HIDDEN INFORMATION DETECTION
BASED ON QUANTIZED LAPLACIAN DISTRIBUTION

Cathel Zitzmann, Rémi Cogranne, Lionel Fillatre, Igor Nikiforov, Florent Retraint and Philippe Cornu

ICD - LM2S - Troyes University of Technology - UMR STMR CNRS
12 rue Marie Curie, BP 2060 - 10010 Troyes Cedex - France
Email: firstname.lastname@utt.fr

ABSTRACT

The goal of this paper is to propose the optimal statistical test based on the modeling of discrete cosine transform (DCT) coefficients with a quantified Laplacian distribution. This paper focuses on the detection of hidden information embedded in bits of the DCT coefficients of a JPEG image. This problem is difficult, in terms of statistical decision, for two main reasons: first, the number of DCT coefficients used to conceal the hidden bits is random; second, the JPEG image compression induces a strong quantization of DCT coefficients. The proposed test explicitly takes into account the randomness of the number of DCT coefficients used. It maximizes the probability of hidden information detection by ensuring a prescribed level of false alarm.

***Index Terms*—** Hypothesis Testing, Steganalysis, JPEG, DCT coefficients.

1. INTRODUCTION AND CONTRIBUTION

Detecting the presence of hidden information in a digital media is a major challenge for law enforcement. In this context, it is necessary that the probability of false alarm and non detection are controlled. This article aims to study the detection of the Least Significant Bits (LSB) replacement method in natural images compressed in JPEG format.

During JPEG compression, the image is cut into blocks of 8×8 pixels and the discrete cosine transform (DCT) is applied to each block. Each block contains 64 coefficients corresponding to different frequencies. The DCT coefficients are then quantized using a quantization matrix that is present in the header of the file. When inserting a message, the embedding algorithm replaces the LSB of the quantized DCT coefficients with bits of the message. To build the detector, a model of the quantized DCT coefficients distribution and a statistical model of the embedding method are required. Thus, a detector based on the Laplacian model for the distribution of DCT coefficients is proposed. In the literature, several approaches exist to detect this kind of embedding [1, 2, 3], and some of them use all the DCT coefficients regardless of the frequencies [4]. Generally, statistical performances of existing detectors are not theoretically established.

The contribution of this paper is threefold: first, the theoretical basis for a Jsteg-embedding detector based on the likelihood ratio test is proposed; second, an optimal test is designed taking into account the Laplacian distribution parameter and the quantization matrix; finally, the analysis of the detector adapted to a random number of DCT coefficients based on a Laplacian model is proposed. The test is applied to a large database of natural images. It must be noted that this approach could be generalized for more complex algorithms.

The rest of the paper is organized as follows. Section 2 describes the embedding algorithm Jsteg. The statistical model of quantized DCT coefficients is exposed in Section 3. In Section 4 is presented the optimal detector designed for a random number of coefficients. The experimental results are presented in Section 5. The conclusion and future works are discussed in Section 6.

2. JSTEG DESCRIPTION

Jsteg algorithm [5] inserts a message in the DCT coefficients of an image in JPEG format. The principle of the algorithm (Fig. 2) is to replace the LSB by bits of the message.

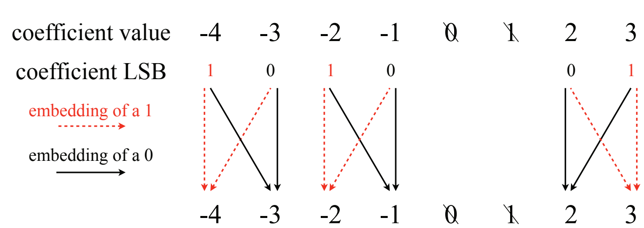

Fig. 2. Embedding into the DCT coefficients $v_{k,i}$.

Let an image be divided into n blocks of 64 coefficients and the vector $V_k = \{v_{k,1}, \ldots, v_{k,n}\}$ of length n consist of the value of k^{th} coefficient of each block (see Fig. 1) with

Thanks to the French National Agency for Research (ANR) through ANR-CSOSG program (project ANR-07-SECU-004) and to the COLOMBO program of Troyes University of Technology for supporting this work.

