IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 15, Issue 4 (Nov. - Dec. 2013), PP 57-64
www.iosrjournals.org

Uncompressed Image Steganography using BPCS: Survey and Analysis

Vipul J Patel[1], Ms. Neha Ripal Soni[2]
[1, 2] *Dept of computer engineering Sardar Vallabhbhai Patel Institute of Technology, India*

***Abstract*:** *Steganography is the art and science of hide secret information in some carrier data without leaving any apparent evidence of data alternation. In the past, people use hidden tattoos, invisible ink or punching on papers to convey stenographic data. Now, information is first hide in digital image, text, video and audio. This paper discusses existing BPCS (Bit Plane Complexity Segmentation) steganography techniques and presences of some modification. BPCS technique makes use of the characteristics of the human visible system. BPCS scheme allows for large capacity of embedded secret data and is highly customized. This algorithm offers higher hiding capacity due to that it exploits the variance of complex regions in each bit plane. In contrast, the BPCS algorithm provided a much more effective method for obtaining a 50% capacity since visual attacks did not suffice for detection.*

Keywords: *BPCS, Data security, Information hiding, Steganography, Stego image*

I. Introduction

Nowadays, data transfer and data sharing is part of high speed Internet technology. Intruders or any third parties try to access the secret information even if existing of data communication expert. So, information security needs to apply and modify exponentially. Cryptography and steganography are the part of information security that most widely used to reduce intruder accessing.

Cryptography [1] wildly uses to encrypt data and make unreachable for unauthorized person. Encryption process clearly marks a message as "private" information, and encrypted message becomes subject to attack. Steganography [2] [5] become another approach to data security. Steganography is also called data hiding. We hide private data in some carrier data which is in form of image, audio, video or text [3]. Imperceptibility and capacity are important parameters that we have consider making any steganography techniques. Capacity and imperceptibility denote reverse relationship that means in general more information hiding and less distortion in carrier information. Result of steganography is stego data. Stego data looks like original carrier data and human can't identify any part on hidden data. A block diagram of a generic image stenographic system is given in fig. 2.

Fig.1. Types of Steganography

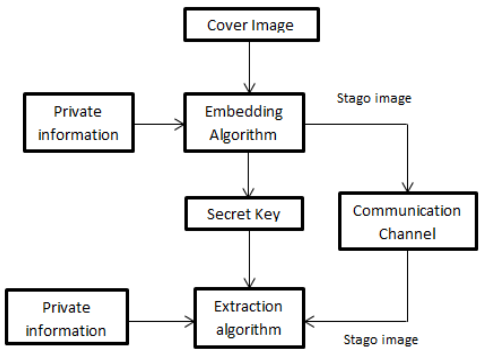


Fig.2. Generic form of image Steganography

In fig. 2, embedding algorithm is one of the data hiding techniques that takes input private information and secret key and embed data in cover image. The resulting stego image is transmitted over a channel and at receiver side authorized person extract private information using extraction algorithm.

For good information hiding technique in image, we need to consider following parameter [4]:

(1) Security: unauthorized person can't access the original information for stego image.

(2) Hiding Capacity: maximum amount of information can be embedded in the carrier information. It is based on image type and method applies for embedding information on it.

(3) Imperceptibility/ perceptual Transparency: difference between stego image and original one must be very slight such that the unauthorized person can't detect original information. If more information is hidden inside the carrier image, that results in degradation of stego image.
(4) Robustness: stego image remain unchanged even if the stego image undergoes transformation, sharping, filtering, scaling, blurring, cropping and other modification.

This paper will first outline the BPCS embedding and extraction technique for digital images. It will also contain different modification in BPCS data hiding techniques for more reliable and hiding information. In this paper, uncompressed image is used as a carrier image or base image to hide any information. Any compressed image may be lossy or lossless compression. If compressed image is lossy then we lost our nearly 25% of hidden information because of result in a significant reduction of the file size [14]. The amount of compression can be specified, and the amount of compression affects the visual quality of the result. When not too great, the compression does not noticeably detract from the image's quality, but JPEG files suffer generational degradation when repeatedly edited and saved. Due to this reason, we take uncompressed image like BMP image format.

