" value. Arbitrarily changing pixels on binary images cause very noticeable artifacts [1], as can be noted in figure (1).

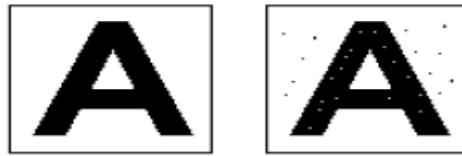


Figure (1) Effect of arbitrarily changing pixel values on a binary image [1],

## 2. Related Work:

There was a limited work on data hiding in binary images because of the bi-level constraints. Some of those researches in such field are reviewed below:

Mei,Q.,Wong,E.K., and Memon,N., use an algorithm just implemented on the text documents [1].They found the connected components of five-pixel long in the image and compare it with 100 pairs of five-pixel long boundary and each pair consists of two patterns, the first is addition of a foreground pixel adjacent to the center pixel and the second requires deletion of the central foreground pixel.

Lu,H.;Kot,C.A., and Cheng,J., use an algorithm based on (Distance–Reciprocal Distortion Measure)(DRDM) [2] which correlates well with human visual perception. A weighted matrix is used and then the distortion is measured for each pixel in a block. At last the pixel with the lowest DRDM value in the block is chosen to flip.

Pamboukian,S.;V.,D., and Kim,H.Y., use an algorithm which divides the image into (3x3 ) non-overlapping blocks and then arrange these blocks in an increasing order according to the visibility of their

central pixels [3] and then compares these blocks with templates which are set up previously and also arranged by the same manner. In an extraction stage, the two templates that differ only in their central pixels must have the same visibility score so the central pixel of that block will take the same value of the central pixel of the template.

Pan,G.;Wu,Z.; and Pan,Y.,(2002) introduce the concept of subblock which can be decomposed to several subblocks [4] and put a rank for each subblock according to the pattern-color and nominated –color: pattern-color is the color of the most pixels in the pattern and nominated color may be other color in the subblock except for pattern –color.

Tzeng,C.; and Tsai,W., use an algorithm that defines a variable ( Surrounding Edge Count) (SEC) [5] which represents the existing edges between pixel p and its eight neighbors and then SEC is found which is the difference between SEC<sup>+</sup> and SEC<sup>-</sup> and then the pixel p in block B flippable if and only if SEC is less than or equal to predefined threshold in order to restrict the distortion, and P and its neighbors are not visited yet in block B in order to make flippable pixels disconnected and this will not make visual artifacts.

Wu,M.; and Liu,B., identify those pixels that can be flipped from white to black and vice versa without introducing irregularities [6]. This is done through a set of rules which find a score to each pixel in a block. They find that the distribution of these flippable pixels is highly uneven over the image so they use the shuffle technique which permutes the pixels so that these flippable can occur more evenly across the image.

Wu,M.;Fridrich,J.; and Goljan,M.,H., clarify the benefits of the shuffling technique in increasing the capacity of embedding [7], while enhancing security since the receiver side only needs the shuffling table or key to correctly extract the hidden data.

In this paper, a proposed algorithm is introduced, the effect of the immediate neighbors (8-way connected) of the central pixel on a local window is taken as a major consideration in the design of the algorithm.

### **3. Data Embedding in Binary Images:**

Data –hiding can be defined as a technique that enables us to secretly embed extra data into digital media such as image, movie, audio and text files [8], or it is a technique used to embed a sequence of bits in a host with small visual deterioration and the means to extract it afterwards

[3]. However, there is a limitation to hide information behind binary images and there are many reasons for that like:

- 1- Embedding data in a binary image will cause obvious image content changes because of (black and white) nature of the image.
- 2- Binary images are more fragile to disturbances or attacks like channel noise, such a property makes authentication of recovered hidden information a required work. [5 ]

In particular flipping white or black pixels that are not on the boundaries are likely to introduce visible artifacts in binary images and in this type of images the human perceptual factor can be taken into account by studying every pixel and its immediate neighbors as shown in figure (2).[6]

