Color Image Steganalysis Based on Steerable Gaussian Filters Bank

Hasan ABDULRAHMAN
Montpellier University
Ecole des Mines d'Alès, LGI2P
30035 Nîmes Cedex 1, France.
Hasan.abdulrahman@mines-ales.fr

Philippe MONTESINOS
Ecole des Mines d'Alès, LGI2P
30035 Nîmes Cedex 1, France.
Philippe.montesinos@mines-ales.fr

Marc CHAUMONT
Nîmes University, Montpellier University
Place Gabriel Péri, 30000 Nîmes
UMR5506-LIRMM, 34095 Montpellier Cedex 5, France.
Marc.Chaumont@lirmm.fr

Baptiste MAGNIER
Ecole des Mines d'Alès, LGI2P
30035 Nîmes Cedex 1, France.
Baptiste.magnier@mines-ales.fr

ABSTRACT

This article deals with color images steganalysis based on machine learning. The proposed approach enriches the features from the Color Rich Model by adding new features obtained by applying steerable Gaussian filters and then computing the co-occurrence of pixel pairs. Adding these new features to those obtained from Color-Rich Models allows us to increase the detectability of hidden messages in color images. The Gaussian filters are angled in different directions to precisely compute the tangent of the gradient vector. Then, the gradient magnitude and the derivative of this tangent direction are estimated. This refined method of estimation enables us to unearth the minor changes that have occurred in the image when a message is embedded. The efficiency of the proposed framework is demonstrated on three stenographic algorithms designed to hide messages in images: S-UNIWARD, WOW, and Synch-HILL. Each algorithm is tested using different payload sizes. The proposed approach is compared to three color image steganalysis methods based on computation features and Ensemble Classifier classification: the Spatial Color Rich Model, the CFA-aware Rich Model and the RGB Geometric Color Rich Model.

Keywords

Steganalysis; steganography; steerable Gaussian filters.

1. INTRODUCTION

Steganography is the art and science of hiding messages inside a digital medium in such a way that only the sender and the receiver, can view the hidden message. The goal of steganalysis is to detect the presence of hidden messages in digital media. The practical performance (i.e. security) of different steganographic techniques is rated by the detectability (for example the average testing error of the steganalyser) for a given relative payload size; additionally the computational complexity could be taken into account [17].

In 2015, color steganalysis has been studied by integrating the modern adaptive embedding method in an experimental evaluation. Previous steganalysis methods did not use recent grey-level embedding algorithms [7], [25], [15], [19] or did not use a machine learning approach with Rich-Model features [11], [24].

Three color image steganalysis methods, based on a machine learning approach fed with rich model features are now well established. For these three methods, the machine learning algorithm is the Ensemble Classifier [16], and the rich features are: the Spatial Color Rich Model abbreviated to **CRM** [10], the CFA-aware Rich Model abbreviated to **CFARM** [9], and the RGB Geometric Color Rich Model abbreviated to **GCRM** (RGB for the Red, Green and Blue channels) [2]. Note that among these three features, the GCRM which is an extension of the CCRM (Correlation Color Rich Model [1]) seems to be the equivalent or better than the CRM and CFARM when recent adaptive grey-level embedding algorithms are used for embedding independently in each color channel. Additionally, note that a recent independent study has shown that CCRM was the most reliable method in every case, even when the algorithm was embedding adaptively with a synchronization of the embedding between the color channels [23].

At the beginning of 2016, **GCRM** is the natural choice for computing features that will be used for a color steganalysis (machine learning approach) for modern color embedding algorithms.

The rest of this paper is organized as follows. Section 2 describes the latest recent methods in color steganalysis by recalling the CRM [10], the CFARM [9] and the GCRM [2]. Then, we present a detailed description of our proposed method in Section 3; the steerable Gaussian filters computation is demonstrated and the feature set is presented. Experimental results and comparisons are given in Section 4. In this section, we also present the databases and show the performance of the proposed method. Finally, Section 5 gives some conclusions and perspectives.

2. COLOR STEGANALYSIS METHODS

Over the last ten years, most of the steganalysis methods are dealing with grayscale images [14], there exist very few steganalysis methods dealing with color images. The most recent methods

in color image steganalysis are explained in detail in the three following subsections.

2.1 Color rich model method

Goljan *et al.* [10] have introduced very efficient color image features, the CRM, which is an extension of the Spatial Rich Model [6], produced from two different sets of features.

