



**HAL**  
open science

## Color Image Steganalysis Using Correlations between RGB Channels

Hasan Abdulrahman, Marc Chaumont, Philippe Montesinos, Baptiste Magnier

► **To cite this version:**

Hasan Abdulrahman, Marc Chaumont, Philippe Montesinos, Baptiste Magnier. Color Image Steganalysis Using Correlations between RGB Channels. IWCC: International Workshop on Cyber Crime, Aug 2015, Toulouse, France. 10.1109/ARES.2015.44 . lirmm-01234229

**HAL Id: lirmm-01234229**

**<https://hal-lirmm.ccsd.cnrs.fr/lirmm-01234229v1>**

Submitted on 26 Nov 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# COLOR IMAGE STEGANALYSIS USING CORRELATIONS BETWEEN RGB CHANNELS

Hasan ABDULRAHMAN<sup>2,4</sup>, Marc CHAUMONT<sup>1,2,3</sup>, Philippe MONTESINOS<sup>4</sup> and Baptiste MAGNIER<sup>4</sup>

<sup>1</sup> Nîmes University, Place Gabriel Péri, F-30000 Nîmes Cedex 1, France.

<sup>2</sup> Montpellier University, UMR5506-LIRMM, F-34095 Montpellier Cedex 5, France.

<sup>3</sup> CNRS, UMR5506-LIRMM, F-34392 Montpellier Cedex 5, France.

<sup>4</sup> Ecole des Mines d'Alès, LGI2P, Parc Scientifique G.Besse, F-30035 Nîmes Cedex 1, France.

## ABSTRACT

Digital images, especially color images, are very widely used, as well as traded via Internet, e-mail and posting on websites. Images have a large size which allows embedding secret messages of large size, so they are a good medium for digital steganography. The main goal of steganalysis is to detect the presence of hidden messages in digital media. In this paper, we propose a steganalysis method based on color feature correlation and machine learning classification. Fusing features with features obtained from color-rich models allows to increase the detectability of hidden messages in the color images. Our novel proposition uses the correlation between different channels of color images. Features are extracted from the channel correlation and co-occurrence correlation. In this work, all stego images are created with a range of different payload sizes using two steganography S-UNIWARD and WOW algorithms. To validate the proposed method, its efficiency is demonstrated by comparison with color rich model steganalysis.

**Index Terms**— Steganalysis, color spatial rich model, channel correlation, co-occurrences, ensemble classifiers.

## 1. INTRODUCTION

Steganography is the art of secret communication, by embedding messages within media such as audio, image or video files, in a way to complicate the detection of these hidden messages.

In general, steganography requires two parts, messages are embedded in the first part and they are extracted in the second part [1]. Several steganography techniques are used for different purposes, such as, for the most prominent, hiding illegal activities or secret communications [2]. Many steganography methods have been developed to enable secret communication. So the real-world needs to develop methods and algorithms rapidly in order to detect the presence of hidden messages and, thus, identify this secret communication. This method is called steganalysis.

---

We Thank the Iraqi Ministry of Higher Education and Scientific Research for funding this work.

Embed messages inside the digital medium involves some slight changes in this medium. Steganalysis research aims to develop some methods that are effectively able to detect these modifications. Although the real-world uses significantly more color images than grayscale images, there is high development in steganalysis to handle grayscale images, but, while overlooking color images [3].

In this paper, we propose new features to enrich the color rich model [4], which is formed by co-occurrences of residuals taken across the color channels. Section 2 is dedicated to methods related to color image steganalysis. Section 3 describes the proposed method by recalling the color rich model, color filter array and color channel correlation gradients. Experimental results an comparisons with its tables and curves are given in section 4. Section 5 concludes this paper.

## 2. LITERATURE REVIEW

In recent years, only a few methods involving color steganalysis have been proposed.

Westfeld, in [5], developed a detector of Last Significant Bit (*LSB*) matching in color images using the 3D color cube representation of the image. The detection method combines the three color components of each pixel to recognize an increased number of neighbor colors. This method estimates a relative frequency of a number of neighbor pixels as a steganalysis feature. *LSB* typically introduces characteristic tails in the histograms that are not observed in cover images. The authors have shown that this algorithm is efficient for decompressed JPEG images. However, it performs poorly in images with a high number of unique colors, such as never color compressed or resized images.

The detection of *LSB* Matching from grayscale images was enhanced by Ker, in [6], to detect *LSB* Matching in color images. By modifying the Harmsen method [7] that used the Center Of the Mass (*COM*) of the Histogram Characteristic Function (*HCF*) for the detection of additive noise based steganography, including *LSB* Matching. This method especially detects hidden messages in JPEG images but is not very efficient concerning bitmap images which have never been compressed.