II. BPCS Steganography Concept

BPCS [6] steganography was introduced by Eiji Kawaguchi and Richard O. Eason. In traditional techniques such as Least Significant Bit (LSB) technique, transformation technique, perceptual masking technique, have limited data hiding capacity and it can hide up to 10 – 15 % of the vessel data amount. BPCS is to overcome the short coming of traditional steganography techniques. This technique makes use of the characteristics of the human vision system whereby a human can't perceive any shape information in a very complicated binary pattern. First, the vessel image is divided into "informative region" and "noise-like region" then we can replace the entire noise-like region" in the bit-planes of the vessel image with secret data without destroying the image quality. BPCS steganography is same like LSB technique but difference is LSB technique hide data in last four bits i.e. only in the 4 LSB bits and BPCS technique hide data in MSB plane along with the LSB planes provided more storage and embedding data.
The merits of BPCS steganography are as follows:
(1) Approx. information hiding capacity of color image is 50%.
(2) A sharpening operation on the carrier image increases the embedding quite a bit.
(3) Canonical Gray Coded (CGC) bit planes are more suitable for BPCS steganography than Pure Binary Coded (PBC) bit planes.
(4) Data compression and encryption operation on secret data makes the embedded data more intangible.

2.1 Basic Principle of BPCS Steganography

Digital images are categorized as either binary or multi-valued pictures. The first step in the BPCS steganography is splitting the image into 'bit planes. Each bit plane is a binary image which contained the i^{th} bit of each pixel where i is the plane number. Ordinary slicing planes are represented by a Pure Binary Coding system (PBC) but in some case the Canonical Gray Coding system (CGC) is much better.

Example: Let P is n-bit gray scale image say n=8. After bit-slicing operation, we get P= [P7 P6 P5 P4 P3 P2 P1 P0] planes where P7 is the MSB bit plane and P0 is the LSB bit plane. Now each bit plane can be divided into "informative" and "noise" region. Noise-looking region consist complex pattern and we replace each noise-looking region with another noise-looking region without changing the overall image quality [11].

2.2 Binary number codding system

In BPCS-Steganography embedding operation is executed after the vessel image has been transformed from PBC to CGC. This is because CGC is better than PBC in producing a "better looking" stego image. The reason is as ordinary image planes are represented by Pure Binary Coded (PBC) provided much greater region for hiding. But PBC suffer from "Hamming cliff", wherein a small change in color affects many bits of color value.

Example: Gray scale image contain two different gray level are 127 and 128. These two gray levels affect in small change in gray image but representation of 127 as 01111111 and 128 as 10000000 effects a large difference in pixel representation. Both the pixels appear identical to human eye but differ greatly in bit representation. This is called "Hamming Cliff" concept.

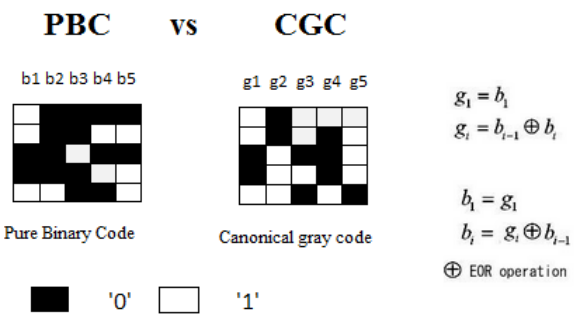


Fig.3. PBC vs CGC in Binary Image

2.3 The complexity of binary images

There is no standard definition of image complexity. Niimi and Kawaguchi discussed this problem in connection with the image thresholding problem, and proposed three types of complexity measures [7] [8] [9]. The different methods to find complexity of binary images are used to create segment between "informative" and "noise-like" image. There is no standard definition of image complexity. In BPCS, first we divide image in bit-planes and generate binary image plane as in fig 4. Different complexity methods are applied to find complexity in accurate way. In general, BPCS-steganography adopts black-and-white border complexity method explained in section 1. Two new complexity measures are then presented in section 2 and section 3 [10].