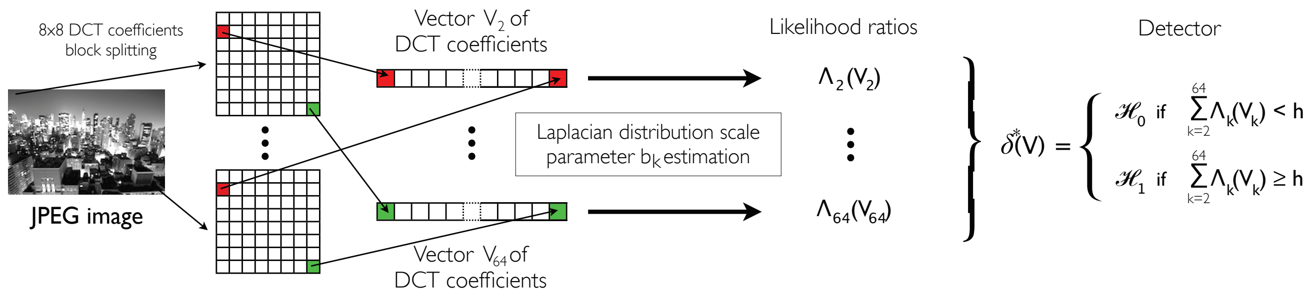

Fig. 1. Towards an optimal detector.

$v_{k,i} \in \{\lfloor\frac{-1024}{\Delta_k}\rfloor, \ldots, \lceil\frac{1023}{\Delta_k}\rceil\}$, $\forall k = 2, \ldots, 64$ and Δ_k the quantization associated with the k^{th} coefficient.

JPEG compression tends to create many 0 in the high frequencies. If a 1 was inserted in a 0, it would take the value 1 and the presence of a message could be easily detected. As a result, the message will not be inserted into the 0. As the insertion is done in pairs of values (see Fig. 2), the 1 are not used either. Moreover, the first DCT coefficient is not used for insertion, as its change would visually impact the image. Thus V_1 is not used for the detection and 0 and 1 are forbidden values. Therefore, the number of relevant coefficients of each vector is represented by a random variable $n_k \leq n$, which depends on the analyzed image (mainly its content) and the quantization matrix. It can be shown that n_k is a positive integer-valued random variable that follows a binomial distribution $\mathcal{B}(n, p_k^*)$, where p_k^* is the probability that a coefficient $v_{k,i}$ of the vector V_k is usable, i.e. $v_{k,i} \notin \{0, 1\}$. By convention, the n_k first components of the vector are usable and the $n - n_k$ last values are forbidden.

3. STATISTICAL MODEL OF QUANTIZED DCT COEFFICIENTS

To build a detector, we need to know the distribution of the DCT coefficients of a natural image. This is an interesting and open problem in many areas. Indeed, Reininger and Gibson, Lam and Goodman [6] and Smoot and Rowe proposed a Laplacian distribution, Srinath and Eggerton a Cauchy distribution and Müller a generalized Gaussian distribution. In this paper, as in other articles [7] dealing with the detection of hidden information, it was chosen to model the DCT coefficients of a natural image with a Laplacian distribution. Continuous Laplacian distribution with mean a and scale parameter b is given by:

$$f_{a,b}(x) = \frac{1}{2b} \exp\left(\frac{-|x-a|}{b}\right), \forall x \in \mathbb{R}.$$

It has been shown [6] that the distribution of DCT coefficients is symmetrical with mean $a = 0$. Thus, only the scale parameter b of the Laplacian distribution has to be estimated.

Let D be a discrete random variable representing DCT coefficients. The probability density $\text{Lap}\,(b_k, \Delta_k)$ of the k^{th} DCT coefficient modeled by a discrete Laplacian distribution quantized with the quantization step Δ_k is obtained by:

$$\begin{aligned} P\left(D = v_{k,i} | D \notin \{0,1\}\right) &= q_{b_k,\Delta_k}(v_{k,i}) \\ &= q_{b_k,\Delta_k}(-v_{k,i}) \\ &= \exp\left(\frac{-\Delta_k |v_{k,i}|}{b_k}\right) \sinh\left(\frac{\Delta_k}{2b_k}\right). \end{aligned}$$

After insertion, the probability density of a DCT coefficient denoted $\text{Lap}\,(b_k, \Delta_k, R)$ by the method of LSB replacement with an insertion rate R is:

$$\begin{aligned} P^R\left(D = v_{k,i} | D \notin \{0,1\}\right) &= q^R_{b_k,\Delta_k}(v_{k,i}) \\ &= \left(1 - \frac{R}{2}\right) q_{b_k,\Delta_k}(v_{k,i}) + \frac{R}{2} q_{b_k,\Delta_k}(\bar{v}_{k,i}), \end{aligned}$$

where $\bar{v}_{k,i} = v_{k,i} + (-1)^{v_{k,i}}$ is the coefficient $v_{k,i}$ with flipped LSB (see Fig. 2). The insertion rate is the probability that a LSB is replaced by a bit of the message.

