Fast Additive Noise Steganalysis

Jeremiah J. Harmsen[a], Kevin D. Bowers[b] and William A. Pearlman[a]

[a]Center for Image Processing Research,
Electrical Computer and Systems Engineering Department,
Rensselaer Polytechnic Institute, Troy, NY
[b]Computer Science Department,
Carnegie Mellon University, Pittsburgh, PA

ABSTRACT

This work reduces the computational requirements of the additive noise steganalysis presented by Harmsen and Pearlman. The additive noise model assumes that the stegoimage is created by adding a pseudo-noise to a coverimage. This addition predictably alters the joint histogram of the image. In color images it has been shown that this alteration can be detected using a three-dimensional Fast Fourier Transform (FFT) of the histogram. As the computation of this transform is typically very intensive, a method to reduce the required processing is desirable. By considering the histogram between pairs of channels in RGB images, three separate two-dimensional FFTs are used in place of the original three-dimensional FFT. This method is shown to offer computational savings of approximately two orders of magnitude while only slightly decreasing classification accuracy.

Keywords: Steganalysis, steganography, additive noise

1. INTRODUCTION

Typically, steganalysis systems are evaluated on their detection rates. The real world application of steganalysis techniques imposes a number of additional judging criteria. One of the foremost is the capacity of a system. Clearly a good system must not only reliably detect steganographic content, but also do so in an efficient manner. The volume of traffic on many networks can present an issue for detection schemes that require extensive computations.

With this goal of efficiency in mind, we present an extension to the Additive Noise Steganalysis.[1] The method uses reduced dimensionality transforms to greatly improve the speed of the detection system.

2. ADDITIVE NOISE STEGANALYSIS

Steganography is the practice of hiding information. The goal in steganography is much more than keeping the message contents private, it is to conceal the very presence of communication. A usable hiding technique must be able to embed a large amount of information into a medium, such as an image, without suspiciously altering the medium. One method of accomplishing this is disguising a message as noise and combining it with a coverimage. Operating on the premise that noise is an inherent part of signal processing, a secret message is hidden in a pseudo-random sequence called the stegonoise. The stegonoise is combined with the coverimage to produce the stegoimage.

The additive noise model[1] of steganography defines such a class of embedding methods with the following properties: the stegoimage is created by the *addition* of a stegonoise to a coverimage, the stegonoise is *independent of the coverimage*, and the stegonoise is *independent and identically distributed*. A block diagram of the model is shown in Figure 1. There are a number of schemes that are described by these properties, such as spread spectrum image steganography (SSIS)[2] and more generally stochastic modulation.[3]

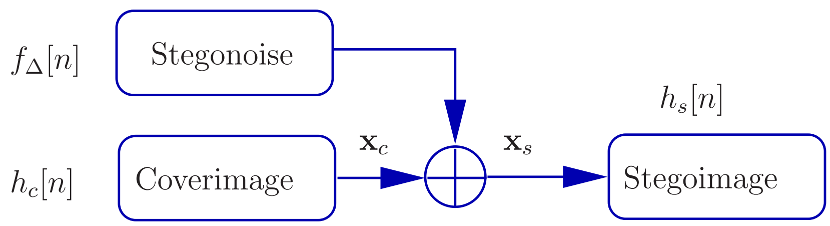

Figure 1. Additive Noise Steganography Model

2.1. Additive Noise Notation

Let $\mathbf{x} = (x_1 \ \dots \ x_{hw})$ denote a list of pixels in an $h \times w$ pixel image. Each pixel $x_i \in \mathbf{Z}^d$ is a length-d vector of the intensities from the ith position of each channel. The coverimage and the stegoimage are denoted $\mathbf{x}_c$ and $\mathbf{x}_s$ respectively.

Let the stegonoise be defined as, $\mathbf{f}_{\boldsymbol{\Delta}} = [f_1 \ \dots \ f_{hw}]$, with $f_a \in \mathbf{Z}^d$. Here f_a is a vector of d i.i.d. realizations of a random variable with probability mass function (PMF) $f_\Delta(n)$.

The stegoimage is created by summing the coverimage and the stegonoise,

$$\mathbf{x}_s = \mathbf{x}_c + \mathbf{f}_{\boldsymbol{\Delta}} = [x_1 + f_1, \dots, x_{hw} + f_{hw}]. \tag{1}$$

The joint histogram of the image is,

$$h(n_1, \dots, n_d) = \sum_{i=1}^{hw} I(x_i, n_1, \dots, n_d) \tag{2}$$

where

$$I(\mathbf{a}, n_1, \dots, n_d) = \begin{cases} 1, & \mathbf{a} = (n_1, \dots, n_d) \\ 0, & \text{else.} \end{cases}$$

We use the notational simplification $h(n_1, \dots, n_d) = h(\mathbf{n})$ and will denote the histograms of the coverimage and stegoimage as $h_c(\mathbf{n})$ and $h_s(\mathbf{n})$ respectively.