In order to compute these features, firstly they extracted the noise residual from each color channel separately: the red (R), green (G), and blue (B) channels. On each channel they applied the same computation, as a grayscale image I. Let us note $I(x,y)$ a pixel value of an 8-bit grayscale cover image at coordinates (x,y). Then, the noise residual is computed using the following formula:

$$\mathbf{R}(x,y) = \hat{I}(x,y)(\mathcal{N}(x,y)) - c \cdot I(x,y), \qquad (1)$$

where: $c \in \mathbb{N}$, is the residual order, $\mathcal{N}(x,y)$ is a local neighborhood of pixel $I(x,y)$ and $\hat{I}(x,y)(\cdot)$ represents a predictor of $c \cdot I(x,y)$, with $I(x,y) \notin \mathcal{N}(x,y)$ and $I(x,y) \in \{0,....,255\}$. All of the submodels $\mathbf{R}(x,y) \in \mathbb{R}^{n_1 \times n_2}$ are formed from noise residual images using a size of $m \times m$, with n_1, n_2 and $m \in \mathbb{N}^*$, by applying a rounding and a truncation:

$$\mathbf{R}(x,y) \leftarrow tranc_T\left(round\left(\frac{R(x,y)}{q}\right)\right), \qquad (2)$$

where,

- $R(x,y) = \begin{cases} tranc_T(u) &= u, \quad \text{for } u \in [-T,T], u \in \mathbb{R} \\ tranc_T(u) &= T \cdot sign(u) \text{ otherwise.} \end{cases}$

- q represents the quantization step,

- $round$ is a function for rounding to an integer value.

Moreover, the color noise residuals are computed as in Eq.1 on the demosaiced image for the CRM features. Secondly, spatial co-occurrence are computed and the inter-channel co-occurrence matrices, as a feature set.

On the one hand, the Spatial Rich Model (SRMQ1) [6] with a fixed quantization $q = 1$ and a truncation $T = 2$ yields a dimensionality of $12,753$ features. These features are computed from each R, G and B color channel separately. On the other hand, from the same noise residuals (i.e. SRMQ1), a collection of 3D color co-occurrence matrices are built (CRMQ1), taking three color values at the same position and computing co-occurrence matrices across the three channels. With a fixed truncation $T = 3$ and a quantization $q = 1$, CRMQ1 produces 5404 features. Finally, for this method, the two sets of features are gathered in a one dimensional vector to produce $18,157$ features as a final set of features.

2.2 CFA-aware features method

Goljan *et al.* introduced in [9] the CFA-aware Color Rich Model, noted CFARM, for color image steganalysis. The features are made up of two parts, the first one is the CRM explained in the previous Section, with $T \in \{2,3\}$. The second part is the CFA-aware feature, consisting of three combinations: *RB/GG split, R/B/GG split* and *NII/INI split*. Then, in order to capture the inter-channel and intra-channel dependencies, four 3D co-occurrence matrices are built to extract features from and between the color channels, according to the structure of the Bayer CFA. In the CFARM approach [9], the authors assume that the upper left pixel corresponds to a non-interpolated pixel from the blue channel of the Bayer CFA.

Four index sets are introduced, corresponding to the geometric structure of the CFA map:

$$\mathcal{I}_B = \{(x,y)|x \text{ even}, y \text{ even}\}, \mathcal{I}_R = \{(x,y)|x \text{ odd}, y \text{ odd}\},$$
$$\mathcal{I}_{G1} = \{(x,y)|x \text{ odd}, y \text{ even}\}, \mathcal{I}_{G2} = \{(x,y)|x \text{ even}, y \text{ odd}\}.$$

Three combinations of features are generated to form the total number of features with the CRM set:

1) *RB/GG split* produces 4146 features,
2) *R/B/GG split* produces 10,323 features,
3) *NII/INI split* produces 5514 features.

Finally, all four sets of features are gathered in a one dimensional vector, which are ready to enter to the classifier.

2.3 RGB channel geometric method

Abdulrahman *et al.* [2] proposed the RGB Geometric Color Rich Model (GCRM). The authors show that if one channel has been affected by a steganography method, the inter channel correlation will measure the local modifications. On this basis, they proposed two types of features computed between color image channels. The first types of features reflects local Euclidean transformations through the computation of the cosine of the angle between channel gradients called ∇R, ∇G and ∇B (for red, green and blue). For example, the cosine of the rotation angle, between the red and the green channel gradients:

$$\mathcal{C}_{RG} = \frac{\nabla R \cdot \nabla G}{|\nabla R| \, |\nabla G|}. \qquad (3)$$

The second gives complementary information on the angle through the computation of the *sine* of the angle between channel gradients. Indeed, computing the horizontal and vertical image derivatives of all channels allows us to increase the steganalysis by computing local deformations between channels. For the red and green channels, let us note these derivatives R_x, G_x for the horizontal and R_y, G_y for the vertical ones. Thus, for the red and green channels, the sine is given by the following formula:

$$\mathcal{S}_{RG} = \frac{R_x \cdot G_y - R_y \cdot G_x}{|\nabla R| \, |\nabla G|}. \qquad (4)$$

At the end, they obtained 4 geometrical measures: $\mathcal{C}_{RG}$, cosine between R and G, $\mathcal{C}_{RB}$, cosine between R and B, $\mathcal{S}_{RG}$, sine between R and G, and $\mathcal{S}_{RB}$ between R and B. Those geometrical measurements allow to generate 6000 features, based on local Euclidean and mirror transformations, when using co-occurrence matrices with a fixed truncation $T=1$ and different values for the quantization $q \in \{0.1, 0.3, 0.5, 0.7, 0.9, 1\}$. Concatenate these features with those from CRM [10] increases the detectability of hidden messages in color images [2], [23].