Thiyagarajan et al. [8] proposed a universal image steganalysis method based on color model conversion. Indeed, considering color images, they convert Red, Green and Blue (*RGB*) channels of the images to the Hue Saturation and Intensity (*HSI*) color model to detect hidden messages. Stego images are then generated by implementing different color image formats, using the last significant bit steganography method. Finally, cover and stego images are recognized using a threshold value which depends on the correlation between pixel pairs in terms of color components.

Lyu et al. [9] described a universal steganalysis algorithm that exploits the inherent statistical regularities of natural images. The statistical model consists of first and higher order color wavelet statistics of noise residuals obtained using predictors of coefficients in Quadratic Mirror Filter (*QMF*) decomposition of the image from all three color channels. Finally, they estimate that the addition of color statistics provides considerable improvement in overall detection accuracy.

Krichner et al. [10] suggested a steganalysis method to detect LSB replacement steganography in color images. These stego images are produced via processing which reveals some traces of Color Filter Array (*CFA*) interpolations. Here the authors have enhanced the Weighted Stego (*WS*) image steganalysis method [11] by replacing the cover predictor in *WS* with position specific predictors. This technique puzzles out the local predictability of pixels, depending on their position in the *CFA* interpolation to account for differences between cover (common) images and suspect images.

Fridrich et al. [12] proposed a general methodology for steganalysis of digital images based on the rich model concept consisting of a high number of diverse submodels. Each submodel considers various types of relationships among neighboring samples of noise residuals obtained by linear filters with compact supports. The submodels are formed from noise residuals, which are computed using high pass filters, then features are extracted using co-occurrence matrices.

From two different components, Goljan et al. [4] developed a color image model which is called the Spatio-Color Rich Model (*SCRMQ1*, where 'C' represents the color version of the Spatial Rich Model -*SRMQ1*- [12], with 'Q1' representing fixed quantization  $q = 1$ ) produced from two different components. The first component is the *SRMQ1* with a single quantization step  $q = 1$  and dimensionality of 12753 features. The *SRMQ1* feature is computed for each color channel separately, then the three features are merged to keep the same dimensionality, as for grayscale images. The second component is a collection of 3D co-occurrences with a larger threshold  $T = 3$  and  $q = 1$ . The final dimensional features obtained from this component are 5404 features. The color rich model is built from the same pixel noise residuals as those used in the *SRMQ1*, but is formed across the three channels of each pixel.

### 3. PROPOSED METHOD

Our proposition is to enrich the *SCRMQ1* with an inter-channel correlation which is composed of two sets of features. The first set, produced by [4], gives 18157 features. The second set gives 3000 features obtained from the correlation of different *R, G, B* channel gradients. The following section recalls the color rich model, then section 3.2 explains the basics of color filter array, and section 3.3 gives an explanation about our proposition.

#### 3.1. Color Rich Model

The following steps describe the process by which the first set of features is computed. We used the spatial color rich model for color image steganalysis in [4] which produced 18157 features. This method extracts the noise residual from each color channel separately by the following formula:

$$\mathcal{R}_{ij} = \hat{X}_{ij}(\mathcal{N}_{ij}) - c \cdot X_{ij}. \quad (1)$$

where  $c \in \mathbb{N}$  is the residual order,  $\mathcal{N}_{ij}$  is a local neighborhood of pixel  $X_{ij}$  at coordinates  $(i, j)$ ,  $\hat{X}_{ij}(\cdot)$  is a predictor of  $c \cdot X_{ij}$ ,  $X_{ij} \notin \mathcal{N}_{ij}$ ,  $X_{ij} \in \{0, \dots, 255\}$ .  $\hat{X}_{ij}$  represents a pixel value of an 8-bit grayscale cover image and its corresponding stego image.

All of the submodels  $(\mathcal{R}_{ij}) \in \mathbb{R}^{n_1 \times n_2}$  are formed from noise residual images of size  $n_1 \times n_2$  computed using high pass filters of the following form:

$$\mathcal{R}_{ij} \leftarrow \text{tranc}_T \left( \text{round} \left( \frac{R_{ij}}{q} \right) \right), \quad (2)$$

where  $\text{tranc}_T$  represents the truncation function with  $T > 0$  defined for any  $x \in \mathbb{R}$ ,  $\text{tranc}_T(x) = x$  for  $x \in [-T, T]$  and  $\text{tranc}_T(x) = T \cdot \text{sign}(x)$  and otherwise  $q$  is the quantization step, and  $\text{round}$  is a function for rounding to an integer value.

