Steganalysis of LSB Matching Based on the Sum Features of Average Co-occurrence Matrix Using Image Estimation

Yan-qing Guo[1,2], Xiang-wei Kong[1], Bo Wang [1], and Qian Xiao[1]

[1] Dalian University of Technology, Dalian, China, 116023
[2] State Information Center, Beijing, China, 100045

Abstract. A new LSB matching steganalysis scheme for gray images is proposed in this paper. This method excavates the relevance between pixels in the LSB matching stego image from the co-occurrence matrix. This method can acquire high accuracy near to 100% at high embedding rate. In order to increase the accuracy at low embedding rate, we strengthen the differences between the cover image and the stego image to improve the performance of our scheme. Two 8 dimensional feature vectors are extracted separately from the test image and the restoration image, and then the combining 16 dimensional feature vector is used for steganalysis with the FISHER linear classification. Experimental results show that the detection accuracy of this method is above 90% with the embedding rate of 25%; even when the embedding rate is 10%, the detection accuracy reaches 80%.Experiments show that this method is more reliable than other state-of-art methods.

Keywords: LSB matching, steganalysis, co-occurrence matrix, image restoration, feature classify.

1 Introduction

Digital steganography, an important branch of information hiding, is the art of invisible communication which can enhance the security of the information, and at the same time increase the difficulty for the network safety supervision. Steganalysis is the art of attacking steganography, which is useful for network safety supervision and for intercepting unsafe digital multimedia information. The cover objects for digital steganograpy are various, such as texts, audio clips, video clips, digital images etc. Because digital image has large redundancy allowance, small storage capacity etc, it is widely used in digital steganography. The embedding methods for images are various. According to the embedding domain, steganography can be classified into space domain and transform domain and so on. Among all the embedding methods, steganography in space domain is drawing the attention of researchers for its simple operation and large capacity, especially for the improvement algorithm of LSB——LSB matching, and our method is focusing on LSB matching.

Y.Q. Shi, H.J. Kim, and F. Pérez-González (Eds.): IWDW 2012, LNCS 7809, pp. 34–43, 2013.

From the beginning of 2003, the researchers have been interested in LSB matching steganalysis. The most representative method was proposed by Harmsen and Pearlman etc [1] using the histogram characteristic function (HCF), depending on whether the function value of cover image is equal or greater than that of the stego image to distinguish the cover and stego images, but this approach is only effective for the BMP color image, and fail for the BMP gray image. Then Ker etc [2] did two improvements with Harmen's method: First, down sample the images; second, use the adjacency histogram instead of the normal histogram. The detection accuracy of LSB matching for gray images can reach 96%. In [3], a steganalysis method based on the correlation of pixel difference is proposed. The image histogram, smoothness of the difference histogram, gradient energy, the 1 and high dimension statistical distribution of pixel difference are used as the features for classification. Fridich [4] put forward a steganalysis method for LSB matching with good performance. This method called WAM extracts statistical moment of noise from wavelet domain to form the 27 dimensional feature vector. The detection accuracy of WAM reaches as high as 99% when the embedding rate is 1. However, the detection accuracy for low embedding rate is low; the method in this paper is to solve the problem of low detection accuracy at low embedding rate.

2 LSB Matching Model

LSB matching and LSB replacement are two widely used steganograpy method based on least significant bits of image pixels. Comparing with LSB replacement, LSB Matching is more secure. It can randomly add and subtract 1 to the image pixel's LSB, thus eliminate the effect of pair in LSB replacement, reduce the distortion of the cover image, and maintain the correlation between the adjacent pixels. It is also known as random ±1 LSB steganograpy, and the specific embedding method can be shown as follow:

$$I_s = \begin{cases} 1 & b \neq LSB(I_c)\ \&\ I_c = 0 \\ I_c \pm 1 & b \neq LSB(I_c)\ \&\ 0 < I_c < 2^L - 2 \\ I_c & b = LSB(I_c) \\ 2^L - 2 & b \neq LSB(I_c)\ \&\ I_c = 2^L - 1 \end{cases} \tag{1}$$

Where, I_C is the cover image, I_S is the stego image, b is one bit, L is the bits number of the image.

The LSB matching model can be expressed as $I_S = I_C + \eta$, where η is the secret message. The steganography method can be modeled as adding noise, and then the corresponding steganalysis method can be modeled as denoising.

LSB Matching method operates on the space domain of images, i.e. directly using the image pixels. From the human visual system, the images are nearly the same, people cannot find the differences before and after the secret message is embedded only with the human eyes; With respect to the image histogram, when a lot of information is embedded, image histogram becomes smoother. This can be seen as an image going through a low-pass filter, and the filter removes the high-frequency coefficients; With respect to the relevance of image pixels, the secret message weak the correlation and relationship between the pixels.

