STEGANALYSIS USING NOISE VARIANCE ESTIMATION

Christopher B. Smith

Southwest Research Institute, San Antonio, Tx, 78238
Department of Electrical Engineering, University of Texas San Antonio, San Antonio, Tx 78249
E-mail:cbsmith@ieee.org

ABSTRACT

Steganography is the art and science of hiding information in an innocuous medium. Digital imagery is a medium with specific noise characteristics. The devices used to capture a digital image, such as the charge coupled device (CCD) in a digital camera is designed to have a relatively small noise characteristic. This noise characteristic is often reduced by the compression used such as the JPEG standard, resulting in a very low noise media. The addition of steganography to the image has the effect of introducing changes to this "natural" image noise. This paper presents an approach to estimate and model the noise present in an image. Using this estimation, it is shown how steganography introduces detectable changes to this natural noise. This approach is demonstrated on three freely available but difficult to detect embedding techniques, F5, JSteg, and Model-based embedding, and show that it results in features that serve as statistically significant discriminators.

***Index Terms*—** Blind steganalysis, Noise variance estimation, Steganography

1. INTRODUCTION

Steganography is the art of hiding information in other information. A classic example of steganography appears in Herodotus' *Histories*. The Greek king Histaeus shaved the head of his most trusted slave and tattooed a message on his scalp. After the slave's hair had grown the message was hidden. The slave was then sent to an ally with the purpose of revealing the hidden message and thereby instigating a revolt against the Persians.

Modern steganography uses digital media as camouflage. With digital media, sophisticated mathematics and digital signal-processing can be employed to hide information. With modern digital communications, these covert messages can be spread through cyberspace, where distribution is effortless; duplication is perfect; and nearly anonymous. This covert communication poses a serious challenge to the field of information security and forensics.[1]

The sophistication of the modern steganographic technique creates a system that is a significant challenge for the potential steganalyst. To formulate the problem of steganalysis, we begin with the model based ideas presented by Sallee in [2]. In Sallee's approach to steganalysis a statistical model of an image is developed. Under this approach, a set of statistics, θ, is developped in which any change in the image due to steganography introduces a change in θ, $\theta \rightarrow \theta'$, such that θ and θ' are significantly different. Steganography and steganalysis then become an antagonistic system of competing models.

Ideally, the statistic θ should express sufficiency, that is, the set of statistics $\{\theta\}$ captures all the information about the data. In image based steganography, the lack of an accurate general model for the information contained in the image has been a problem for steganalysis and the more fundamental problem of image denoising. In the spirit of model based steganalysis, this paper builds a model of the image, but the approach used is based on modeling the statistics of naturally occurring noise in the image rather than the information contained in the image. Harmsen and Pearlman worked with a similar concept, looking for the effects of noise on a specific image feature, but did not focus on the noise itself. Here, a wavelet-based method is used to estimate image noise variance, allowing for the development of statistics of noise variance in both clean and stego-embedded images.

The remainder of this paper includes, first a review of the properties of image noise, focusing on CCD based devices. Second, a discussion of noise variation between natural and stego-embedded images is developed and a robust noise variance estimator is introduced. Then an analysis of this approach is presented, testing for the significance of the results across a wide range of images, and comparing to Harmsen and Pearlman's noise based methods. Finally, the paper is concluded with comments on remaining work and future directions.

2. IMAGE NOISE

As with any sensor, an imaging device is subject to noise. Imaging devices exhibit two basic types of noise, fixed and stochastic. The first, fixed noise includes as dark noise and bias noise. The stochastic types of noise are generally attributed to two effects, first photon or Poisson noise, and second electronic noise. These forms of noise are present in varying degrees in both CCD and CMOS style photo sensors. In this article it is only important to form a rough understanding of the noise present in a digital image.

First, the fixed effects generated by the imaging device are dark noise and bias noise. Dark noise or dark current is an accumulation of heat generated electrons within the photo sensor. Instead of only measuring photo induced electrons from the sensor, other electrons from the device itself are also measured. This effect is highly repeatable given a reasonably constant temperature over a similar period of time. Readout noise is generated by errors in reading electrons from the photon detector. This is largely a function of the amplifier design so is also highly repeatable. The constant nature of these effects has led to design improvements that measure and remove these fixed effects negating the need for an explicit model.

