TRANSACTIONS ON DATA PRIVACY 1 (2009) 140 —161

Close Color Pair Signature Ensemble Adaptive Threshold based Steganalysis for LSB Embedding in Digital Images

S.Geetha*, Siva S.Sivatha Sindhu*and N.Kamaraj**

* Department of Information Technology
Thiagarajar College of Engineering, Anna University
Madurai, Tamil Nadu, India.
E-mail: sgeetha@tce.edu, sivathasindhu@tce.edu

** Department of Electrical and Electronics Engineering
Thiagarajar College of Engineering, Anna University
Madurai, Tamil Nadu, India.
E-mail: nkeee@tce.edu

Abstract We present a novel technique for effective steganalysis of high-color-depth digital images that have been subjected to embedding by LSB steganographic algorithms. The detection theory is based on the idea that under repeated embedding, the disruption of the signal characteristics is the highest for the first embedding and decreases subsequently. That is the marginal distortions due to repeated embeddings decrease monotonically. This decreasing distortion property exploited with Close Color Pair signature is used to construct the classifier that can distinguish between stego and cover images. For evaluation, a database composed of 1200 plain and stego images (at 10% and 20% payload and each one artificially adulterated with 20% additional data) was established. Based on this database, extensive experiments were conducted to prove the feasibility of our proposed system.

Our main results are (i) a 90%+ positive-detection rate; (ii) Close Color Pair ratio is not modified significantly when additional bit streams are embedded into a test image that is already tampered with a message.; (iii) an image quality metric Czenakowski Measure, that is substantially sensitive to LSB embedding is utilized to derive the effective image adaptive threshold; (iv) capable of detecting stego images with an embedding of even 10% payload while the earlier methods can achieve the same detection rate only with 20% payload.

1 Introduction

In the past decade, digital technology has accelerated the development of network multimedia systems and introduced many advanced multimedia services. One prominent feature of digital technology is that editing, storage, transmission, and access of multimedia are easily done by any subject. For secure transmission, conventional methods exploited cryptographic techniques to thwart unauthorized access and tampering of secret messages. However, the encrypted form may draw special attention of network wardens and is thus not completely secret. Current information hiding techniques are developed to deceive wardens by embedding messages into multimedia in an imperceptible manner, but still preserve their original formats and quality. Actually, information hiding may offer negative effects in the aspects of personal privacy, business activity, and national security. It may be abused for covert communication between criminals. For example, commercial spies or traitors may thieve confidential trading or technical messages and deliver them to competitors for a great benefit by using hiding techniques. Terrorists may also use related techniques to cooperate for international attacks (like the 9/11 event in the U.S.) and prevent themselves from being traced. Some others may even think of the possibility of conveying a computer virus or Trojan horse programs via data hiding techniques. Thus, it raises the concerns of enhancing wardens' capability and lessening these negative effects by developing the techniques of "steganalysis".

From the above-mentioned, the primary goal of steganography is to set up a covert communication channel in a completely undetectable manner. This implies that the warden should be capable of discriminating suspicious objects from a large number of innocuous ones (passive steganalysis). In contrast to passive steganalysis, the goal of active steganalysis is to retrieve, modify, and even fabricate the embedded messages for destroying or interfering with covert communications and rendering hidden data useless. Applications of steganalysis then include, for example, an inlet/outlet content-monitoring program that inspects and intercepts suspected multimedia data transmitted on the network. In addition, steganalysis techniques can also be utilized to evaluate the security of covert communication channels under construction.

While the number of freeware packages available for steganography is increasing each year, the detection of most of these methods is neither satisfactory nor fully automated. While it is possible to hide messages within a variety of data file types, image data is likely to be the medium of choice for cyber criminals for several reasons. First, because of the high level of redundancy in image data, it is possible to embed a great deal of hidden information. Second, innocuous-looking images are commonplace on every computer and arouse little suspicion. Virtually all computer-users keep digital photos of friends and family, vaca-

tions, special events, etc. on their hard drives. Many web sites use images as a way to add interest and break up the monotony of text. By contrast, audio or video files posted on web sites are prone to be examined for copyright infringement.

The shear volume of image data available online makes it difficult to identify suspicious content. Thus, automated stego detection systems that can accurately detect hidden data can bring great benefits to the cyber forensic community in terms of quick and accurate detection.

Generally speaking, making decisions about the presence or absence of embedded messages in cover media is essential to steganalysis. Although it is simple to inspect suspicious objects and extract hidden messages by comparing them to the original versions, the restricted portability and accessibility of original cover-signals generally make blind steganalysis more attractive and feasible in many practical applications.

To be practical, a steganalysis algorithm is required to possess other properties such as low complexity and low classification risk. A low-complexity algorithm makes the system capable of inspecting objects at a high throughput rate. An algorithm of low classification risk generally makes tradeoffs between costs resulting from missing errors (i.e., false negative) and from false alarms (i.e., false positive). This motivates our current research: devising a threshold-based algorithm to classify images as being with or without hidden data. Our objective is not to extract the hidden messages or to identify the existence of particular information (as it is in watermarking applications), but only to determine whether an image was modified by LSB embedding technique. Once classified, the suspicious objects can then be inspected in detail by any particular data embedding/retrieving algorithms. This preprocess would particularly save time in active steganalysis.

The structure of the paper is as follows. In Section 2 related works done in LSB steganalysis are discussed. Section 3 elaborates the proposed methodology. The overview of the algorithm is illustrated under Section 4. In Section 5, we describe the experimental setup established for the performance evaluation of the proposed method and present the results obtained. Section 6 concludes the paper with representing a number of issues for future research.