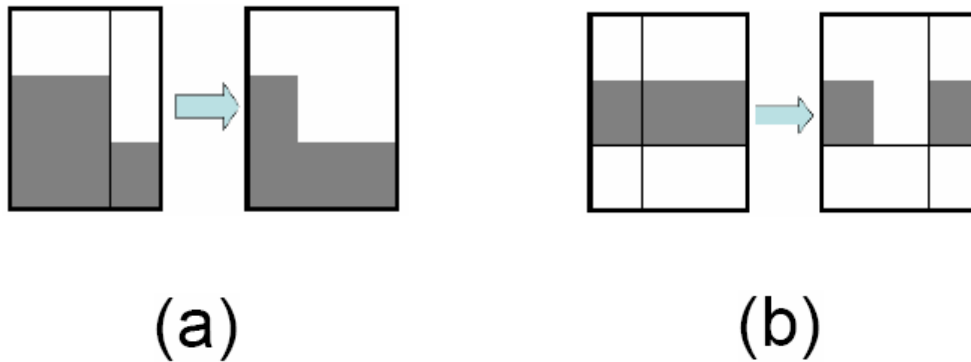


Figure (2): The flipping of the center pixel in (b) is more noticeable than that in (a) [6].

Application of data hiding like authentication and annotation requires high-rate embedding with blind detection which does not need the original image and first the flippable pixels whose change would not incur substantial visual distortion must be identified and to enforce properties of a group of pixel via local manipulation of a small number of flippable pixels. [1]

Embedding and extraction mechanism should be chosen precisely because any small mistake will propagate to the entire image in embedding and then in extraction method and directly encoding the hidden information in flippable pixel (e.g., set to black if to embed a "0" and to white if to

embed a "1") may not allow the extraction of embedded data without the original image.

The reason is that the embedding process may change a flippable pixel in the original image to a pixel that may no longer be considered flippable [6] as it is shown in figure (3).

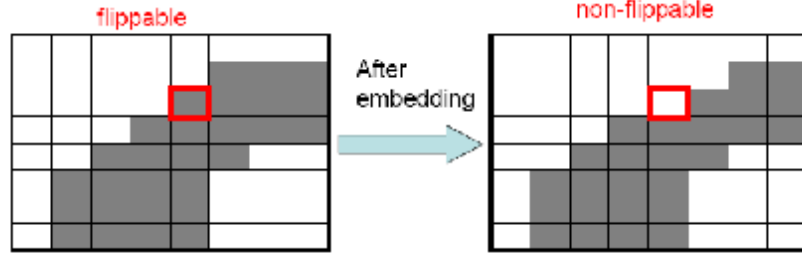


Figure (3): It can be seen that after embedding the marked pixel it is no longer considered flippable if applying the same rule.[6]

This example shows the difficulties for the detector to carry hidden information without using the original image.

#### **4. The Proposed Algorithm:**

The effect of the immediate neighbors (8-way connected) of the central pixel in a local window is taken as a major consideration in the design of the proposed algorithm. It can be noted in figure (4) that the flipping of the central pixels in both cases does not introduce major visual irregularities especially if the flipping is done on the boundaries. This property is used in the design of the templates which are set up previously and used later in the algorithm in the embedding and extraction process, and in figure (5) the four basic templates which are used in this algorithm are shown.

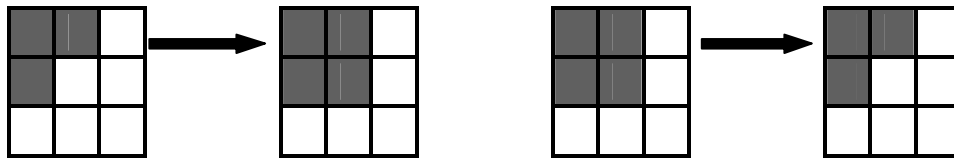


Figure (4): the flipping of the central pixels in both cases does not cause visual artifacts especially if the flipping is on the boundary.