2.2. Convolution

As the stegonoise is independent of the coverimage and i.i.d., the stego-histogram is simply related to the cover-histogram and stegonoise PMF. That is, the stego-histogram is equal to the convolution of the cover-histogram and the stegonoise PMF along each of the histogram dimensions,

$$h_s(\mathbf{n}) = h_c(n_1, \dots, n_d) \bigotimes_{i=1}^{d} f_\Delta(n_i), \tag{3}$$

where $f_\Delta(n)$ is the stegonoise probability mass function. Intuitively, the addition of the stegonoise affects the cover-histogram as a FIR filtering with tap weights defined by $f_\Delta(n)$.

2.3. Characteristic Functions

The convolution of (3) may also be viewed in the frequency domain. We denote a j-dimensional DFT as $DFT_j(\cdot)$. This gives us the following characteristic functions,

$$F_\Delta(k) \triangleq DFT_1(f_\Delta(n)), \tag{4a}$$

$$H_c(\mathbf{k}) \triangleq DFT_d(h_c(\mathbf{n})), \tag{4b}$$

$$H_s(\mathbf{k}) \triangleq DFT_d(h_s(\mathbf{n})). \tag{4c}$$

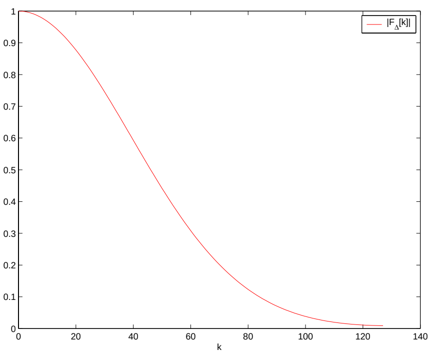

Figure 2. $|F_{\Delta}(k)|$

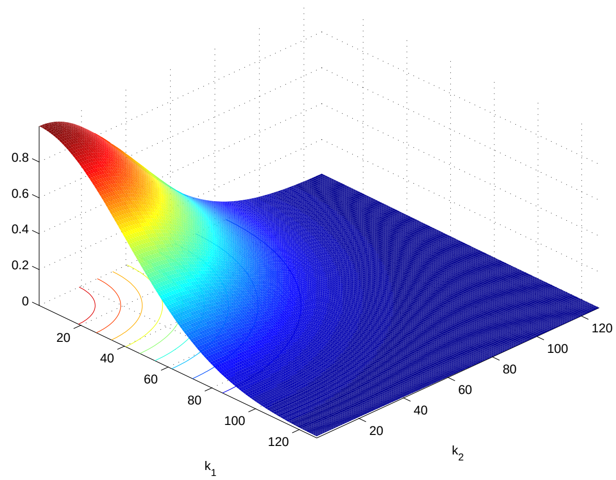

Figure 3. $|F_{\Delta}(k_1)F_{\Delta}(k_2)|$

As the transforms of (4) are of real data they will exhibit symmetries.[4] We restrict $\mathbf{k}$ to $\mathcal{K}_d$, where $\mathcal{K}_d$ is the set of non-redundant coordinates in a d-dimensional transform.

Using these characteristic functions we rewrite (3) in the frequency domain as,

$$H_s(\mathbf{k}) = H_c(\mathbf{k}) \prod_{i=1}^{d} F_{\Delta}(k_i). \tag{5}$$

The d-dimensional transform of an image histogram will be referred to at the Histogram Characteristic Function ($\mathcal{HCF}$). Note that the stegoimage-$\mathcal{HCF}$ is the multiplication of the coverimage-$\mathcal{HCF}$ with the DFT of the stegonoise PMF. Figures 2 and 3 show the DFT of a $\mathcal{N}(0,1)$ stegonoise for the one and two dimensional cases respectively.

2.4. Center of Mass

The embedding of stegonoise affects the coverimage-$\mathcal{HCF}$ by (5). This modification can be detected through the center of mass (COM) statistic of the $\mathcal{HCF}$. The COM gives a general description of the energy distribution in the $\mathcal{HCF}$.

