QUANTIZED GAUSSIAN EMBEDDING STEGANOGRAPHY

Mehdi Sharifzadeh, Mohammed Aloraini, and Dan Schonfeld

University of Illinois at Chicago
Electrical and Computer Engineering
Chicago, IL, USA 60607

ABSTRACT

In this paper, we develop a statistical framework for image steganography in which the cover and stego messages are modeled as multivariate Gaussian random variables. By minimizing the detection error of an optimal detector within the generalized adopted statistical model, we propose a novel Gaussian embedding method. Furthermore, we extend the formulation to cost-based steganography, resulting in a universal embedding scheme that works with embedding costs as well as variance estimators. Experimental results show that the proposed approach avoids embedding in smooth regions and significantly improves the security of the state-of-the-art methods, such as HILL, MiPOD, and S-UNIWARD.

***Index Terms*—** Steganography, Optimal detector, Hypothesis testing, Gaussian embedding

1. INTRODUCTION

Steganography is the art of covert communication through a cover medium without raising any suspicion from steganalysis [1]. In this paper, we focus on the most studied cover medium, digital images. Non-adaptive image steganography approaches [2, 3] are easily detectable as they neglect pixel to pixel dependencies [4]. Therefore, to achieve a better security, steganography should be content adaptive, in which message embedding is done while minimizing the caused distortion formulated to a source coding problem [5–9].

There are two types of content adaptive image steganography. We call the first type cost based methods where cost of embedding in each pixel is computed then embedding is done while minimizing the distortion based on calculated costs. In **S**patial **UNI**versal **WA**velet **R**elative **D**istortion (S-UNIWARD) [10], embedding costs are calculated using directional filter banks. In **HI**gh-pass, **L**ow-pass, and **L**ow-pass (HILL), they are calculated using a high-pass filter to find noisy parts, and subsequently smoothing the costs using two low-pass filters [11]. Although, these methods achieve superior results, there is no theoretical relation between statistical security measures and their derived costs [12].

This has been addressed in the second type of image steganography approaches which we call statistical model based approaches. In [13], by modeling the cover as independent Gaussian random variables and assuming a ternary message, embedding is done while minimizing the Kullback-Leibler divergence between the cover and stego messages. The result was further improved in [14], using the same framework but with a generalized Gaussian statistical model for the cover utilizing a better variance estimator and embedding quinary message. Building upon the result of these two works, **Mi**nimizing the **P**ower of **O**ptimal **D**etector (MiPOD) [15] was proposed reaching state-of-the-art performance. In MiPOD, the cover was modeled as independent Gaussian random variables and assuming a ternary message, embedding was done while maximizing the error of a hypothesis testing detector that utilizes a likelihood ratio test.

In all the-state-of-the-art statistical model based steganography approaches, the formulation is derived for a fixed embedding scenario such as ternary (embedding ± 1) or quinary (embedding ± 1,± 2). In other words, for every embedding scenario, the problem requires to be reformulated. The generalization of the problem formulation is a missing piece which is addressed in our work for the first time. Furthermore, the former methods embed a significant portion of the payload in smooth regions with high cost or near zero variance. In this sense, security can be further improved by embedding less in these regions.

In this work, we propose a novel Gaussian embedding technique utilizing a similar cover model to MiPOD. However, for the first time, we model the hidden message as a continuous random variable with Gaussian distribution. This allows us to do q-ary embedding for any q by only changing the quantization levels without requiring problem reformulation for each q. The second contribution is that the explained formulation is also extended to distortion minimization framework. Thus, the proposed method works with any embedding cost calculator as well as any variance estimator.

The statistical model for the cover and stego images are shown in Sec. 2. A hypothesis testing steganalysis framework is developed in Sec. 3. A novel Gaussian embedding method is proposed in Sec. 4. Then, the method is extended to cost based steganography framework in Sec. 4.1. The results and conclusions are provided in Sec. 5 and 6 respectively.

