HPS: HISTOGRAM PRESERVING STEGANOGRAPHY IN SPATIAL DOMAIN

Yinan Wang[1], *Weirong Chen*[1], *Yue Li*[1], *Wei Wang*[1,*] *and ChangTsun Li*[2]

1 College of Software, Nankai University, Tianjin, China
2 Department of Computer Science, University of Warwick, UK

ABSTRACT

Minimization of perceptual and statistical distortions is one of the main challenges facing steganographic schemes. The most common approach to minimizing perceptual distortion is Least Significant Bit embedding. However, the statistical features (e.g., histogram) of the stego-images may be changed significantly even with selection of the regions of interest in the embedding process. This opens security gaps for steganalysis. To address this issue in this paper, we propose a novel steganographic algorithm, which is capable of preserving the histogram of the cover images in the stego-images and minimizing visual distortions. We partition the gray intensity scale into small segments and only allow the embedding process to modulate the gray intensity within the same segment. Randomization is employed to ensure the preservation of the histogram of cover images.

Index Terms— Steganography, Histogram Preserving, Segment, Random Number

1. INTRODUCTION

Steganography is an art of covert communications by embedding secret messages in a perceptually undetectable manner in cover mediums, like images, videos, audios etc. An effective steganographic algorithm should not give rise to suspicions on the stego-mediums. Therefore, minimizing perceptual and statistical distortions due to the embedding of the secret message is the key to the success of steganography. Current researches on steganography focused on the Least Significant Bit embedding algorithm. LSB replacement was once the most popular stego method studied in spatial domain. Since this scheme only change the LSBs of the pixels of interest, its advantage, low level of visual distortion, is obvious. Yet due to the inevitable symmetry of the histogram in the LSBs, that kind of methods are easy to analysis and were broke soon after proposed. Ker et al. proposed an effective LSB replacement attack algorithm by revisiting approach[1] estimating the embedding rate of a stego image.

To make up the vulnerability of LSB replacement, LSB Matching, which is also called ± 1 embedding, is proposed. LSB matching is a sophisticate version of LSB replacement, thus the degree of difficulty in detection is much higher than LSB replacement.

To minimize the distortion of the stego noise in LSB embedding, most methods choose to select undetectable region, e.g. edges, for embedding. The basic idea of region selection is to take advantage of the complexity of some region in the cover image and let the image noise mix up with the stego noise. And this method is always used based on the well-known fact that edge area of an image varies more, in other words, with a larger complexity. Pevny et al. has proposed a novel selective embedding scheme called HUGO[2], which evaluate each pixel's distortion to the steganalysis feature and select proper pixels to replace.

The attack scheme, on the contrary side of steganography, is called steganalysis. What steganalysis faces is to detect possible modified images among non-modified ones, which is also called blind attack. To do so, steganalysis always focuses on detecting stego noise caused by the embedding phase. Feature extraction, which is the most crucial part of steganalysis, mainly targets at magnifying the stego noise. Several feature extraction methods in spatial domain have been proposed currently. Take the most general LSB algorithm as example. For LSB replacement, Ker et al. proposed a revisiting algorithm[1] which can estimate the embedding rate of the stego image. For LSB matching, which is much more difficult to detect, histogram characteristic function (HCF) modeling the distortion that the image's histogram made is first used by Harmsen et al.[3] and has been proved to be promising. HCF was further improved by Ker who proposed two novel algorithms of applying HCF: calibrating the output using a down-sampled image and computing the adjacency histogram instead of the usual histogram[4]. Lyu et al. described an approach to detect hidden messages in images by using a wavelet-like decomposition to build high- order statistical models of natural images[5] for general detection.

Most algorithms on steganalysis is implemented following a pipeline[4, 5, 6]: extract stego-noise features, train a classifier, i.e. Support Vector Machine(SVM) by a set of noted image database and determine the unknown image by the trained classifier.

Currently, researches are normally constrained by the method of LSB embedding and minimize the stego noise by select region, trying to make the noise undetectable. Although

978-1-4799-4370-8/14/$31.00

the visual distortion is considerably minimized with the selective embedding scheme, this kind of algorithm cannot minimize the statistic distortion: the modification of statistic feature which may fluctuate significantly by little change of the cover image. Without the consideration of such distortion, current steganography algorithms are always vulnerable to steganalyzers which focus on magnifying the little visual distortion. Compared with the visual distortion, statistic one is also and even more important.

We believe that the maintenance of the high-level feature (the statistic feature) would significantly improve the imperceptibility. In this paper, we propose a steganography algorithm, which is capable of preserving the cover image's histogram of intensities. Due to the preserved histogram, the proposed algorithm strengthens its security against current attack algorithms.