2 Steganalysis of LSB Encoding

Current trend in steganalysis [3] seems to suggest two extreme approaches (a) little or no statistical assumptions about the image under investigation where statistics are learnt using a large database and (b) a parametric model is as-

sumed for the image and its statistics are computed for steganalysis detection. The messages embedded into an image are often imperceptible to human eyes. But there exists some detectable artifacts in the images depending on the steganographic algorithm used [4][5]. The steganalyst uses these artifacts for the detection of the steganography.

By far the most popular and frequently used steganographic method is the Least Significant Bit embedding (LSB). It works by embedding message bits as the LSB's of randomly selected pixels. Several techniques for the steganalysis of the images for LSB embedding are present. Fridrich and Long [6] proposed an algorithm for stego only attack. They analyzed the security of the most common steganographic technique - the LSB encoding in 24-bit color images. They have introduced a powerful steganalysis technique that enables us to reliably detect the presence of a pseudorandom binary message randomly spread in a color image. The method is based on statistical analysis of the image colors in the RGB cube. It is shown that even for secret message capacities of 0.1 -0.3 bits per pixel; it is possible to achieve a high degree of detection reliability.

Johnson and Jajodia [7] present a careful analysis of fingerprints introduced by current steganographic software packages. They point out that most techniques for palette images with a small number of colors can be easily broken by analyzing the palette for close pairs of colors.

Westfeld and Pfittzmann [8] introduced a method based on statistical analysis of Pairs of Values (PoVs) that are exchanged during message embedding. This method, which became known as the chi-square attack, is quite general and can be applied to many embedding paradigms besides the LSB embedding. It provides very reliable results when the message placement follows sequential embedding. However, their technique will not be effective for raw high-color images and for messages that are randomly scattered in the image (unless the capacity of the stego-technique is close to 1 bit per pixel).

Fridrich et al. [9] developed a steganographic method for detecting LSB embedding in 24 bit color images-the Raw Quick Pairs (RQP) method. The new method is based on analyzing close pairs of colors created by LSB embedding. On the condition that the number of unique colors in the cover image will be less than 30 percent that of the total pixels, it works reasonably well. When the number of unique colors exceeds about 50 percent that of total pixels, the results gradually become unreliable. This frequently happens for high resolution raw scans and images taken with digital cameras stored in an uncompressed format. Another disadvantage of the RQP method is that it can't be applied to grayscale images.

Sorina et al. [2] have introduced statistical sample pair approach to detect LSB steganography in digital signals such as images and audio. It is shown that the

length of hidden messages embedded in the LSBs of signal samples can be estimated with relatively high precision

Trivedi and Chandramouli [10] present a steganalysis method that estimates the secret key used in sequential embedding. Stationary and non-stationary host signals with low, medium, and high signal-to-noise ratio (SNR) embedding are considered. For non-stationary digital image data hiding in the DCT domain, the secret key estimation accuracy is good when the embedding is done in mid and high frequency DCT coefficients. Its performance suffers for low frequency embedding in the DCT domain.

Avcibas et al. [14] proposed the idea that any image may incur quality degradation after smoothing or low-pass filtering and this degradation (reacting on image quality) depends on the type of the test image, especially in categories with or without embedded information. That is, by observing quality difference between a test image and its smoothed version, it is possible to discriminate images with and without hidden messages. They hence utilized a regression analysis with several quality measuring operators for steganalysis. Ker [17] proposes a more accurate attack on LSB embedding through a weighted stego image detector for finding the sequential image replacement.

Mitra et al. [11] have described a detection theory based on statistical analysis of pixel pairs using their RGB components to detect the presence of hidden messages in LSB steganography. It is a stego -only attack in LSB insertion for uncompressed color images encoded in 24-bit BMP format. They have employed a fixed threshold method that resulted in poor detection rates and false alarm rates. The variable threshold value depends on the correlation between pixel pairs in terms of color components. The algorithm is able to detect a hidden message of size 20%. A quantitative steganalysis method to detect hidden information embedded by flipping pixels along boundaries in binary images is presented in [12]. For random embedding, these techniques provide an accurate estimate of the message length when the embedding rate is less than 50 percent.

Raja et al [13] explain a LSB Steganalysis method CPAVT based on variable threshold Color Pair Analysis. They have employed "Color Density" as the measure to derive the variable threshold. There are some shortcomings with the classification for some group of images as the color density measure is not highly sensitive to the LSB embedding.

In this paper, a particular stego-only attack in LSB insertion for high-density color image format using the close color pair signature is identified. Stego-only attack is applied when only the stego-image is available and the attacker has no idea about the original cover image, stego key or encoding algorithm. It is probably the most feasible attack that occurs in real world. In the current paper, the goal is to inspect a set of images for statistical artifacts due to message embedding in color images using the LSB insertion method and to find out, which im-

ages out of them are likely to be stego. The decision is based on a threshold value and the image are graded as stego bearing or not. Various image quality metrics [14] that are substantially sensitive to typical LSB data hiding are studied and tried for threshold value. Among them the Czenakowski Distance metric proved to be the most effective for threshold derivation. The experimental results also prove the same and its superior reliability than the CPAVT method.

3 Close Color Pair Analysis with Adaptive Threshold (CCPAAT) Steganalysis Method

In a natural uncompressed image (like 24 bit BMP) each pixel is represented by three-color channels (Red, Green and Blue), each of the channels is 8 bits wide. The LSB of any color channel of a typical scanned real image taken with a digital camera contains least information about the image and is most random in nature. Hence, most of the methods for hiding information in an uncompressed natural image are based on replacing the LSB of color channels by message bits. Thus, on the average only half of the LSB's are changed and it is assumed that, embedding messages in this way will not hamper the statistics of the cover-image and in turn no detectable signature will be generated. This assumption is true if and only if the number of unique colors in the cover-image is comparable to the total number of pixels in the image. However, it is observed that, in a natural uncompressed image, the ratio of the number of unique colors to the total number of pixels is approximately 1:6. Hence after LSB embedding, which is equivalent of introducing noise, the randomness of LSB pattern will increase. This increase in randomness is reflected in increase in the number of close color pairs, which is utilized as the distinguished signature for these types of images.