2. STATISTICAL MODELS

Cover images are denoted by $\mathbf{c} = [c_1, \ldots, c_n] \in \mathcal{P} = \{0, \ldots, 255\}^n$, where $\mathcal{P}$ is the set of all vector representation of 8-bit gray-scale images of size $n_1 \cdot n_2 = n$. Each c_i is modeled as a predictable part plus a residual, x_i, that has a Gaussian distribution, $\mathcal{N}(0, \sigma_i^2)$, where σ_i^2 includes both the pixel's and estimation error's variance. By assuming σ_i is much greater than 1, the quantization step size, the probability distribution of the i^{th} cover pixel residual is

$$p_{x_i}(k) \propto \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left(\frac{-k^2}{2\sigma_i^2}\right) \tag{1}$$

Refer to [15] for more information regarding this cover model. We also model all the stego message elements, m_i, as Gaussian random variables with variance β_i. Thus, the probability distribution of m_i is given by

$$p_{m_i}(k) = \frac{1}{\beta_i\sqrt{2\pi}} \exp\left(\frac{-k^2}{2\beta_i^2}\right). \tag{2}$$

The stego image is the summation of the cover image with the stego message elements, i.e. $\mathbf{s} = \mathbf{c} + \mathbf{m}$. Thus, the i^{th} stego pixel residual is $y_i = x_i + m_i$. Based on (1) and (2), the probability distribution of the i^{th} stego pixel residual is:

$$p_{y_i}(k) \propto \frac{1}{\sqrt{2\pi(\sigma_i^2 + \beta_i^2)}} \exp\left(\frac{-k^2}{2(\sigma_i^2 + \beta_i^2)}\right), \tag{3}$$

which is also derived with the assumption of unbounded quantization levels and $\sqrt{\sigma_i^2 + \beta_i^2} \gg \Delta$, for simplicity. The next section is devoted to explain how steganalyzer distinguishes the cover image from the stego image.

3. STEGANALYSIS

We assume the steganalyzer utilizes a likelihood ratio test to do a binary hypothesis testing between $\mathcal{H}_0$ and $\mathcal{H}_1$, representing the cases of receiving a cover or a stego image respectively. Suppose that $\mathbf{r} = [r_1, \ldots, r_n]$ are statistically independent residuals of the received image's pixels. Thus, the likelihood ratio for the whole image can be written as $\prod_{i=1}^{n} \Lambda_i$, in which Λ_i is the likelihood ratio for the i^{th} pixel. In the worst case scenario of an omniscience steganalyzer who knows all the message and cover variances (β_i and σ_i), based on (1) and (3), the likelihood ratio for the i^{th} pixel, Λ_i, is given by

$$\Lambda_i = \frac{p_{y_i}(r_i)}{p_{x_i}(r_i)} = \sqrt{\frac{\sigma_i^2}{\sigma_i^2 + \beta_i^2}} \exp\left(\frac{r_i^2}{2} \frac{\beta_i^2}{\sigma_i^2(\sigma_i^2 + \beta_i^2)}\right). \tag{4}$$

As a result the natural logarithm of the likelihood ratio is

$$\ln \Lambda_i = \ln\left(\sqrt{\frac{\sigma_i^2}{\sigma_i^2 + \beta_i^2}}\right) + \frac{\beta_i^2}{2\sigma_i^2(\sigma_i^2 + \beta_i^2)} r_i^2, \tag{5}$$

which is a constant plus a Gamma distributed random variable, $\Gamma(k_i, \theta_i)$, since r_i has a normal distribution. k_i and θ_i are the shape and scale parameters respectively. For deriving these parameters for both hypotheses, the following approximations, based on $\beta_i^2/\sigma_i^2 < 1$ and Taylor series of $\ln(1 + x)$, are utilized. If $x = \beta_i^2/\sigma_i^2$,

$$\ln\left(\frac{\sigma_i^2}{\sigma_i^2 + \beta_i^2}\right) = -\ln\left(1 + \frac{\beta_i^2}{\sigma_i^2}\right) \approx -\frac{\beta_i^2}{\sigma_i^2} + \frac{1}{2}\left(\frac{\beta_i^2}{\sigma_i^2}\right)^2. \tag{6}$$