The center of mass (COM) for a d-dimensional $\mathcal{HCF}$ is a vector of d elements,

$$C(H(\mathbf{k})) = \left[\mathcal{C}^1 \ \ldots \ \mathcal{C}^d\right], \tag{6}$$

with

$$\mathcal{C}^i \triangleq \frac{\sum_{\mathbf{k}\in\mathcal{K}_d} k_i |H(\mathbf{k})|}{\sum_{\mathbf{j}\in\mathcal{K}_d} |H(\mathbf{j})|}. \tag{7}$$

Using the discrete Čebyšev inequality,[5] it is shown that for a stegonoise with a non-increasing $|F_{\Delta}(k)|$, $\mathcal{C}^i$ will always decrease or remain the same.[1]

This COM can be used as the feature for a variety of pattern recognition tasks.

3. LOWER DIMENSIONAL TRANSFORMS

As the number of channels in an image, d, increases, the complexity of finding the d-ary COM increases. For an RGB image with 8 bits/channel the DFT operates on a size 256^3 data structure. To reduce the number of computations, a marginal histogram may be used. By considering the joint histogram of a subset of the channels, a feature vector is produced using reduced resources.

3.1. Marginal Histogram

To consider a subset of the full histogram, we define a set, $\mathcal{A}$, that contains the indices of the components we wish to study. The number of components listed in the subset is s, i.e. $|\mathcal{A}| = s$.

The *marginal histogram* is defined as,

$$h^{\mathcal{A}}(n_1, \ldots, n_s) \triangleq \text{joint histogram of components listed in } \mathcal{A}. \tag{8}$$

Taking the s-dimensional DFT of the marginal-histogram we have the *marginal-$\mathcal{HCF}$*,

$$H^{\mathcal{A}}(\mathbf{k}) \triangleq DFT_s\left(h^{\mathcal{A}}(\mathbf{n})\right). \tag{9}$$

We create a vector by concatenating the COMs of each length s subset. This vector is called the *composite-COM* and defined as,

$$C^s_{comp} \triangleq \left[C\left(H^{\mathcal{A}_1}(\mathbf{k})\right) \;\; \ldots \;\; C\left(H^{\mathcal{A}_g}(\mathbf{k})\right)\right], \tag{10}$$

where $A_1, \ldots, A_g$ is an enumeration of all the sets with $|A_k| = s$, and $g = \begin{pmatrix} d \\ s \end{pmatrix}$.

Thus there are g marginal histograms of length s and the composite-COM is a vector of length,

$$s \begin{pmatrix} d \\ s \end{pmatrix} = \frac{d!}{(d-s)!(s-1)!}. \tag{11}$$

Note that $\ell\left(C^s_{comp}\right) \geq d$ with equality if $s = \{1, d\}$.

3.2. Complexity Comparison

For a DFT of elements on support $[0, \ldots, p-1]^d$ the computational complexity is $O(p^d \log p)$. As the number of channels, d, increases the complexity increases exponentially. It is for this reason we consider using the marginal histogram.

The complexity finding the composite-COM using a size s marginal is,

$$O\left(\begin{pmatrix} d \\ s \end{pmatrix} p^s \log p\right). \tag{12}$$

In Appendix A it is shown that for $1 \leq s \leq d \leq p$ the composite-COM of size s is less computationally intensive than the full d-dimensional DFT.

Table 1. Classification Results

s	False Positive	Stegoimage Correct
1 (3×1D)	28.3%	77.15%
2 (3×2D)	13.1%	96.45%
3 (1×3D)	10.15%	99.5%

4. RESULTS

The above methods are now applied to RGB images using the Spread Spectrum Image Steganography (SSIS).[2] The embedding process of this hiding method is simulated by the addition of an i.i.d. Gaussian stegonoise with zero mean and unit variance ($\mathcal{N}(0,1)$). The test images are color images in the RGB format,[6] thus $d = 3$. So $\mathbf{x}_c = [x_1 \ \dots \ x_{hw}]$ with $x_a \in \mathbf{Z}^3$. Here $x_a = [r_a \ g_a \ b_a]$, the respective red, green and blue component intensities.

For embedding we assume that the stegonoise is generated using a continuous probability density function. After initial embedding a rounding procedure is used.

The stegonoise PMF is a discretization of the continuous Gaussian distribution,

$$f_\Delta(n) = \int_{n-0.5}^{n+0.5} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx. \tag{13}$$

For evaluation, three Bayesian[7] classifiers are trained and tested in parallel, allowing for direct comparison of results. The first is a classifier for the $s = 1$ case. This is equivalent to finding the one-dimensional COM for each color component individually and combining them into the composite-COM of length three. The second classifier uses $s = 2$, thus three, two-dimensional COMs are found. The COMs generated by the RG, GB, and BR components produce a composite-COM of length six. Finally, the $s = 3$ is equivalent to the original method, wherein a single three-dimensional COM is found.