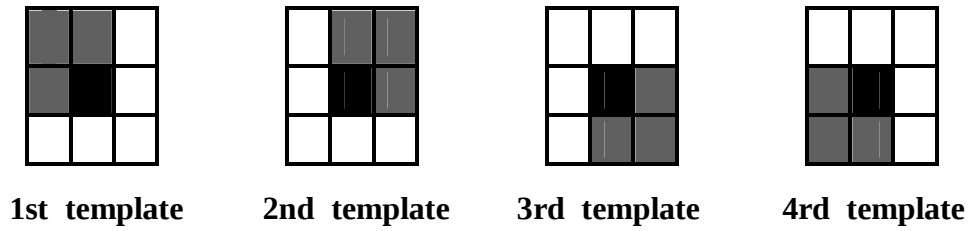


Figure (5): 4 Basic templates used in the proposed algorithm, the central ( black ) pixels only will be flipped from “0” to “1” or vice versa.

The four templates in figure (5), represent the starting point for the algorithm and by them the comparisons begin between them and the specified block to find the correct template and according to the value of the central pixel in that template the value of the central pixel in the block will be the same.

In figure (6), the extension of the first and fourth basic templates is shown and every time the number of the black pixels (suppose that the background is white and foreground is black) that surrounds the central pixel in a block is increased in the used templates and then if this central pixel only is changed from black to white or vice versa that will not cause a visual distortion in the marked image (stego image). The extension also is done for the second and third templates, and then the detection of the hidden data will be easy depending also on the same procedure and the same templates, that means a blind detection without need of the original image.

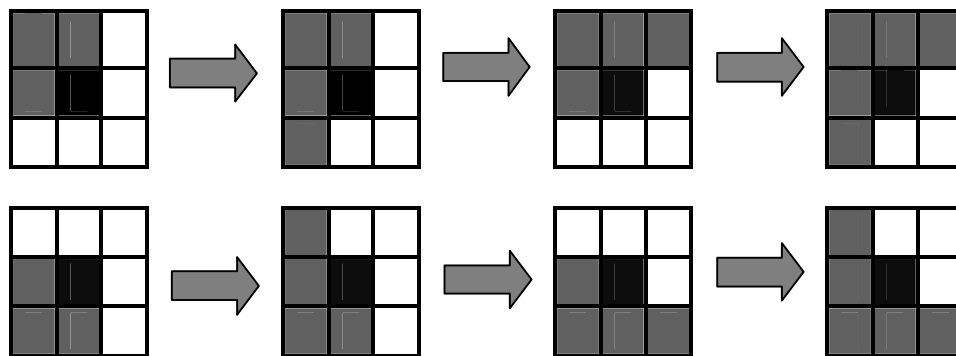


Figure (6): The increase of the black (gray here) pixels around the central pixels (black here) in the first and fourth templates.

#### **4.1. The Embedding and Extraction Process:**

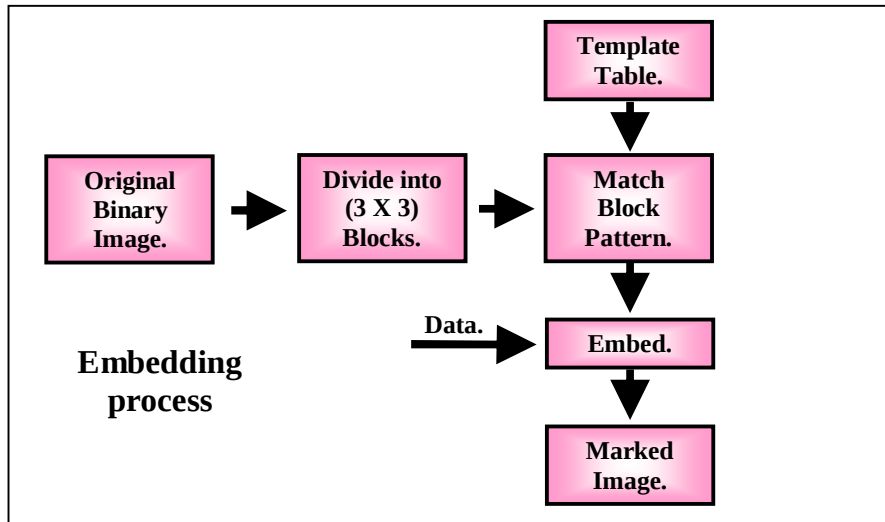
Embedding and extraction data processing can be summarized as follows:

##### **A) Embedding process:**

1. Divide the original image (cover image) into non-overlapping blocks of size (3x3).
2. Check the non-uniform blocks if they match one of the templates which are set up previously and when this happens only the central pixel of these non-overlapped blocks will be changed (either from “0” to”1” or from “1” to”0”) according to that template and the permutation of the pixels that surround the central pixels in that block and this change represents the embedding data and this is clarified in figure (7).

##### **B) Extraction process:**

1. Divide the marked image (stego image which carries the hidden data) into non-overlapping blocks of size (3x3).
2. Find the true template which will differ only in the central pixels in the examined block and this value of the central pixel in the block in the marked image represents the hidden data that will be extracted and this is clarified in figure (7).



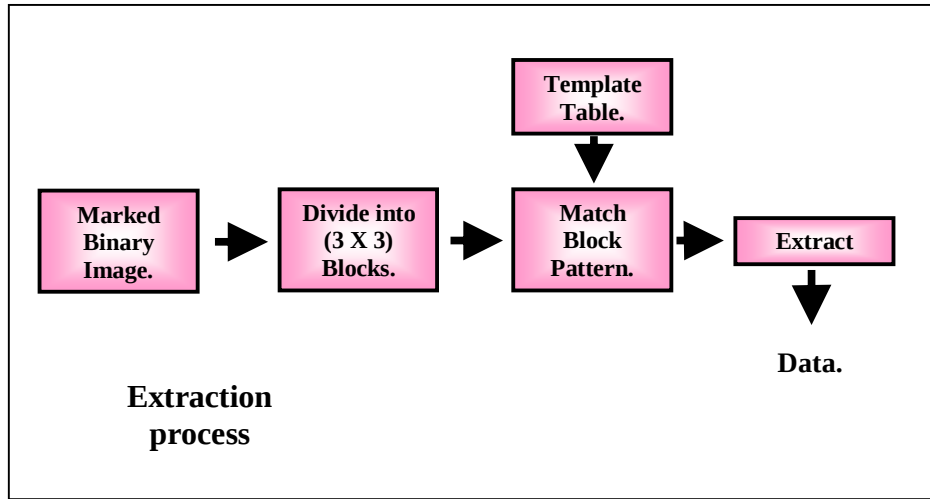


Figure (7): Block diagram of embedding and extraction processes

## 5. Results:

The algorithm is tested to embed a number of bits in binary images of different types (computer-generated text, cartoon-like images, scanned documents images ) and of different sizes. The correlation factor between the marked images and the original images was approximately 0.99 for all the tested images. The marked images have good robustness to printing and scanning, also they have pleasant visual quality.

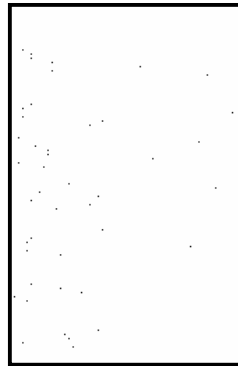
In figure (8), a scanned text document of size (171 x 234) is tested and modified by 44 bits.

√PC based app  
adding process  
ings show that  
-written/hand-c  
a defined sha  
acters in scanne  
ld be adjusted :

(a)

√PC based app  
adding process  
ings show that  
-written/hand-c  
a defined sha  
acters in scanne  
ld be adjusted :

(b)



(c)

Figure (8) : (a) original scanned text image (171 x 234), (b) marked image with 44 bits, (c) modified pixels.