4. OPTIMAL DETECTOR FOR A RANDOM NUMBER OF DCT COEFFICIENTS

The purpose of the detection problem is to construct a statistical test that detects hidden information from the set of vectors of DCT coefficients assuming they are independent. The impact on the detector of the phenomena of correlation (between the vectors and in the vectors) will be studied later. Now, all calculations will be conditioned by the fact that $v_{k,i}$ values must be usable.

When analyzing an image, two scenarios are possible: $\mathcal{H}_0 = \{$the image is a cover media$\}$ and $\mathcal{H}_1 = \{$the image is a stego media with an embedding rate $R\}$. The hypothesis test considered is written as follows:

$$\begin{aligned} &\mathcal{H}_0 : \{v_{k,i} \sim \text{Lap}(b_k, \Delta_k), \forall k{=}2, \ldots, 64, \forall i{=}1, \ldots, n_k\} \\ &\mathcal{H}_1 : \{v_{k,i} \sim \text{Lap}(b_k, \Delta_k, R), \forall k{=}2, \ldots, 64, \forall i{=}1, \ldots, n_k\}. \end{aligned} \tag{1}$$

We look for the most powerful test in the class K_{α_0}, defined as $K_{\alpha_0} = \{\delta^* : P_{\mathcal{H}_0}(\delta^*(V) = \mathcal{H}_1) < \alpha_0\}$, where $V =$

$\{V_2, \dots, V_{64}\}$ to solve the problem (1). The parameter b_k is estimated and Δ_k is given in the header, this is why we consider these parameters known. This test is given by the Neyman-Pearson lemma and consists of the following decision rule:

$$\delta^*(V) = \begin{cases} \mathcal{H}_0 & \text{if} \quad \Lambda^*(V) = \sum_{k=2}^{64} \Lambda_k(V_k) < h \\ \mathcal{H}_1 & \text{if} \quad \Lambda^*(V) = \sum_{k=2}^{64} \Lambda_k(V_k) \geq h \end{cases} \tag{2}$$

where

$$\Lambda_k(V_k) = \Lambda_k(v_{k,1}, \dots, v_{k,n_k})$$
$$= \sum_{i=1}^{n_k} \log\left(1 - \frac{R}{2} + \frac{R}{2} \exp\left(\frac{\Delta_k}{b_k}\left(\underbrace{\operatorname{sign}(v_{k,i})(v_{k,i} - \bar{v}_{k,i})}_{=\zeta_{k,i}}\right)\right)\right). \tag{3}$$

Random variable $\zeta_{k,i}$ represents the impact of the insertion and equals to 1 or -1 depending on the parity and the sign of $v_{k,i}$. It is due to the natural asymmetry of Jsteg embedding (see Fig. 2): even coefficients could only be incremented, and odd coefficients decremented, by replacing the LSB and also due to the LSB value that are inverted for positive and negative value. For instance if $v_{k,i}$ is positive and odd, the LSB is 0, whereas the LSB is 1 if $v_{k,i}$ is negative and odd. The threshold h is the solution of the equation $P_0(\Lambda^*(V) \geq h) = \alpha_0$.

The following theorem establishes the statistical probabilities of the Neyman-Pearson test.

Theorem 1 *The test (2) reaches the power β_{δ^*} when n tends to infinity, for a decision threshold h set by α_0:*

$$\beta_{\delta^*} \simeq 1 - \Phi\left(\frac{\sigma_0^*}{\sigma_1^*}\Phi^{-1}(1-\alpha_0) + \frac{\mu_0^* - \mu_1^*}{\sigma_1^*}\right)$$

where $h \simeq \sigma_0^* \, \Phi^{-1}(1-\alpha_0) + \mu_0^*$, and where μ_j^* and σ_j^{*2} are defined in the proof.

The ratio $\frac{\mu_0^* - \mu_1^*}{\sigma_1^*}$ measures the separability between the two hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$. It depends on R, b_k and Δ_k. This theorem is asymptotic as n is proportional to the number of pixels in the image, and for a large image, n tends to infinity.

SKETCH OF THE PROOF - To take into account the presence of prohibited values in the vectors of DCT coefficients, it is necessary to consider the random variable n_k in the calculation of detector performance.