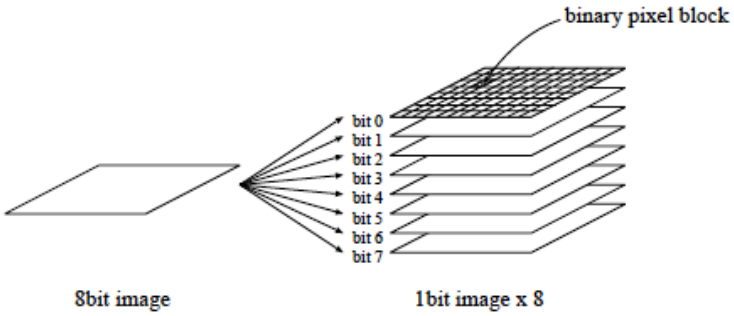


Fig.4. Binary pixel blocks on bit-planes

2.3.1 Black-and-White Border Complexity Measure

The length of the black-and-white border in a binary image is a good measure for image complexity. If the border is long, the image is more complex. The total length of the black-and-white border equals to the summation of the number of color-changes along the row and columns in an image.

We define the image complexity as

$$\alpha = \frac{\text{k}}{\text{The max. possible B} - \text{W pixel changes in image}}$$

Where, k is the total length of B-W border in the image. So, the value range of α over $0 \leq \alpha \leq 1$.

For binary image, minimum border length is 0. The equation for maximum length of the border for $2^n \times 2^n$ binary image is given by $2 \times 2^n \times (2^n - 1)$.

Thus, image complexity is also given by

$$\alpha = \frac{k}{2 \times 2^n \times (2^n - 1)}$$

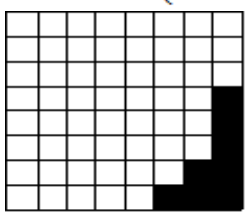


Fig.5. a simple block

For example, $2^3 \times 2^3$ (8×8) block in fig. 5 contain maximum border = $2 \times 2^3 \times (2^3 - 1) = 112$ and total border of image = 8. Thus, $\alpha = \frac{8}{112}$ is complexity of image block.

2.3.2 Run-Length Irregularity

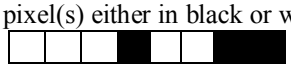
If the distribution of the black and white pixels in a block has a regular periodicity, it should not be used for embedding. The run-length irregularity is a new complexity measure introduced to evaluate the non-uniformity of the distribution of the black and white pixels in a block. The run-length irregularity is defined based on the histogram of the run-lengths of both black and white pixels along a row or a column.

Suppose that we have a binary pixel sequence shown in figure. It consists of a run of three white pixels, a run of one black pixel, a run of two white pixels, and a run of two black pixels.

Here we find that

h[1] = 1, h[2] = 2, and h[3] = 1

Where, h[i] is the frequency of runs of i pixel(s) either in black or white.



The following h_s is now introduced to measure the irregularity of a binary pixel sequence:

$$h_s = -\sum_{i=1}^{n} h[i]\log_2 p_i \qquad p_i = \frac{h[i]}{\sum_{j=1}^{n} h[j]}$$

Where n is the longest run-length possible, i.e., the length of the pixel sequence. This h_s evaluates the inequality of the run-length distribution in the binary sequence.

Let the block size be $n \times n$. Let r_i and c_j be i^{th} row and j^{th} column of a block respectively. The run-length irregularity β of a block is now defined as follows:

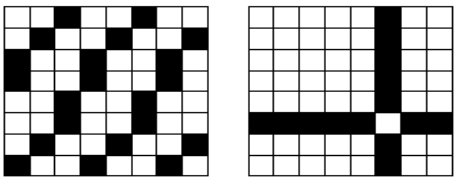
$$\beta = \min\{\overline{H_s(r)}, \overline{H_s(c)}\}$$

Where,

$$H_s(r) = \{h_s(r_0), \dots, h_s(r_{n-1})\}, H_s(c) = \{h_s(c_0), \dots, h_s(c_{n-1})\}$$

and $\bar{X}$ is the average of all the element of X.

The run-length irregularity β alone generally works well as a block complexity measure. Unfortunately, it can fail to reject some simple blocks. Fig. 6 shows examples. Although they look simple, their run-length irregularity βs are not small. This kind of exceptions can happen since we cannot evaluate similarities between adjacent rows or columns by the run-length irregularity.