3.1 Decreasing Distortion Property

The close color pair (P) and unique color (U) is defined as follows:

- Two colors $(R1, G1, B1)$ and $(R2, G2, B2)$ are close if

$$|R1 - R2| = 1,\ |G1 - G2| = 1 \text{ and } |B1 - B2| = 1$$

or

$$(R1 - R2)^2 + (G1 - G2)^2 + (B1 - B2)^2 \leq 3; \quad (1)$$

- Two colors $(R3, G3, B3)$ and $(R4, G4, B4)$ are unique if any one of the following is true

$$|R3 - R4| = 1 \text{ or } |G3 - G4| = 1 \text{ or } |B3 - B4| = 1. \quad (2)$$

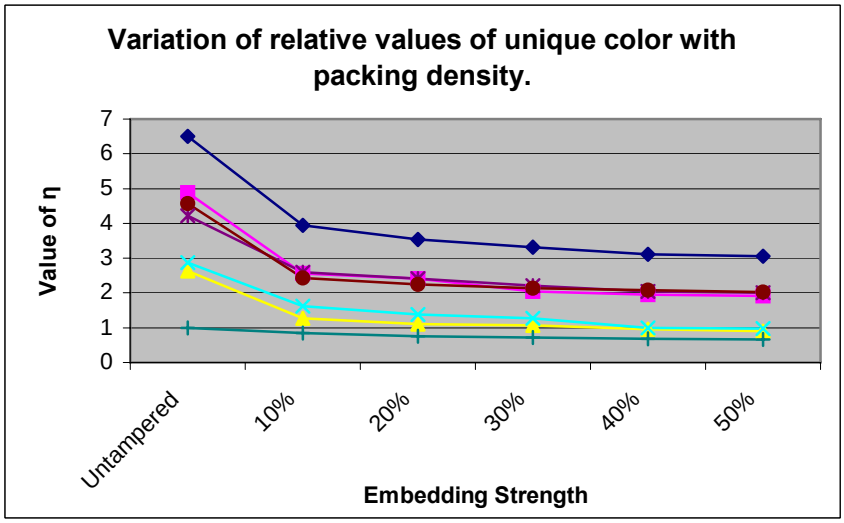
For any uncompressed real image, the ratio $\eta = \frac{P}{U}$ gives us an idea about the relative number of close color pairs with that unique colors.

Table 1. Experimental data to show the variation of the relative values of unique color with packing density.

% of message bit insertion	Mean Value of η						
	Class of Image						
	Animals	Birds	Buildings	Nature (Sky and cloud)	Flowers	Fruits	Faces
Untampered	6.51	4.89	2.63	2.88	4.22	4.56	0.99
10%	3.95	2.56	1.27	1.63	2.6	2.44	0.85
20%	3.54	2.41	1.11	1.39	2.42	2.25	0.76
30%	3.31	2.04	1.06	1.27	2.21	2.13	0.72
40%	3.11	1.95	0.96	1.00	2.04	2.08	0.69
50%	3.06	1.92	0.90	0.98	2.00	2.02	0.66

Now, it is observed that, for an untampered image, which does not have any embedded message, the value of η is greater in comparison with an image which has a message already embedded in it. This happens when an embedded message behaves as a random noise, which increases the number of unique colors U abruptly. As an example, we have taken 24 bit BMP images (Figure 2) having wide variation in color composition and have experimented with tampered images when different length of message bits are embedded by LSB insertion. The average values of the ratio η for both untampered and tampered images of various categories are compared in Table 1. It is noticed that, due to wide variation in U, i.e. the number of unique colors in different images, it is almost impossible to find a universal threshold for η efficient for all images to differentiate uniquely a stego-image from a non-stego one. The graphical representation of η with different percentage of data embedded in different nature of images is shown in Figure 1.

Figure 1. Variation of relative values of unique color with packing density for various image categories.



The initial value of each curve gives the cardinality of the unique color set in the untampered image. The rate of change of relative values of the unique color depends on the nature of the image. After prolonged testing with different kinds of images having wide color variation, a particular property is observed which enables us to reliably distinguish a tampered image from an untampered one. It is noticed that, if any test image is already tampered with a message, embedding it further with additional bit streams will not modify the η value significantly. Alternately, if the test image is an untampered one, the ratio η decreases significantly when it is further tampered by additional bit streams. i.e., under repeated embedding, the highest disruption of the signal characteristics is for the first embedding and then decreases steadily. This principle of decreasing distortion is used to derive a steganalysis tool that detects the presence of hidden messages in an uncompressed twenty-four bits BMP image.

To explore the decreasing property we have artificially packed the test image with data through the standard steganographic software S-Tools [15]. If U' and P' are the number of unique colors and close color pairs, respectively, then,

$$\eta' = \frac{P'}{U'} \tag{3}$$

gives the relative number of close color pair in the artificially tampered image. The change in the ratio η is measured in terms of μ where, μ is the percentage of change in η defined as:

$$\mu = \frac{(\eta - \eta')}{\eta} \mathbf{X}\, 100\%. \quad (4)$$

μ can now be properly thresholded to distinguish a tampered image from an untampered one. Earlier methods [11] chose a fixed threshold based on the observation of the image database. The results were feeble and not promising for some categories of images. Judicious selection of the threshold value determines the robustness of the software in terms of false positives and negatives. The various image quality metrics [14] are analysed and they were tried for the threshold value. From the experimental results, the Czenakowski Distance measure, a first order image statistics, is found to be the most promising feature to be chosen as the threshold. The improvement achieved in the performance has been shown using experimental data.

