



HAL
open science

Steganography via Cover-Source Switching

Patrick Bas

► **To cite this version:**

Patrick Bas. Steganography via Cover-Source Switching. WIFS, Dec 2016, Abu Dabi, United Arab Emirates. hal-01386190

HAL Id: hal-01386190

<https://hal.science/hal-01386190>

Submitted on 23 Oct 2016

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.

Steganography via Cover-Source Switching

Patrick Bas

Univ. Lille, CNRS, Centrale Lille, UMR 9189

CRISTAL, F-59000 Lille, France

Email: patrick.bas@ec-lille.fr

Abstract—This paper proposes a new steganographic scheme relying on the principle of “cover-source switching”, the key idea being that the embedding should switch from one cover-source to another. The proposed implementation, called Natural Steganography considers the sensor noise naturally present in the raw images and uses the principle that, by the addition of a specific noise the steganographic embedding tries to mimic a change of ISO sensitivity. The embedding methodology consists in 1) designing the stego-signal that enable to switch from one source to another the raw domain, 2) computing the statistical distribution of the stego-signal in the processed domain, 3) embedding the payload in the processed domain. We show that this methodology is easily tractable whenever the processes are known and enables to embed large and undetectable payloads.

I. INTRODUCTION

Image steganography consists in embedding an undetectable message into a cover image to generate a stego image, the application being the transmission of sensitive information. As Cachin proposed in [1], one theoretical approach proposed for steganography is to minimize a statistical distortion, and the author proposes to use the Kullback-Leibler divergence. It is interesting to note that this line of research has been rarely used as a steganographic guideline with few notable exceptions such as model-based steganography [2] which mimics the Laplacian distributions of DCT coefficients during the embedding, HUGO [3] whose model-correction mode tries to minimize the difference between the model of the cover image and the stego image, and more recently the mi-pod steganographic scheme [4] which minimizes a statistical distortion (a deflection coefficient) between normal distributions of cover and stego content.

The large majority of steganographic algorithms are currently based on the use of a distortion (also called a cost) which is computed for each pixel, and which is combined with a coding scheme that minimizes the global distortion while embedding a given payload. Classical distortions functions such as the ones proposed by S-UNIWARD [5] or by HILL [6] try to infer the detectability of each pixel by assigning small costs to pixels that are difficult to predict (usually textural parts of the image) and by assigning large costs to pixels that are easy to predict (belonging to homogeneous areas and to some extent to edges).

Once the distortion is computed, a steganographic scheme can either simulate the embedding by sampling according to the modifications probabilities π_k , $k \in [1, \dots, Q]$ for a Q-arry embedding, or can directly embed the message using Syndrome Trellis Codes (STCs) [7] or multilayer STCs [7], [8]. The size of the embedding payload N is computed as $N = \sum \pi_k \log \pi_k$ for each pixel of the image, and in practice the STCs succeed to reach 90% to 95% of the capacity [7] and consequently are close to optimal.

Another ingredient to tend to undetectable steganography is to use the information contained in a “pre-cover”, i.e. the high resolution image that is used to generate the cover at a lower resolution, in order to weight the cost w.r.t the rounding error. For quantization or interpolation operations, a pixel of the pre-cover at equal distance between two quantization cells will have a lower cost than a pre-

cover pixel very close to one given quantization cell. This strategy has been used in Perturbed-Quantization [9] but also adapted in more recent schemes using side information [10].

The proposed paper uses similar ingredients shared by modern steganographic methods, namely model-based steganography, Q-arry embedding and the associated modification probabilities π_k , and side-information. The main originality of this paper relies on the possible definitions of cover sources and the use of *cover-source switching* to generate stego content whose statistical distributions are very close to cover content.

A. Steganography via cover-source switching

The key idea of this paper is to propose a steganographic scheme where the message embedding will be equivalent to switching from one source \mathcal{S}_1 to another source \mathcal{S}_2 ; this practically can be done by designing an embedding that, when applied on \mathcal{S}_1 , mimics the statistical properties of \mathcal{S}_2 . More specifically in this paper we have decided to use the sensor noise to model a given source because its statistical model is rather simple, and we perform the embedding in such a way that the statistical properties of stego images mimic the sensor noise of source \mathcal{S}_2 . As we shall see in Sections II and III, the difference between \mathcal{S}_1 and \mathcal{S}_2 will come from the ISO sensitivity. One can argue that this parameter is reported in the EXIF file of the image, but EXIF information can be easily edited or even removed using software such as `exiftool` [11].