The expression of the log-likelihood ratio for a frequency:

$$\Lambda_{k,j}(v_{k,1}, \dots, v_{k,n_k}) = \sum_{i=1}^{n_k} X_{k,i}$$
$$= \sum_{i=1}^{n_k} \log\left(1 - \frac{R}{2} + \frac{R}{2} \exp\left(\frac{\Delta_k}{b_k}(\zeta_{k,i})\right)\right)$$

with $\zeta_{k,i}$ the random variable that is equal to 1 with probability $p_{k,j}$ which is -1 with probability $1 - p_{k,j}$ under the hypothesis $\mathcal{H}_j$ and $p_{k,j}$ is the probability that $v_{k,i}$ is positive or negative and odd or even under the assumption $\mathcal{H}_j$.

$$\begin{aligned} \Lambda_{k,j}(v_{k,1}, \dots, v_{k,n_k}) &= n_k a_k + (c_k - a_k) \sum_{i=1}^{n_k} Y_{k,i} \\ &= n_k a_k + (c_k - a_k) S_{n_k}, \end{aligned}$$

with $a_k = \log\left(1 - \frac{R}{2} + \frac{R}{2}\exp\left(-\frac{\Delta_k}{b_k}\right)\right)$, and $c_k = \log\left(1 - \frac{R}{2} + \frac{R}{2}\exp\left(\frac{\Delta_k}{b_k}\right)\right)$.

The random variables $Y_{k,i} = \frac{X_{k,i} - a_k}{c_k - a_k}$ are independent, identically distributed and follow a Bernoulli distribution of mean $p_{k,j}$ and variance $p_{k,j}(1 - p_{k,j})$.

Let $S_{n_k} = \sum_{i=1}^{n_k} Y_{k,i}$. Since $Y_{k,i}$ follows the Bernoulli $\mathcal{B}(1, p_{k,j})$ and the random variable n_k follows the binomial distribution $\mathcal{B}(n, p_k^*)$, it can be shown that S_{n_k} follows the binomial distribution $\mathcal{B}(n, p_k^* p_{k,j})$.

As n is large and $p_k^* p_{k,j}$ is not too close to 0 or 1, the binomial distribution is very well approximated by a normal distribution. Thus, under hypothesis $\mathcal{H}_j$, for each frequency, we obtain:

$$\frac{\frac{\Lambda_{k,j} - n_k a_k}{c_k - a_k} - n p_k^* p_{k,j}}{\sqrt{n p_k^* p_{k,j}(1 - p_k^* p_{k,j})}} \xrightarrow[n \to \infty]{\mathcal{L}} \mathcal{N}(0,1)$$

and

$$\frac{\Lambda_{k,j} - \mu_{k,j}}{\sigma_{k,j}} \xrightarrow[n \to \infty]{\mathcal{L}} \mathcal{N}(0,1),$$

with $\mu_{k,j} = n p_k^* \left[a_k(1 - p_{k,j}) + c_k p_{k,j}\right]$ and $\sigma_{k,j}^2 = n p_k^* \left[a_k^2(1 - p_k^*) + (c_k - a_k)^2 p_k^* p_{k,j}(1 - p_k^* p_{k,j})\right]$, where $\xrightarrow[n \to \infty]{\mathcal{L}}$ denotes the convergence in law.

The sum of several independent normal distributions are normally distributed, the optimal detector, assuming $\mathcal{H}_j$ yields to:

$$\frac{\Lambda^*(V) - \mu_j^*}{\sigma_j^*} \xrightarrow[n \to \infty]{\mathcal{L}} \mathcal{N}(0,1),$$

where $\mu_j^* = \sum_{k=2}^{64} \mu_{k,j}$ and $\sigma_j^{*2} = \sum_{k=2}^{64} \sigma_{k,j}^2$.

5. NUMERICAL EXPERIMENTATION

To illustrate the results of previous sections, we use the image `mandrill`that is part of a reference database[1]. This is a color image of size 512×512 in the TIFF format. It has been JPEG compressed using `imagemagick` with compression quality 70 and converted in shades of gray. Fig. 3 represents the empirical distribution of the 2nd DCT coefficient and the estimation of the corresponding Laplacian distribution.