2. STEGANOGRAPHY STRATEGY

According to the current steganalysis research, general steganalysis methods are based on the statistic features. Therefore, an ideal security steganography algorithm is supposed to maintain the consistency of features extracted by the steganalysis algorithms. Yet, though the feature preserving method is effective, it is hard to handle all the features: there would always be a new feature proposed in the analysis phase but we failed preserving in embedding. A statistic measurement whose preservation can minimize the distortion of steganalysis features, comes into our sight.

After reviewing most of the statistic features, we have recognized that most of the image statistic features in spatial domain is the combination of p_i and $I_{x,y}$. p_i is the probability of the appearance of the ith gray level and $I_{x,y}$ is the (x, y)th entry of the gray scale matrix of a $m \times n$ image: $\left\{ I_{x,y} | I_{x,y} \in \{0, 1, \ldots, 255\}, x \in \{1, \ldots, m\}, y \in \{1, \ldots, n\} \right\}$ For instance, the moment features, such as mean, variance, skewness, and kurtosis used by Lyu et al.[5], HCF feature used by Ker [4] etc. are all computed by the two variables. Further more, the high-level feature in wavelet domain is partially related to the intense-level probability p_i. Thus p_i is a good measurement to analysis.

One of the most important instance of p_i is the histogram of intensities. Any modification on the images, if only it distorts the distribution of appearance of occurrences (n_i for intensity i), could make the existence of the secret message easily detected by the statistic-based methods[4, 5]. Inspired by this fact, we propose a intensity histogram preserving steganography algorithm designed to be immune to most of the steganalysis algorithms in spatial domain. And like most researches, we will only discuss the steganography of gray image to simplify the problem. However, the proposed algorithm is entirely feasible in the color image.

2.1. Data Embedding

Let the region of interest for data embedding be defined as $\{R_{x,y}\}$, where (x_i, y_i) is the position of the pixel in which the secret data X will be embedded. And the histogram H could be calculated on R based on the gray intensities,

A certain interval of gray-scale space α (α is not necessary to fulfill entire scale space of $[0, 255]$) is firstly separated to several segments S_α. Each segment could either be separated in a regularly same width or a variable width based on specific determination. A regular and small width will decrease the distortion and complexity of both the segmentation and embedding, but weaken in the capacity and security. For simplification reasons, a regular width $\delta = 3$ is chosen for algorithm presentation and experiment discussion.

Let S_α be the segment studied. In the embedding phase, the algorithm ensures that the pixels to be embedded will only be changed to another intensity in the same segment or remain the same. In other words, the embedding modification will be limited in a certain segment where the pixel's intensity lies. Therefore, for simplification reasons, the subscript α, representing the index of segments, is omitted, and S_α is re-symboled as S.

Within the studied segment, the appearance of occurrences of each intensity is h^k, where k is the index of intensities and the intensities of h^k belongs to S . Then the local probability p^k can be calculated as follow:

$$p^k = \frac{h^k}{\sum_i h^i} \tag{1}$$

where k denotes the index of intensities in the segment.

The within segment probability refers to the local relationship between intensities in S. If p^k is preserved after the permutation for embedding, the local relationship among intensities will be preserved at the same time, so as the global relationship of $\{p^k\}$ is preserved as well. Then the secret bit x is inserted into the image by modifying the pixels whose intensity is located in S and ensure each p^k is preserved after embedding. Firstly, all p^k in their segment S are required to be sorted in descent order and if $\max_{\arg k} p^k \geq 0.5$, this S is defined as 'unembedable' segment, all pixels of the intensities lie in S will not be used for embedding. On the other hand, if $\max_{\arg k} p^k < 0.5$, then S is defined as embeddable segment and every pixel whose intensity locates in S can be used for embedding. The examination of embedment is designed to avoid the extraction error which will be discussed in Section 2.2.

Let $\{P^k\}$ be the sorted set of $\{p^k\}$ in descend order, and $k \in \{1, 2, \ldots, n\}$ where n is the number of members in S. According to the combination theory, there are $n!$ arrangements of $\{P^k\}$. For each arrangement, the $[0, 1)$ real interval can be separated into the following intervals based on any

combinations of P:

$$[0, P^{k_1}), [P^{k_1}, P^{k_1}+P^{k_2}), \ldots, [\sum_{i=1}^{n-1} P^{k_i}, \sum_{i=1}^{n} P^{k_i}) \quad (2)$$

Since $\sum_{i=1}^{n} P^{k_i} = 1$ and $P_i^k \geq 0$, the separation above is the exact separation of the interval of $[0, 1)$.