3.2 Image Adaptive Threshold Selection

Our goal is to design a threshold based classifier that can discriminate between stego and clean images. The threshold we employ for the classifier should be such that they reflect the distortions an image suffers from a LSB manipulation. We focused on content independent image features, which are predominant at the pixel level (LSB substitution) that are sensitive to image manipulations and could act as good candidates for threshold estimation. This is due to the fact that in any feature based classification method, there is the risk that the variability in the image content itself may eclipse image alterations present from the detector.

More specifically, let C denote a clean test image and $\varepsilon + C$ be its processed version (i.e., adulterated with 20% payload), and similarly let $\hat{C}$ and $\omega + \hat{C}$ indicate the test stego image and its processed version. Furthermore, consider a generic distortion function $M(a,b)$ between two signals a and b. A simple example of which being the well known mean square distortion function, $M(a,b) = E[(a-b)^2]$ with E being the expectation operator. The threshold we prefer will be based on the statistics of the difference of the distortions. The statistics which is sensitive to LSB encoding will act as a good signature for steganalysis.

In fact [14] discovered that the Czenakowski similarity measures are suitable distortion metrics because of their appreciable sensitivity to pixel level manipulations. We discuss these measures in the following section.

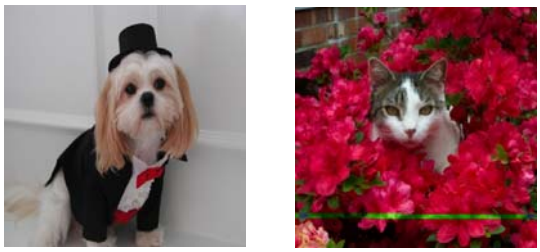
3.3 Czenakowski Similarity Measure as Threshold

We denote the color components of a three band (R, G, B) color image of dimension N X N, at pixel position i,j and in band k as $C_k(i, j)$, where k=1, …,3 and i,

Face Category

Nature Category

Animal Category



Building Category

Fruit Category

Flower Category

Figure 2. Sample cover images used in performance evaluation.

j=1,…,N. The symbols $C(i, j)$ and $\hat{C}(i,j)$ indicate the color pixel vectors, respectively , of the original and the stego image. C itself denotes a color image. The norm and inner product of the vectors are defined as

$$\|C(i, j)\| = \sqrt{C_1(i, j)^2 + C_2(i, j)^2 + C_3(i, j)^2} \quad (5)$$

and

$$\left\langle C(i,j),\hat{C}(i,j)\right\rangle = C_1(i,j)\,\hat{C}_1(i,j) + C_2(i,j)\,\hat{C}_2(i,j) + C_3(i,j)\,\hat{C}_3(i,j) \quad (6),$$

respectively. The pixels can take values from the set $(0, \ldots, G)$ in any spectral band. The actual color images we considered had G=255 in each band.

For steganalysis we prefer a distortion metric that is sensitive to the presence of a hidden message and whose reaction is proportional to the embedding strength. The Czenakowski distance is a metric useful for comparing vectors with strictly non-negative components, like in the case of color images, and is given by

$$CZD = \frac{1}{N^2}\sum_{i,j=0}^{N-1}\left(1 - \frac{2\sum_{k=1}^{K}\min\left[C_k(i,j),\hat{C}(i,j)\right]}{2\sum_{k=1}^{K}\left[C_k(i,j)+\hat{C}_k(i,j)\right]}\right) \quad (7)$$

This metric is very sensitive to noise [14]. A small distortion can result in a significant distance between two objects. However, in steganalysis, the main issue under consideration is not the content of an image file but the minor distortions introduced during the data- hiding process. As a result, this characteristic of Czenakowski distance makes it very helpful in the steganalysis.

The Czenakowski distance also called the Percentage of Similarity, measures the similarity among different samples, communities, and quadrates. Obviously as the difference between two images tends towards zero $\varepsilon = C - \hat{C} \to 0$ the value of the Czenakowski distance tend towards 100%, while as ε^2 increases the Czenakowski distance tends towards 0%.

Other first order Statistics based on Czenakowski similarity measures are:

$$\chi_{ij} = \frac{2\left\langle C(i,j),\hat{C}(i,j)\right\rangle}{\left\|C(i,j)\right\| + \left\|\hat{C}(i,j)\right\|} \quad (8)$$

$$d_1 = \mu_\chi = \frac{1}{N^2}\sum_{i,j=0}^{N-1}\left|\chi_{ij}\right| \quad (9)$$

$$d_2 = \left[\frac{1}{N^2}\sum_{i,j=0}^{N-1}\left(\chi_{ij}-\mu_\chi\right)^2\right]^{1/2} \quad (10)$$

The threshold δ is calculated based on these three metrics namely, CZD, d_1 and d_2.

$$\delta = \frac{\mu}{(CZD + d_1 + d_2)} \tag{11}$$

The values for these metrics are appreciably different for a clean image and for a stego image vis-à-vis their artificially adulterated version. The classification result is also provided in Table 2 to show the effectiveness of these metrics in revealing pixel manipulations.

4 Detection Algorithm Overview

Given a test image C, the objective is

(a) To analyze the test image to detect, whether it is a stego-image or not.
(b) To facilitate Close Color Pair Analysis ensemble image Adaptive Threshold based on First Order Czenakowski Distortion Metrics.

Assumptions:
(i) The test image C and payload image I are 24-Bit depth BMP images of arbitrary size.
(ii) The stego only attack is on images embedded with LSB steganography.