The spatio-color rich model consists of two different components. On one hand, the spatial rich model (*SRMQ1*) [12] with a fixed quantization  $q=1$  and truncation  $T=2$  yields a dimensionality of 12753 features. These features are computed from each *R, G* and *B* color channel separately. Finally, the three dimensionality features are added together to keep the same dimensionality as for grayscale images. On the other hand, from the same noise residuals (i.e. *SRMQ1*), the *CRMQ1* builds a collection of 3D color co-occurrences, taking three color values at the same position (across the three channels of each pixel). Thus, with fixed truncation  $T=3$  and quantization  $q=1$ , *CRMQ1* produces 5404 features.

#### 3.2. Color Filter Array

Digital cameras capture color images using a single sensor in conjunction with a Color Filter Array (*CFA*). The *CFA* allows only one part of the spectrum to pass to the sensor so that only one color is measured at each pixel (red, blue or green)

and such images are called mosaic images. To construct a color image, a demosaicking algorithm is used in order to interpolate each color plane (i.e. *CFA* interpolations). Several patterns exist for the color filter array, with the most common being Bayer *CFA* [13]. During this process, the green color channel is the most important factor which determines the luminance of the color image, 50% of the pixels in the Bayer *CFA* structure is assigned to the green channel, while 25% is assigned to the red channel and 25% to the blue color channel.

The interpolation introduces specific correlations between the samples of a color image. Hence, a subset of samples, within a color channel, is correlated with their neighboring samples of the same color. Since the color filters in a *CFA* are typically arranged in a periodic pattern, these correlations are also periodic [14].

### 3.3. Channel Correlation

In this section, we introduce an inter-channel correlation measure, and demonstrate that it can be listed to first order Euclidean invariants (see Hilbert [15] for the invariant theory). Such invariants have mainly been used for stereo-matching [16]. In this paper, we show that the information provided by them can enhance steganography detection.

Starting from the local correlation of red and green channels (similarly, correlation of red and blue channels) :

$$Corr_{R,G}(i, j, k, l) = \sum_{(i', j') \in \mathcal{W}_{i,j}} X_{i', j'}^{(R)} X_{k+i', l+j'}^{(G)} \quad (3)$$

with:

- $X_{i', j'}^{(R)} \in [0, 255]$ , being a pixel value at position  $(i, j)$  in the red channel,
- $X_{k, l}^{(G)} \in [0, 255]$ , being a pixel value at position  $(k, l)$  in the green channel,
- $\mathcal{W}_{i,j}$ , being a small window centred in  $(i, j)$ .

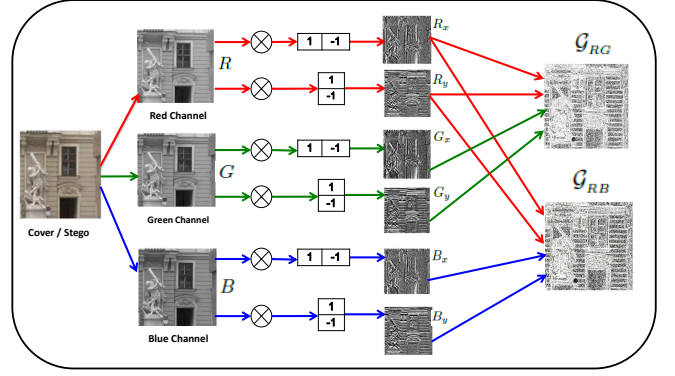
Considering  $(k, l) = (0, 0)$  and a limited development of  $X^{(R)}$  and  $X^{(G)}$  around  $(i, j)$ , then :

$$Corr_{R,G}(i, j, 0, 0) = \sum_{\substack{\mathbf{h} = (i'-i, j'-j) \\ (i', j') \in \mathcal{W}_{i,j}}} \left( X_{i,j}^{(R)} + \nabla X_{i,j}^{(R)} \cdot \mathbf{h} \right) \left( X_{i,j}^{(G)} + \nabla X_{i,j}^{(G)} \cdot \mathbf{h} \right). \quad (4)$$

Developing this equation leads to four terms. Three of which are constant or not informative, then there is only one informative term :

$$\nabla X_{i,j}^{(R)} \cdot \nabla X_{i,j}^{(G)}. \quad (5)$$

If only one channel has locally been altered, the gradient in this channel is modified. Consequently, the scalar product of



**Fig. 1:** Diagram illustrating the preprocessing steps to obtain correlations between gradients in each channel.

two channel gradients reflects the change in the cosine of the difference between the two gradient angles.

