## IEEE COPYRIGHT AND CONSENT FORM

To ensure uniformity of treatment among all contributors, other forms may not be substituted for this form, nor may any wording of the form be changed. This form is intended for original material submitted to the IEEE and must accompany any such material in order to be published by the IEEE.Please read the form carefully and keep a copy for your files.

### Author Online Use

6. Personal Servers. Authors and/or their employers shall have the right to post the accepted version of IEEE-copyrighted articles on their own personal servers or the servers of their institutions or employers without permission from IEEE, provided that the posted version includes a prominently displayed IEEE copyright notice and, when published, a full citation to the original IEEE publication, including a link to the article abstract in IEEE Xplore.Authors shall not post the final, published versions of their papers.

7. Classroom or Internal Training Use. An author is expressly permitted to post any portion of the accepted version of his/her own IEEE-copyrighted articles on the authors personal web site or the servers of the authors institution or company in connection with the authors teaching, training, or work responsibilities, provided that the appropriate copyright, credit, and reuse notices appear prominently with the posted material. Examples of permitted uses are lecture materials, course packs, e-reserves, conference presentations, or in-house training courses.

# STATISTICAL DETECTION OF JSTEG STEGANOGRAPHY USING HYPOTHESIS TESTING THEORY

*Tong Qiao, Cathel Zitzmann, Florent Retraint, Rémi Cogranne*

ICD - LM2S - Université de Technologie de Troyes (UTT) - UMR STMR CNRS,
12 rue Marie Curie - CS 42060 - 10004 Troyes cedex - France.
E-mail : {tong.qiao, cathel.zitzmann, florent.retraint, remi.cogranne}@utt.fr

**ABSTRACT**

This paper investigates the statistical detection of Jsteg steganography. The approach is based on the statistical model of Discrete Cosine Transformation (DCT) coefficients. The hidden information detection problem is case in the framework of Hypothesis testing theory. In an ideal context where all model parameters are perfectly known, the Likelihood Ratio Test (LRT) is presented and its performance are theoretically established. The statistical performance of LRT serve as an upper bound of the detection power. For a practical use, when the distribution parameters are unknown, a detector based on estimation of those parameters is designed. The loss of power of the proposed detector, compared with the optimal LRT is small, which shows the relevance of the proposed approach.

***Index Terms—*** Hypothesis testing theory, Jsteg steganalysis, DCT distribution model, hidden information detection.

## 1. INTRODUCTION

### 1.1. State of the Art

Steganalysis has received a great focus in the past decades. As of December 2011, WetStone declared that about 70 percent of the available steganographic software are based on the least significant bit (LSB) replacement algorithm [1] which consists in replacing LSB plane, in the spatial domain or frequency domain, of the cover media by the bits of the secret message (see [2–4]). Although the LSB replacement steganalysis method (see [5–10]) has been studied for many years, it can be noted that most of the prior-art detectors are designed to detect data hidden in the spatial domain and for only a few statistical properties are established, referred to as the optimal detectors in [11].

In 2004, the weighted stego-image (WS) method [12] and the test proposed in [13] for LSB replacement steganalysis changed the situation open the way to optimal detectors. Driven by this pioneer works, the enhanced WS algorithm have been improved in [14]. Inspired by the prior studies [12, 14], the WS steganalyser for JPEG covers was proposed in [15]. However, the WS steganalyser does not allow to get high detection performance for low False Alarm Rate (FAR), see [7], and its statistical properties remains unknown, which prevents the guarantee of a prescribed FAR. In practical forensic cases, since a large database of images needs to be processed, the getting of a very low FAR is crucial.

### 1.2. Contributions of the Paper

In this context, the detector proposed in [16] is an interesting alternative; however it is based on the assumption that Discrete Cosine Transformation (DCT) coefficients are independent and identically distributed (i.i.d) within a subband and have a zero expectation which might be inaccurate and hence make the detection performance poor in practice. On the opposite, this paper proposes a statistical model that assumes that each DCT coefficient has a different expectation and difference variance. The use of this model, together with hypothesis theory, allows us to optimize the most powerful Likelihood Ratio Test (LRT) when the distribution parameters (expectation and variance) are known. Then in the practical case of not knowing those parameters, estimations have to be used instead; this leads to the design of the proposed detector with estimated parameters. By taking into account those distribution parameters as nuisance parameters and using accurate estimation, it is shown that the loss of power compared with the optimal detector is small.