Let C be the test image of size M X N, I be the payload used for creating stego object for the algorithm Close Color Pair Analysis with Adaptive Threshold (CCPAAT). Let $\hat{C}$ denote stego object after embedding payload I into test image C. By using the standard steganography software tool S-Tools, which uses LSB algorithm, a stego object $\hat{C}$ is created by embedding I in C. Then unique color pair (U), close color pair (P) and ratio η for the test image C are computed. Similarly U', P' and η' for stego object $\hat{C}$ are also computed. The percentage variation in of η and η' denoted as μ is evaluated. The threshold δ is found out and is compared with μ. Based on comparison of μ and δ a decision whether the test image C is stego image or a non-stego image is arrived.

Algorithm CCPAAT:
Input: Test image C.
Output: Classification Result Stego or Clean Image.

Step 1: Artificially adulterate the test image C with 20 % payload, to create a stego image $\hat{C}$, using standard steganographic tool.

Step 2: For all the M X N pixels, count the total number of Unique colors U and Close Color Pairs P in the image C according to the Equations 1 and 2.

Step 3: Compute the value of $\eta = \frac{P}{U}$

Step 4: For all the M X N pixels, count the total number of Unique colors U′ and Close Color Pairs P′ in the adulterated image $\hat{C}$ according to the Equations 1 and 2.

Step 5: Compute the value of $\eta' = \frac{P'}{U'}$

Step 6: Compute the value of $\mu = \frac{(\eta - \eta')}{\eta}$ **X** 100%

Step 7: Compute the first order Czenakowski distance using the equations.

$$CZD = \frac{1}{N^2}\sum_{i,j=0}^{N-1}\left(1 - \frac{2\sum_{k=1}^{K}\min\left[C_k(i,j), \hat{C}(i,j)\right]}{2\sum_{k=1}^{K}\left[C_k(i,j) + \hat{C}_k(i,j)\right]}\right)$$

$$\chi_{ij} = \frac{2\left\langle C(i,j), \hat{C}(i,j)\right\rangle}{\left\|C(i,j)\right\| + \left\|\hat{C}(i,j)\right\|}$$

$$d_1 = \mu_\chi = \frac{1}{N^2}\sum_{i,j=0}^{N-1}\left|\chi_{ij}\right|$$

$$d_2 = \left[\frac{1}{N^2}\sum_{i,j=0}^{N-1}\left(\chi_{ij} - \mu_\chi\right)^2\right]^{1/2}$$

Step 8: Compute the threshold $\delta = \frac{\mu}{(CZD + d_1 + d_2)}$

Step 9: Compute the value of $\beta = \eta / \eta'$

Step 10: If ($\beta < 100\%$) or ($\mu < \delta$)

Image Type="**Stego Image**";

else

Image type="**Clean Image**"

end

5 Performance Analysis

To test the performance of the proposed method a large database of 200 color images taken from [18] with categories like Animals, Birds, Buildings, Nature (Sky and cloud), Flowers, Fruits and Face is used. The data set comprised of decompressed JPEG images with diverse nature having various degrees of texture, color, brightness and intensity. Some of the sample images are shown in Figure. 2.

5.1 Experimental Setup

This database is augmented with the stego versions of these images using the popular LSB embedding software, S-Tools [15], and two different payload strength were employed i.e., 10% and 20%. So there are 200 clean images and 400 stego images (200 under each hiding capacity; overall 600 images). All these images are adulterated artificially with a payload of strength 20% using the S-Tools to analyze the purity of the images (600*2=1200 images). S-Tools is used only for adulterating the images artificially. Any LSB based data hiding algorithm would equally work well. One obvious reason for selecting S-Tools was it is freely available on the Internet and it permits adjusting the payload insertion strength, which was instrumental to probe the sensitivity of the Czenakowski distance measure.

5.2 Embedding methods and message lengths

The LSB technique operates in the spatial domain with the least significant bit of each pixel value is flipped. We assumed that p bits could be embedded in each pixel value, where p is a fraction $0 < p < 1$. Thus the message length consists of a percentage point of the total number of pixels, and the length is independent of the type of image format, but proportional to the size of the image.

5.3 Classifier Results

Statistics from the original unmarked images as well as the stego images were obtained by computing the number of Unique Colors, Close Color Pairs, the Czenakowski distance, η, η', μ and the threshold δ, introduced in Section

IV. The classification is done as per the detection algorithm in Section V by comparing the values of μ and δ.

Table 2. A section of the experimental results showing the value of μ with variable threshold μ at 20% payload.

Image Name	Value of η/η'	Value of μ	Value of the adaptive threshold δ	Classification
Crocodile_Ut	1.006	0.063	0.0343	Clean
Crocodile_Ut_20				
Cat_Ut	1.004	0.041	0.0251	Clean
Cat_Ut_20				
TajMahal_Ut	1.0014	0.143	0.0917	Clean
TajMahal _Ut_20				
LesireRoom_Ut	1.002	0.024	0.0133	Clean
LesireRoom _Ut_20				
Crocodile_St	0.9997	-0.027	-0.0146	Stego
Crocodile_St_20				
Cat_St	1.000	-0.003	-0.0016	Stego
Cat_St_20				
TajMahal _St	0.9999	-0.014	-0.0008	Stego
TajMahal _St_20				
LesireRoom _St	1.001	-0.007	-0.0003	Stego
LesireRoom _St_20				
Sunflower_Ut	0.9952	0.4993	0.2314	Clean
Sunflower_Ut_20				
Rose_Ut	1.0001	0.0073	0.0031	Clean
Rose_Ut_20				
Estate_Ut	1.0984	0.0597	0.2636	Clean
Estate _Ut_20				
Amazon_Ut	1.0064	0.6318	0.3366	Clean
Amazon _Ut_20				
Sunflower_St	1.005	0.499	0.533	Stego
Sunflower_St_20				
Rose_St	0.9999	-0.006	-0.0027	Stego
Rose_St_20				
Estate _St	0.9992	-0.079	-0.0456	Stego
Estate _St_20				
Amazon _St	1.000	-0.005	-0.00266	Stego
Amazon _St_20				

It is observed that, the percentage change in η i.e., the value of μ, is quite small in case of stego images in comparison with its value for clean images. If we select threshold at a fixed value, then the performance of Building and Flower categories is satisfactory while for Nature and Face categories, the probability of erroneous detection is very high. The selection of fixed value is done on the basis of trial and error method on a large set of database.