(a) (b)

Fig.6. Block those are not complex, but have large run-length irregularities: (a) $\beta = 0.745$ (b) $\beta = 0.694$

2.3.3 Border Noisiness

If we hide information in blocks on the boundary of noisy regions and informative regions of a container image, the noisy regions would grow after embedding. As a result, we would find noticeable changes on the container image. The border noisiness is the other new complexity measure introduced to check if many black-and- white pixel borders are in a block and if they are well-distributed over the block. If the border noisiness of a block is large enough, it cannot be on the boundary of a noisy region and an informative region.

The border noisiness complexity measure is computed based on the differences between adjacent binary pixel sequences in a block .Let the block size be $n \times n$ (n > 1). Let r_i and c_j be i^{th} row and j^{th} column of a block respectively. The border noisiness γ of a block is defined as follows:

$$\gamma = \frac{1}{n}\min\{E_f(P_x(r)), E_f(P_x(c))\}$$

Where,

$$P_x(r) = \{p(r_0 \oplus r_1), \dots, p(r_{n-2} \oplus r_{n-1})\}$$
$$P_x(c) = \{p(c_0 \oplus c_1), \dots, p(c_{n-2} \oplus c_{n-1})\}$$

$\oplus$ denotes bitwise exclusive-or, ρ(x) is the number of ones in a binary sequence x, and

$$E_f(X) = \left(\frac{1.0 - V(X)}{\max\{V(X)\}}\right)\bar{X}$$

Where,

$$X = \{x_0, \dots, x_{n-1}\},\ V(X) = the\ varience\ of\ X \text{ and } \bar{X} = the\ averange\ of\ X$$

2.4 Conjugation of a binary image

If figure 7, W and B denote all-white and all-black patterns, respectively. Another two checkerboard patterns Wc and Bc are introduced in this method, where Wc has a white pixel at the upper-left position, and Bc is its complement, i.e. the upper-left pixel is black.

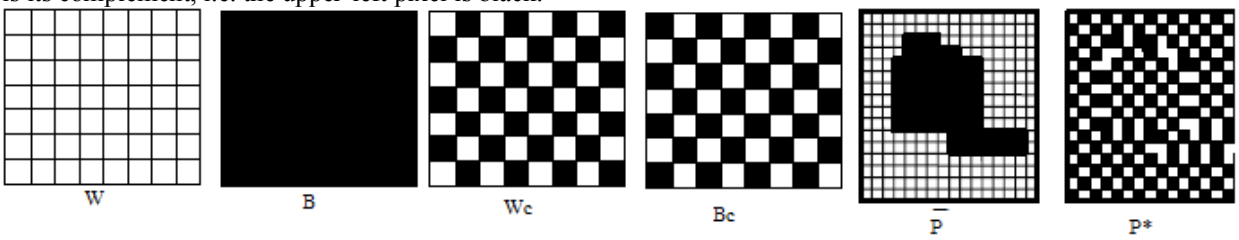


Fig.7. Binary plane patterns

Conjugation operation is kind of XOR operation image with Wc and Bc. Let P is binary image and conjugation operation with Wc create P*. Correspondence between P and P* is one-to-one. There are certain property of P and P* as follow.

1) P* = P⊕ Wc
2) (P*)* = P
3) P*≠ P
4) $\alpha(P^*) = 1 - \alpha(P)$

2.5 Criterion to segment a bit-plane into informative and noise-like regions

An investigation on the relation between α and noise-like region was performed by such a practical approach that follows. Random binary patterns having the size of $2^3 \times 2^3$ were generated as many times as possible (4,096,000 times actually). This is to simulate all the $2^{8\times8}$ possible 8×8 patterns in an image. In order to check statistical properties of 8×8 regions, we made a histogram as in figure of α. This histogram shapes almost exactly fit the normal distribution curve. The average value of the complexity in this histogram was exactly 0.5. The standard deviation was 0.047 in a. We denote this deviation by s.

The most important result from this experiment was as follows. If the secret data, which we want to encrypt, can be treated as a random (i.e., noise-like) binary image, and $\alpha$of each local area satisfies $0.5 - 4\sigma \leq \alpha$, we can embed the secret image into these areas of a dummy image.

