DETECTION OF LSB STEGANOGRAPHY BASED ON IMAGE SMOOTHNESS[*]

Tao Zhang[1,2]   *Yan Zhang*[1] *Xijian Ping*[1] *Mingwu Song*[3]

[1] Dept. of Information Science, Zhengzhou Information Science and Technology Institute, P.R.China
[2] NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing, China
[3] Dept. of Training, Zhengzhou Information Science and Technology Institute, P.R.China
{deltd, zlittleyan, pingxijian}@yahoo.com.cn, rbfn@sina.com

ABSTRACT

The detection of LSB steganography is a question of common interest in the research of steganalysis techniques. In this paper, the distribution of the difference between the current pixel value and its neighborhood average pixel value is statistically modeled, and then the variance of this statistical distribution is defined as a measurement of image smoothness. Based on the analysis of the effects on the image smoothness brought by message embedding and LSB plane flipping, a new steganalytic technique capable of reliable detection of spatial LSB steganography is proposed. The algorithm can exactly estimate the amount of hidden messages and detect the existence of hidden messages embedded in the image simultaneously. Experimental results show that the proposed algorithm is effective.

1. INTRODUCTION

Detection of embedded hidden messages in images, also known as image steganalysis techniques, is a hot spot in the research field of information hiding. Petitcolas et al. pointed out that the goal of stegoanalyst is to detect, find out, remove or jam the secret messages [1]. Johnson et al. further stated that once the existence of hidden messages is detected, the goal of steganography is defeated [2]. Johnson also introduced some detection methods by identifying unique signatures of steganography tools which will reveal the existence of hidden messages.

Westfeld et al. introduced a method based on statistical analysis of PoVs (Pairs of Values) that are exchanged during message embedding [3]. This method provides very reliable results for steganography based on sequential LSB replacement. However, we can only detect randomly scattered messages with this method when the message length becomes comparable with the number of pixels in the image.

Fridrich et al. proposed an effective steganalysis technique called RS analysis method [4]. By inspecting the

---

[*] Project 60473022 supported by NSFC

lossless capacity in the LSB and shifted LSB plane the secret message length is derived. RS analysis is reliable to detect LSB non-sequential embedding in digital images.

After that several similar steganalysis algorithms are proposed, such as sample pair analysis method proposed by Dumitrescu et al.[5], difference image histogram based method proposed by Zhang et al.[6], log likelihood ratio test (LLRT) based method proposed by Sullivan et al.[7], contaminated distribution based method proposed by Guan et al.[8]. All of those methods are designed to detect hidden messages embedded in images in the way of spatial LSB replacement.

In this paper, we refer to the idea of "contaminated distribution" proposed by Guan et al. in [8], and statistically model the distribution of the difference between the current pixel value and its neighborhood average pixel value. Then we define the image smoothness as the variance of this statistical distribution. Based on the analysis of the effects on the image smoothness brought by message embedding and LSB plane flipping, a new steganalytic technique capable of reliable detection of spatial LSB steganography is proposed. Experimental results show that the proposed algorithm is effective.

2. BACKGROUND

2.1. Steganography based on spatial LSB replacement

LSB steganography is the most classic and simplest steganographic techniques, which embeds secret messages in a subset of the LSB plane of the image. A large number of popular steganographic tools, such as S-Tools 4, Steganos and StegoDos, are based on LSB replacement in the spatial domain.

LSB steganography can be described as follows: if the LSB of the pixel value $I(i, j)$ is equal to the message bit m to be embedded, $I(i, j)$ remain unchanged; if not, set the LSB of $I(i, j)$ to m. The message embedding procedure can be described using Equation (1) as follows,

$$I_S(i, j) = \begin{cases} I(i, j) - 1 & \text{LSB}(I(i, j)) = 1 \ and \ m = 0 \\ I(i, j) & \text{LSB}(I(i, j)) = m \\ I(i, j) + 1 & \text{LSB}(I(i, j)) = 0 \ and \ m = 1 \end{cases} \quad (1)$$

where **LSB($I(i, j)$)** stands for the LSB of $I(i, j)$ and m is the next bit to be embedded.

Secret messages can be embedded in LSB plane by sequential or random LSB replacement. Sequential LSB replacement can be implemented more conveniently but has a much serious security problem in that there is an obvious statistical difference between the modified part and the unmodified part of the stego-image. By random LSB replacement secret messages can be randomly scattered in stego-images, so the steganographic security is improved.

Let us assume that we have a cover-image **C** with $M{\times}N$ pixels. Obviously the maximum data hiding capacity of LSB steganography is $M{\times}N$ bits. Define the embedding ratio **p** as the ratio of the length of embedded messages to the maximum capacity, where $0 \leqslant p \leqslant 1$.