If $x = -\beta_i^2/(\sigma_i^2 + \beta_i^2)$, the approximation is

$$\ln\left(\frac{\sigma_i^2}{\sigma_i^2 + \beta_i^2}\right) \approx -\frac{\beta_i^2}{\sigma_i^2 + \beta_i^2} - \frac{1}{2}\left(\frac{\beta_i^2}{\sigma_i^2 + \beta_i^2}\right)^2. \tag{7}$$

In addition, further simplification, based on assuming $\beta_i^2/\sigma_i^2 < 1$ and Taylor series of $x/(1+x)$, can be done as shown below.

$$\frac{\beta_i^2}{\sigma_i^2 + \beta_i^2} \approx \frac{\beta_i^2}{\sigma_i^2} \tag{8}$$

Given $\mathcal{H}_0$, the Gamma distribution parameters are $k = 1/2$ and $\theta_i = \beta_i^2/(\sigma_i^2 + \beta_i^2)$. The resulted mean and variance of the natural logarithm of the likelihood ratio are:

$$\begin{cases} \mathrm{E}^{\mathcal{H}_0}_{r_i|\sigma_i,\beta_i}[\ln \Lambda_i] = \ln\left(\sqrt{\frac{\sigma_i^2}{\sigma_i^2+\beta_i^2}}\right) + k\theta_i \approx \frac{-1}{4}\left(\frac{\beta_i^2}{\sigma_i^2}\right)^2 \\ \mathrm{Var}^{\mathcal{H}_0}_{r_i|\sigma_i,\beta_i}[\ln \Lambda_i] = k\theta_i^2 \approx \frac{1}{2}\left(\frac{\beta_i^2}{\sigma_i^2}\right)^2 \end{cases} \tag{9}$$

where the approximations are based on (7) and (8). Given $\mathcal{H}_1$, the Gamma distribution parameters are $k = 1/2$ and $\theta_i = \beta_i^2/\sigma_i^2$. The resulted mean and variance are:

$$\begin{cases} \mathrm{E}^{\mathcal{H}_1}_{r_i|\sigma_i,\beta_i}[\ln \Lambda_i] = \ln\left(\sqrt{\frac{\sigma_i^2}{\sigma_i^2+\beta_i^2}}\right) + k\theta_i \approx \frac{1}{4}\left(\frac{\beta_i^2}{\sigma_i^2}\right)^2 \\ \mathrm{Var}^{\mathcal{H}_0}_{r_i|\sigma_i,\beta_i}[\ln \Lambda_i] = k\theta_i^2 = \frac{1}{2}\left(\frac{\beta_i^2}{\sigma_i^2}\right)^2 \end{cases} \tag{10}$$

where approximation is based on (6). Summation of Gamma distributed variables with equal shape and bounded scale parameters converges to normal distribution with mean and variance equal to summation of means and variances respectively [16]. Thus, based on (9) and (10), for large enough n, probability distribution of $\sum_{i=1}^{n} \ln(\Lambda_i)$, can be approximated as:

$$\begin{cases} \mathcal{N}(\frac{-1}{4}\alpha,\ \frac{1}{2}\alpha) & \text{if } \mathcal{H}_0 \text{ is true,} \\ \mathcal{N}(\frac{+1}{4}\alpha,\ \frac{1}{2}\alpha) & \text{if } \mathcal{H}_1 \text{ is true,} \end{cases} \tag{11}$$

where α is as follows:

$$\alpha = \sum_{i=1}^{n} \left(\frac{\beta_i^2}{\sigma_i^2}\right)^2. \tag{12}$$

This result is also consistent with the shift hypothesis stating that embedding affects only the mean of the detector's output [17]. The detector's decision is given by

$$\sum_{i=1}^{n} \ln(\Lambda_i) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma, \tag{13}$$

where γ is the decision threshold. To find γ, we employ minimax criterion, one of the most common optimality criteria for hypothesis testing, which does not require the hypothesis prior probabilities and assumes the least favorable ones. The least favorable priors are the solution of the equalizer rule, and they are $1/2$ for this problem assuming symmetric costs.