### 1.3. Organization of the paper

The paper is organized as follows. Section 2 recalls principle of Jsteg Steganography. Then, the DCT distribution model is discussed. In Section 3, based on the proposed model, the LRT for Jsteg detection is proposed. Section 4 presents the proposed practical detector with estimated parameters. Finally, Section 5 presents numerical results of the proposed steganalysers on the simulated and real images and Section 6 concludes this paper.

The Ph.D thesis of Tong Qiao is funded by the China Scholarship Council (CSC) program.

## 2. JPEG STEGANOGRAPHY AND DISTRIBUTION OF DCT COEFFICIENTS

### 2.1. Jsteg Steganography Overview

Jsteg is a steganographic algorithm based on LSB replacement method for hiding data in DCT coefficients of JPEG images [17]. The principle of the algorithm is to replace the LSB of DCT coefficients by bits of the message to be hidden. If the embedding bit, 0 or 1, matches the LSB of DCT coefficient no change is made; if not, the LSB of DCT coefficient is flipped to 1 or 0.

In the JPEG compression scheme, the DCT transformation is applied to blocks of $8 \times 8$ pixels. Hence there is 64 different subbands all containing $N$ coefficients. Then let us denote the vector $V_k = \{v_{k,1}, \cdots, v_{k,N}\}$, where $n \in \{1, \cdots, N\}$ for the index of blocks and $k \in \{2, \cdots, 64\}$ for the index of 63 AC coefficient subbands. Since the change of the first DCT subband (DC components) modifies the image visually, it is not used for inserting secret bits. Let $v_{k,n}, 1 \leq n \leq N$ represents the value of DCT coefficient of $k$-th subband in the $n$-th block. Since the DCT coefficients are quantized, let $\triangle_k$ be the quantization step of the $k$-th subband. The Jsteg algorithm does not use the DCT coefficient with value 0 or 1 to hide bits because such changes can be detected easily. Thus let us denote $N_k$ as the number of the usable coefficients of $k$-th suband $V_k$, where $N_k \leq N$, hence $N - N_k$ represents the number of unusable coefficients. In the framework of optimal detectors, it is crucial to study the statistical distribution of DCT coefficients before and after secret bits embedding.

### 2.2. Statistical Model of DCT Subband

In the prior researches [16, 18], it is considered that the DCT coefficients within a subband follow a Laplacian distribution. This model is a good trade-off between complexity and accuracy. The probability density function (PDF) of the continuous Laplacian distribution can be written as:

$$f_{e,b}(x) = \frac{1}{2b} \exp\left(\frac{- \mid x - e \mid}{b}\right), \forall x \in \mathbb{R} \quad (1)$$

where the location parameter $e$ represents the expectation of the Laplacian distribution and $b$ the scale parameter. In practice, the distribution of DCT coefficients by using the continuous Laplacian distribution (1) can not express the discrete DCT coefficients. Each DCT subband has its own quatization step $\triangle_k$. In prior researches, it is generally admitted that, within the $k$-th subband, the DCT coefficients are i.i.d: they share the same location and scale parameter, respectively denoted $e_k$ and $b_k$. Hence, the discrete Laplacian distribution of each DCT subband can be represented by [16]:

$$\mathcal{L}(X=x)=\begin{cases}1-\exp(\frac{-\triangle_k}{2b_k}) & \text{if } x=e_k\\ \exp(\frac{-\triangle_k|x-e_k|}{b_k})\sinh\left(\frac{\triangle_k}{2b_k}\right) & \text{if } x\neq e_k\end{cases} \quad (2)$$

where $\mathcal{L}(X = x)$ represents the probability that the value of a DCT coefficient equals $x$ with $x \in \mathbb{Z}$. On the opposite, in this paper, it is not assumed that DCT coefficients within a subband are i.i.d. Hence, in the following section, based on the model (2), a new distribution model of DCT coefficients is proposed.