2.6 Hiding Capacity

Hiding capacity and complexity of image have linear relation, the more complex image, the higher capacity of the image. The capacity of the image could be determined by accessing each plane and subsequently each segment within the plane to determine if it was complex. If the segment was complex, a counter was incremented. To demonstrate the capacity, and investigate the claim that 50% of the size of the image could be used, an evaluation was carried out on ten testing images.

2.7 Effectiveness Analysis

The effectiveness was measured in terms of how noticeable any alterations to the image were. In order to assess the user's perception, a Likert scale form was used. This was to determine not only if they suspected something wrong with the image, but also to place this on a scale for a more complete view of their interpretations. The scale used allowed the users the extreme choices of "No alteration" and "Definitely Altered" with the middle options being "probably no alteration", "cannot tell", and "a little alteration".

III. BPCS Steganography Algorithm

In BPCS-steganography, uncompressed image file like BMP file format is used for carrier image. We segment each secret file to be embedded into a series of blocks having 8 bytes of data each. These blocks are regarded as 8×8 image patterns. We call such blocks the secret blocks.

The steps for encoding algorithm (i.e. to hide private information in carrier image) in BPCS-steganography:

1. The carrier (color) image is divided into 24 different bit-planes, which create binary image for all 24-bits.

2. Transform all 24 bit-planes of carrier image from PBC to CGC system. Then all the bit-planes are divided into small pieces of the same size, which is called bit-plane blocks, such as 8×8 bits.
3. Segment each bit-plane of the carrier image into "informative" and "noise-like" regions by using a threshold value (α_0).
4. Group the bytes of the secret file into a series of secret blocks.
5. Embed each secret block into the noise-like regions of the bit-planes.
6. If a block (let say P) is less complex than the threshold (α_0), than conjugate it to make it a more complex block (P*). The conjugated block must be more complex than α_0.
7. If the block is conjugated, then record this fact in a "conjugation map". This Make a record of the blocks that have taken conjugate processing, and this information also need to be embedded into the carrier.
8. Also embed the conjugation map as was done with the secret blocks.
9. Convert the embedded carrier image from CGC to PBC.

The decoding algorithm (i.e. to extract original private information from stego image) is just the reverse procedure of the embedding steps. The process of secret information extraction is simple. Firstly, pick up all the pieces of the carrier data whose complexity is greater than α_0, and then pick up the extra embedded information mentioned in step (7) to confirm the blocks that have taken conjugate processing. These blocks need take XOR operation with tessellated chock to get the recovery of secret.

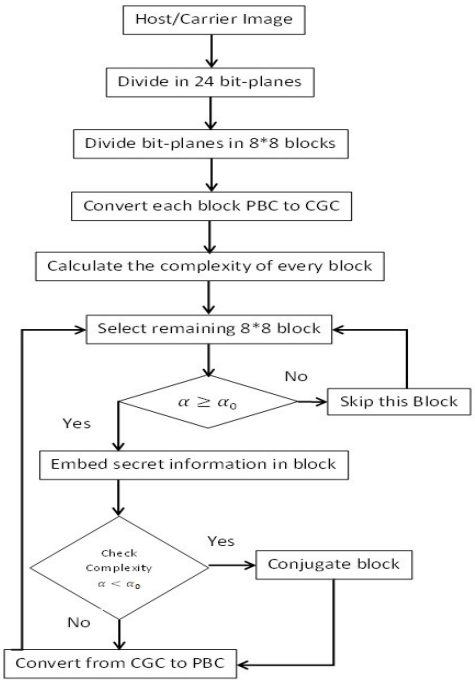


Fig.8. Flowchart- BPCS steganography

IV. Customization of BPCS-Steganography

BPCS-steganography is highly customizable technique for different user. We can modify algorithm as per user's requirement. It is very easy for a single BPCS steganography program to allow the user to customize parameters such as below, producing a very large number of possible customized programs. Several parameters of BPCS, which can modify for customization, are as follow:

1. The embedding threshold, α_0.
2. Size of binary image plane blocks
3. Encryption parameter of the secret file.
4. Compression parameter of the secret file.
5. Conjugation map

In this paper, we represent one modified BPCS-steganography by Smita P. Bansod, Vanita M. Mane and Leena R. Ragha [12].