Note that this idea of steganography based on mimicking sensor noise is far from being new. In 1999 Franz and Pfitzmann [12] propose a paradigm for a stego-system “simulating a *usual process* of data processing” where the usual process is defined by the scan process, in this paper the authors study the properties of scanning noise coming from different scanners. A practical implementation of this concept is proposed in 2005 by Franz and Schneidewind [13], where the authors model the sensor noise for each pixel by a Normal distribution and perform the embedding by first estimating the noiseless scan, and secondly adding a noise mimicking the sensor noise. The algorithm was benchmarked using features derived from wavelet higher order statistics [14] and showed relatively good performance compared with naive noise addition. It is important to notice that contrary to the work presented here, if the idea of mimicking the sensor noise is present in [14], it does not rely on neither cover-source switching nor a sharp physical model of the noise in the RAW domain. The scheme presented here aims at generating stego images that are very close to a given source of cover images, the same way the ASO embedding scheme [15], with the distinction that the source is different than the one used during embedding.

An important requirement in order to achieve practical embedding is to be able to compute the probability of embedding changes π_k in the developed domain, this in order to perform the practical embedding but also in order to compute the embedding rate. This particular aspect will be addressed in Section III and IV.

Because the embedding scheme relies on natural statistical noise of the sensor, we decided to call this steganographic scheme ‘‘Natural Steganography’’ (NS).

II. SENSOR NOISE ESTIMATION

We present in this section the different noise sources affecting the sensor during a capture and then explain how to estimate the sensor noise.

A. Sensor noise model

Camera sensor noise models have been extensively studied in numerous publications [16], [17], [18] and have already been used in image forensics for camera device identification [19], [20]. These models can only be applied to linear sensors such as CDD or CMOS sensors, but this encompasses the majority of modern digital cameras at the date the paper is written. A camera sensor is decomposed into a 2D array of photo-sites and the role of each photo-site is to convert k_p photons hitting its surface during the exposure time into a digit. The conversion involves the quantum efficiency of the sensor measuring the ratio between k_p and the number of charge units k_e accumulated by the photo-site during the exposure time. k_e is then converted into voltage, which is amplified by a gain K (where K is referred as the system overall gain [18]) and then quantized.

For each photo-site at location (i, j) , the converted signal $x(i, j)$ originates from two components:

- The ‘‘dark’’ signal $x_d(i, j)$ with expectation $E[X_d(i, j)] = \mu_d$ which accounts for the number of electrons present without light and depends on the exposure time and ambient temperature,
- The ‘‘electronic’’ signal $x_e(i, j)$ with expectation $E[X_e(i, j)] = K\mu_e$, which accounts for the number of electrons originating from photons coming from the scene which is captured.

The expectation μ of each photo-site response is equal to:

$$\mu_{i,j} = E[X(i, j)] = E[X_d(i, j)] + E[X_e(i, j)] = \mu_d + K\mu_e. \quad (1)$$

Beside the signal components, there are three types of noise affecting the acquisition:

- The ‘‘shot noise’’ associated with the electronic signal with accounts for the fluctuation of the number of charge units. Because the electronic signal comes from the variation of counting events, it has a Poisson distribution $X_e(i, j) \sim \mathcal{P}(\mu_e)$ and can be approximated in a continuous setting by a normal distribution $\mathcal{N}(\mu_e, \sigma_e^2)$ with $\sigma_e^2 = \mu_e$, hence the noise associated to the electronic signal is distributed as $\mathcal{N}(0, \mu_e)$. This noise is independently distributed for each photo-site.
- The noise related to the ‘‘read-out’’ and the amplifier circuit. The read-out noise associated to the dark signal is independent and normally distributed as $\mathcal{N}(0, \sigma_d^2)$ and σ_d^2 is constant for a given camera.
- The quantization noise, which is independent and uniformly distributed with variance $\sigma_q^2 = \Delta^2/12$ where Δ denotes the quantization step.

Since these noise sources are mutually independent, the variance of the sensor noise can then be expressed as [18]:

$$\sigma_s^2 = K^2\sigma_d^2 + \sigma_q^2 + K(\mu - \mu_d). \quad (2)$$

In the sequel, we make the following approximations for a given cover-source: we assume that the system gain K is constant for a given *ISO* setting, that the dark signal is constant with

negligible variance ($\sigma_d^2 = 0, \mu_d = \text{cst}$), and that the quantization noise is negligible w.r.t. the shot noise ($\sigma_q^2 = 0$). As we shall see in V-D, the two first assumptions have negligible impact on the performance of the scheme and the last assumption does not impact the performance of the steganographic system whenever 16-bit quantization is considered as side-information. Finally, we also assume that the spatial non-uniformity of the sensor, which is associated with the photo response non-uniformity (PRNU) and the dark signal non-uniformity (DSNU), is negligible.

For a given *ISO* setting ISO_1 , the global sensor noise $N_{i,j}^{(1)}$ can be approximated using both Eq. (2) and the above-mentioned assumptions, as normally and independently distributed. We have consequently a linear relation between the sensor noise variance and the photo-site expectation μ :

$$N_{i,j}^{(1)} \sim \mathcal{N}(0, a_1\mu_{i,j} + b_1). \quad (3)$$

The acquired photo-site sample $x_{i,j}^{(1)}$ is given by:

$$x_{i,j}^{(1)} = \mu_{i,j} + n_{i,j}^{(1)}, \quad (4)$$

and $X \sim \mathcal{N}(\mu_{i,j}, a_1\mu_{i,j} + b_1)$.