Similarly, we can apply the same computation for the red and blue channel and then obtain :

$$\nabla X_{i,j}^{(R)} \cdot \nabla X_{i,j}^{(B)} \quad (6)$$

As stated by Gouet et al. [16] (and following the Hilbert theory [15]), it is unnecessary to investigate the  $\nabla X_{i,j}^{(G)} \cdot \nabla X_{i,j}^{(B)}$  term, as it is already implicitly contained in the first two expressions (Eq. 5 and 6).

Our proposition in this paper is to add two new features depending on the normalized correlation between red and green channels, and red and blue channels as :

$$\mathcal{G}_{RG} = \frac{\nabla X_{i,j}^{(R)} \cdot \nabla X_{i,j}^{(G)}}{|\nabla X_{i,j}^{(R)}| |\nabla X_{i,j}^{(G)}|} \quad (7)$$

$$\mathcal{G}_{RB} = \frac{\nabla X_{i,j}^{(R)} \cdot \nabla X_{i,j}^{(B)}}{|\nabla X_{i,j}^{(R)}| |\nabla X_{i,j}^{(B)}|} \quad (8)$$

Fig. 1 illustrates our preprocessing steps to obtain normalized correlations between gradients of each channel. Note that gradients of each channel are estimated by convolution with a  $[-1; 1]$  mask (horizontal and vertical).

Then, our features are computed from these correlations  $\mathcal{G}_{RG}$  and  $\mathcal{G}_{RB}$ , by computing the co-occurrences as in the rich model [12]. We used different values of the quantization  $q \in \{0.1, 0.3, 0.5, 0.7, 0.9, 1\}$  with fixed truncation  $T=1$ . The reason for using these different values of quantization  $q$  is that  $\mathcal{G}_{RG}$  and  $\mathcal{G}_{RB}$  belong to  $[-1, 1]$ . Moreover, the use of these values gives more accurate features and avoids the generation of too many zero values caused by the truncation step in the co-occurrence vector.

For each quantization, we obtain 12 submodels, one symmetrized spam14h and one spam14v with 25 features

each, and one of minmax22h, minmax22v, minmax24, minmax34h, minmax34v, minmax41, minmax34, minmax48h, minmax48v, and one minmax54, with 45 features for each. All submodels are gathered in a one dimension vector to erect a dimensionality of  $(10 \times 45 + 2 \times 25) \times 6 = 3000$  features which are added to 18157 features of color rich models with fixed  $q = 1$  and  $T = 3$  in order to collect a final set of 21157 features.

### 3.4. Ensemble Classifiers

Modern steganalysis methods for digital images are based on extracted features and these methods need machine learning techniques to detect hidden messages. In our work, we choose ensemble classifiers [17] because of their efficient classification performance for large scale learning.

Kodovsky's et al. [17] proposed ensemble classifiers<sup>1</sup>, which is a new machine learning tool for steganalysis, consisting of many classifier  $L$  base learners ( $B_l$ ).

Each classifier is built from randomly selected subspace  $d_{sub}$ -dimensionals from the entire full  $d$ -dimension feature space. The authors use Fisher Linear Discriminants ( $FLD$ ) as base learners and the final decision is made by aggregating the decision of individual base learners.

Let  $d$  be a full dimension feature space. First, the classifiers construct a number  $L$  of  $FLD$  base learners ( $B_l$ ) with  $l \in \{1, \dots, L\}$ . Each one performs its learning on a subspace of  $d_{sub}$  dimension, where  $d_{sub} < d$ . The  $i^{th}$  image formed  $f_i \in \mathbb{R}^d$ , will be mapped  $\mathbb{R}^d \rightarrow \{0, 1\}$ , with 0 being a cover image and 1 a stego image.

In the learning phase, each classifier learns to map a feature vector  $f_i$ , to the correct class number:

$$FLD_l : \mathbb{R}^d \rightarrow \{0, 1\}$$

$$f_i \rightarrow FLD_l(f_i)$$

Each classifier uses the training database to compute the vector orthogonal to the hyperplane separating the two classes. For a test feature, the  $l^{th}$  base learner reaches its decision by computing a projection and comparing it to a threshold. After collecting all  $L$  decisions, the final classifier selected is the one which has received the most votes. Then the decision threshold of each base learner is adjusted to minimize the total detection error under an equal prior on the training data [17]:

$$P_E = \min_{P_{FA}} \frac{1}{2} [P_{FA} + P_{MD}(P_{FA})].$$

where  $P_{FA}$  represents the false alarm probability and  $P_{MD}$  the missed detection probability.

<sup>1</sup>Ensemble classifiers is available at <http://dde.binghamton.edu/download/ensemble>.