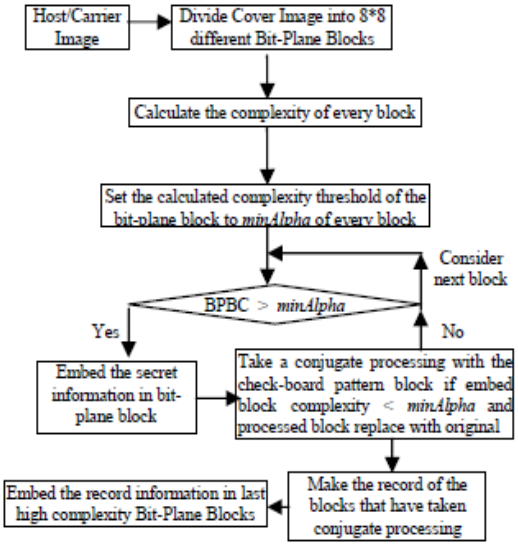


Fig.9. Flowchart- Modified BPCS Steganography

The flowchart of BPCS steganography is shown in fig. 9, described as follows:
1. The carrier image is divided into 24 different Bit- Planes. All the bit-planes are divided into small pieces of the same size, which is called bit-plane blocks, such as 8 × 8 bits.
2. Calculate the complexity α of every block. The complexity is defined as the amount of all the adjacent pixels that get different values (one pixel is 0, and the other is 1).
3. Setting the complexity threshold of the bit-plane block is max minAlpha (customization parameter). Here α is a parameter. The image complexity α is defined by the following.
α is calculated over the whole image area. It gives us the global complexity of a binary image. However, we can also use α for a local image complexity (e.g. an 8 × 8 pixel-size area). The bit-plane block whose complexity is larger than minAlpha is used to embed secret information. The smaller the value of minAlpha, the more secret information can be embedded.
4. Secret information is formed into bit-plane blocks. The bit-plane block can replace the original one straightly if its complexity is greater than minAlpha. Yet, it needs to conjugate processing with the white checkerboard pattern block if the complexity of embedded block is less than or equal to minAlpha, then take the new block to replace the original one.
5. Make a record of the blocks that have taken conjugate processing and this information also need to be embedded into the cover image. The embedding of this extra information cannot produce an effect on the embedded secrets, and it must be correctly picked up.
The basic steganography uses for bit 0, 1, 2 and 3. For bit 4, 5, 6 and 7; a new technique is used as: Apart from basic values of alpha (that is minimum complexity threshold), a new value (say gamma) is considered that indicate change in complexity from original 8*8 block of image to same stego image block. For 4, 5, 6 and 7 bit planes, first calculate alpha and if it is greater than minAlpha, then generate the bit pattern to be embedded from secret file and calculate alpha of the bit pattern as well. Now after recalculating alpha for generated pattern and compare it with minAlpha, if smaller the alpha value, then take stego image and complex conjugate as in previous algorithm. Now calculate change in pattern from original image. If this value (gamma) is less than minGamma, hide data in that calculated 8x8 blocks. If value is greater than minGamma of the block, then ignore that block for hiding purpose. The data can be hided in the block and use first two bits of block to indicate whether the bit pattern is conjugated and whether a valid data is indeed hidden or not. This way we can make use of entire image and increase the size of the data that can be hidden.

V. Evaluation Criteria

5.1 Peak signal to Noise Ratio (PSNR)

PSNR measure the quality of the image by comparing the original image or cover image with the stego image. Let cover image C(i.j) and stego image S(I,j) contain $n \times n$ pixels.
Mean squared error (MSE) formula as follow:

$$MSE = \frac{1}{[N \times N]} \sum_{i=1}^{N} \sum_{j=1}^{N} [C(i,j) - S(i,j)]^2$$

The PSNR formula as follow:

$$PSNR = \frac{10\log_{10} 255^2}{MSE} db$$

5.2 Capacity Measure

Capacity in data hiding indicates the maximum amount of information that can be hidden and successfully recovered by the steganography system [13]. Because of that the number of hidden bits varies depending on cover image size, to measure the hidden capacity, we use bitper- pixel (bpp) given as follow:

$$bpp = \frac{hidden\ bits}{Numpix(I_c)}$$

Where, Numpix(Ic) is total píxels number of pixels in the cover image.