In practice, we want to analyze an image, i.e. decide whether a hidden message was inserted into the image with

[1] http://sipi.usc.edu/database/

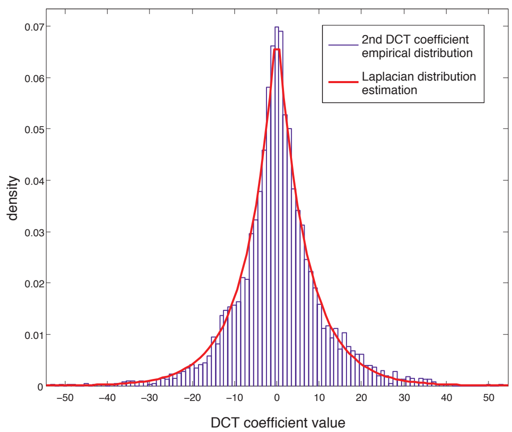

Fig. 3. Empirical distribution of the 2[nd] DCT coefficient and estimation of the corresponding Laplacian distribution.

a given insertion rate. The method relies on different stages. First, the vectors $V_k = \{v_{k,1}, \dots, v_{k,n}\}$ containing the DCT coefficients of each frequency k and the quantization step Δ_k are extracted. Then, the Laplacian parameters b_k are estimated, for each image and frequency, using the maximum likelihood estimator (Fig. 3):

$$\hat{b}_k = \frac{1}{n} \sum_{i=1}^{n} |v_{k,i}|.$$

This estimator is unbiased and asymptotically efficient. It can be noticed that for a cover image or a stego one, the estimates of b_k are very close to each other.

The detectors were tested on the simulated images and on the real images from Boss[2] database containing 10,000 images of size 512×512 in PGM format. The real images were converted with `imagemagick` in JPEG format with a compression quality of 70. For simulated images, the estimated parameter $\hat{b}_k$ has been used to generate the vectors V_k by using a pseudorandom generator.

As it follows from Figure 4, the proposed detector performs much better than Zhang and Ping's algorithm [4] for the simulated data. It can be explained by the fact that experimental data coincide with the theoretical model (see Section 3), however it's not the case for real data. It is important to stress a considerable difference between the performance of the algorithm on simulated and real images. This means that the proposed model should be optimized. It is also assumed that the statistical model of the vectors V_k will be improved in order to reduce the difference between the power of the proposed test on real data and on the simulated data.

[2]ftp://mas22.felk.cvut.cz/PGMs/BossBase-1.0-cover.tar.bz2

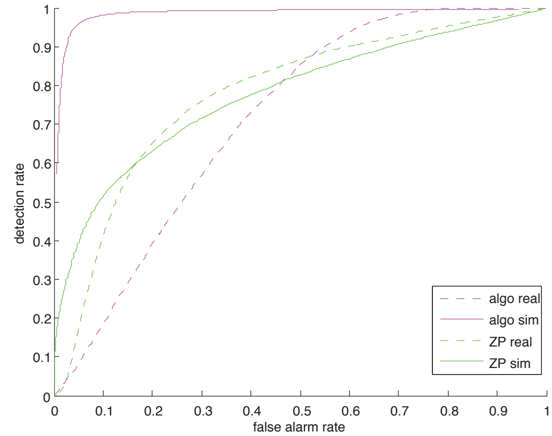

Fig. 4. The power of the proposed and ZP detectors is given as a function of the probability of false alarm for simulated (solid line) and real (dotted line) images. The power of proposed detector is shown in magenta and the power of ZP detector in green. The embedding rate is $R = 0.05$.

6. CONCLUSION AND OUTLOOK

This article lays the theoretical basis of a detector of hidden information in LSB of the DCT coefficients of a JPEG image. This work will be continued by studying the impact of correlations between the DCT coefficient vectors on the detector, and statistically modeling more complex embedding algorithms.

7. REFERENCES

[1] A. Westfeld, "Generic adoption of spatial steganalysis to transformed domain," in *IH*. 2008.

[2] R. Böhme, "Weighted Stego-Image Steganalysis for JPEG Covers," in *IH*. 2008.

[3] J. Kodovský and J. Fridrich, "Quantitative Structural Steganalysis of Jsteg," *IEEE TIFS*, vol. 5, no. 4, 2010.

[4] T. Zhang and X. Ping, "A Fast and Effective Steganalytic Technique Against JSteg-like Algorithms," in *SAC*, 2003, pp. 307–311.

[5] D. Upham, "Jsteg (1993)," http://www.nic.funet.fi/pub-/crypt/steganography/jpeg-jsteg-v4.diff.gz.

[6] E.Y. Lam et al., "A Mathematical Analysis of the DCT Coefficient Distributions for Images," *IEEE Trans. on IP*, vol. 9, no. 10, 2000.

[7] M. Wu et al., "A New Steganalytic Algorithm for Detecting Jsteg," in *ICCNMC*, 2005, pp. 1073–1082.