B. Sensor noise estimation

In order to estimate the model of the sensor noise (i.e. the couple of parameters (a, b)) for a given camera model and a given *ISO* setting, we adopt a similar protocol as the one proposed in [17].

We first capture a set of N_a raw images of a printed photo picturing a rectangular gradient going from full black to white. The camera is mounted on a tripod and the light is controlled using a led lightning system in a dark room. The raw images are then converted to PPM format (for color sensor) or to PGM format (for B&W sensor) using the `dcraw` open-source software [21] using the command:

```
dcraw -k 0 -4 file_name
```

which means that the dark signal is not automatically removed (option `-k = 0`), and that the captured photo-sites are not post-processed and plainly converted to 16-bit (option `-4`).

In order to have a process independent of the quantization, the photo-site outputs are first normalized by dividing them by $y_{max} = 2^{16} - 1$. The range of possible outputs is divided into $1/\delta$ segments of width δ . Each normalized photo-site location is assigned to one subset of photo-sites \mathcal{S}_ℓ according to its empirical expectation over the acquired images $\hat{\eta}_{i,j} = \left(\sum_{l=1}^{N_a} y_{i,j}^{(l)} / y_{max} \right) / N_a$. The subset index is $\ell = \lceil \hat{\eta}_{i,j} / \delta \rceil$ where $\lceil \cdot \rceil$ denotes the integer rounding operation. Once the segmentation into subsets is performed, the empirical mean is:

$$\hat{\mu}_\ell = \frac{1}{|\mathcal{S}_\ell|} \sum_{i=1}^{|\mathcal{S}_\ell|} \mathcal{S}_\ell(i), \quad (5)$$

where $\mathcal{S}_\ell(i)$ denotes the value of a photo-site belonging to the subset \mathcal{S}_ℓ and $|\cdot|$ denotes the cardinal of a set.

The unbiased variance associated to each subset as:

$$\hat{\sigma}_\ell^2 = \frac{1}{|\mathcal{S}_\ell| - 1} \sum_{i=1}^{|\mathcal{S}_\ell|} (\mathcal{S}_\ell(i) - \hat{\mu}_\ell)^2. \quad (6)$$

As an illustration, Figure (1) plots the relation in solid lines between $\hat{\mu}_\ell$ and $\hat{\sigma}_\ell^2$ for $N_a = 20$ raw images captured with a Leica M Monochrome Type 230 at 1000 *ISO* and 1250 *ISO* for $\delta = 5 \cdot 10^{-5}$.

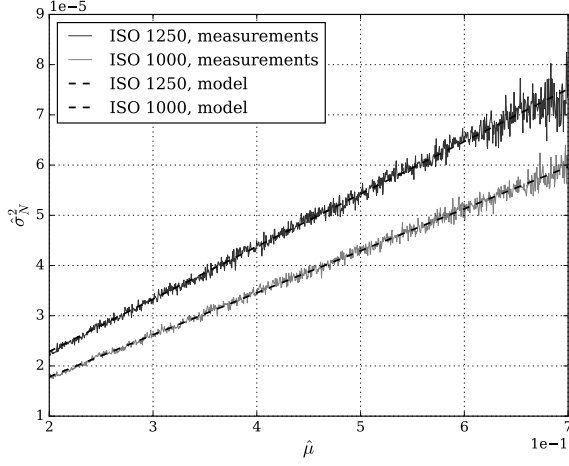


Fig. 1. Sensor noise estimation for the Leica Monochrome camera and 1000 ISO and 1250 ISO on normalized images. The estimated coefficients of the linear model are respectively $(a_1, b_1) = (8.36 \cdot 10^{-5}, 1.11 \cdot 10^{-6})$ and $(a_2, b_2) = (10.46 \cdot 10^{-5}, 1.95 \cdot 10^{-6})$ for this setup.

The last step consists of estimating the parameters (\hat{a}, \hat{b}) , this is done by linear regression $\hat{\sigma}_N^2 = f(\hat{\mu}) = \hat{a}\hat{\mu} + \hat{b}$. We see on the same figure that the linear relation, depicted by the dashed lines, is rather accurate.

III. EMBEDDING FOR OOC MONOCHROME PICTURES

We first propose in this section a steganographic system practically working for a basic developing setup, this system is realistic for a monochrome sensor where nor demosaicing neither colour transform is possible. We also assume that the developed images do not undergo gamma correction or further processing and only subjected to 8-bit quantization. We can call this type of images “Out Of Camera” (OOC) Pictures.