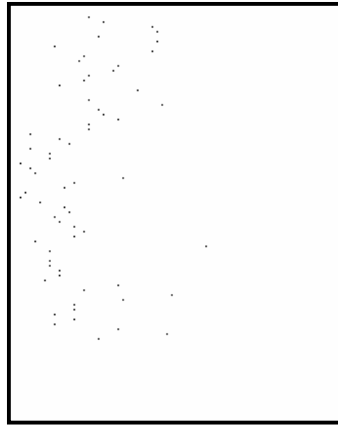
In figure (9), the second tested “Hunter” image of size( 165 x 213) is used and modified by 65 bits, this image also was used by [6] and the number of the embeded bits was 70 bits.



(a)



(b)



(c)

Figure (9) : (a) original scanned “Hunter” image (165 x 213), (b) marked image with 65 bits, (c) modified pixels.

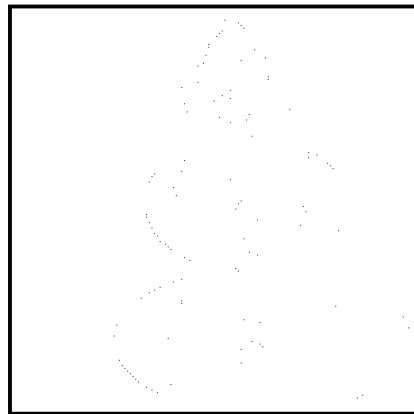
In figure (10 ), a computer generated image by microsoft “painter” program of size (382 x 456 ) is used and then marked by the algorithm with 105 bits.



(a)



(b)



(c)

Figure (10) : (a) original “snow man” image (382 x 456), (b) marked image with 105 bits, (c) modified pixels.

## **6. Conclousion:**

A new algorithm for embedding data in binary images has been proposed. It depends in it's design on the effect of immediate neighbors for a specified pixel. The algorithm has been tested on different types of binary images. The extraction of the hidden data does not need the original image (cover image) also the marked images have pleasant visual quality .

**REFERENCES**

- [1] Mei,Q.,Wong,E.K.,Memon,N.,”Data Hiding in Binary Text Documents “, Department of Computer and Information Science Polytechnic University, 5 Metrotech Center , Brooklyn,Ny11201,[wong@poly.edu](mailto:wong@poly.edu) .
- [2] Lu,H.,Kot,C.A.,Cheng,J.,”Secure Data Hiding In Binary Document Images for Authentication “, School of Electrical and Electronic Engineering Nanyang Technological University, Singapore 639798, [ehplu@ntu.edu.sg](mailto:ehplu@ntu.edu.sg).
- [3] Pamboukian,S.,V.,D.,Kim,H.Y.,”Reversible Data Hiding and Reversible Authentication Watermarking for Binary Images”,Universidade Presbiteriana Mackenzie,Brazil, [Sergiop@mackenzie.com.br](mailto:Sergiop@mackenzie.com.br).
- [4] Pan,G.,Wu,Z.,Pan,Y.,2002, ” A Data Hiding Method for Few-color Images”, Zhejiang University, Department of Computer Science and Engineering, Hangzhon,310027,P.R. China, {gpan,wzh}@cs.zju.edu.cn.
- [5] Tzeng,C.,Tsai,W., ”Hiding Authenticable General Digital Information Behind Binary Images with Reduced Distortion “, Department of computer and Information Science, National Chiao Tung University , Hsinchu, Taiwan 300, Republic of China , [chtzeng@cis.nctu.edu.tw](mailto:chtzeng@cis.nctu.edu.tw).
- [6] Wu,M.,Liu,B., ”Data Hiding in Binary Images for Authentication and Annotation “, 2004, IEEE Trans.on Multimedia.
- [7] Wu,M.,Fridrich,J.,Goljan,M.,H., ”Handling Uneven Embedding Capacity in Binary Images:A Revisit “, ECE Department,University of Maryland,USA.
- [8] Hirohisa,H.,” A Modified CPT Scheme For Embedding Data Into Binary Images”, Center for Promotion of Excellence in Higher Education, Kyoto University, [hioki@i.h.kyoto-u.ac.jp](mailto:hioki@i.h.kyoto-u.ac.jp).