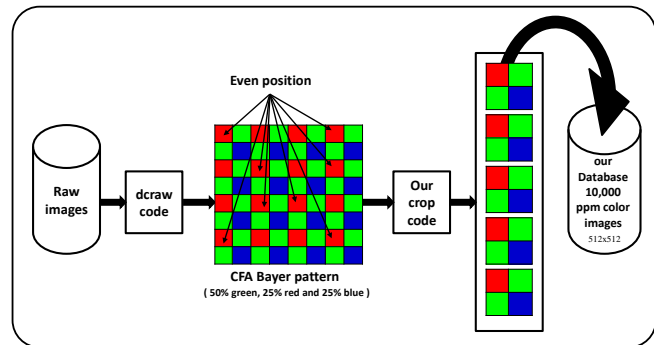


Fig. 2: The preprocessing steps for building our database depend on the  $CFA$  idea.

## 4. EXPERIMENTAL RESULTS

### 4.1. Building Database

A raw image is a class of computer files containing untouched pixel information coming from the digital camera sensor. These files hold a large amount of meta-information about the image generated by the camera [18].

In our work, the color image database is very carefully built depending on the  $CFA$  idea. We collected raw images from two subsets which are the most standard, and have the highest number of images captured (i.e. the Dresden Image Database [19] 3500 full-resolution Nikon digital camera raw color images and the Break Our Steganographic System (BOSSbase<sup>2</sup>) 1000 Canon digital camera raw color images).

In order to obtain color images in Portable Pixel Map ( $PPM$ ) format of size  $512 \times 512$ , all images take the same  $CFA$  map layout, as illustrated in Fig. 2. Two steps are required. The first step consists of using a demosaicking algorithm to convert raw images in demosaicked images. The second step consists of cropping five sub-images from one image. Fig. 3 shows sample images produced by the cropping step.

First we use demosaicking algorithm Patterned Pixel Grouping (PPG) in the dcrw platform<sup>3</sup> to be able to read numerous raw formats used to convert raw images in  $PPM$  and  $TIFF$  formats. As illustrated in Fig.2, the obtained images are such that the Bayer Pattern is always of type  $RGBG$  (red channel pixel is placed at even position). We wrote a spatial code in C language under Linux to start the crop from the red channel position. Indeed, from one image, this code randomly selected the red channel position and cropped five images of size  $512 \times 512$ , so that all blocks share the same  $CFA$  map layout. The final number of images used 10000  $PPM$  color images of size  $512 \times 512$ .

<sup>2</sup>BOSSbase can be accessed at <http://www.agents.cz/boss/BOSSfinal>.

<sup>3</sup>dcrw code is available at <http://www.cybercom.net/defin/dcrw>.

## 4.2. Experimental Protocol

Stego images are obtained using two steganography algorithms: the Spatial- UNiversal Wavelet Relative Distortion (S-UNIWARD<sup>4</sup>) steganography algorithm [20] and Wavelet Obtained Weights (WOW<sup>5</sup>) steganography algorithm [21]. As shown in [22], these two algorithms are highly adapted because they are selected for this strong resistance to steganalysis.

The payload is embedded in equal proportion in the three  $R, G$  and  $B$  channels. The different payload sizes are  $\{0.01, 0.05, 0.1, 0.2, 0.3, 0.4 \text{ and } 0.5\}$  Bit Per Channel ( $BPC$ ). Two sets of features are extracted from each image, the first set is the color rich model vector which consists of 18157 features of residuals among neighboring pixels (see section 3.1). The second set of features consists of 3000 proposed features from correlations between channels by applying gradient derivative filters on each channel, as explained in section 3.3. Then we add the two sets of features in a one dimensional vector to get a result of 21157 features of 10000 covers and 10000 stegos which are ready to enter in the classifier. The images are randomly divided into two halves for training and testing. A random subset of images, 5000 covers and 5000 stegos is used to train and test the classifier. The result of the classification is the average testing error over 10 splits  $\bar{P}_E$ . The rest of the cover and stego images are tested against the ensemble classifiers [17]. Then the decision values are collected for each. Given the decision values,  $ROC$  curves are obtained. The area under the  $ROC$  curves is calculated as the accuracy of the ensemble classifiers.