The sources of random noise are the vital quantities for this research. The first random effect is the stochastic nature of the photon. Photons do not arrive uniformly on any surface. This non-uniform spread of electrons can lead to small variations in intensity across an image. This effect follows a Poisson distribution.[4] The final source of noise is due to variations in electronic components. In modern digital cameras the effect is generally small.

3. NOISE VARIANCE ESTIMATION

3.1 Image noise variance and steganography

TThe development of a model of an image with steganography begins with a model of the image itself,

$$y_i = s_i + n_i,$$

where y_i is a noisy image, s_i is a statistically deterministic, or information containing component of the image, and n_i is a noise component. Steganography is then modeled as an additive change.

$$y_s = y_i + n_s = s_i + n_i + n_s$$

Whether the steganography is additive, such as ±k-based embedding, or simply replaces portions of the image, such as with LSB embedding, either can be modeled by an additive component as long as some portion of the original noise remains. Taking the simple statistical model, n_i is assumed to be i.i.d., white, and Gaussian. Using an additive noise model, since n_i is Gaussian and s_i is deterministic, y_i is taken to follow a Gaussian distribution. This Gaussian has a mean contributed by the signal and a variance contributed by the noise, or $y_i \sim n\left(s_i, \sigma_i^2\right)$. Given the assumption that the stego-embedding follows a Gaussian distribution, such as in the transform domain cases shown in Chapter 2, $n_s \sim n\left(0, \sigma_s^2\right)$, the resulting stego-image would follow

$$y_s \sim n\left(s_i, \sigma_i^2 + \sigma_s^2\right).$$

Taking a different image model, n_i is a mixture of a Poisson random variable of unknown expected value, and a Gaussian of small, almost negligible variance.

$$\begin{array}{c} y_i \sim n\left(s_i, \sigma_i^2\right) \approx s_i \\ s_i \sim poisson(\lambda_i) \end{array},$$

where λ_i is the count of photons due to the light intensity at a pixel. This Poisson due to arriving photons is acted on by an integrator or other function $f(\cdot)$, due to the measurement device electronics.

$$y_i \sim f\left(poisson(\lambda_i)\right),$$

This results in the following model of steganography,

$$\begin{array}{c} y_s \sim n\left(s_i, \sigma_s^2\right) \\ s_i \sim f\left(poisson(\lambda_i)\right) \end{array}. \qquad (1)$$

If the resulting variation of s_i is small, the noise variance is dominated by σ_s^2. From pixel to pixel the intensity varies, so each pixel has a different expected value, λ. This is a non-uniform Poisson process. The action of the function $f(\cdot)$, such as an integrator, can reduce the inherent variance of s_i to result in a process driven by the variance of the stego-embedding,

$$y_s \sim n\left(s_i, \sigma_s^2\right). \qquad (2)$$

3.2. Estimator

Estimating the noise level from a single image is a daunting task. Variations in the image itself often mimic noise-like behavior. In 1994, Donoho introduced an estimator for use in wavelet based denoising. In this approach the median absolute deviation of the wavelet coefficients at its finest scale is used as a robust estimate of noise variance, i.e.,

$$\hat{\sigma}^2 = \frac{median(|y_i|)}{0.6745}, y_i \in subband\ HH. \qquad (3)$$

Figure 1, shows a wavelet decomposisiton with the appropriate subband highlighted. Donoho found these coefficients to be almost exclusively noise. This estimator is robust, but is biased high by the signal energy that is also in this subband. [3] At the end of this paper are a set of clean images that show the estimated noise variance, computed in blocks across the image.