## 3. LIKELIHOOD RATIO TEST FOR TWO SIMPLE HYPOTHESIS

### 3.1. Problem Statement

Let us assume that the location parameter $e_{k,n}$, and the scale parameter $b_{k,n}$ are known. It immediately follows from (2) that the $n$-th coefficient of $k$-th subband follows a distribution $\mathcal{P}_{e_{k,n},b_{k,n}}$ whose Probability Mass Function (PMF) is:

$$P_r(x)=\begin{cases}1-\exp(\frac{-\triangle_k}{2b_{k,n}}) & \text{if} x=e_{k,n}\\ \exp\left(\frac{-\triangle_k|x-e_{k,n}|}{b_{k,n}}\right)\sinh\left(\frac{\triangle_k}{2b_{k,n}}\right) & \text{if} x\neq e_{k,n}\end{cases} \quad (3)$$

where $P_r(x)$ is the probability that a given coefficient equals $x \notin \{0, 1\}$, and the quantization step $\triangle_k$ is obtained from the JPEG header file. Then, after embedding a secret information using Jsteg with payload $R$, a simple calculation shows that distribution of stego-image DCT coefficients, denoted $\mathcal{Q}^R_{e_{k,n},b_{k,n}}$ whose PMF is defined by:

$$P^R_r(x, R) = \Big(1 - \frac{R}{2}\Big)P_r(x) + \frac{R}{2}P_r(\bar{x}) \quad (4)$$

where $x \notin \{0, 1\}$, and $\bar{x} = x + (-1)^x$ denotes the integer $x$ with the flipped LSB. The problem of detecting steganographic images consists in choosing between the following hypotheses:

$$\begin{cases}\mathcal{H}_0=\{v_{k,n}\sim \mathcal{P}_{e_{k,n},b_{k,n}} \forall v_{k,n}\notin 0,1\\ \mathcal{H}_1=\{v_{k,n}\sim \mathcal{Q}^R_{e_{k,n},b_{k,n}} \forall v_{k,n}\notin 0,1\, R\in(0,1]\end{cases} \quad (5)$$

where $\forall k=(2, ..., 64), \forall n=(1, ..., N_k)$.

For solving statistical detection problem such as (5), it follows from the Neyman-Pearson lemma [19, Theorem 3.2.1] that the LRT is optimal in the sense described below. For definition, let

$$\mathcal{K}_\alpha = \{\delta : \mathbb{P}_0[\delta(\mathbf{V}) = \mathcal{H}_1] \leq \alpha\} \quad (6)$$

be the class of tests, solving problem (5), with an upper-bounded FAR $\alpha$. Here $\mathbb{P}_j[\cdot]$ is the probability under $\mathcal{H}_j, j \in \{0, 1\}$, $\mathbf{V} = \{V_2, \cdots, V_{64}\}$. Among all tests in $\mathcal{K}_\alpha$ the LRT is the most powerful test, it maximizes the detection power

$$\beta_\delta = \mathbb{P}_1[\delta(\mathbf{V}) = \mathcal{H}_1]. \quad (7)$$

From the expression of alternative hypothesis $\mathcal{H}_1$ (4), the Likelihood Ratio (LR) value for the $n$-th coefficient on $k$-th

DCT subband is given by:

$$\Lambda_{k,n}(v_{k,n}) = 1 - \frac{R}{2} + \frac{R}{2} \cdot \frac{P_r(\bar{v}_{k,n})}{P_r(v_{k,n})}. \tag{8}$$

Then, it follows from the statistical independence of vectors $v_{k,n}$ that the LRT for all the DCT coefficients is given by:

$$\delta(\mathbf{V}) \;=\; \begin{cases} \mathcal{H}_0 \text{ if } \Lambda(\mathbf{V}) = \sum_{k=2}^{64} \Lambda(V_k) \leq \tau_\alpha \\ \mathcal{H}_1 \text{ if } \Lambda(\mathbf{V}) = \sum_{k=2}^{64} \Lambda(V_k) > \tau_\alpha \end{cases} \tag{9}$$

$$\text{with } \Lambda(V_k) \;=\; \sum_{n=1}^{N_k} \log(\Lambda_{k,n}) \tag{10}$$

where the decision threshold $\tau_\alpha$ is the solution of equation $\mathbb{P}_0[\Lambda(\mathbf{V}) > \tau_\alpha] = \alpha$ to guarantee that $\delta(\mathbf{V}) \in \mathcal{K}_\alpha$.