3. THE PROPOSED METHOD

In order to be less visible, most of the steganographic methods modify the pixel values in the texture/edge areas [22], [13], [18]... Our proposition is to enrich the CRM method by introducing new sets of features obtained by applying steerable Gaussian filters and then computing the co-occurrence of pixel pairs in eight different directions.

The proposed features are composed of two distinctive sets. The first set, produced by [10], is made of $18,157$ features. The second is made of 4406 features. In the first step, we computed a *tangent vector* to contour for each pixel and for each channel. This *tangent vector* corresponds to the edge direction and is orthogonal to the gradient vector. Then, in the second step, the co-occurrence matrices are computed, firstly, on the three gradient magnitude images and afterwards, on the three derivative images related to the *tangent vectors*.

3.1 Steerable Gaussian filters

In the domain of image analysis, the estimation of a precise gradient is crucial, and is often based on the computation of local derivatives.
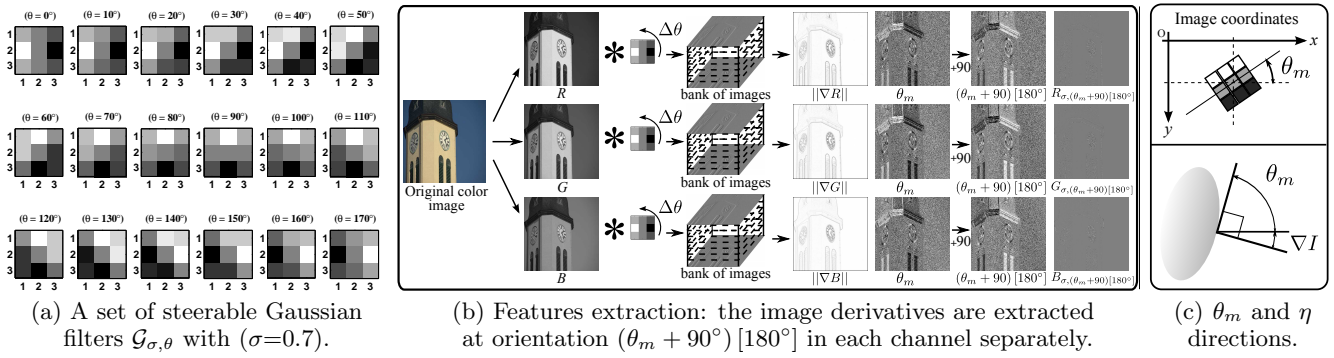
(a) A set of steerable Gaussian filters $\mathcal{G}_{\sigma,\theta}$ with $(\sigma=0.7)$.

(b) Features extraction: the image derivatives are extracted at orientation $(\theta_m + 90°)\,[180°]$ in each channel separately.

(c) θ_m and η directions.

Figure 1: Steerable filters are used to compute a gradient and to estimate precise edges directions.

When using filters with image derivatives in only two directions, the x and y (i.e $0°$ and $90°$), the gradient estimation is not accurate enough to describe the geometrical structures in the image. Using an orientation filter bank can improve the quality of the gradient estimation; indeed, its orientation and its magnitude are far more accurate.

Due to multiple orientations, a filter bank allows us to better detect image features such as edges. One of the most popular filter banks is the steerable filters. As a solution to the above stated problem, Freeman and Adelson [5] introduced an elegant way for steerable filters that can be directed at specific angles using a linear combination of isotropic filters like Gaussian derivatives. Let us note the basic derivatives of Gaussian filters $\partial\mathcal{G}_\sigma/\partial x$ and $\partial\mathcal{G}_\sigma/\partial y$ along the x-axis and y-axis respectively, for example:

$$\begin{cases} \dfrac{\partial\mathcal{G}_\sigma(x,y)}{\partial x} = \dfrac{-x}{2\pi\sigma^4}\cdot e^{-\frac{x^2+y^2}{2\sigma^2}} \\[3mm] \dfrac{\partial\mathcal{G}_\sigma(x,y)}{\partial y} = \dfrac{-y}{2\pi\sigma^4}\cdot e^{-\frac{x^2+y^2}{2\sigma^2}}, \end{cases} \quad (5)$$

with σ the standard-deviation of the Gaussian filter.