Therefore, the decision boundary, γ, is zero which has the minimum expected risk over all possible prior distributions. The expected value of the detection error, summation of false alarm and missed detection, for such a detector is given by

$$\mathrm{P_E} = \phi(\frac{-\frac{\alpha}{4}}{\sqrt{\frac{\alpha}{2}}}) = \phi(-\sqrt{\frac{\alpha}{8}}), \tag{14}$$

where ϕ is the cumulative distribution of standard normal distribution. Eq. (14) is a monotonically decreasing function of α. Therefore, the steganographer can minimize α instead of maximizing the detector's error, $\mathrm{P_E}$.

4. GAUSSIAN EMBEDDING METHOD

In this section, a novel image steganography method is introduced based on maximizing the detection error of the optimal detector shown in previous section. The problem of optimal embedding for a fixed payload can be written as:

$$\begin{cases} \underset{(\beta_1,\ldots,\beta_n)}{\arg\max}\ \mathrm{P_E} \equiv \underset{(\beta_1,\ldots,\beta_n)}{\arg\min}\ \alpha \\ \sum\limits_{i=1}^{n} \mathrm{H}(p_{m_i}) = np \end{cases} \tag{15}$$

where p is the relative payload per pixel in nats. Shannon entropy of the hidden message elements, m_i, a Gaussian random variable with variance β_i^2, is given by

$$\mathrm{H}(p_{m_i}) = \frac{1}{2}\ln(2\pi \mathrm{e}\beta_i^2) \tag{16}$$

The solution of (15) using (12) and Lagrangian multipliers is the solution of the following equation

$$\frac{\partial}{\partial \beta_i}\left(\sum_{j=1}^{n}\left(\frac{\beta_j^2}{\sigma_j^2}\right)^2 + \lambda\Big(np - \frac{1}{2}\sum_{j=1}^{n}\ln(2\pi \mathrm{e}\beta_j^2)\Big)\right) = 0, \tag{17}$$

for $j = 1, \ldots, n$, where λ is the Lagrangian multiplier that is calculated using the payload constraint in (15), and thus will be shown as a function of the payload, p. Solution of (15) is

$$\beta_i^* = \frac{\sqrt[4]{\lambda(p)}}{\sqrt{2}}\sigma_i \quad \text{for } i = 1, \ldots, n \tag{18}$$

This solution guarantees more embedding in noisy and textured regions with high residual variances, σ_i, and less embedding in smooth regions with low residual variances.

By assuming a $(2q+1)$-ary embedding scenario, the message is a Gaussian random variable with variance β_i truncated and quantized to $\mathcal{Q} = \{-q, \ldots, -1, 0, 1, \ldots, +q\}$. The probability distribution of m_i, is given by

$$p_{m_i}(k) = \frac{\phi(\frac{k+0.5}{\beta_i}) - \phi(\frac{k-0.5}{\beta_i})}{\phi(\frac{q+0.5}{\beta_i}) - \phi(\frac{-q-0.5}{\beta_i})} \quad \forall k \in \{\text{-}q, \ldots, \text{+}q\} \tag{19}$$

which is probability of changing the i^{th} pixel by k. They are computed by solving the following system of equations with $n+1$ equations and variables, $\beta_1, \ldots, \beta_n$ and λ, using Newton-Raphson method.

$$\begin{cases} \beta_i^* = \frac{\sqrt[4]{\lambda(p)}}{\sqrt{2}}\sigma_i \quad \text{for } i = 1, \ldots, n \\ -\sum\limits_{i=1}^{n}\sum\limits_{k=-q}^{q}(p_{m_i}(k)\ln p_{m_i}(k)) = np \end{cases} \tag{20}$$

For implementing the proposed embedding technique by syndrome-trellis codes [18], we need to find the embedding costs for all the pixels. These costs are calculated by solving the following system of equations, having Gibbs form [19].

$$p_{m_i}(k) = \mathrm{e}^{-\rho_i(k)} / \sum_{d=-q}^{q}\mathrm{e}^{-\rho_i(d)}, \tag{21}$$

for $\forall i \in \{1, \ldots, n\},\ \forall k \in \{-q, \ldots, q\}$, where $\rho_i(k)$, indicates the amount of distortion added to image by changing the i^{th} pixel by k. There are $n \times q$ equations and variables by assuming symmetric costs, and $\rho_i(0) = 0,\ \forall i \in \{1, \ldots, n\}$.