A. Principle of the embedding

We propose to model the stego signal $S_{i,j}$ in such a way that it mimics the model of images captured at $ISO_2 > ISO_1$. Note that this construction only simulate the embedding by computing the embedding probabilities for each symbol. These probabilities can be either used to simulate optimal embedding, or additive costs $\rho_{i,j}$ can be derived and used to feed a multilayered Syndrome Trellis Code using a Q-ary version of the “flipping lemma” [7] (see also Section VI of [7] for Q-ary embedding and multi-layered constructions).

Based on the assumptions made in II-A, the equivalent of (3) and (4) for a camera sensitivity parameter ISO_2 are $N_{i,j}^{(2)} \sim \mathcal{N}(0, a_2\mu_{i,j} + b_2)$ and $x_{i,j}^{(2)} = \mu_{i,j} + n_{i,j}^{(2)}$.

Since the sum of two independent signals normally distributed is normal with the variances summing up, we can write that $x_{i,j}^{(2)} = \mu_{i,j} + n_{i,j}^{(1)} + s'_{i,j} = x_{i,j}^{(1)} + s'_{i,j}$ with $S'_{i,j} \sim \mathcal{N}(0, (a_2 - a_1)\mu_{i,j} + b_2 - b_1)$ representing the noise necessary to mimic image captured at ISO_2 .

Assuming that the observed photo-site is very close to its practical expectation, i.e. that $\mu_{i,j} \simeq x_{i,j}^{(1)}$, $x_{i,j}^{(2)}$ can be approximated by:

$$x_{i,j}^{(2)} \simeq x_{i,j}^{(1)} + s_{i,j} \triangleq y_{i,j}, \quad (7)$$

with:

$$S_{i,j} \sim \mathcal{N}(0, (a_2 - a_1)x_{i,j}^{(1)} + b_2 - b_1). \quad (8)$$

Adopting the following notations $a' \triangleq a_2 - a_1$, $b' \triangleq b_2 - b_1$, $\sigma_S^2 \triangleq a'x_{i,j}^{(1)} + b'$, and the photo-site of the stego image is distributed as:

$$Y_{i,j} \sim \mathcal{N}(x_{i,j}^{(1)}, \sigma_S^2). \quad (9)$$

Note that equation (7) shows explicitly the principle of cover-source switching which is simply represented in this case by adding an independent noise on each image photo-site to generate the stego photo-site $y_{i,j}$. The distribution of the stego signal in the continuous domain (see (8)) takes into account the statistical model of the sensor noise estimated for two ISO settings using the procedure presented in Section II-A.

B. 16-bit to 8-bit quantization

For OOC images, the only developing process lies in the 8-bit quantization function, consequently the goal here is to compute the embedding changes probabilities $\pi_{i,j}(k) = \Pr[\bar{S}_{i,j} = k]$ after this process.

We use the high resolution continuous assumption given by (10) and then we compute the discretized probability mass function after a quantization step of size Δ (typically $\Delta = 256$ by quantizing from 16-bit resolution to 8-bit resolution).

The embedding probabilities are directly linked to the 8 bits quantized value $x_{8B} = Q_\Delta(x_{16B}) = \text{round}((x_{16B} + 1)/\Delta)$ - where $Q_\Delta(\cdot)$ denotes the quantization function - and the pdf of the Normal distribution $f(\cdot)$:

$$\begin{aligned} \pi(k) &= \int_{u_k}^{u_{k+1}} f(y|x = x_{16B})dy, \\ &= \frac{1}{2} \left(\text{erf} \left(\frac{u_{k+1} - x_{16B}}{2\sigma_S} \right) - \text{erf} \left(\frac{u_k - x_{16B}}{2\sigma_S} \right) \right), \end{aligned} \quad (10)$$

with $u_k = x_{8B} - (0.5 - k)\Delta$.

Once the embedding probabilities are computed for each pixel, it's possible to derive the payload size using the entropy formula:

$$H(\bar{S}) = - \sum_{i,j,k} \pi_{i,j}(k) \log_2 \pi_{i,j}(k). \quad (11)$$

IV. GAMMA CORRECTION

The gamma correction is a sample-wise operation defined by $y_\gamma \triangleq \Gamma(y) = y_{max}(y/y_{max})^{1/\gamma}$ with $y_{max} = 2^{16} - 1$, its inverse transform given by $\Gamma^{-1}(y) = y_{max}(y_\gamma/y_{max})^\gamma$.

In order to compute the distribution of the stego signal after gamma correction, one can simply compute the distribution of the transform of a continuous variable as [22]:

$$\begin{aligned} f_{Y_\gamma}(y_\gamma) &= f_Y(y_\gamma) \frac{d}{dy} \Gamma^{-1}(y), \\ &= \frac{1}{\sqrt{2\sigma_S^2\pi}} \exp \left(-\frac{(y_\gamma - x^{(1)})^2}{2\sigma_S^2} \right) \gamma (y_\gamma/y_{max})^{\gamma-1}. \end{aligned} \quad (12)$$

However, since in practice $\sigma_S \ll x^{(1)}$ we can use a first order Taylor expansion of the gamma correction, given by:

$$\begin{aligned} y_\gamma &\simeq \Gamma(x^{(1)}) + (y - x^{(1)})(x^{(1)}/y_{max})^{1/\gamma-1}/\gamma, \\ &= \Gamma(x^{(1)}) + \alpha s, \end{aligned} \quad (13)$$

with $\alpha = (x^{(1)}/y_{max})^{1/\gamma-1}/\gamma$. This means that the gamma correction acts as an affine transform on the stego signal.