Freeman and Adelson have shown that the first order directional Gaussian derivative $\mathcal{G}_{\sigma,\theta}$ at an angle θ can be generated by a linear combination of a rotation of the basic derivatives of isotropic Gaussian filters (illustrated in Fig. 1 (a) and (b)):

$$\mathcal{G}_{\sigma,\theta}(x,y,\sigma) = \cos(\theta)\cdot\frac{\partial\mathcal{G}_\sigma}{\partial x}(x,y) + \sin(\theta)\cdot\frac{\partial\mathcal{G}_\sigma}{\partial y}(x,y). \quad (6)$$

The image derivative $I_{\sigma,\theta}$ is obtained by convolving the original grayscale image I with the oriented Gaussian kernels $\mathcal{G}_{\sigma,\theta}$:

$$I_{\sigma,\theta}(x,y) = (I * \mathcal{G}_{\sigma,\theta})(x,y). \quad (7)$$

Finally, the gradient magnitude $\|\nabla I(x,y)\|$ is calculated as the maximum absolute value response to the oriented operator $\mathcal{G}_{\sigma,\theta}$:

$$\|\nabla I(x,y)\| = \max_{\theta\in[0,180[}(|I_{\sigma,\theta}(x,y)|), \quad (8)$$

$$\theta_m = \arg\max_{\theta\in[0,180[}(|I_{\sigma,\theta}(x,y)|). \quad (9)$$

Note that θ_m, represents the *kernel angle* and it differs from the *gradient angle* which is equal to $(\theta_m + 90°)\,[180°]$.

In this work, the Gaussian filters are angled in different directions to compute the more precise gradient magnitude $\|\nabla I\|$ and its associated kernel angle θ_m. Thus, $\|\nabla I\|$ corresponds to the absolute value of the image derivative for the kernel angled at θ_m, as illustrated in Fig. 1(c). Note that $\|\nabla I\|$ and θ_m are different for each pixel of I. These techniques are applied to the three color channels R, G and B to obtain three gradient magnitude images $\|\nabla R\|$, $\|\nabla G\|$ and $\|\nabla B\|$ (see Fig. 1(b)).

As pointed out previously, the steganographic methods essentially modifies the pixel values in the textures and edge areas. For the edge areas, the embedding modifications have to be detected along the "isophote" lines i.e. along the curves of constant intensity when considering an image as a surface. This led us to consider the orthogonal vector to the gradient named the *tangent vector*, instead of the gradient; this means that the derivative along the edge must be computed. This derivative corresponds to the result of the convolution of the image with the steerable kernel angled at $(\theta_m + 90°)\,[180°]$ and is orthogonal to the kernel used for the gradient estimation (as illustrated in Fig. 2). The derivative image is named $I_{\sigma,(\theta_m+90°)[180°]}$.

For a color image, each channel is considered separately. The tangent derivatives[1], as illustrated in Fig. 1 (b), are respectively computed for each pixel at position (x,y) of each channel and named: $R_{\sigma,(\theta_m+90°)[180°]}(x,y)$ for the red, $G_{\sigma,(\theta_m+90°)[180°]}(x,y)$ for the green, and $B_{\sigma,(\theta_m+90°)[180°]}(x,y)$ for the blue channel. Fig. 1 (b) shows an example of steerable Gaussian filters used to compute these new features. In our method, the Gaussian filters are
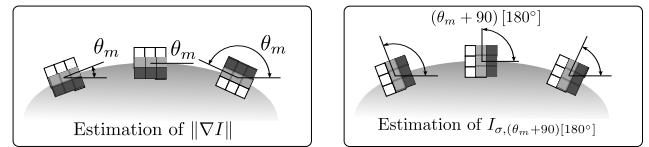

Figure 2: Positions of the steerable filters at the level of the edges to build the gradient image $\|\nabla I\|$ and image derivative $I_{\sigma,(\theta_m+90°)[180°]}(x,y)$.

angled in different directions to compute a precise gradient and a precise derivative along isophote lines. In order to detect the slight changes in the images, our experiments (see Section 4.2) leads to a $\sigma = 0.7$, with a filter support for the size 3×3 pixels, a rotation step for the filters bank $\Delta\theta = 10°$, and a rotation range such as $\theta \in \{0°, ..., 180° - \Delta\theta\}$ (it leads to 18 filter orientations as represented in Fig. 1 (a)).

3.2 Complete feature set

As explained in Fig.1 (b), Eq.8 and detailed above, the co-occurrence matrices are completed from the three gradient magnitude images

---

[1] As these three images are derivatives, pixel values can be positives or negatives.

and also the three derivative images using their edge directions:

$$\begin{cases} \|\nabla R\|, & \|\nabla G\|, & \|\nabla B\|, \\ R_{\sigma,(\theta_m+90)[180°]}, G_{\sigma,(\theta_m+90)[180°]}, B_{\sigma,(\theta_m+90)[180°]}. \end{cases}$$

Before we make the co-occurrence computation, different truncations are applied. For the gradient magnitude images, the truncation $T \in \{2,3\}$, and for the derivative images, $T \in \{1,2,3\}$. Thus this leads to 2 triplets of quantized-truncated gradient magnitude images, and 3 triplets of quantized-truncated derivative images. Then, the pairs co-occurrence matrices are computed, like the Subtractive Pixel Adjacency Model (**SPAM**) [21], with eight directions for the scan images. These direction feature subsets are as follows: $F_h^{\rightarrow}, F_h^{\leftarrow}, F_v^{\uparrow}, F_v^{\downarrow}, F_d^{\nearrow}, F_d^{\swarrow}, F_{md}^{\nwarrow}$ and $F_{md}^{\searrow}$. 2808 features are created by the gradient magnitude and 1598 by the derivative images. All features are gathered in to a one dimension vector to erect a dimensionality of 4406 features. As a final set for the proposed method, $22,563$ features are obtained, by concatenating these features with those obtained from CRM [10].