### 3.2. Statistical Performance of LRT

The performance of the test is evaluated by normalizing the $\Lambda(\mathbf{V})$. One can obtain the expectation and variance of each $\log(\Lambda_{k,n})$ under the hypothesis $\mathcal{H}_0$, the case of the $\mathcal{H}_1$ can be deduced straightforward.

$$\mu_{k,n,0} = \sum_{x \in \mathbb{Z}} \log[\Lambda_{k,n}(x)] P_r(x) \tag{11}$$

$$\sigma^2_{k,n,0} = \sum_{x \in \mathbb{Z}} (\log[\Lambda_{k,n}(x)] - \mu_{k,n,0})^2 P_r(x) \tag{12}$$

where $x$ denotes all the possible value of each DCT coefficient. In the case of $x \in \{0, 1\}$ which is not used for embedding, $\log(\Lambda_{k,n})$ takes value 0. Subsequently, based on Central Limit Theorem (CLT), the summation of $\log(\Lambda_{k,n})$ follows the Gaussian distribution:

$$\Lambda(V_k) \sim \mathcal{N}(\mu_{k,0}, \sigma^2_{k,0}) \tag{13}$$

where the expectation $\mu_{k,0} = \sum_{n=1}^{N} \mu_{k,n,0}$ and the variance $\sigma^2_{k,0} = \sum_{n=1}^{N} \sigma^2_{k,n,0}$. It should be noted that $N$ defines the number of all the coefficients on each subband. Then, the sum of several independent random variable with the Gaussian distribution still follows the Gaussian distribution:

$$\Lambda(\mathbf{V}) \sim \mathcal{N}(\mu_0, \sigma^2_0) \tag{14}$$

where the expectation $\mu_0 = \sum_{k=2}^{64} \mu_{k,0}$ and the variance $\sigma^2_0 = \sum_{k=2}^{64} \sigma^2_{k,0}$. For clarity, let us define the normalized LR as follows:

$$\Lambda^\star(\mathbf{V}) = \frac{\Lambda(\mathbf{V}) - \mu_0}{\sigma_0} \sim \mathcal{N}(0, 1). \tag{15}$$

It is thus straightforward to define the normalized LRT with $\Lambda^\star(\mathbf{V})$ by:

$$\delta^\star(\mathbf{V}) = \begin{cases} \mathcal{H}_0 \text{ if } \Lambda^\star(\mathbf{V}) \leq \tau^\star_\alpha \\ \mathcal{H}_1 \text{ if } \Lambda^\star(\mathbf{V}) > \tau^\star_\alpha. \end{cases} \tag{16}$$

Then, we establish the statistical properties of the proposed test (16) which are given in the following theorems; for clarity, $\Phi(\cdot)$ and $\Phi^{-1}(\cdot)$ represent the Gaussian cumulative distribution function and its inverse respectively.

**Theorem 1.** *Assuming that the model of hypothesis (5) holds, then for any* $\alpha \in (0; 1)$*, the decision threshold:*

$$\tau^\star_\alpha = \Phi^{-1}(1 - \alpha) \tag{17}$$

*guarantees that the LRT* $\delta^\star(\mathbf{V})$ *(16) is in the class* $\mathcal{K}_\alpha$*.*

**Theorem 2.** *Assuming that the model of hypothesis (5) holds, for any decision threshold* $\tau^\star_\alpha \in \mathbb{R}$*, the power function associated with the test* $\delta^\star(\mathbf{V})$ *(16) is given by:*

$$\beta_{\delta^\star} = 1 - \Phi\left(\frac{\tau^\star_\alpha \sigma_0 + \mu_0 - \mu_1}{\sigma_1}\right). \tag{18}$$

## 4. UNKNOWN PARAMETERS ESTIMATION AND DESIGN OF THE PROPOSED DETECTOR

In practice, the location parameter $e_{k,n}$ and scale parameter $b_{k,n}$ are unknown and, hence, must be estimated when analyzing an image of the size $I \times J$. Let us define the matrix $\mathbf{P} = \{p_{i,j}\}$ $i \in \{1, ..., I\}$, $j \in \{1, ..., J\}$ as the value of each pixel obtained by decompressing the JPEG image through inverse DCT, which can be decomposed as [20, 21]:

$$\mathbf{P} = \mathbf{\Theta} + \mathbf{\Xi} \tag{19}$$

where $\mathbf{\Theta} = \{\theta_{i,j}\}$ denotes the expectation of $\mathbf{P}$ and $\mathbf{\Xi} = \{\xi_{i,j}\}$ the random noise corruption. It is should be noted that DCT is a linear transformation which implies:

$$\mathcal{DCT}(\mathbf{P}) = \mathcal{DCT}(\mathbf{\Theta} + \mathbf{\Xi}) = \mathcal{DCT}(\mathbf{\Theta}) + \mathcal{DCT}(\mathbf{\Xi}) \tag{20}$$

where $\mathcal{DCT}(\cdot)$ denotes the DCT transform. Hence, the estimated location parameter of DCT coefficients $\mathcal{DCT}(\hat{\mathbf{\Theta}})$ can be obtained from spatial domain. In the present paper, $\hat{\mathbf{\Theta}}$ is estimated by using the wavelet denoising filter proposed in [22] which operates in two steps: first, the local image variance is estimated, then, the local Wiener filter is applied in the wavelet domain. Hence, we approximately obtain the estimated location parameter matrix:

$$\hat{\mathbf{E}} \approx \mathcal{DCT}(\hat{\mathbf{\Theta}}). \tag{21}$$

For clarity, the estimation $\hat{\mathbf{E}} = \{\hat{e}_{i,j}\}$ is divided into 63 vectors of $N$ components, as the DCT coefficients, denoted $\hat{E}_k = \{\hat{e}_{k,1}, \cdots, \hat{e}_{k,N}\}, k \in \{2, \cdots, 64\}$.

Next, let us denote a quantized DCT matrix extracted from a JPEG image as $\mathbf{C} = \{c_{i,j}\}$ $i \in \{1, ..., I\}$, $j \in \{1, ..., J\}$, where $I$ and $J$ denote the height and the width of the matrix $\mathbf{C}$. Then it is proposed to split $\mathbf{C}$ into nonoverlapping block $\mathbf{c}_{p,q}$ with the dimension of $8 \times 8$, where $(p, q)$ the index of the block (excluding border blocks), denoted by

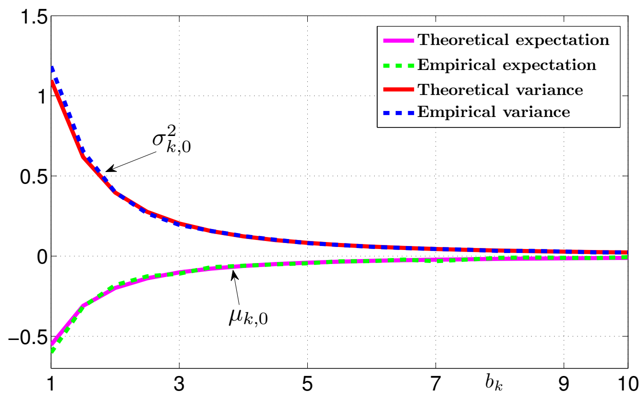


**Fig. 1**: Expectation $\mu_{k,0}$ and variance $\sigma^2_{k,0}$ as a function of the scale parameter $b_k$ theoretically and empirically.

$\{p = 2, \ldots, \frac{I}{8} - 1\}$ and $\{q = 2, \ldots, \frac{J}{8} - 1\}$. Similar to the method for estimating the local variance in [15] it is proposed to estimate the scale parameter locally using the adjacent neighbors blocks, including the center. Using the Maximum likelihood Estimation (MLE), the estimated scale parameter of coefficients in the block $\mathbf{c}_{p,q}$ is defined by:

$$\hat{\mathbf{b}}_{p,q} = \frac{\sum_{x=-l}^{l} \sum_{y=-l}^{l} \mid \mathbf{c}_{p+x,q+y} \mid}{L} \quad (22)$$

where $L = (2l + 1)^2$, $l \in \mathbb{N}^+$, $\{x, y\} \in \mathbb{Z}$. For convenience, let the estimated parameter matrix $\hat{\mathbf{B}}$ represent all the $\hat{\mathbf{b}}_{p,q}$'s. For clarity, it is proposed to denote the vector $\hat{B}_k = \{\hat{b}_{k,1}, \cdots, \hat{b}_{k,N}\}, \forall k = 2, \cdots, 64$, where $\hat{b}_{k,n}$ denotes the $n$-th DCT coefficient of the $k$-th subband.