Consequently, as a first approximation, the stego signal S_γ after gamma correction can be considered as normally distributed:

$$S_\gamma \sim \mathcal{N}(0, \alpha^2 \sigma_S^2), \quad (14)$$

and the distribution of the stego photo-site is given by:

$$Y_\gamma \sim \mathcal{N}(\Gamma(x^{(1)}), \alpha^2 \sigma_S^2). \quad (15)$$

Because gamma correction is a sample-wise operation, the stego signal is independently distributed, and the embedding probabilities after 8-bit quantization can be directly computed as:

$$\begin{aligned} \pi_\gamma(k) &= \int_{u'_k}^{u'_{k+1}} f(y_\gamma | x = \Gamma(x_{16B})) dy, \\ &= \frac{1}{2} \left[\operatorname{erf} \left(\frac{u'_{k+1} - \Gamma(x_{16B})}{2\alpha^2 \sigma_S^2} \right) - \operatorname{erf} \left(\frac{u'_k - \Gamma(x_{16B})}{2\alpha^2 \sigma_S^2} \right) \right], \end{aligned} \quad (16)$$

with $u'_k = x'_{8B} - (0.5 - k)\Delta$, $x'_{8B} = Q_\Delta(\Gamma(x_{16B}))$. The payload size is consequently given as $H(\bar{S}_\gamma) = -\sum_{i,j,k} \pi_{\gamma,i,j}(k) \log_2 \pi_{\gamma,i,j}(k)$.

V. EXPERIMENTAL RESULTS

The goal of this section is to benchmark the detectability of NS, to compare it with other steganographic schemes using same embedding payload, but also to analyze the effects of developing operations w.r.t. both detectability and embedding rates.

A. Generation of MonoBase

In order to benchmark the concept of embedding using cover-source switching, we needed to acquire different sources providing OOC images. To do so we conducted the following experiment: using a Leica M Monochrome Type 230 camera, we captured two sets of 172 pictures taken at 1000 ISO or 1250 ISO. In order to have large diversity of content most of the pictures were captured using a 21mm lens in an urban environment, or a 90mm lens capturing cluttered places.

The exposure time was set to automatic, with exposure compensation set to -1 in order to prevent over-exposure. A tripod was used so that pictures for the two sensitivity settings correspond to the same scene. Each RAW picture was then converted into a 16-bit PGM picture using the same conversion operation as the one presented in Section II-B and each 5212×3472 picture was then cropped into $6 \times 10 = 60$ PGM pictures of size 512×512 to obtain two sets of 10320 16-bit PGM pictures. We consequently end up with a database of a similar size than BOSSBase, with pictures of same size that contrary to un-cropped pictures can be quickly processed either for embedding or feature extraction. This database called MonoBase can be downloaded here [23].

B. Benchmark setup

For all the following experiments, we adopt the Spatial Rich Model feature sets [24] combined with the Ensemble Classifier (EC) [25] and we report the average total error $P_E = \min((P_{FA} + P_{MD})/2)$ obtained after training the EC on 10 different training/testing sets divided in 50/50.

Using NS we use a non-standard but realistic benchmark built to test if the cover source switching operation is undetectable: the stego database consists of images captured at 1000 ISO perturbed with a stego signal mimicking 1250 ISO, and the cover database consists of images directly captured at 1250 ISO.

In order to have an effect equivalent with the principle of training using pairs of cover and stego images, and to ease the steganalysis since this is a worst-case scenario for the steganographer, the pairs are constructed using one couple of images capturing the same scene.

The parameters of the stego signal are denoted a'' and b'' with the relations $a'' = a'(2^{N_b} - 1)$ and $b'' = b'(2^{N_b} - 1)^2$, where a' and b' are computed using normalized image values in order to be resolution independent (see Section III-A). $N_b = 16$ when the cover image is coded in 16-bit is used and $N_b = 8$ when the stego image is directly generated from the 8-bit representation of the cover image.