4.1. Extension to Cost-based Methods

In cost-based methods, steganographer tries to minimize expected value of a distortion function, $\mathrm{D}(\mathbf{s}, \mathbf{c})$ ($\mathbf{s}$ and $\mathbf{c}$ are stego and cover images respectively). We define distortion as the expected value of absolute difference of pixel intensities between cover and stego same as the prior arts. Thus, steganography problem for a payload limited sender can be written as:

$$\begin{cases} \underset{(\beta_1,\ldots,\beta_n)}{\arg\min}\ \mathrm{E}[\mathrm{D}(\mathbf{s}, \mathbf{c})] = \underset{(\beta_1,\ldots,\beta_n)}{\arg\min} \sum\limits_{i=1}^{n} \mathrm{E}_{m_i|\beta_i}\left[\rho_i|s_i\text{-}c_i|\right] \\ \sum\limits_{i=1}^{n} \mathrm{H}(p_{m_i}) = np \end{cases} \tag{22}$$

where ρ_i is cost of embedding ± 1 in i^{th} pixel calculated by any of mentioned algorithms [10, 11, 20]. Assuming the same Gaussian embedding scenario with $m_i \sim \mathcal{N}(0, \beta_i^2)$, the expected value of the defined distortion is

$$\mathrm{E}_{m_i|\beta_i}\left[\rho_i|s_i - c_i|\right] = \mathrm{E}_{m_i|\beta_i}\left[\rho_i|m_i|\right] = \rho_i\beta_i\sqrt{\frac{2}{\pi}}. \tag{23}$$

Using Lagrangian multipliers approach, solution of (22) is

$$\beta_j^* = \frac{\lambda(p)}{\rho_j}\sqrt{\frac{\pi}{2}}, \tag{24}$$

where $\lambda(p)$ is the Lagrangian multiplier calculated using the payload constraint in (22). The rest of the embedding approach is similar to (20) and (21).

5. EXPERIMENTS AND DISCUSSIONS

We use BOSSbase 1.01 database including 10k gray-scale 512×512 pixels images [21]. To evaluate the performance of each method, the average detection error, the average false positive and negative rates, is reported. It is evaluated by an ensemble classifier steganalyzer [22] with a 10-fold cross validation, trained on SRM features [23]. 4096 and 4096 images are chosen randomly as training/validation set and testing set respectively. Three state-of-the-art content-adaptive image steganography methods, HILL [11], MiPOD [15], and

Table 1. Detection error computed by steganalysis using SRM features in different payloads ranging from 0 to 1 bpp for HILL algorithm and its Gaussian versions with different q's in a $(2q{+}1)$-ary embedding scenario.

Payload	**0.01**	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.75**	**1**
HILL	.500±.0048	.487±.0041	.462±.0045	.409±.0061	.357±.0044	.302±.0046	.252±.0043	.143±.0048	.070±.0052
q=1	.499±.0047	.487±.0040	.467±.0057	.419±.0059	.365±.0059	.310±.0050	.258±.0063	.150±.0064	.075±.0080
q=2	.499±.0062	.487±.0026	.467±.0047	.418±.0037	.367±.0058	.314±.0065	.265±.0059	.159±.0074	.089±.0063
q=3	**.497±.0042**	**.487±.0035**	**.468±.0046**	**.421±.0047**	**.373±.0061**	**.323±.0065**	**.275±.0073**	**.171±.0070**	**.103±.0060**
q=4	.500±.0024	.488±.0043	.469±.0048	.424±.0049	.375±.0038	.320±.0062	.274±.0055	.173±.0062	.107±.0063
q=5	.499±.0036	.489±.0031	.469±.0067	.421±.0040	.373±.0056	.322±.0069	.272±.0071	.175±.0056	.109±.0075

Table 2. Detection error computed by steganalysis using SRM features in different payloads ranging from 0 to 1 bpp .