4. EXPERIMENTAL RESULTS

All the experiments were carried out on $10,000$ color images of size 512×512. All detectors were trained as binary classifiers implemented using the ensemble classifier[2] [16] with default settings.

4.1 Setup database

In order to build our image database, the full-resolution raw images were collected from two subsets which are the most standard (i.e. 3500 full-resolution Nikon digital camera raw color images from the Dresden Image Database [8] and 1000 Canon digital camera raw color images from the Break Our Steganographic System (BOSSbase1[3]) [3]). Afterwards, the RGB color images are obtained by using the Patterned Pixel Grouping (PPM) demosaicking algorithm named "dcraw"[4]. Finally, from each color RGB image, we randomly extracted five cropped images measuring 512×512. Finally, the cropped images having the higher variation correspond to those exploited in the used steganography algorithms. As a result, the final number of RGB cropped images is $10,000$. Moreover, in this last process, the cropped images are extracted carefully so that the CFA map layout always stays the same, as illustrated in Fig. 3.

Three spatial domain steganography algorithms are used to produce stego images:

- Spatial UNIversal WAvelet Relative Distortion (S-UNIWARD[5]) [13],
- Wavelet Obtained Weights (WOW[6]) [12],
- Synchronizing Selection Channel (Synch-HILL[7]) [4].

As explained in [20], the steganography methods have an important impact on the performance of the general methodology. These algorithms are used to embed messages into color images by decomposing the R, G and B channels like three gray-scale images

[2]The Ensemble classifier is available at http://dde. binghamton.edu/download/ensemble.

[3]BOSSbase can be accessed at http://www.agents.cz/boss/BOSSFinal.

[4]dcraw code is available at http://www.cybercom.net/defin/dcraw.

[5]The Matlab version of S-UNIWARD is available at http://dde.binghamton.edu/download/stego_algorithms.

[6]The Matlab version of WOW is available at http://dde.binghamton.edu/download/stego_algorithms.

[7]The Matlab version of Synch-HILL is available at http://dde.binghamton.edu/download/stego_algorithms.
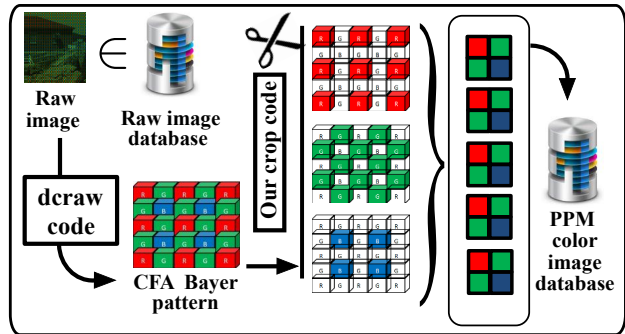

Figure 3: The preprocessing steps for building our database depend on the CFA idea.

and embedding the same proportion payload into each channel. Finally, $10,000$ color images were used to test each of the seven different payload sizes: $\{0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5\}$ Bit Per Channel (**BPC**).

4.2 Performance of the proposed method

To evaluate the performance of the proposed method, all steganalyzers were implemented as binary classifiers using the ensemble classifier [16]. In this paper, the detection accuracy is measured by using the average of testing errors under equal priors: $\bar{P}_E = min_{P_{FA}} \frac{1}{2} [P_{FA} + P_{MD}(P_{FA})]$, where, P_{FA} represents the false alarm probability and P_{MD} the missed detection probability. 5000 cover images are randomly chosen from the database for the training sets and 5000 stego images for the testing sets. The ensemble classifiers apply a vote to estimate the error of detection. This process is repeated 10 times to obtain $\bar{P}_E$ which quantify the detectability and are collected for each method and each payload in order to evaluate the steganalysis method.

As explained in Section 3.1, the experiments were run in such a way as to find the best filter bank parameters: the scale of the steerable filters σ (Eq. 6) and the rotation step ($\Delta\theta$) with angles evenly drawn from $0°$ to $180°$. These experiments have been led using S-UNIWARD with a payload size of 0.3 bpc with the use of the $22,563$ features obtained by concatenating the CRM with the gradient features. Table 1 shows that $\sigma = 0.7$ with $\Delta\theta = 10°$ corresponds to the optimal parameters for the steerable filters operation in this steganalysis work because, compared to the other parameters, they bring the best detection rate.