## 4.3. Results and Discussion

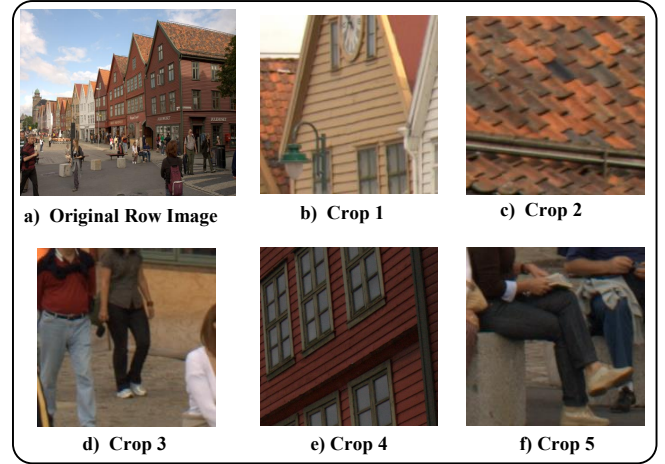
This section contains the experimental results of our proposed method. We illustrate these results in Tables 1 and 2. S-UNIWARD and WOW methods were tested with different relative payloads  $\{0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5\}$  Bits Per channel ( $BPC$ ) against the method of color rich model [4].

We used the same set of payload values with the same embedding methods. Our method, achieved higher performance by registering 87.54% and 86.63% detection rates for S-UNIWARD and WOW respectively with the payload 0.5 bpc. At the same time, the color rich model method [4] is less efficient because it achieved respectively 86.14% and 85.27% detection rates on the same sample. We noted the same thing with the rest of the payload values for both (the proposed and compared method), as shown in Tables 1 and 2.

Curves in Fig. 4 (a) and (b) also illustrate the comparison between the proposed method and the color rich model. As a result, the average testing error of the proposed method is less than that of the color rich model. That proves the importance

<sup>4</sup>S-UNIWARD steganography method is available at [http://dde.binghamton.edu/download/stego\\_algorithms/](http://dde.binghamton.edu/download/stego_algorithms/).

<sup>5</sup>WOW steganography method is available at [http://dde.binghamton.edu/download/stego\\_algorithms/](http://dde.binghamton.edu/download/stego_algorithms/).



**Fig. 3:** Sample images of our database built by random cropping from locations of red channel pixels (even position) in a Bayer pattern :

- a) Original RAW image  $3906 \times 2602$ ,
- b) crop 1 position  $x=2116, y=1928$ ,
- c) crop 2 position  $x=902, y=1182$ ,
- d) crop 3 position  $x=3080, y=436$ ,
- e) crop 4 position  $x=1866, y=1778$ ,
- f) crop 5 position  $x=650, y=1032$ .

of the additional 3000 features derived from the correlation between different channels proposed by our method.

Another experiment involved embedding the entire payload only in one channel of the color image, i.e. with payload 0.2 bpc in the green channel only. In this case, the detection rate records higher than the same payload if it is distributed equally between the three color channels. Table 3 illustrates the comparison of detection rates between the S-UNIWARD and WOW methods with payloads 0.2 bpc and 0.4 bpc embedded in one channel only and in the three channels separately. Fig. 5 (a) and (b), show the  $ROC$  curves, illustrating the performance of our method. This experiment revealed that it is easier to detect a hidden message in only one channel than a message shared in all channels.

Payload bpc	S-UNIWARD		WOW	
	Green	RGB	Green	RGB
0.2	87.92%	76.30%	86.41%	74.98%
0.4	94.10%	85.30%	93.34%	84.26%

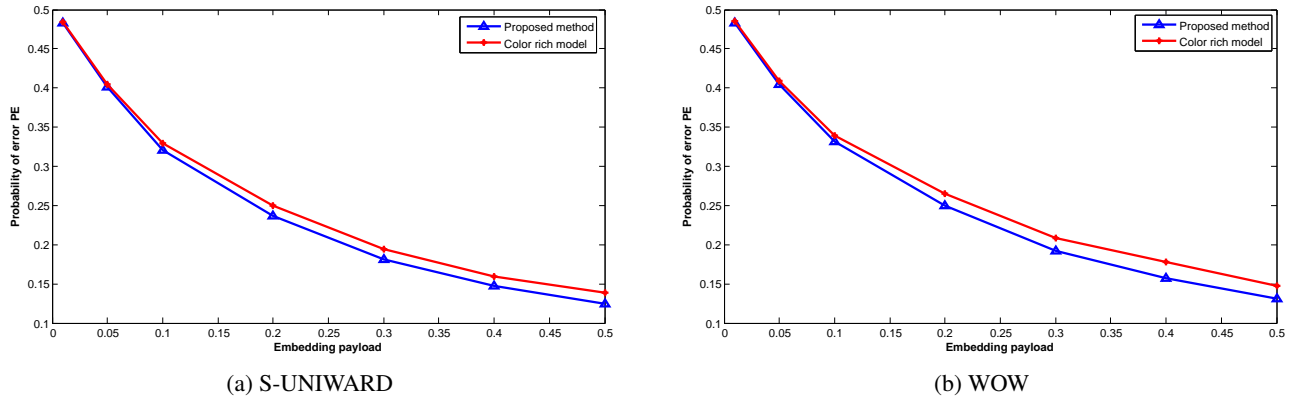
**Table 3:** Our steganalysis detection rate of S-UNIWARD and WOW steganography methods at 0.2 bpc and 0.4 bpc payload embedding in the green channel compares with equal embedding in three channels.