The proposed method with variable threshold outperforms the fixed threshold method [11] as well as the Raja et al. method [13]. The percentage of False Detection Rate (FDR) and False Alarm Rate (FAR) computed for fixed and variable threshold with 20% payload is tabulated in Table 3. It has been observed that percentage of FAR and FDR are appreciably low in the proposed method.

Table 3. Experimental results showing the improvement in FAR and FDR in case of variable threshold compared to fixed threshold with 20% payload (CPAVT).

Image Category	False Alarm Rate %			False Detection Rate %		
	Fixed Threshold	CPAVT[13]	CCPAAT	Fixed Threshold	CPAVT[13]	CCPAAT
Animal	2.94	2.0	1.23	36.6	13.04	8.22
Birds	8.49	6.5	0.02	14.2	6.2	1.05
Buildings	27.1	15.4	1.6	2.3	1.0	0.6
Nature	36.5	0	0	0	0	0
Flower	17.8	8.7	0.2	4.0	3.36	1.63
Fruit	12.35	3.5	1.0	3.9	1.8	0
Face	12.5	0	0	5.6	5.26	2.14

Table 4. Experimental results showing the improvement in FAR and FDR in case of variable threshold compared to fixed threshold with 10% payload (CPAVT).

Image Category	False Alarm Rate %			False Detection Rate %		
	Fixed Threshold	CPAVT[13]	CCPAAT	Fixed Threshold	CPAVT[13]	CCPAAT
Animal	3.5	3.2	1.9	36.6	13.04	8.22
Birds	8.99	6.44	1.3	14.2	6.2	1.05
Buildings	28.0	16.2	2.3	2.3	1.0	0.6
Nature	40	0	0	0	0	0
Flower	20.7	9.4	3.5	4.0	3.36	1.63
Fruit	14.35	3.9	1.1	3.9	1.8	0
Face	48.5	0	0	5.6	5.26	2.14

This classifier is able to detect the stego image even with a payload of 10% where as the earlier methods managed up to 20% payload only. Table 3 and Table 4 display the comparison of FAR and FDR percentages of the proposed method with the methods in [11] and [13]. It is observed from these tables that the FDR percentage is much lower for the CCPAAT method. The FAR percentage with 10% payload does not change as the original cover image used in the stego object is the same as used with 20% payload.

Table 5. Comparison of FDR and FAR for variable threshold of current algorithm CPAVT with previous algorithm CCP with 20% payload.

Image Category	CCP Algorithm		CPAVT Algorithm		CCPAAT Algorithm	
	FAR %	FDR %	FAR %	FDR %	FAR %	FDR %
Animal	2.94	36.6	2.0	13.04	1.23	8.22
Birds	8.49	14.2	6.5	6.2	0.02	1.05
Buildings	27.1	2.3	15.4	1.0	1.6	0.6
Nature	36.5	0	0	0	0	0
Flower	17.8	4.0	8.7	3.36	0.2	1.63
Fruit	12.35	3.9	3.5	1.8	1.0	0
Face	12.5	5.6	0	5.26	0	2.14

The comparison of FAR and FDR percentages for previous algorithms, i.e., CCP (Close Color Pair) algorithm [11], CPAVT algorithm [13], and CCPAAT algorithm is given in Table 5. The FAR and FDR percentages are less in the CCPAAT algorithm than in the other algorithms with 20% payload. It is observed that the FDR with respect to the Face category is only 2.14 while in the earlier algorithm is 5.26, an improvement of around 60%. Similarly there are remarkable improvements in the other categories like Animal, Birds, Buildings also. The overall performance of the CCPAAT classifier is outperforming the earlier methods. This is due to the fact that the threshold value is based on the Czenakowski distance metric, which is very sensitive to data-hiding noise. Even a small distortion due to data hiding results in a significant distance between two objects. Thus it is an effective metric to decide the threshold value. To conclude, our algorithm gives better performance than the earlier work (for comparison see Table 7). Moreover, it also works satisfactorily with a 10% payload (See Table 4 and 6).

Figure 3. Performance Evaluation of CCP, CPAVT and CCPAAT Methods at 20% payload.