Table 1: The probability of error $\bar{P}_E$ to determine the efficient standard deviation (σ) and angle step ($\Delta\theta$) employed for steerable Gaussian filters using S-UNIWARD steganography method payload 0.3 bpc.

$\Delta\theta$	σ	Mask size	$\bar{P}_E$	Detection rate
$10°$	**0.7**	3×3	0.1559 ± 0.0022	84.41 %
$10°$	1	5×5	0.1896 ± 0.0031	81.04 %
$10°$	2	10×10	0.2028 ± 0.0037	79.72 %
$10°$	3	15×15	0.2539 ± 0.0036	76.41 %
$5°$	0.7	3×3	0.1768 ± 0.0026	82.32 %
$10°$	0.7	3×3	0.1559 ± 0.0022	84.41 %
$15°$	0.7	3×3	0.1602 ± 0.0026	83.98 %
$20°$	0.7	3×3	0.1653 ± 0.0019	83.47 %
$30°$	0.7	3×3	0.1854 ± 0.0027	81.46 %
$45°$	0.7	3×3	0.1893 ± 0.0012	81.07 %
$90°$	0.7	3×3	0.1996 ± 0.0031	80.04 %

Table 2: Error probability $\bar{P}_E$ and the detection rate $\mathcal{P}_D\%$ for four steganography methods.

Steganography Method	Payload (bpc)	Proposed Method $\bar{P}_E$	22,563 Dim $\mathcal{P_D}\%$	CRM $\bar{P}_E$	18,157 Dim $\mathcal{P_D}\%$	CFARM $\bar{P}_E$	27,460 Dim $\mathcal{P_D}\%$	GCRM $\bar{P}_E$	24,157 Dim $\mathcal{P_D}\%$
	0.01	0.4664	**53.36**	0.4841	51.59	0.4863	51.37	0.4680	53.20
	0.05	0.3835	**61.65**	0.4045	59.55	0.4072	59.28	0.3859	61.41
	0.1	0.2984	**70.16**	0.3298	67.02	0.3194	68.06	0.3037	69.63
S-UNIWARD	0.2	0.2164	**78.36**	0.2498	75.02	0.2317	67.83	0.2191	78.09
	0.3	0.1559	**84.41**	0.1947	80.53	0.1806	81.94	0.1623	83.77
	0.4	0.1202	**87.98**	0.1599	84.01	0.1429	85.71	0.1289	87.11
	0.5	0.1117	**88.83**	0.1386	86.14	0.1239	87.61	0.1124	88.76
	0.01	0.4687	**53.13**	0.4850	51.50	0.4875	51.25	0.4753	52.47
	0.05	0.3854	**61.46**	0.4092	59.08	0.4174	58.26	0.3906	60.94
	0.1	0.3091	**69.09**	0.3397	66.03	0.3275	67.25	0.3161	68.39
WOW	0.2	0.2269	**77.31**	0.2654	73.46	0.2440	75.60	0.2381	76.19
	0.3	0.1685	**83.15**	0.2081	79.19	0.1895	81.05	0.1793	82.07
	0.4	0.1377	**86.23**	0.1783	82.17	0.1487	85.13	0.1384	86.16
	0.5	0.1206	**87.94**	0.1473	85.27	0.1296	87.04	0.1207	87.93
	0.01	0.4651	**53.49**	0.4893	51.07	0.4843	51.57	0.4687	53.13
	0.05	0.3647	**63.53**	0.3991	60.09	0.4030	59.70	0.3720	62.80
	0.1	0.2946	**70.54**	0.3311	66.89	0.3189	68.11	0.3086	69.14
Synch-HILL	0.2	0.2113	**78.87**	0.2595	74.05	0.2394	76.06	0.2269	77.31
	0.3	0.1536	**84.64**	0.1997	80.03	0.1753	82.47	0.1607	83.93
	0.4	0.1294	**87.06**	0.1684	83.16	0.1478	85.22	0.1311	86.89
	0.5	0.1125	**88.75**	0.1475	85.25	0.1258	87.42	0.1193	88.07