Embedding Payload	Proposed method Average testing error	Detection Rate%	Color rich model Average testing error	Detection Rate %
0.01	0.4830 ± 0.0010	51.70	0.4841 ± 0.0017	51.59
0.05	0.4010 ± 0.0032	59.90	0.4045 ± 0.0026	59.55
0.1	0.3203 ± 0.0023	67.97	0.3298 ± 0.0016	67.02
0.2	0.2370 ± 0.0031	76.30	0.2498 ± 0.0026	75.02
0.3	0.1808 ± 0.0034	81.92	0.1947 ± 0.0023	80.53
0.4	0.1470 ± 0.0025	85.30	0.1599 ± 0.0015	84.01
0.5	0.1246 ± 0.0023	87.54	0.1386 ± 0.0023	86.14

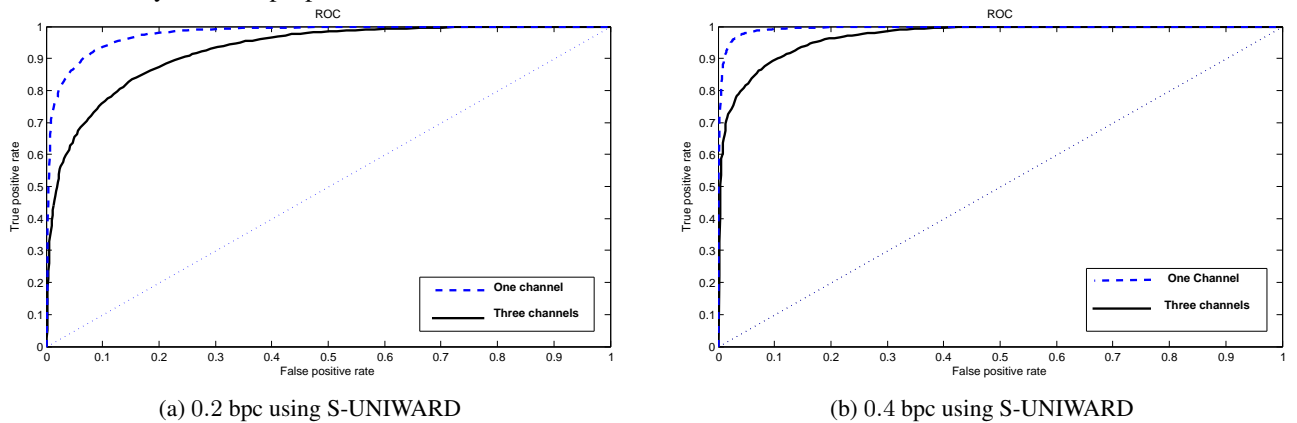
**Table 1:** The comparison between the steganalysis of the proposed method with the color rich model in different relative payloads for the S-UNIWARD steganography method: the average testing error for the proposed is less than for the color rich model method.

Embedding Payload	Proposed method Average testing error	Detection Rate%	Color rich model Average testing error	Detection Rate %
0.01	0.4836 ± 0.0017	51.64	0.4850 ± 0.0014	51.50
0.05	0.4042 ± 0.0027	59.58	0.4092 ± 0.0032	59.08
0.1	0.3317 ± 0.0034	66.83	0.3397 ± 0.0023	66.03
0.2	0.2502 ± 0.0041	74.98	0.2654 ± 0.0025	73.46
0.3	0.1918 ± 0.0013	80.82	0.2081 ± 0.0029	79.19
0.4	0.1574 ± 0.0021	84.26	0.1783 ± 0.0034	82.17
0.5	0.1307 ± 0.0015	86.63	0.1473 ± 0.0025	85.27

**Table 2:** The comparison between the steganalysis of the proposed method with the color rich model in different relative payloads for the WOW steganography method: the average testing error for the proposed is less than for the color rich model method.



**Fig. 4:** Detection error  $\bar{P}_E$  as a function of the payload for (a) S-UNIWARD and (b) WOW steganography methods illustrates good detectability with our proposed features.



**Fig. 5:** ROC curves of the proposed method for the payloads (a) 0.2 bpc (b) 0.4 bpc using the S-UNIWARD steganography method. The steganalysis technique detects easier hidden messages in only one channel more than equal embedding in the three channels.