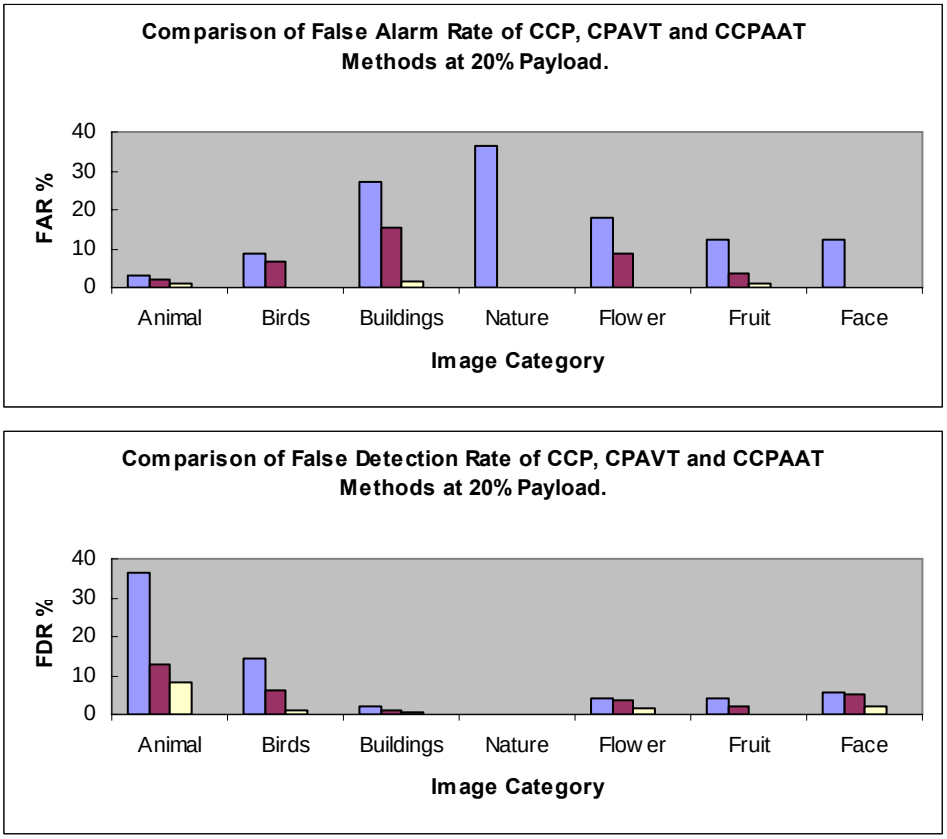


Table 6. Classification rates obtained for various image category by CCPAAT method

Image category	Positive Detection PD	Negative Detection ND	Classification Rate (PD+ND)/2	False Positive FP	False Negative FN	Error Rate (FP+FN)/2
Animal	91.32%	99.23%	95.275%	1.23%	8.22%	4.725%
Birds	99.07%	99.86%	99.465%	0.02%	1.05%	0.535%
Buildings	99.9%	97.9%	98.9%	1.6%	0.6%	1.1%
Nature	100%	100%	100%	0%	0%	0%
Flower	98.72%	99.45%	99.085%	0.2%	1.63%	0.915%
Fruit	99%	100%	99.5%	1.0%	0%	0.5%
Face	100%	97.86%	98.93%	0%	2.14%	1.07%

Figure 4. Performance Evaluation of CCP, CPAVT and CCPAAT Methods at 10% payload.

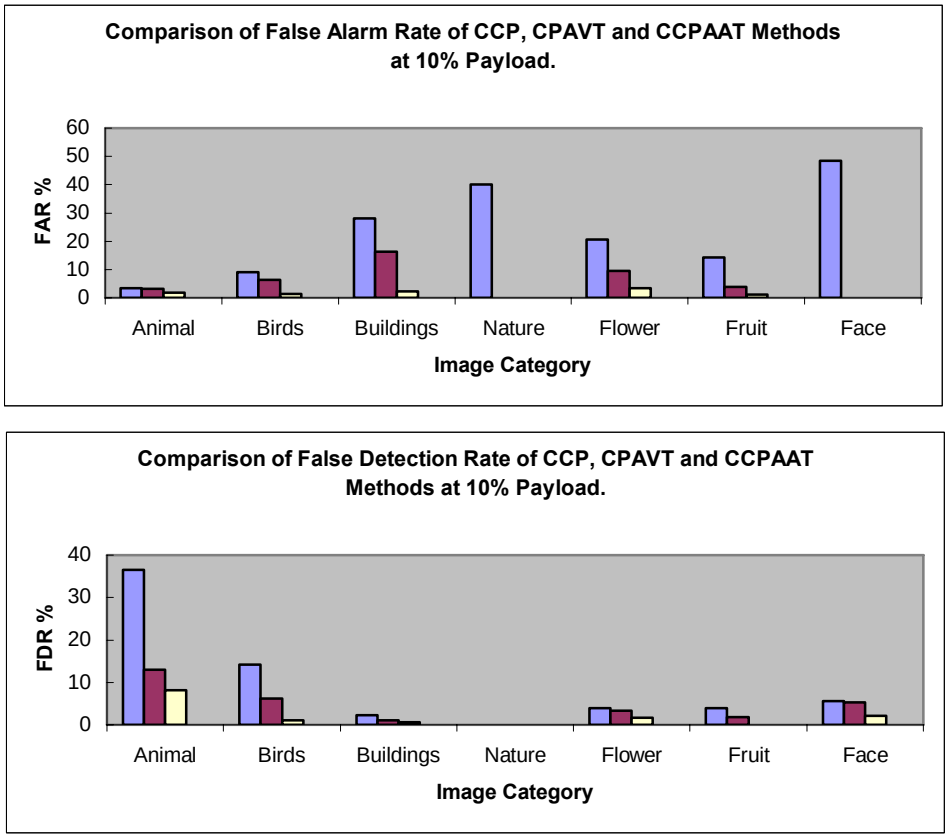


Table 7. Summarization of previous works involving Close Color Pair Analysis and our proposed system

Criterion	[11]	[13]	Proposed System
Size of Training database	400	300	1200
Number of Test images	50	50	300
Threshold type/feature	Fixed	Variable / Color Density	Variable / Czena-kowski Distance
Payload	>20%	>20%	>10%
Average PD Rate	86.83	95.23	98.74
Average ND Rate	13.16	4.77	1.26

6 Discussions and Conclusion

Recently, information hiding techniques have been applied to several fields, e.g., copyright protection, steganography, fingerprinting, digital rights management (DRM), etc. Though much current research is focused on how to embed data transparently and securely, it is, however, interesting to detect the existence of hidden data resulting from any kind of embedding scheme, known as the "steganalysis." In this paper, a classifier based on close color pair signature coupled with image adaptive threshold, is proposed to defeat LSB steganography techniques. In our experiments, a database composed of plain images and stego images generated by using LSB embedding schemes was utilized to evaluate the performance of our proposed features and classifier. Table 7 summarizes and compares the characteristics of our proposed method with those of similar works in the literature. The key findings of this work are as follows.