5.3 Bit Error Rate (BER)

If the communication channel is ideal and there are not attacks, the proposed steganography system successfully recovers the hidden data. However we must consider a real communication scheme, and then we have to measure the bit error rate (BER), which is computed as follow:

$$BER = \frac{Error\ bits}{Message\ bits}$$

VI. Drawback of BPCS-Steganography

1. BPCS-steganography is based on complexity of image. For maximum embedding information, we require more and more complex image.
2. BPCS-steganography is not robust to even small changes in the stego image. I.e.to extract embed data from stego image correctly, there should not any change in stego image.
3. In some application, the presence of the embedded data may be known, but without the customization parameters, the data is inseparable from the image.

VII. Conclusions And Future Work

The main part of this paper is image complexity and BPCS-steganography algorithm. Any small change in image complexity can't detected by human. If image is very complex, then human cannot see any information embedded in bit-planes of a color image. BPCS-steganography provides maximum hiding capacity of image and lower PSNR. We can combine BPCS-steganography with encrypted embedded data for very strong information security. Future research will identify and formalizing the customization parameters and developing new applications.

References

[1] Behroz A. Forouzan, "Cryptography & Network Security", McGraw Hill Publication,2008, New Delhi.
[2] N.F. Johnson and S. Jajodia. "*Exploring Steganography: Seeing the Unseen"*. IEEE Computer, Volume 31: pages 26–34, 1998.
[3] Souvik Bhattacharyya, Indradip Banerjee and Gautam Sanyal. "*A Survey of Steganography and Steganalysis Technique in Image, Text, Audio and Video as Cover Carrier*". Volume 2, No. 4, April 2011 JGRCS.
[4] Hassan Mathkour, Batool Al-Sadoon, Ameur Touir, "*A New Image Steganography Technique*", Wireless Communications, Networking and Mobile Computing, 4th International Conference , 2008 IEEE,pp-1-4 .
[5] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "*Information Hiding-A Survey*", Proc. of the IEEE, Vol.87, No.7, pp. 1062-1078, 1999.
[6] Eiji Kawaguchi and Richard O. Eason, "*Principle and Applications of BPCS-Steganography*", Kyushu Institute of Technology, Kitakyushu, Japan – University of Maine, Orono, Maine.
[7] Kawaguchi, E. and Taniguchi, R., "*Complexity of binary pictures and image thresholding – An application of DF-Expression to the thresholding problem*", Proceedings of 8th ICPR, vol.2, pp.1221-1225, 1986.
[8] Kawaguchi, E. and Taniguchi, R., "*The DF-Expression as an image thresholding strategy*", IEEE Trans. On SMC, vol.19, no.5, pp.1321-1328, 1989.
[9] Kawaguchi, E. and Taniguchi, R., "*Depth-First Coding for multi-valued figures using bit-plane decomposition*", IEEE Trans. On Comm., vol.43, no.5, pp.1961-1995.
[10] HIOKI Hirohisa. "*A data embedding method using bpcs principle with new complexity measures*". http://www.i.h.kyoto-u.ac.jp/~hioki/research/DH/files/abcde_steg02_revised.pdf
[11] Hideki Noda Michiharu Niimi and Eiji Kawguchi. "*A steganography based on region segmentation by using complexity measure.*" Trans. of IEICE, J81-D-II, pp.1132-1140, 1998.
[12] Smita P. Bansod, Vanita M. Mane and Leena R. Ragha, "*Modified BPCS steganography using Hybrid Cryptography for Improving Data embedding Capacity*" IEEE computer, 2011.
[13] M. Goljan, J. Fridrich and R. Du, "*Distortion-free data embedding*", in Proc. of 4th Information Hiding Workshop, 2001, pp. 27-41.
[14] Daniel L. Currie III and Cynthia E. Irvine "*Surmounting the Effects of Lossy Compression on Steganography*." 1996 - DTIC Document.