C. Basic developing and comparison with S-Uniward

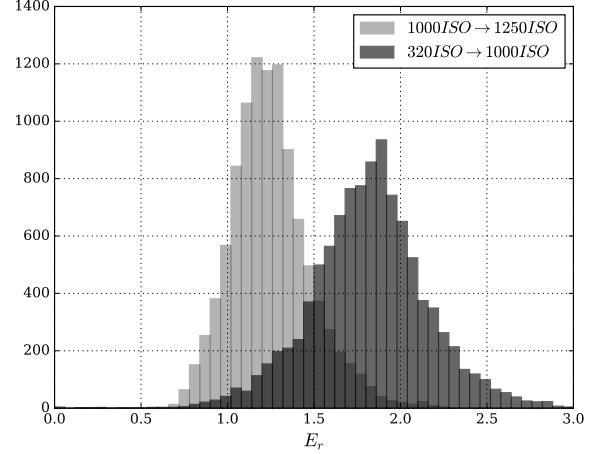


Fig. 2. Histogram of the embedding rates for MonoBase with covers-source switching from 1000 ISO to 1250 ISO. $E[E_r] = 1.24$ bpp. .

We first benchmark the scheme proposed in Section III and generate 8-bit stego images where the stego signal is generated according to the embedding probabilities computed in (10). Like all modern steganographic schemes, we forbid embeddings by attributing wet pixels to cover pixels saturated at 0 or 2^{16} .

The python source codes to perform NS embedding can be downloaded on Github [26].

The two first columns of Table I show the high undetectability of the proposed scheme, and the small improvement associated to wet the dark pixels. We note that we are still around 5% from random guessing, and we think that it can be due to the different assumption presented in Section III, particularly the fact that the quantization noise is ignored.

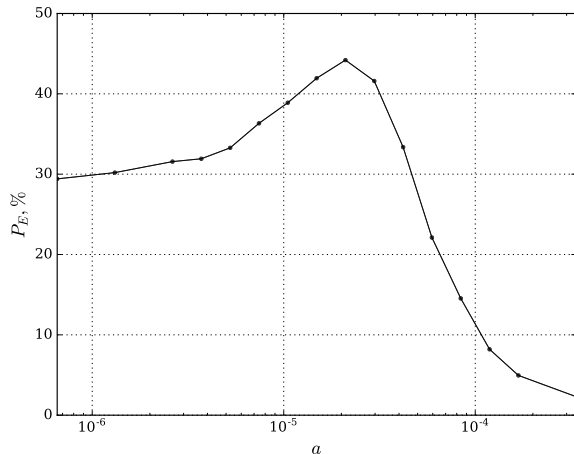
Figure 2 depicts the histogram of the embedding rate (E_r) on MonoBase. We can see that most of the embedding rates are relatively high for steganography with an average of 1.24 bpp for this base. Note that on MonoBase, most of the images are under-exposed, which means that the average embedding rate should be higher for a “typical” database. It is important to point here that contrary to most of the steganographic schemes, the current implementation of NS does not enable an embedding at a constant payload, but this as the price of high undetectability.

The two next columns of Figure 2 compare the performance of NS with S-Uniward [5] (SUni) using ternary embedding. We chose this steganographic scheme because of its excellent performance and because it has recently been tuned to take into account the side-information offered by the pre-cover after 16-bit to 8-bit conversion [10] (SUni-SI). The two implementations of S-Uniward where

	NS	SUni-SI 1000 ISO	SUni 1000 ISO	1000 ISO vs 1250 ISO
P_E	44.8%	18.2%	12.3%	26.0%

TABLE I

RESULTS AND COMPARISON WITH S-UNIWARD ON MONOBASE CODED WITH 8 BPP.

Fig. 3. Sensitivity w.r.t. the model parameters a^n .

benchmarked on MonoBase 1000 ISO with the same embedding rate as NS for each image but we saturate at $\log_2(3)$ bits since S-Uniward is restricted to ternary embedding. If the side-informed implementation provides better performance, the two schemes cannot compete with an embedding scheme like NS which is model-based and which enables cover-source switching. We obtained similar performance on MonoBase 1250 ISO (20.6% for SUni-SI), but also by comparing stego images generated from covers at 1000 ISO with covers at 1250 ISO (23.0% for SUni-SI).

The last column compares our steganalysis task with the classification task of separating images captured at 1000 ISO from images captured at 1250 ISO. We can see that this task using the SRM is not an easy task since the error probability is still rather large.

Figure 3 depicts the sensitivity of our scheme to the estimation of the sensor noise by computing the classification error for different values of a^n . We can see that the estimation of the sensor noise is rather important, going from example from $a^n = 2.1 \cdot 10^{-5}$ to $a^n = 10^{-5}$ increases the detectability by approximately 5%.

D. Results without pre-cover information

Table II presents the results when the sensor noise is estimated from raw images but the embedding is directly performed on 8-bit cover images without using the pre-cover, i.e. $\Delta = 1$ in (10).

We can notice that in this case the scheme becomes highly detectable. We explain this problem by the fact that dark regions, which undergo both stego signal and sensor noise of small variance, are not modified correctly when the embedding does not take into account the pre-cover. Indeed, the dark regions are especially sensitive to steganalysis because they are less noisy than bright regions, and because the initial value of the photo-site highly impacts the sign of the embedding change in this case.