To evaluate the classifiers, the twenty four images are divided into four groups. Each group is created by randomly selecting (without replacement) the appropriate number of images. The groups are,

1. 10 Unaltered image composite-COMs used as positive training exemplars.
2. 10 SSIS $\mathcal{N}(0,1)$ image composite-COMs negative training exemplars.
3. 2 Unaltered image composite-COMs classified.
4. 2 SSIS image COMs composite-COMs classified.

These four groups are used to train and test each of the classifiers. The classification results, averaged over 1000 iterations, can be seen in Table 1. As shown the 1D case ($s = 1$) performs poorly, with missed signals and false positives of over 25%. The 2D case ($s = 2$) performs very closly to the original method, with a difference of approximatly 3%.

In addition the computation time of the COM from the image is shown in Table 2. The times are shown relative to the computation time of the three-dimensional COM. The experimental time is for an implementation of the system on MATLAB 6.5 while the theoretical time is found using (14b). As expected the fast methods have significantly reduced the computation time for an image. The $s = 1$ case leads to an improvement in speed of approximately three orders of magnitude. The $s = 2$ case does particularly well: for this test set the capacity has been increased by a factor of 75, while sacrificing approximately 3% accuracy.

Table 2. Relative Time To Compute Center of Mass

s	**Experimental Time**	**Theoretical Time**
1	0.00058	0.000046
2	0.01320	0.011700
3	1.00000	1.000000

5. CONCLUSION

In a real-world application, the computational efficiency of a detection algorithm plays a significant role in the success or failure of a system. To properly monitor the vast amount of information moving over a network, a detection algorithm must expeditiously process each piece of information and produce an accurate judgment on the nature of that information.

Computational efficiency will increase in importance as new networking technologies facilitate the transfer of greater and greater amounts of information. The extensions presented in this work anticipate this challenge and significantly improve the computational performance of additive noise steganalysis, while incurring only a small decrease in accuracy.

APPENDIX A. COMPLEXITY PROOF

The complexity of a DFT over support $[0, p-1]^d$ is $O(p^d \log p)$. Comparing this to the complexity of generating the composite-COM using marginals of size s we have, (for $1 \leq s \leq d \leq p$),

$$\frac{\binom{d}{s} p^s \log p}{p^d \log p} = \frac{p^s \binom{d}{s}}{p^d} \tag{14a}$$

$$= \frac{p^s d!}{p^d s!(d-s)!} \tag{14b}$$

$$\leq \frac{p^s d^{d-s}}{p^d (d-s)!} \tag{14c}$$

$$\leq \frac{p^s p^{d-s}}{p^d (d-s)!} \tag{14d}$$

$$\leq \frac{1}{(d-s)!} \tag{14e}$$

$$\leq 1 \tag{14f}$$

Where the inequality of (14c) comes from the relation,

$$\frac{d!}{s!} \leq d^{d-s}.$$

ACKNOWLEDGMENTS

The support of the Center for Integrated Transmission and Exploitation (CITE) and the Information Directorate of the Air Force Research Laboratory, Rome, NY is gratefully acknowledged.

REFERENCES

1. J. J. Harmsen and W. A. Pearlman, "Steganalysis of additive noise modelable information hiding," in *Proc. SPIE Electronic Imaging 5022*, (Santa Clara, CA), Jan. 21–24, 2003.
2. L. M. Marvel, C. G. Boncelet, Jr, and C. T. Retter, "Spread spectrum image steganography," *IEEE Trans. Image Processing* **8**, pp. 1075–1083, Aug. 1999.
3. J. Fridrich and M. Goljan, "Digital image steganography using stochastic modulation," in *Proc. SPIE Electronic Imaging*, (Santa Clara, CA), Jan. 21–24, 2003.
4. A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 2 ed., 1999.
5. D. S. Mitrinović, J. E. Pečarić, and A. M. Fink, *Classical and New Inequalities in Analysis*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1993.
6. R. Franzen, "Kodak lossless true color image suite: PhotoCD PCD0992," Mar. 27, 2002. Available: http://sqez.home.att.net/thumbs/Thumbnails.html.
7. R. O. Duda, P. E. Hart, and H. G. Stork, *Pattern Classification*, Wiley-Interscience, New York, NY, 2 ed., 2000.