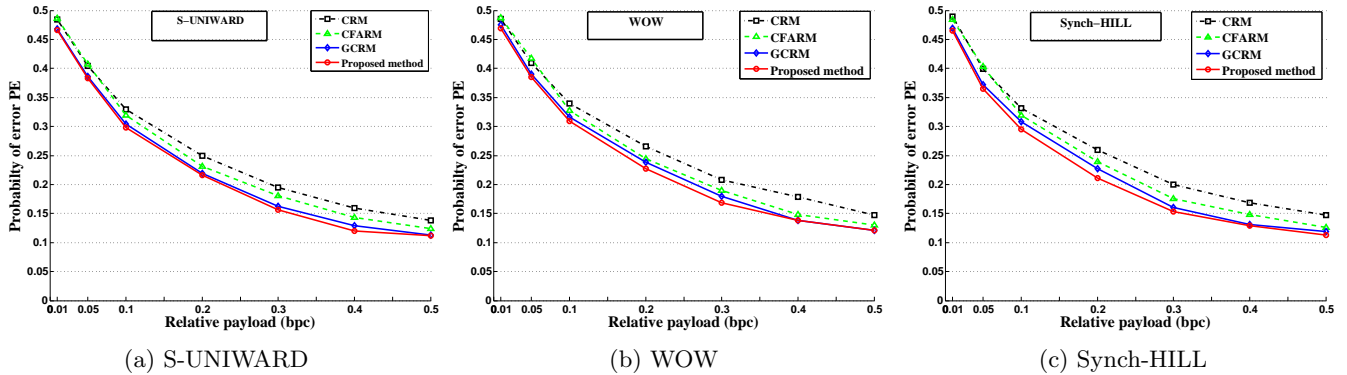(a) S-UNIWARD (b) WOW (c) Synch-HILL

Figure 4: Error probability $\bar{P}_E$ as a function of the payload for four steganography methods.

The experimental results are given in Table 2. Three algorithms have been tested: S-UNIWARD, WOW and Synch-HILL with different relative payloads sizes: $\{0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5\}$. Furthermore, the proposed method is tested against three other approaches: CRM [10], CFARM [9] and GCRM [2].

Table 2 demonstrates that the proposed method registered the highest performance. As an example, detection rates for, at a time, S-UNIWARD, WOW and Synch-HILL for a payload of 0.5 bpc are 88.83%, 87.94% and 88.75% respectively using the proposed approach, to the contrary of the other three compared methods. The CRM method [10] achieved 86.14%, 85.27% and 85.25% respectively. The CFARM method [9] achieved 87.61%, 87.04% and 87.42% respectively. Additionally, the GCARM method [2] achieved 88.76%, 87.93% and 88.07% respectively. This performance is due to the Gaussian filters bank, created by the steerable filters, which allow a more precise estimation of the gradient and its associated tangent vector.

In order to increase the detectability rate of the GCRM method [2], another experiment has been performed by concatenating the GCRM [2] features with the new proposed features in one dimen-

sional vector to produce $28,563$ features. These new dimensional vectors achieved 85.03% for S-UNIWARD, and 85.07% for Synch-HILL steganography methods payload 0.3 bpc as a detection rate. It obtains a difference of 1.26%, 1.14% respectively more than GCRM [2] alone and this result is close (slightly better) to our proposed approach results (less than 1%).

5. CONCLUSIONS

This article presents new features for color image steganalysis. Applying a Gaussian filters bank, to an order of 1, in different directions, enabled us to detect the slight changes in the images which occurred as a result of embedding the message.

The proposed approach treats the three color channels separately. Firstly, as steerable filters estimate precisely the edge directions in images, features correspond to the three image derivatives along the edges and the three gradient magnitude images. Secondly, features are extracted from these six images using the co-occurrence matrices of pixel pairs. Finally, our proposed features are integrated with the CRM features [10] to get the new approach. To evaluate the per-

formance of the new approach, the embedding algorithms used are S-UNIWARD, WOW, and Synch-HILL at different payloads.

Experimental results show that fusing proposed features with those obtained by CRM allows in the majority of cases, the detectability of hidden messages in the color images. Additionally, the new approach achieved higher detection rates than the three recent steganalysis approaches: CRM, CFARM, and GCRM. This observed detection improvement is due to a fine estimation of the tangent vector which is used for the estimation of the image derivatives in the edges directions. The proposed features allow the Ensemble Classifier to reveal the hidden message between the stego and cover images.

Eventually, future works consist of better understanding WOW-CMD-C or HILL-CMD-C [23] embedding algorithms, which synchronize the color selection channel during the embedding process.

6. REFERENCES

[1] H. Abdulrahman, M. Chaumont, P. Montesinos, and B. Magnier. Color image steganalysis using correlations between rgb channels. In *Proc. 10th Int. Conf. on Availability, Reliability and Security (ARES), 4th Int. Workshop on Cyber Crime (IWCC), Toulouse, France*, pages 448–454. IEEE, Aug. 24- 28, 2015.

[2] H. Abdulrahman, M. Chaumont, P. Montesinos, and B. Magnier. Color images steganalysis using rgb channel geometric transformation measures. *Wiley J. on Security and Communication Networks*, (DOI: 10.1002/sec.1427):12 pages, Feb. 2016.

[3] P. Bas, T. Filler, and T. Pevnỳ. "Break our steganographic system": The ins and outs of organizing boss. *Inf. Hiding, 13th Int. Workshop, Lecture Notes in Computer Science, Prague, Czech Republic*, pages 59–70, May 2011.

[4] T. Denemark and J. Fridrich. Improving steganographic security by synchronizing the selection channel. In *Proc. of the 3rd ACM Workshop on Inf. Hiding and Multimedia Security (IH&MMSec), Portland, Oregon*, pages 5–14, June 2015.

[5] W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE Trans. on Pattern Analysis & Machine Intelligence*, Vol.13(9):pp.891–906, 1991.