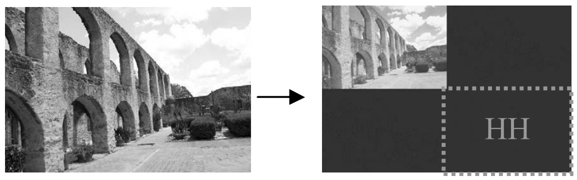

Figure 1 Wavelet decomposition with subband HH highlighted

4. STEGANALYSIS

From the model in (2), it is expected that a stego-embedded image will have a different noise variance than a clean image. To test for any existing differences a number of ststistics of the noise variance were used. The 1st quartile is one example used for demonstration in Figure 2 and 3. For these experiments the Daubechies-8 wavelet was used in the noise variance estimator.

To determine the effectiveness of each statistic, a series of tests were constructed. Three embedding mechanisms were used, each was chosen due to its easy availability on the Internet. This free distribution makes these techniques likely to be in wide use. First, the F5 algorithm, as described in [5], second Sallee's model-based steganography, as described in [2], and third the classic JSteg [6]. Each technique was used to embed a series of 912 images obtained from the Image Science Group at Dartmouth. Each was cropped to 512x512 pixels and embedded with varying levels of data from one to five hundred kilobits. The context of these images was in no way controlled, and ranged widely from wedding photos, to cars, animals, and other natural scenes.

To test for significance in each feature, an analysis of variance (ANOVA) was performed. ANOVA is a fundamental tool in statistics to determine if measurements taken from different groups are statistically different. To provide a reference, the results are compared with Harmsen and Pearlman's technique of measuring the center of mass of the histogram characterization function (HCF-COM) as described in [7]. In [7], the HCF-COM method is used to detect spread-spectrum embedding methods. Figure 2 shows the sample means and standard error for the proposed noise estimation method. Form this plot this method statistically separates clean images from all steganography techniques, further it separates F5 from JSteg and Sallee's method, suggesting a method of classifying steganography methods beyond detection. In Figure 3, the HCF-COM of each of the 3 layers demonstrates no statistical difference between Jsteg and Sallee's embedded and non-embedded images for this wide class of images.

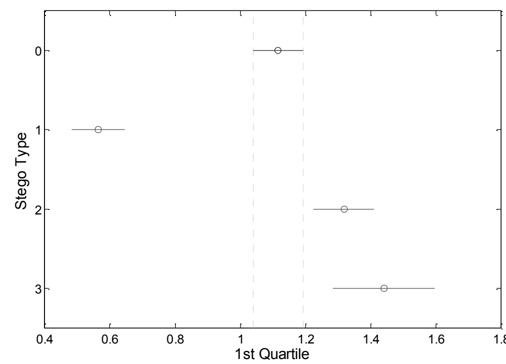

Figure 2 Plot of population mean and standard errors resulting from ANOVA analysis of 1st Quartile of the noise variance. Type 0: clean image, 1: F5, 2: Sallee, 3: JSteg. Non-overlapping regions show statistically significant differences between the types of stego-embedded files.

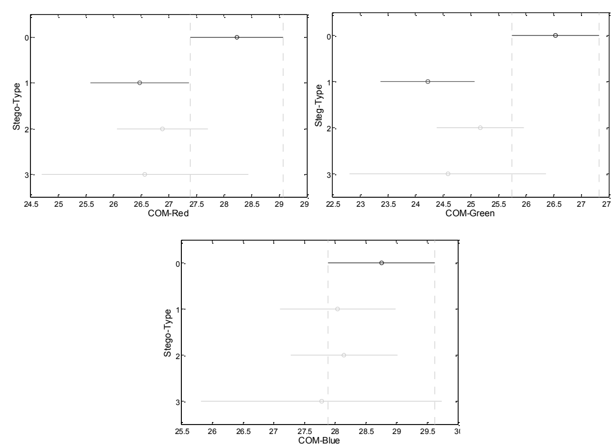

Figure 3 Plot of population mean and standard errors resulting from ANOVA analysis of HCF-COM for each RGB color layer. Type 0: clean image, 1: F5, 2: Sallee, 3: JSteg. The large overlapping regions show no statistically significant differences between the types of stego-embedded files.