## 5. CONCLUSION AND FUTURE WORK

Here we proposed and implemented a new method for color image steganalysis. The proposed method estimates correlations between the gradients of red, green and blue channels. Then these correlations are incorporated in the rich model using a co-occurrence matrix in order to obtain 3000 features. These features are added to those that are obtained from the color rich model in order to build a vector of a total of 21157 features. All feature vectors are fed to the ensemble classifiers. We used a quantization step with a set of values that differed from the color rich models. The ensemble classifiers is used to detect steganographic messages. Images for cover and stego are collected using the Dresden database and BOSSbase 1.0. Multiple comparisons are made between the proposed method with the color rich model using WOW and S-UNIWORD steganography methods in different payloads, reflecting the efficiency of the proposed features.

Our future work will focus on developing a new steganalysis method for digital color images, enhance the feature vector and calculate a new correlation between all channels of the color image.

## REFERENCES

- [1] J. Fridrich, *Steganography in digital media: principles, algorithms, and applications*, Cambridge University Press, 2009.
- [2] C. Hosmer and C. Hyde, *Discovering covert digital evidence*, In Digital Forensic Research Workshop (DFRWS), 2003.
- [3] J. Fridrich and M. Long, *Steganalysis of lsb encoding in color images*, In IEEE International Conference on Multimedia and Expo (ICME), pp. 12791282 vol.3, 2000.
- [4] M. Goljan, J. Fridrich, and R. Cogranne, *Rich model for steganalysis of color images*, In IEEE Workshop on Information Forensic and Security, GA, 2014.
- [5] A. Westfeld, *Detecting low embedding rates*, In 5th International Workshop on Information Hiding, pp. 324339, Springer, 2003.
- [6] A. D. Ker, *Resampling and the detection of lsb matching in color bitmaps*, In International Society for Optics and Photonics, pp. 115, 2005.
- [7] J. J. Harmsen and W. A. Pearlman, *Steganalysis of additive-noise modelable information hiding*, In Electronic Imaging International Society for Optics and Photonics, pp. 131142, 2003.
- [8] P. Thiyagarajan, G. Aghila, and V. P. Venkatesan, *Steganalysis using color model conversion*, arXiv preprint, arXiv:1206.2914, 2012.
- [9] S. Lyu and H. Farid, *Steganalysis using color wavelet statistics and one-class support vector machines*, In SPIE Symposium on Electronic Imaging, pp. 3545, 2004.
- [10] M. Kirchner and R. Bohme, *Steganalysis in technicolor boosting ws detection of stego images from cfainterpolated covers*, In IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 39823986, 2014.
- [11] J. Fridrich and M. Goljan, *On estimation of secret message length in lsb steganography in spatial domain*, In International Society for Optics and Photonics in Electronic Imaging, pp. 2334, 2004.
- [12] J. Fridrich and J. Kodovsky, *Rich models for steganalysis of digital images*, In IEEE Transactions on Information Forensics and Security, vol. 7, no. 3, pp. 868882, 2012.
- [13] B. E. Bayer, *Color imaging array*, US Patent 3,971,065, July 20 1976.
- [14] J. Wang, C. Zhang, and P. Hao, *New color filter arrays of high light sensitivity and high demosaicking performance*, In IEEE International Conference on Image Processing (ICIP), pp. 31533156, 2011.
- [15] D. Hilbert, *Theory of algebraic invariants*, Cambridge University Press, 1993.
- [16] V. Gouet, P. Montesinos, and D. Pel, *A fast matching method for color uncalibrated images using differential invariants*, In BMVC, pp. 110, 1998.
- [17] J. Kodovsky, J. Fridrich, and V. Holub, *Ensemble classifiers for steganalysis of digital media*, In IEEE Transactions on Information Forensics and Security, vol. 7, no. 2, pp. 432444, April 2012.
- [18] L. Yuan and J. Sun, *High quality image reconstruction from raw and jpeg image pair*, In IEEE International Conference on Computer Vision (ICCV), pp. 21582165, 2011.
- [19] T. Gloe and R. Bohme, *The dresden image database for benchmarking digital image forensics*, In ACM, pp. 15841590, 2010.
- [20] V. Holub, J. Fridrich, and T. Denemark, *Universal distortion function for steganography in an arbitrary domain*, In EURASIP J. Information Security, pp. 113, 2014.
- [21] V. Holub and J. Fridrich, *Designing steganographic distortion using directional filters*, In IEEE International Workshop on Information Forensics and Security (WIFS), pp. 234239, 2012.
- [22] V. Holub and J. Fridrich, *Digital image steganography using universal distortion*, In ACM Workshop on Information Hiding and Multimedia Security, pp.5968, 2013.