Let us define $\hat{\Lambda}_{k,n}(v_{k,n})$ the LR value with estimated parameters on each DCT coefficient. Then, in the case of not knowing the distribution parameters, the proposed test for all the DCT coefficient subbands is given by:

$$\hat{\delta}(\mathbf{V}) = \begin{cases} \mathcal{H}_0 \text{ if } \hat{\Lambda}(\mathbf{V}) = \sum_{k=2}^{64} \hat{\Lambda}(V_k) \leq \hat{\tau}_\alpha \\ \mathcal{H}_1 \text{ if } \hat{\Lambda}(\mathbf{V}) = \sum_{k=2}^{64} \hat{\Lambda}(V_k) > \hat{\tau}_\alpha \end{cases} \quad (23)$$

$$\text{with } \hat{\Lambda}(V_k) = \sum_{n=1}^{N_k} \log(\hat{\Lambda}_{k,n}) \quad (24)$$

with, again, $\hat{\tau}_\alpha$ the solution of equation $\mathbb{P}_0[\hat{\Lambda}(\mathbf{V}) > \hat{\tau}_\alpha] = \alpha$.

## 5. NUMERICAL SIMULATIONS

To verify the sharpness of the theoretically established results, we generate 1000 sets of 4000 random variable (a Monte-Carlo simulation) following the Laplacian distribution, where $R = 0.05$, $e_{k,n} = 0$ and $b_k$ distributed from 1 to 10 with a step of 0.5. Then, the expectation and variance values are calculated empirically and theoretically. As shown in Fig. 1, the empirically calculated moments are almost equal to the analytically established ones.

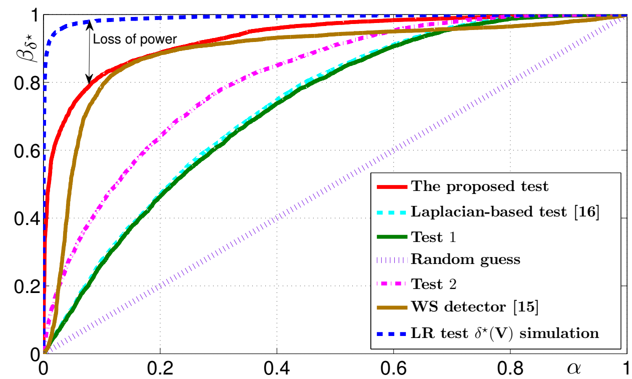


**Fig. 2**: ROC curves comparison for 10000 JPEG image from BossBase with quality factor 70, detection power as a function of FAR $\alpha$.

For fairly comparing with previous algorithms [15, 16], it is proposed to perform a numerical simulation over the 10000 images from BossBase [23] which have been compressed in JPEG with quality factor 70. The payload, or embedding rate, $R$ is set at 0.05 for Jsteg algorithm. To show the improvement provided by the proposed model the test proposed in [16] has been modified by using only the proposed approach for estimating the location parameter; this shows the improvement due to the use of locally estimated scale parameter. On the opposite, it is also proposed to estimate the scale parameter as described above, while assuming the expectation of DCT coefficients is zero. Those two tests are referred to as "Test 1" and "Test 2" respectively in Figure 2 and show the improvement due to estimation of each parameter alone. As Fig. 2 illustrates, the proposed test (23) outperforms the detector proposed in [16], based on the Laplacian i.i.d model. Similarly, the WS detector [15], based on a Gaussian model of DCT coefficients, performs well, but for very low FAR, the proposed test achieves much higher detection rates. Finally, It is proposed to give the performance of our detector on 10000 simulated images in which a DCT subband is generated by strictly following the Laplacian distribution. Then a comparison with simulations of the LR test shows the loss of power due to the estimation of expectation and scale parameters.

## 6. CONCLUSION

This paper aims at improving the detection of data hidden within the DCT coefficients of JPEG images. Based on the statistical model of DCT coefficients, the problem is cast in the framework of hypothesis testing theory. Assuming the location parameter and the scale parameter are previously known, the statistical performance of the LRT is analytically established. In the practical case, the parameters are unknown. Numerical results show that by estimating those parameters, the Laplacian model allows the designing of a statistical test which outperforms the prior-art detectors.

## 7. REFERENCES


[1] J. Fridrich and J. Kodovskỳ, "Steganalysis of lsb replacement using parity-aware features," in *Information Hiding*. Springer, 2013, pp. 31–45.