2.2. Image Smoothness

In view of the characteristic of LSB steganography, we introduce the concept of the image smoothness to evaluate the effect on the image quality of message embedding. Denote the intensity value of the image **I** at the position (i,j) as $I(i,j)$, and the local neighbor is defined as a $K*K$ window centered on the current pixel where K is an odd number. Define the difference variable $X_I(i,j)$ as the difference between the current pixel value and the average value over its local neighbor,

$$X_I(i, j) = I(i, j) - \frac{1}{K^2 - 1} \sum_{\substack{m,n=-(K-1)/2 \\ (m,n)\neq(0,0)}}^{(K-1)/2} I(i+m, j+n)$$

$$(2)$$

Consider the probability distribution of the difference variable X_I, and denote the probability density function of X_I as $f(x)$. Obviously, $f(x)$ has the following properties:

(1) The integral over its domain is equal to 1, that is

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \quad (3)$$

(2) The expectation of the random variable X_I is equal to 0, that is

$$\int_{-\infty}^{+\infty} f(x) \cdot x dx = 0 \quad (4)$$

Figure.1 shows the standard image "Lena" and the statistical distribution of the difference variable X_I for "Lena".
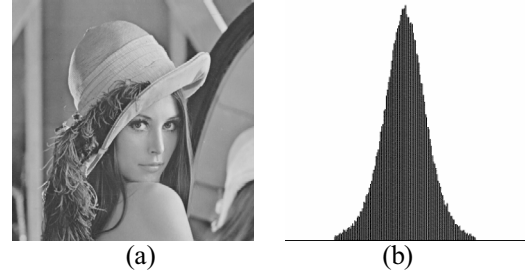

Fig.1 *(a) standard image "Lena";*
(b) the distribution of the difference variable X_I

The variance of the difference variable X_I reflects the degree to which the pixel value deviates from its neighborhood average pixel value, and can be used as a measurement of the image smoothness. Thus we can define the image smoothness of a natural image **I** as

$$SM_I = \sigma^2 = \int_{-\infty}^{+\infty} f(x) \cdot x^2 dx \quad (5)$$

The larger the value of the image smoothness SM_I is, the more intensively the pixel value deviates from its neighborhood average pixel value, and the weaker the smoothness of the image is.

3. DETECTION OF LSB STEGANOGRAPHY BASED ON IMAGE SMOOTHNESS

For the LSB Embedding reduces the smoothness of the image, we expect that the stego-image with secret message embedded have a larger image smoothness value under the definition of Equation (5) than that of the original cover-image.

3.1. Analysis on the image smoothness of the stego-image

For the original cover-image **C** and the stego-image **S**, consider the probability density function of the difference variable X_C and X_S, denoted by $f_C(x)$ and $f_S(x)$ respectively. For the cover-image **C**, the image smoothness is represented as:

$$SM_C = \sigma_C^2 = \int_{-\infty}^{+\infty} f_C(x) \cdot x^2 dx = \int_{-\infty}^{+\infty} f(x) \cdot x^2 dx \quad (6)$$

For the stego-image **S** containing secret messages with the embedding ratio p, compared with the original cover-image, there are $(1-p/2)*M*N$ pixels with its pixel value remaining unchanged, $p/4*M*N$ pixels with its pixel value increasing by 1, and $p/4*M*N$ pixels with its pixel value decreasing by 1. However, we noted that the neighborhood average pixel value remain approximately unchanged after message embedding since the number of pixels with its pixel value increasing by 1 in the local neighborhood area is approximate equal to the number of pixels with its pixel value decreasing by 1. For the pixel set with its pixel value remaining unchanged, the probability density function of X_S is approximate equal to $f(x)$; for the pixel set with its pixel value increasing by 1, the probability density function of X_S

is approximate equal to $f(x-1)$; for the pixel set with its pixel value decreasing by 1, the probability density function of X_S is approximate equal to $f(x+1)$. Thus for the stego-image S, the probability density function of X_S can be represented as:

$$f_S(x) \approx (1 - \frac{p}{2}) \cdot f(x) + \frac{p}{4} \cdot f(x-1) + \frac{p}{4} \cdot f(x+1) \qquad (7)$$