Payload	**0.01**	**0.05**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.75**	**1**
G-HILL	.497±.0042	.487±.0035	.468±.0046	.421±.0047	.373±.0061	.323±.0065	.275±.0073	.171±.0070	.103±.0060
HILL	.500±.0048	.487±.0041	.462±.0045	.409±.0061	.357±.0044	.302±.0046	.252±.0043	.143±.0048	.070±.0052
G-MiPOD	.498±.0027	.482±.0037	.454±.0050	.404±.0047	.347±.0053	.292±.0050	.243±.0068	.153±.0072	.090±.0059
MiPOD	.498±.0038	.468±.0044	.434±.0041	.373±.0046	.322±.0050	.269±.0049	.225±.0061	.128±.0057	.061±.0068
G-S-UNIWARD	.499±.0047	.485±.0030	.457±.0030	.400±.0038	.338±.0069	.279±.0072	.229±.0072	.135±.0070	.077±.0048
S-UNIWARD	.498±.0030	.479±.0036	.448±.0051	.375±.0048	.309±.0062	.251±.0049	.203±.0057	.105±.0057	.048±.0058

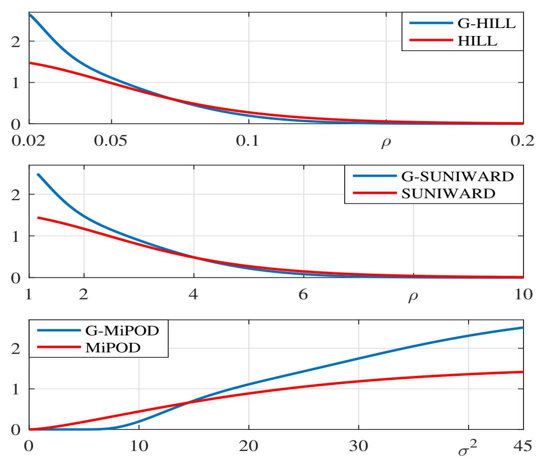

Fig. 1. Bits embedded in pixels of "1.pgm" against their embedding costs or variances for embedding 0.3 bpp for HILL, S-UNIWARD, and MiPOD.

S-UNIWARD [10,24], are used for evaluations with their best performing settings. In addition, in all the methods, we avoid embedding in saturated pixels due to performance drop [25].

5.1. Determining Maximum Pixel Change (q)

To find the optimal q in (19), we have evaluated the performance of HILL and its Gaussian versions derived in (24) for $q = 1, \dots, 5$. Results are shown in Table 1. It is observed that larger q results in higher security, but no significant improvement is seen for $q > 3$. In addition, the complexity of the coding algorithm increases as q increases [18]. Therefore, we choose $q = 3$ (septenary) for the rest of the experiments.

5.2. Comparison with Prior Arts

In Table 2, the security of three stat-of-the-art image steganography methods, HILL [11], MiPOD [15], and S-UNIWARD [10], is compared with their Gaussian versions with $q = 3$ shown with a prefix of G. In G-HILL and G-S-UNIWARD, the message variances are calculated by (24) using the embedding cost, ρ, computed by HILL and S-UNIWARD respectively. In G-MiPOD, the message variances are calculated by (18) using the variance estimator of MiPOD to compute pixel residual variances. It is observed that utilizing the Gaussian embedding scheme, security of all the algorithms are significantly improved for every payload (0-1 bits per pixel). We believe that the improvement is due to the fact that the proposed Gaussian method embeds more bits in textured areas (pixels with low embedding costs, ρ, or equivalently high residual variances, σ^2) and less in smooth areas (pixels with high embedding costs or equivalently low residual variances). To verify this, in Fig. 1, we plotted the distribution of the message for original and Gaussian versions of three state-of-the-art methods in pixels of one of the images of the BOSSbase 1.01 database (1.pgm) versus their embedding costs computed by HILL and S-UNIWARD, and also the pixels residual variances computed by MiPOD. It shows more bits are hidden in pixels with small ρ or high σ for all three prior arts compared to their Gaussian version.

6. CONCLUSIONS

A statistical framework is developed for steganography problem. The cover and the stego messages are modeled by independent Gaussian random variables. Then, a novel Gaussian embedding technique is proposed by minimizing the detection error of an optimal hypothesis testing detector and it works with pixel embedding costs as well as residual variances. We achieve better performance comparing to prior arts against advanced steganalysis due to better concentration of the payload in textured regions and less embedding in smooth regions. In future, we will extend this method to batch steganography.