Among the $n!$ arrangements of $\{P^k\}$, we specifically choose the following two:

$$\begin{cases} \mathcal{P}_I = \left\{P^1, P^2, \ldots, P^n\right\},\ if\ x = 0 \\ \mathcal{P}_{II} = \left\{P^2, \ldots, P^n, P^1\right\},\ if\ x = 1 \end{cases} \quad (3)$$

A random number r is then generated in $[0, 1)$ by a certain secret key for the random number generator. Then one bit of the secret message x is read, if x is 0, arrangement $\mathcal{P}_I$ is selected, else arrangement $\mathcal{P}_{II}$ is selected. After the determination of the arrangement, r can be projected to the $[0, 1)$ interval separated by the selected arrangement, and an interval $[\sum_{k=1}^{m-1} \mathcal{P}^k, \sum_{k=1}^{m} \mathcal{P}^k)$ can be determined. Further, the index of intensity $\mathcal{K}$ can be deducted. Finally, the identical intensity can be mapped and will be used as the output intensity of the pixel of interest.

After embedding, the probability of every intensity in the modified image will be highly similar to p^k according to the Monte Carlo method[7]. In this method, if a random variable, following the uniform distribution, is projected into intervals with the width of $\{p^k\}$, the probability of the appearance in each interval remain the same to p^k.

2.2. Data Extraction

The extraction phase is actually the inverse process of the embedding phase.

In the embedding phase, the secret bit x is embedded into original cover image $I_{x,y}$, to obtain $I'_{x,y}$. On the contrary, in the extraction process, the modified pixel $I'_{x,y}$ is given, and the secret bit x is required to be decoded. The preserved histogram, which is supposed to be the same as original image, is the potential parameter in extraction. And the two sides of communication only need a constant random number generator to keep the random number the same as the original one.

For a pixel of focus, its identity $I'_{x,y}$ is given, then the segment S and the histogram H' of S could also be determined. Through equation (1), the probability of appearance of occurrence $p^{k\prime}$ is generated. Notice that both $H' = H$ and $p^{k\prime} = p^k$. And the two arrangements $\mathcal{P}_I$,$\mathcal{P}_{II}$ can also be constructed.

Then the extraction has become a matching task. Different from the embedding phase, the random number r scaled in $[0, 1)$ is required to be projected to both combinations because of the mystery of the secret bit x. By the two determined intervals $[\sum_{k=1}^{K_1-1} \mathcal{P}_I^k, \sum_{k=1}^{K_1} \mathcal{P}_I^k)$ and $[\sum_{k=1}^{K_2-1} \mathcal{P}_{II}^k, \sum_{k=1}^{K_2} \mathcal{P}_{II}^k)$ where $K_{1,2}$ is the index of interval of each arrangement, the corresponding intensities $\mathcal{K}_{1,2}$ of both $\mathcal{P}_I^{K_1}$ and $\mathcal{P}_{II}^{K_2}$ can be further deducted. Since $I_{x,y}$ is modified by the same method of the previous steps, $I'_{x,y}$ equals exactly to either $\mathcal{K}_1$ or $\mathcal{K}_2$. Thus, to solve the secret bit, only matching $I'_{x,y}$ to $\mathcal{K}_1$ and $\mathcal{K}_2$ is needed. The decode protocol follows equation (4)

$$\begin{cases} x = 0,\ if\ I_{x,y} = \mathcal{K}_1 \\ x = 1,\ if\ I_{x,y} = \mathcal{K}_2 \end{cases} \quad (4)$$

Since the embedding phase follows the embedment examination referred to section 2.1, no probability of an intensity in the identical segment equals or more than 0.5. There is, thus, no interval that has a length of 0.5 or more in any combinations. Thus, $\{\mathcal{P}_I^k\}$ and $\{\mathcal{P}_{II}^k\}$ have no intersections for any $I'_{x,y}$ and r. And the solution of secret bit can never be ambiguous if former protocol is followed.

3. EXPERIMENT & ANALYSIS

To evaluate the proposed scheme, the Break Our Steganography System Base (BOSSBase) is selected. Such dataset contains 10000 512×512 grayscale images which is used in the BOSS contest[8]. Meanwhile, to demonstrate the performance of HPS, we take the general idea of LSB as the peer-art for comparison. In this section of experiment discussion, a set of distortion data with the measurement of the Peak Signal to Noise Ratio (PSNR) and the attack of a Revisiting Weighted Stego-Image Steganalysis scheme[1], which is used to estimate embedding rate, are experimented and discussed.

3.1. Analysis of Distortion

With the character of constrained modification, the distortion of HPS is considerably low. Although the distortion of the proposed algorithm is higher than the LSB replacement, which is one of the least distort schemes, the difference is in acceptable area.

The distortion is measured by PSNR computed by the following equation:

$$PSNR = 10 \cdot \log_{10} \frac{255^2 \cdot W \cdot H}{\sum_{x=1}^{W} \sum_{y=1}^{H} \left(I_o(x,y) - I_s(x,y)\right)^2} \quad (5)$$

where W and H are the width and height of the image, respectively. I_o is the original image and I_s is the image where the secret bits are embedded.