	NS	NS 8-bits
P_E	44.8%	18.4%

TABLE II

IMPACT OF USING THE PRE-COVER FOR NS.

Figure 4 shows the embedding change for a portion of a cover image having dark areas and we can see that for the 8-bit embedding, the number of embedding changes are less important since the dithering effect offered by the use of the pre-cover image is lost here (the sensor noise is in this case centered directly on the quantization cell). Trying to improve NS in this practical setup is left for future research.

E. Gamma correction

Table III shows the detectability results of NS once gamma correction is performed during the developing step both on cover and on stego images. 16-bit cover images are used. Since the model of the stego signal is adapted to fit the model of the sensor noise after the gamma correction (see IV) we can check that the undetectability of NS is still high.

We see also that on MonoBase the embedding rate increases w.r.t. the parameter γ . This is because for $\gamma > 1$ the variance of the stego signal increases for small photo-site values and decreases for large photo-site values. The opposite occurs for $\gamma < 1$. In the first case this is due to the convexity of the transform, in the second case to the concavity of the transform. On a database only composed of bright images, the effect would be the opposite.

γ	2.5	2	1.5	1	0.5
P_E	44.4%	44.5%	43.7%	44.8%	46.2%
E_r (bpp)	1.61	1.62	1.55	1.24	0.5

TABLE III

PERFORMANCE OF NS AFTER GAMMA CORRECTION.

VI. CONCLUSION

We have proposed in this paper a new methodology for steganography based on the principle of cover-source switching, i.e. the fact that the embedding should mimic the transition from one cover-source to another. The presented scheme (NS) uses the sensor noise to model each source, and message embedding is performed by generating a suited stego signal which enables the transition between the first and the second source. This method, in order to provide good undetectability performance while proposing high embedding rates, has to use RAW images as inputs. Beside quantization and gamma correction, we also show in a longer version [27] how to handle different steps of image developing, including color transforms, demosaicing and rescaling operations.

In future works we want also to investigate other setups for NS steganography, such as choosing other ISO parameters and different camera models. It will also be important to try to improve direct embedding on 8-bit images and to address more practical implementation such as embedding in the JPEG-domain. From the adversary point of view, we would like to see if more appropriate feature could be designed for this category of schemes, this new kind of features should not be only sensitive to image variation, but also to the sensor noise whose variance is function of the pixel luminance. Finally we hope that this methodology will pave the road for new directions in steganography.

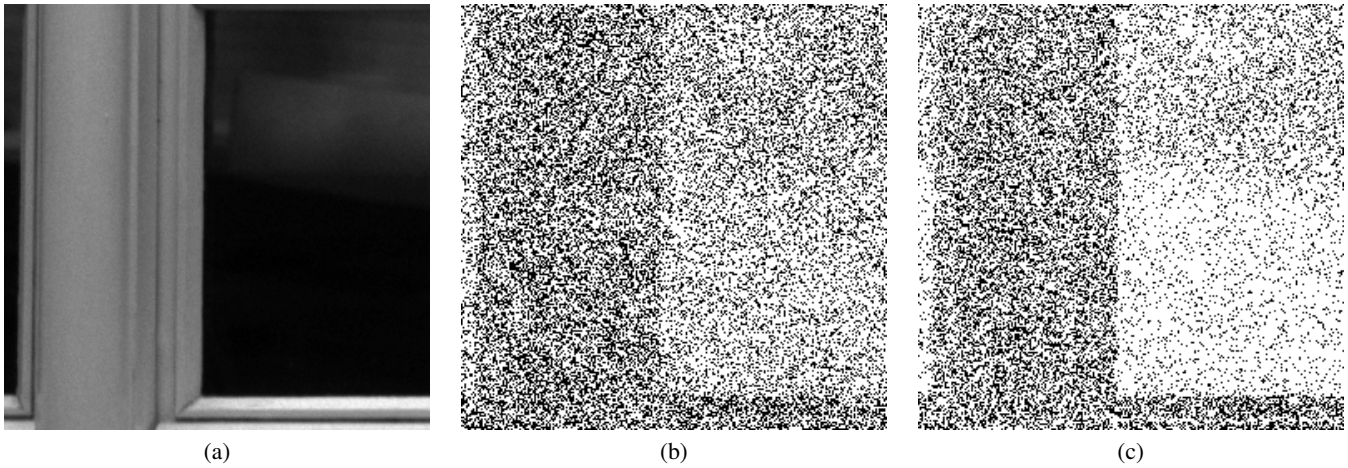


Fig. 4. Portion of an image (a) and locations of embedding changes when the pre-cover is used (see section III-B (b) and when it is not used (c) (for better rendering, inactivate interpolation on your pdf viewer).

VII. ACKNOWLEDGMENTS

The author would like to thank Boris Valet for his work on sensor noise estimation, Cyrille Toulet and Matthieu Marquillie for their help on the Univ-lille HPC, Remi Bardenet for his help on sampling strategies, Tomas Pevny and Andrew Ker for their inspiring conversations of the definition of the source, the conferences reviewers for their valuable comments and corrections, and CNRS for a supporting grant on cyber-security.