7. REFERENCES

[1] G. J Simmons, "The prisoners problem and the subliminal channel," in *Advances in Cryptology*. Springer, 1984, pp. 51–67.

[2] A. Cheddad, J. Condell, K. Curran, and P. Mc Kevitt, "Digital image steganography: Survey and analysis of current methods," *Signal processing*, vol. 90, no. 3, pp. 727–752, 2010.

[3] N. F Johnson and S. Jajodia, "Exploring steganography: Seeing the unseen," *Computer*, vol. 31, no. 2, 1998.

[4] 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.

[5] Jessica Fridrich, "Minimizing the embedding impact in steganography," in *Proceedings of the 8th workshop on Multimedia and security*. ACM, 2006, pp. 2–10.

[6] J. Fridrich and T. Filler, "Practical methods for minimizing embedding impact in steganography," in *Electronic Imaging 2007*. International Society for Optics and Photonics, 2007, pp. 650502–650502.

[7] C. E Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec*, vol. 4, no. 142-163, pp. 1, 1959.

[8] Mehdi Sharifzadeh, Chirag Agarwal, Mohammed Aloraini, and Dan Schonfeld, "Convolutional neural network steganalysis's application to steganography," in *2017 IEEE Visual Communications and Image Processing (VCIP)*. IEEE, 2017, pp. 1–4.

[9] M. Sharifzadeh, C. Agarwal, M. Salarian, and D. Schonfeld, "A new parallel message-distribution technique for cost-based steganography," *arXiv preprint arXiv:1705.08616*, 2017.

[10] 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.

[11] B. Li, M. Wang, J. Huang, and X. Li, "A new cost function for spatial image steganography," in *2014 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2014, pp. 4206–4210.

[12] Rainer Böhme, *Advanced statistical steganalysis*, Springer Science & Business Media, 2010.

[13] Jessica J Fridrich and Jan Kodovskỳ, "Multivariate gaussian model for designing additive distortion for steganography.," in *ICASSP*, 2013, pp. 2949–2953.

[14] Vahid Sedighi, Jessica Fridrich, and Rémi Cogranne, "Content-adaptive pentary steganography using the multivariate generalized gaussian cover model," in *Media Watermarking, Security, and Forensics 2015*. International Society for Optics and Photonics, 2015, vol. 9409, p. 94090H.

[15] V. Sedighi, R. Cogranne, and J. Fridrich, "Content-adaptive steganography by minimizing statistical detectability," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 2, pp. 221–234, 2016.

[16] AM Mathai, "Storage capacity of a dam with gamma type inputs," *Annals of the Institute of Statistical Mathematics*, vol. 34, no. 1, pp. 591–597, 1982.

[17] A. D. Ker, "Batch steganography and pooled steganalysis," in *Information Hiding*. Springer, 2006, vol. 4437, pp. 265–281.

[18] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 920–935, 2011.

[19] T. Filler and J. Fridrich, "Gibbs construction in steganography," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 4, pp. 705–720, 2010.

[20] T. Pevnỳ, T. Filler, and P. Bas, "Using high-dimensional image models to perform highly undetectable steganography," in *International Workshop on Information Hiding*. Springer, 2010, pp. 161–177.

[21] P. Bas, T. Filler, and T. Pevnỳ, " break our steganographic system: The ins and outs of organizing boss," in *International Workshop on Information Hiding*. Springer, 2011, pp. 59–70.

[22] J. Kodovsky, J. Fridrich, and V. Holub, "Ensemble classifiers for steganalysis of digital media," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 432–444, 2012.

[23] J. Fridrich and J. Kodovsky, "Rich models for steganalysis of digital images," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 868–882, 2012.

[24] T. Denemark, J. Fridrich, and V. Holub, "Further study on the security of s-uniward," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2014, pp. 902805–902805.

[25] V. Sedighi and J. Fridrich, "Effect of saturated pixels on security of steganographic schemes for digital images," in *Image Processing (ICIP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 2747–2751.