Figure 1 shows the average PSNR of the proposed embedding algorithm. It can be seen that when the capacity reaches 1 bpp (bit per pixel), the distortion is 47.1 dB, while as the bpp degrades to 0.1, PSNR respectively grows to 57.2 dB. In comparison, for LSB, PSNR is 51.0 dB, which is slightly higher than HPS. However, in the security comparison, HPS is much effective than LSB.

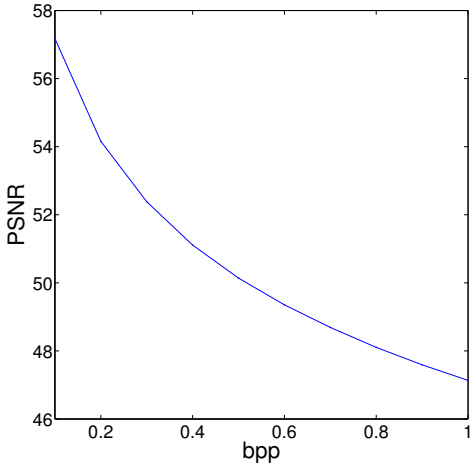
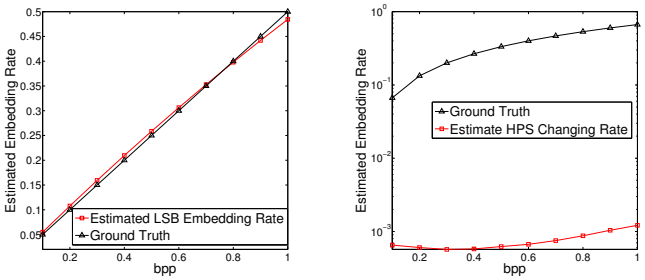


Fig. 1: Distortion of Proposed Scheme

3.2. Attacking Test

As one of advantages of the proposed algorithm, the symmetry of LSB replacement is eliminated. To test the character of that, [1] is used which focuses on the detection of symmetry.



(a) LSB (b) HPS

Fig. 2: Estimated Embedding Rate of LSB and HPS

As shown in Figure 2, the estimate algorithm can almost perfectly detect the LSB embedding rate while failing to detect ours. Notice that the $y - axis$ of Figure 2a is linearly scaled, however, the $y - axis$ of Figure 2b is in log. The estimation algorithm is almost unrecognizable to our algorithm considering the fact that the estimation and the ground truth in Figure 2b have huge difference and the estimate algorithm almost consider non-embedded in our algorithm.

4. CONCLUSION AND FUTURE WORK

In this paper, a Histogram Preserving Steganographic(HPS) scheme is proposed to strengthen the security of the commonly used LSB embedding methods. Our scheme is capable of preserving the histogram of the cover images due the novel use of randomization. Visual distortion to the stego-images is also minimized through the partitioning of the intensity scale into small segments and by confining the intensity modulation within the same segment. The proposed scheme is immune to several steganalysis methods and has considerable low distortion. However, the distortion due to our embedding method requires further reduction when compared against the some LSB algorithms. This constitutes the agenda of our future research.

5. REFERENCES

[1] Andrew D Ker and Rainer Böhme, "Revisiting weighted stego-image steganalysis," in *Electronic Imaging 2008*. International Society for Optics and Photonics, 2008, pp. 681905–681905.

[2] Tomáš Pevnỳ, Tomáš Filler, and Patrick Bas, "Using high-dimensional image models to perform highly undetectable steganography," in *Information Hiding*. Springer, 2010, pp. 161–177.

[3] Jeremiah Harmsen and William Pearlman, "Higher-order statistical steganalysis of palette images," in *Proc. SPIE Security Watermarking Multimedia Contents*, 2003, vol. 5020, pp. 131–142.

[4] Andrew D Ker, "Steganalysis of lsb matching in grayscale images," *Signal Processing Letters, IEEE*, vol. 12, no. 6, pp. 441–444, 2005.

[5] Siwei Lyu and Hany Farid, "Steganalysis using higher-order image statistics," *Information Forensics and Security, IEEE Transactions on*, vol. 1, no. 1, pp. 111–119, 2006.

[6] Tomas Pevny, Patrick Bas, and Jessica Fridrich, "Steganalysis by subtractive pixel adjacency matrix," *Information Forensics and Security, IEEE Transactions on*, vol. 5, no. 2, pp. 215–224, 2010.

[7] Nicholas Metropolis, "The beginning of the monte carlo method," *Los Alamos Science*, vol. 15, no. 584, pp. 125–130, 1987.

[8] Patrick Bas, Tomáš Filler, and Tomáš Pevnỳ, " break our steganographic system: The ins and outs of organizing boss," in *Information Hiding*. Springer, 2011, pp. 59–70.