5. CONCLUSIONS

This paper has introduced a method of blind steganalysis based on estimating noise variance in images and changes in noise variance statistics due to three different embedding techniques. Here the focus was determining if this noise estimating approach is a viable approach to steganalysis. Much work remains in this area. A limited comparison has been made with one other noise based method of steganalysis. A comparison with a wider group of steganalysis methods should be conducted. Further, noise variance in a natural image is a function of many different factors. More accurately modeling these factors, beyond the simple estimation used here, has the potential to detect many current steganography techniques. The implication of this work is that there are measurable features of an image that modern steganography techniques do not take into account. As long as there are measurable features of an image which are not modeled by the embedding process, steganalysis will be successful.

6. ACKNOWLEDGEMENTS

Thanks to Hani Farid and the *Image Science Group at Dartmouth* for providing a number of the images used in this study.

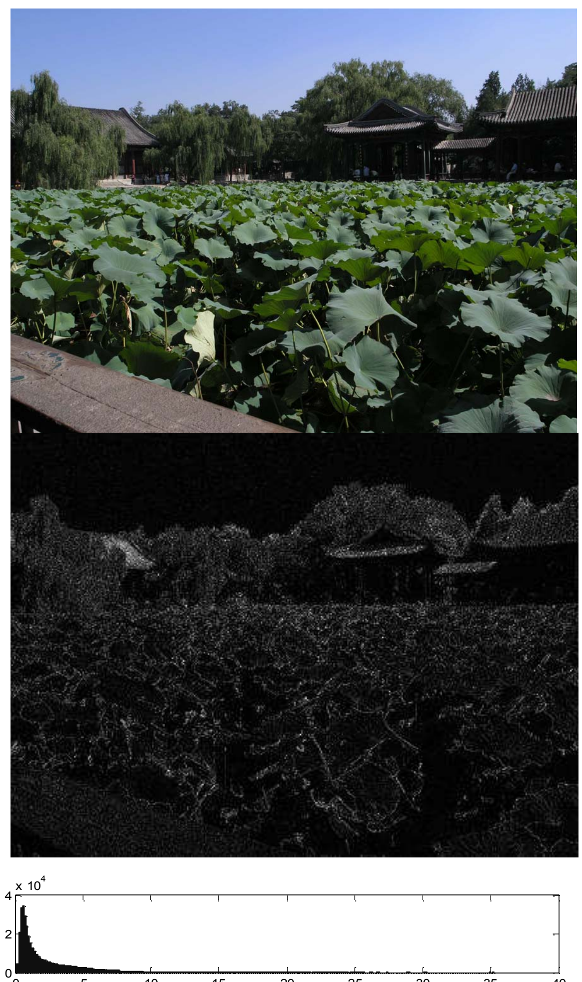

Figure 4 Top: Original image, Middle: noise variance estimate in 8x8 pixel blocks. Bottom: Histogram of the noise variance estimated for each 8x8 block.

7. REFERENCES

[1] Fabien A. P. Petitcolas, Ross J. Anderson and Markus G. Kuhn, Information Hiding – A Survey, Proceedings of the IEEE, special issue on protection of multimedia content, 87(7):1062-1078, July 1999.

[2] P. Sallee, "Model Based Steganognphy", In Proceedings of International Workshop on Digital Watermarking 03, Seoul, Korea, September, 2003.

[3] D.L. Donoho and I.M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage", *Biometrika* , vol. 81, pp. 425-455, 1994.

[4] G. E. Healey and R. Kondepudy. "Radiometric CCD camera calibration and noise estimation". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(3):267–276, March 1994.

[5] A. Westfeld and A. Pfitzmann, "High Capacity Despite Better Steganalysis (F5–A Steganographic Algorithm)", Lecture Notes in Computer Science, vol.2137, Springer-Verlag, Berlin, 2001.

[6] JPEG-JSteg-V4, http://u?nv.runet.fi/pub/crypt/steganognphy/jpeg-jstcg-v4.diff.gz.

[7] J. J. Harmsen and W. A. Pearlman, "Steganalysis of Additive Noise Modelable Information Hiding," Proc. SPIE Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents V, pp. 131–142, 2003.

Figure 5 Top: Original image, Middle: noise variance estimate in 8x8 pixel blocks. Bottom: Histogram of the noise variance estimated for each 8x8 block.