REFERENCES

- [1] C. Cachin, "An information-theoretic model for steganography," in *Information Hiding: Second International Workshop IHW'98*, Portland, Oregon, USA, April 1998.
- [2] P. Sallee, "Model-based steganography," in *International Workshop on Digital Watermarking (IWDW)*, LNCS, vol. 2, 2003.
- [3] T. Pevny, T. Filler, and P. Bas, "Using high-dimensional image models to perform highly undetectable steganography," in *Information Hiding 2010*, 2010.
- [4] V. Sedighi, R. Cogranne, and J. Fridrich, "Content-adaptive steganography by minimizing statistical detectability," *Information Forensics and Security, IEEE Transactions on*, vol. 11, no. 2, pp. 221–234, 2016.
- [5] V. Holub, J. Fridrich, and T. Denemark, "Universal distortion function for steganography in an arbitrary domain," *EURASIP Journal on Information Security*, vol. 2014, no. 1, pp. 1–13, 2014.
- [6] B. Li, M. Wang, J. Huang, and X. Li, "A new cost function for spatial image steganography," in *Image Processing (ICIP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 4206–4210.
- [7] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," *Information Forensics and Security, IEEE Transactions on*, vol. 6, no. 3, pp. 920–935, 2011.
- [8] P. Wang, H. Zhang, Y. Cao, and X. Zhao, "Constructing near-optimal double-layered syndrome-trellis codes for spatial steganography," in *ACM workshop on Information hiding and multimedia security*. ACM, 2016.
- [9] J. Fridrich, M. Goljan, and D. Soukal, "Perturbed quantization steganography with wet paper codes," in *Proceedings of the 2004 workshop on Multimedia and security*. ACM, 2004, pp. 4–15.
- [10] T. Denemark and J. Fridrich, "Side-informed steganography with additive distortion," in *Information Forensics and Security (WIFS), 2015 IEEE International Workshop on*. IEEE, 2015, pp. 1–6.
- [11] "<https://sourceforge.net/projects/exiftool/>."
- [12] E. Franz and A. Pfitzmann, "Steganography secure against cover-stego-attacks," in *Information Hiding*. Springer, 1999, pp. 29–46.
- [13] E. Franz and A. Schneidewind, "Pre-processing for adding noise steganography," in *Information Hiding, 7th International Workshop*, 2005, pp. 189–203.
- [14] T. Holotyak, J. Fridrich, and S. Voloshynovskiy, "Blind statistical steganalysis of additive steganography using wavelet higher order statistics," in *Communications and Multimedia Security*, vol. 3677, 2005, pp. 273–274.
- [15] S. Kouider, M. Chaumont, and W. Puech, "Adaptive steganography by oracle (aso)," in *Multimedia and Expo (ICME), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1–6.
- [16] 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, 2008.
- [17] A. Foi, S. Alenius, V. Katkovnik, and K. Egiazarian, "Noise measurement for raw-data of digital imaging sensors by automatic segmentation of nonuniform targets," *IEEE Sensors Journal*, vol. 7, no. 10, pp. 1456–1461, 2007.
- [18] E. M. V. Association *et al.*, "Standard for characterization of image sensors and cameras," *EMVA Standard*, vol. 1288, 2010.
- [19] T. Qiao, F. Retraint, R. Cogranne, and T. H. Thai, "Source camera device identification based on raw images," in *Image Processing (ICIP), 2015 IEEE International Conference on*, Sept 2015, pp. 3812–3816.
- [20] T. H. Thai, R. Cogranne, and F. Retraint, "Camera model identification based on the heteroscedastic noise model," *IEEE Transactions on Image Processing*, vol. 23, no. 1, pp. 250–263, Jan 2014.
- [21] "<http://www.cybercom.net/dcoffin/dcrawl/>."
- [22] A. Papoulis and U. Pillai, *Probability, Random Variables and Stochastic Processes*. Mac Graw Hill, 2002.
- [23] P. Bas, "Monobase," <http://patrickbas.ec-lille.fr/MonoBase/>, July 2016.
- [24] J. Fridrich and J. Kodovsky, "Rich models for steganalysis of digital images," *Information Forensics and Security, IEEE Transactions on*, vol. 7, no. 3, pp. 868–882, 2012.
- [25] J. Kodovsky, J. Fridrich, and V. Holub, "Ensemble classifiers for steganalysis of digital media," *Information Forensics and Security, IEEE Transactions on*, vol. 7, no. 2, pp. 432–444, 2012.
- [26] P. Bas, "Source Codes for Natural Steganography," <https://github.com/patrickbas/NaturalSteganography>, October 2016.
- [27] —, "Natural Steganography: Cover-source Switching For Better Steganography," Jul. 2016, working paper or preprint. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-01360024>