Obviously, $f_S(x)$ is a linear combination of three distribution with the same form and can be viewed as a special kind of contaminated distribution. The image smoothness of the stego-image S can be derived from the Equation (7) as follows:

$$SM_S = \sigma_S^2 = \int_{-\infty}^{+\infty} f_S(x) x^2 dx$$

$$\approx \int_{-\infty}^{+\infty} \left((1-\frac{p}{2})f(x) + \frac{p}{4} f(x-1) + \frac{p}{4} f(x+1) \right) x^2$$

$$= (1-\frac{p}{2}) \int_{-\infty}^{+\infty} f(x) x^2 dx + \frac{p}{4} \int_{-\infty}^{+\infty} f(x-1) x^2 dx + \frac{p}{4} \int_{-\infty}^{+\infty} f(x+1) x^2 dx$$

$$= (1-\frac{p}{2})\sigma_C^2 + \frac{p}{4} \int_{-\infty}^{+\infty} f(y)(y+1)^2 dy + \frac{p}{4} \int_{-\infty}^{+\infty} f(y)(y-1)^2 dy$$

$$= (1-\frac{p}{2})\sigma_C^2 + \frac{p}{4}(\sigma_C^2 + 1) + \frac{p}{4}(\sigma_C^2 + 1)$$

$$= \sigma_C^2 + \frac{p}{2}$$

$$(8)$$

3.2. Analysis on the image smoothness of the stego-image after LSB plane flipping

By flipping the LSB plane in the stego-image S, that is, performing the logic operation NOT for all the least significant bits ($0 \rightarrow 1$, $1 \rightarrow 0$), we get an image T. Compared with the original cover-image, there have $p/2*M*N$ pixels remaining unchanged in the image T, $(1/2-p/4)*M*N$ pixels with its pixel value increasing by 1 and $(1/2-p/4)*M*N$ pixels with its pixel value decreasing by 1. Therefore, for the image T, the probability density function of the difference variable X_T can be represented as:

$$f_T(x) \approx \frac{p}{2} f(x) + (\frac{1}{2} - \frac{p}{4}) f(x-1) + (\frac{1}{2} - \frac{p}{4}) f(x+1) \qquad (9)$$

Therefore, the image smoothness of T can be derived as follows:

$$SM_T = \sigma_T^2 = \int_{-\infty}^{+\infty} f_T(x) \cdot x^2 dx$$

$$\approx \int_{-\infty}^{+\infty} \left(\frac{p}{2} f(x) + (\frac{1}{2} - \frac{p}{4}) f(x-1) + (\frac{1}{2} - \frac{p}{4}) f(x+1) \right) x^2$$

$$= \frac{p}{2} \sigma_C^2 + (\frac{1}{2} - \frac{p}{4})(\sigma_C^2 + 1) + (\frac{1}{2} - \frac{p}{4})(\sigma_C^2 + 1)$$

$$= \sigma_C^2 + 1 - \frac{p}{2}$$

$$(10)$$

3.3. The estimation of the embedding ratio

From Equation (8) and (10), we get an estimation of the length of the embedded message in a given stego-image S. The estimation of the embedding ratio can be described as

$$\hat{p} = 1 + SM_S - SM_T = 1 + \sigma_S^2 - \sigma_T^2 \qquad (11)$$

It should be noticed that while calculating SM_T the neighborhood average value in Equation (2) can be replaced by the neighborhood average value on the same position in the stego-image S.

4. EXPERIMENTAL RESULTS

4.1. Test results for standard images

We selected 5 standard grayscale test images with 512×512 pixels from USC-SIPI Image database (See Fig.1(a) and Fig.2) and tested the new algorithm [9]. First, we created a series of stego-images by embedding secret messages into the five images using random LSB replacement method with embedding ratios of 0, 2%, 5%, 8%, 10%, 20%, 50%, 80%, 100%. Then, we estimated the embedding ratio from those stego-images using our steganalytic technique where K is equal to 3. Table I list the estimated results and indicate that the new algorithm is effective and reliable.
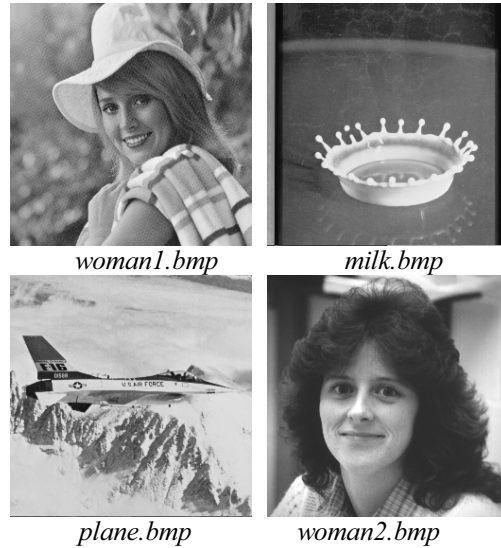

woman1.bmp   *milk.bmp*

plane.bmp   *woman2.bmp*
Fig.2 *standard test images*

Table I *Test results for standard images (in percent)*

	Lena	Wom1	Milk	Plane	Wom2
0%	-4.76	-4.96	-3.24	0.56	-1.12
2%	-2.65	-3.64	-1.25	3.04	1.11
5%	0.92	1.70	2.32	4.44	3.81
8%	3.11	4.46	4.63	10.36	6.37
10%	5.78	7.46	7.48	10.49	9.98
20%	17.07	14.71	18.66	20.98	19.81
50%	49.44	49.14	48.49	50.54	50.31
80%	78.66	81.95	81.45	80.72	79.30
100%	96.58	102.35	98.35	102.28	100.16

To compare the reliability and accuracy of our algorithm with RS analysis method, we did the same experiments on the CBIR image database [10]. A random message embedding process for each image is performed and then the embedding ratio is estimated using the two kinds of methods. The mask used in RS analysis is [0,1,1,0]. Table II lists statistical data for both methods.