1) The principle of decreasing distortion is the basic aspect of this work. i.e., under repeated embedding, disruption of the signal characteristics is the highest for the first embedding and decreases subsequently.

2) Close Color Pair signature is a key artifact that is disturbed by LSB embedding. The distortion is more significant for clean images than for stego image.

3) The earlier systems use a fixed threshold or a variable threshold based on color density. The proposed approach is much more effective.

4) It seems that the classification performance is necessarily proportional to the threshold used.

5) Our training and test databases collect a large number of stego and nonstego image samples, which were generated by using LSB steganographic scheme at different embedding rates (10% - 20% payload).

6) The average classification rate (98%, including the PD and ND rates) for our proposed system is superior to the others in the literature.

To make the system more suitable for practical purposes, this work can be expanded towards

a) Calculation of the Close Color Pair signature of the image incurs high time complexity. The situation becomes worse for higher dimension images. To make the system effective for online real time applications, this part of the algorithm should be optimized.

b) Fitting the proposed system to classify compressed images or videos. Our algorithm is easily applied to uncompressed color images/videos in standardized format (e.g., BMP, TIFF, PNG).

c) Locating the image regions exploited to hide secret messages. In this case, we may be able to locate, retrieve, and analyze the embedded messages to infer the conveyed information.

Acknowledgement

This work was supported by grants from National Technical Research Organization of Government of India, as a part of "Smart and Secure Environment" project. The authors sincerely thank the Management, Principal and Head of the Department of Information Technology of Thiagarajar College of Engineering, Madurai, India, for their support and encouragement. Authors would like to thank the anonymous reviewers for the constructive comments which helped to improve the clarity and presentation of the paper.

References

[1] Fridrich, J., "Applications of Data Hiding in Digital Images," *Tutorial for the ISPACS'98 Conference*, Melbourne, Australia, November 1998, pp. 1-3.

[2] Sorina, D., Xiaolin, W. and Zhe, W., "Detection of LSB Steganography via Sample Pair Analysis", IEEE Transactions on Signal Processing, Vol.51, No. 7, July 2003, pp. 1995 – 2007.

[3] Chandramouli, R. and Subbalakshmi, K.P., "Current trends in steganalysis: a critical survey", *International Control, Automation, Robotics and Vision Conference 2004*, Volume 2, December 2004, pp.964 – 967

[4] Johnson, N.F. and Jajodia, S., "Steganalysis of images created using current steganography software", *Lecture Notes in Computer Science*, vol.1525, Springer, Berlin, April 1998, pp.273-289.

[5] Johnson, N.F. and Jajodia, S., "Steganalysis: The Investigation of Hidden Information", *IEEE Information Technology Conference*, September 1998, pp.113-116.

[6] Fridrich, J. and Long, M., "Steganalysis of LSB encoding in Color images," *IEEE International Conference on Multimedia and Expo*, vol.3, 2000, pp. 1279 – 1282.

[7] Johnson, N.F. and Jajodia, S., "Steganalysis: The Investigation of Hidden Information", *Proceedings of IEEE Information Technology Conference*, NY, Sept 1998.

[8] Westfeld, A. and Pfittzmann, A., "Attacks on steganographic systems", *Information hiding, Third International Workshop, IH'99*, Dresden, Germany, 29 September-1 October, 1999.

[9] Fridrich, J., Goljan, M. and Du, R., "Reliable detection of LSB Steganography in grayscale and color images", *Proceeding of ACM, Special Session on Multimedia Security and Watermarking*, Ottawa, Canada, 2001, pp. 27-30.

[10] Trivedi, S. and Chandramouli, R., "Secret key estimation in sequential steganography" *IEEE Transactions on Signal Processing*, Volume 53, Issue 2, Part 2, pp. 746 - 757, Feb.2005.

[11] Mitra, S., Roy, T.K., Mazumdar, D. and Saha, A.B., "Steganalysis of LSB Encoding in Uncompressed Images by Close Color Pair Analysis" *IITK-HACK04*, 23 - 24 Feb 2004.

[12] Ming J., Memon, N., Wong, E. and Xiaolin W., "Quantitative steganalysis of binary images", *International Conference on Image Processing*, 2004, Vol. 1, 24-27 Oct. 2004, pp. 29 – 32.

[13] Raja, K.B., Shankara, N, Venugopal, K.R., and Patnaik, L.M., "Steganalysis of LSB Embedded Images Using Variable Threshold Color Pair Analysis", *International Journal on Information Processing*, Vol. 1, 2007.

[14] Avcıbas, I., Memon, N. and Sankur, B., "Steganalysis using image quality metrics," *IEEE Trans. Image Processing*, Vol. 12, No. 2, 2003, pp. 221–229.

[15] Stools, A. Brown, S-Tools version 4.0, Copyright C., http://members.tripod.com/steganography/stego/stools4. html.

[16] Fridrich,J., Goljan, M. and Du, R., "Detecting LSB Steganography in Color and Grayscale Images", *IEEE*, vol.8(4) *Multimedia*, October-December 2001, pp.22-28.

[17] Andrew D. Ker.,"A Weighted Stego Image Detector for Sequential LSB Replacement",in Proc. 2007 International Workshop on Data Hiding for Information and Multimedia Security attached to IAS 07. IEEE Computer Society Press, 2007.

[18] Images were obtained from: http://philip.greenspun.com/.