[6] J. Fridrich and J. Kodovskỳ. Rich models for steganalysis of digital images. *IEEE Trans. on Inf. Forensics and Security*, vol.7(no.3):pp.868–882, Jun. 2012.

[7] J. Fridrich and M. Long. Steganalysis of lsb encoding in color images. In *IEEE Int. Conf. on Multimedia and Expo (ICME) 2000, New York, NY, USA*, volume Vol.3, pages 1279–1282, July 2000.

[8] T. Gloe and R. Böhme. "The dresden image database" for benchmarking digital image forensics. *In Proc. ACM Symp. on Applied Computing, Sierre, Switzerland*, Vol.2:pp.1584–1590, Mar. 2010.

[9] M. Goljan and J. Fridrich. CFA-aware features for steganalysis of color images. In *Proc. IS&T/SPIE Electronic Imaging, Int. Society for Optics and Photonics (SPIE), San Francisco, CA, USA*, volume 94090V, page (13), Feb. 2015.

[10] M. Goljan, J. Fridrich, and R. Cogranne. Rich model for steganalysis of color images. *In Proc. IEEE Int. Workshop on Inf. Forensics Security, Atlanta, GA, USA*, pages 185–190, Dec. 2014.

[11] J. J. Harmsen and W. A. Pearlman. Steganalysis of additive-noise modelable information hiding. In *Proc. SPIE Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents V, Santa Clara, CA, USA*, pages 131–142, Jan. 2003.

[12] V. Holub and J. Fridrich. Designing steganographic distortion using directional filters. In *Proc. IEEE Int. Workshop on Inf. Forensics and Security (WIFS), Tenerife, Spain*, pages 234–239, Dec. 2012.

[13] V. Holub, J. Fridrich, and T. Denemark. Universal distortion function for steganography in an arbitrary domain. *EURASIP J. on Inf. Security*, Vol.2014(no.1):pp.1–13, Jan. 2014.

[14] A. D. Ker, P. Bas, R. Böhme, R. Cogranne, S. Craver, T. Filler, J. Fridrich, and T. Pevnỳ. Moving steganography and steganalysis from the laboratory into the real world. In *Proc. 1st ACM workshop on Inf. hiding and multimedia security (IH&MMSec), Montpellier, France*, pages 45–58, June 17-19, 2013.

[15] M. Kirchner and R. Bohme. "Steganalysis in technicolor" boosting ws detection of stego images from CFA-interpolated covers. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy*, pages 3982–3986, May 2014.

[16] J. Kodovskỳ, J. Fridrich, and V. Holub. Ensemble classifiers for steganalysis of digital media. *IEEE Trans. on Inf. Forensics and Security*, Vol.7(no.2):432–444, Apr. 2012.

[17] B. Li, J. He, J. Huang, and Y. Q. Shi. A survey on image steganography and steganalysis. *J. of Inf. Hiding and Multimedia Signal Process.*, Vol.2:pp.142–172, Apr. 2011.

[18] B. Li, M. Wang, J. Huang, and X. Li. A new cost function for spatial image steganography. *In Proc. IEEE, Int. Conf. Image Processing (ICIP), Paris, France*, pages 4206–4210, Oct. 2014.

[19] Q. Liu, A. H. Sung, B. Ribeiro, M. Wei, Z. Chen, and J. Xu. Image complexity and feature mining for steganalysis of least significant bit matching steganography. *J. of Information Sciences*, Vol.178(1):21–36, Jan. 2008.

[20] Y. Miche, P. Bas, A. Lendasse, C. Jutten, and O. Simula. Reliable steganalysis using a minimum set of samples and features. *EURASIP J. on Inf. Security*, vol.2009, article ID 901381:(13), 2009.

[21] T. Pevnỳ, P. Bas, and J. Fridrich. Steganalysis by subtractive pixel adjacency matrix. *IEEE Trans. on Inf. Forensics and Security (TIFS)*, Vol.5(no.2):215–224, June 2010.

[22] T. Pevnỳ, T. Filler, and P. Bas. Using high-dimensional image models to perform highly undetectable steganography. *In Proc. 12th Int. Workshop Inf. hiding, Calgary, AB, Canada*, vol.6387:161–177, Jun. 2010.

[23] W. Tang, B. Li, W. Luo, and J. Huang. Clustering steganographic modification directions for color components. *Signal Processing Letters,IEEE*, Vol.23(No.2):197–201, Feb. 2016.

[24] P. Thiyagarajan, G. Aghila, and V. P. Venkatesan. Steganalysis using color model conversion. *Int. J. of Signal and Image Processing (SIPIJ)*, Vol.2(No.4), Dec. 2011.

[25] A. Westfeld and A. Pfitzmann. Attacks on steganographic systems. In *Proc. on Information Hiding, In: Pfitzmann A. (eds.): 3rd Int. Workshop Lecture Notes in Computer Science, Springer-Verlag, Berlin Heidelberg*, volume 1768, pages 61–76, 2000.