[2] R. Böhme, *Advanced statistical steganalysis*, Springer, 2010.

[3] J. Fridrich, *Steganography in digital media: principles, algorithms, and applications*, Cambridge University Press, 2009.

[4] I. Cox, M. Miller, J. Bloom, J. Fridrich, and T. Kalker, *Digital watermarking and steganography*, Morgan Kaufmann, 2007.

[5] T. Zhang and X. Ping, "A fast and effective steganalytic technique against jsteg-like algorithms," in *Proceedings of the 2003 ACM symposium on Applied computing*. ACM, 2003, pp. 307–311.

[6] R. Cogranne, F. Retraint, C. Zitzmann, I. Nikiforov, L. Fillatre, and P. Cornu, "Hidden information detection using decision theory and quantized samples: Methodology, difficulties and results," *Digital Signal Processing*, vol. 24, no. 0, pp. 144 – 161, 2013.

[7] R. Cogranne, C. Zitzmann, F. Retraint, I. Nikiforov, P. Cornu, and L. Fillatre, "A local adaptive model of natural images for almost optimal detection of hidden data," *Signal Processing* In press., 2014.

[8] C. Zitzmann, R. Cogranne, F. Retraint, I. Nikiforov, L. Fillatre, and P. Cornu, "Statistical decision methods in hidden information detection," in *Information Hiding*. Springer, 2011, pp. 163–177.

[9] J. Fridrich, M. Goljan, and R. Du, "Reliable detection of LSB steganography in color and grayscale images," in *Proceedings of the 2001 workshop on Multimedia and security: new challenges*. ACM, 2001, pp. 27–30.

[10] S. Dumitrescu, X. Wu, and Z. Wang, "Detection of lsb steganography via sample pair analysis," *Signal Processing, IEEE Transactions on*, vol. 51, no. 7, pp. 1995–2007, 2003.

[11] 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 *Proceedings of the first ACM workshop on Information hiding and multimedia security*. ACM, 2013, pp. 45–58.

[12] J. Fridrich and M. Goljan, "On estimation of secret message length in lsb steganography in spatial domain," in *Electronic Imaging 2004*. International Society for Optics and Photonics, 2004, pp. 23–34.

[13] O. Dabeer, K. Sullivan, U. Madhow, S. Chandrasekaran, and B.S. Manjunath, "Detection of hiding in the least significant bit," *Signal Processing, IEEE Transactions on*, vol. 52, no. 10, pp. 3046 – 3058, oct. 2004.

[14] A. D Ker and R. Böhme, "Revisiting weighted stego-image steganalysis," in *Electronic Imaging 2008*. International Society for Optics and Photonics, 2008, pp. 681905–681905.

[15] R. Böhme, "Weighted stego-image steganalysis for jpeg covers," in *Information Hiding*. Springer, 2008, pp. 178–194.

[16] C. Zitzmann, R. Cogranne, L. Fillatre, I. Nikiforov, F. Retraint, and P. Cornu, "Hidden information detection based on quantized laplacian distribution," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012, pp. 1793–1796.

[17] D. Upham, "Jsteg steganographic algorithm," *Available on the internet http://www.filewatcher.com/m/jpeg-jsteg-v4.diff.gz.8878-0.html*, 1999.

[18] E. Y Lam and J. W Goodman, "A mathematical analysis of the dct coefficient distributions for images," *Image Processing, IEEE Transactions on*, vol. 9, no. 10, pp. 1661–1666, 2000.

[19] J.P Romano and E. Lehman, "Testing statistical hypotheses," in *Second Edition*. Springer, 2005.

[20] T. H Thai, R. Cogranne, and F. Retraint, "Camera model identification based on the heteroscedastic noise model," *Image Processing, IEEE Transactions on*, vol. 23, no. 1, pp. 250–263, 2014.

[21] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian, "Practical poissonian-gaussian noise modeling and fitting for single-image raw-data," *Image Processing, IEEE Transactions on*, vol. 17, no. 10, pp. 1737–1754, Oct. 2008.

[22] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *Information Forensics and Security, IEEE Transactions on*, vol. 1, no. 2, pp. 205–214, 2006.

[23] P. Bas, T. Filler, and T. Pevný, "Break our steganographic system — the ins and outs of organizing boss," in *Information Hiding, 13th International Workshop*, T. Filler, Ed., 2011.