Table II *Comparison with RS analysis method (in percent)*

	Our Algorithm		RS Analysis	
	Mean	Std. Dev.	Mean	Std. Dev.
0%	0.294	4.434	0.036	1.8467
10%	10.673	4.309	9.380	2.2616
20%	21.143	4.308	19.166	2.7507
30%	31.119	4.188	29.941	2.8677
40%	41.124	4.345	41.486	2.2116
50%	51.242	4.143	54.282	2.0935
60%	61.513	3.988	67.805	4.0104
70%	71.140	3.823	77.393	4.9523
80%	80.898	3.851	83.452	3.3432
90%	90.459	3.692	94.275	4.8150
100%	100.147	3.277	98.202	4.4164

From Table II we know that the mean value of the estimated embedding ratio using the new algorithm is closer to actual value, while the standard deviation using the new algorithm is larger than that using RS analysis when the embedding ratio is smaller than 50%. This indicates that for a large embedded message the new algorithm has a better performance than RS analysis method, while for a short message the new algorithm is not as good as RS analysis method.

On a PC with Intel Pentium IV 2.4GHz processor, the detection speeds of our algorithm are 3.119MBps (Mega-Bytes per second), while the speeds of RS analysis method are 0.621 MBps. This indicates that our algorithm runs much faster than RS analysis method and is more favorable to realizing real-time detection.

The accuracy of estimation on the embedding ratio of secret messages is influenced by two main factors: (1) Equation (7) and (9) is not strictly satisfied especially when the embedding ratio p is close to 0 or 1; (2)we supposed that the neighborhood average pixel value remain approximately unchanged after message embedding, but that is not tenable when the number of embedding bits is very small; (3) the diversity of image data, the randomness of both secret messages and the embedding process can also bring error to the estimation of the embedding ratio.

5. CONCLUSION

In this paper, we refer to the idea of "contaminated distribution" proposed by Guan et al., and statistically model the distribution of the difference between the current pixel value and its neighborhood average pixel value. Then we define of the image smoothness as the variance of this statistical distribution. Based on the analysis of the effects on the image smoothness brought by message embedding and LSB plane flipping, a new steganalytic technique capable of reliable detection of spatial LSB steganography is proposed. Experimental results show that the proposed algorithm is effective. It has a distinct physical meaning, smaller computational cost, higher computational speed, lower computational complexity and is more suitable for real-time detection.

Moreover, the method proposed in this paper can be used as a reference for the detection of steganography based on pixel value increment/decrement.

6. REFERENCES

[1] Pfitzmann B. Information hiding terminology. Proceedings of International Workshop on Information Hiding, Cambridge, UK, 1996, pp347-350.

[2] Johnson N, Jajodia S. Steganalysis of images created using current steganography software. Proceedings of International Workshop on Information Hiding, Berlin, Germany, 1998, pp273-289.

[3] Westfeld A, Pfitzmann A. Attacks on steganographic systems. Proceedings of International Workshop on Information Hiding, Dresden, Germany, 1999, pp61-76.

[4] Fridrich J, Goljan M, Du R. Detecting LSB steganography in color and gray-scale images. IEEE Multimedia, 2001, Vol.8, No.4, pp22–28.

[5] Dumitrescu S, Wu X, Wang Z. Detection of LSB steganography via sample pair analysis. IEEE Transaction on Signal Processing, 2003, Vol.51, No.7, pp1995-2007.

[6] Tao Zhang, Xijian Ping. A New Approach to Reliable Detection of LSB Steganography in Natural Images. Signal Processing, 2003, Vol.83, No.10, pp2085-2093.

[7] Sullivan K, Dabeer O, Madow U, Manujunath B, Chandrasekaran S. LLRT based detection of LSB Hiding. Proceedings of IEEE International Conference on Image Processing, Barcelona, 2003, pp497-500.

[8] Guan Wei, Zhang Weiming, Liu Wenfen. Steganalysis based on contamination data analysis. Acta Scientiarum Naturalium Universitatis Sunyatseni, 2004, Vol.43, supp.2, pp62-65 (in Chinese)

[9] University of South California. The USC-SIPI Image Database. http://sipi.usc.edu/services/database/Database.html.

[10] University of Washington. CBIR Image Database. http://ww.cs.washington.edu/research/magedatabase/ roundtruth/.