3 Sum Features of Average Co-occurrence Matrix Based Steganalysis

This section analyzes the correlation between the image pixels. After steganography, the changes of the pixel pairs' values in the edge area are less obvious than those in the smooth area, so we can use co-occurrence matrix to describe pixel difference in local area of the image, and construct sum features of average co-occurrence matrix for steganalysis. In order to strengthen the differences between cover images and stego images and improve the detection accuracy at low embedding rate, we combine with the technology of image estimation for steganalysis.

3.1 Analysis of Correlation between Image Pixels

Natural images are modeled as stationary source of local area by researchers. Objects have similar reflective characteristics of electromagnetic waves, which makes the pixels of local area have strong correlation. In the analysis of image, two hypotheses are admitted, one is the Markov assumption, that is a pixel value and the pixel values of its certain space neighborhood is correlated; the other is the translation invariance hypothesis ,that is the distribution of pixels in the neighborhood is independent on the absolute position of the neighborhood in the image [9].

In terms of the content of the image, the distribution of the pixels in a meaningful image is regular, and this regularity constitutes the content of the image. From the macroscopic view, image is sights that can be seen; from the microscopic view, it is a series of point sets or point pairs. These pixel sets or pairs have the same or close pixel value. There is a small critical region for the pixels in the smooth area of the image; but the distribution of the pixels in the edge area of the image fluctuates widely, the pixel differences in this part have large value.

Adjacent pixels are correlated in the image. LSB steganagraphy is modeled as adding noise. Due to the existence of noise, LSB steganagraphy will reduce the correlation between adjacent pixels. LSB Matching method changes pixel values in the range of [-1, 0, 1].The changes of the pixel pairs' values in the edge area are less obvious than those in the smooth area. Therefore, we mainly consider the changes in the smooth area during steganalysis. In the smooth area, the change of the correlation between pixels is more obvious, thus more suitable for steganalysis. For the relative smooth neighborhood, the distribution of the pixels in the neighborhood can be described using eq (2):

$$p-\varepsilon < (f(x,y) \mid x \in [x-\Delta x, x+\Delta x], y \in [y-\Delta y, y+\Delta y]) < p+\varepsilon \tag{2}$$

Where, ε value is integer, $f(x,y)=p$ is the pixel value at location (x,y), (x,y) is the index of the image, Δx and Δy are the index increments.For more accurate depiction of the difference before and after steganography, in this paper ε takes 3.

3.2 Gray Level Co-occurrence Matrix

Gray level Co-occurrence Matrix (GLCM) was proposed in 1973 by Haralick. Firstly it was applied in the texture feature extraction, and had good superiority in texture analysis. Since then he put forward 14 gray statistical features for texture analysis, which were later widely used in image texture extraction, edge detection, and analysis of remote sensing image. Many researchers use co-occurrence matrix as a feature for steganalysis. Sullivan [5] used the co-occurrence matrix for the first time in spread spectrum steganalysis.129 features were selected from the co-occurrence matrix elements for classification. G. Xuan [6] also took the co-occurrence matrix in LSB and DCT domain for steganalysis, a total of 1029 elements in the main diagonal line and its top two diagonals were chosen as the features. In order to reduce dimensions, the CNPCA analysis method was adopt for classification. In [7] and [8], Fridrich et al employed co-occurrence matrix, histogram characteristics of DCT coefficients for detection of the secret message. In [10], a space domain steganalysis method was proposed.180 elements from the co-occurrence matrix of the difference matrix are used to detect the hidden message. As can be seen from the above examples, co-occurrence matrix has a lot of advantages in steganalysis. The brief introduction of the gray level co-occurrence matrix is as follow.

GLCM describes the correlation of the two pixels in the θ angle direction with the distance of d , denoted as $p(i,j,d,\theta)$, θ is 0° , 45° 90° and 135° respectively. Fig.1 shows the gray level co-occurrence matrix. GLCM not only reflects the distribution characteristics of the luminance, but also reflects the distribution characteristics of the locations with the same or close luminance, including the comprehensive information about the direction, the adjacent interval, the amplitude of change. A digital image can be denoted as (x, y), the largest gray level is 255, gray level co-occurrence matrix meeting certain space requirements can be expressed by eq(3) :

$$p(i,j,d,\theta)=\begin{cases} \left[(x,y),(x+\Delta x, y+\Delta y)\right] | f(x,y)=i, \\ f(x+\Delta x, y+\Delta y)=j; \\ x=0,1,2,...,N_x-1; y=0,1,2,...,N_y-1 \end{cases} \tag{3}$$

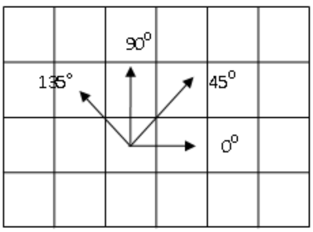

Fig. 1. Gray level co-occurrence matrix

Def 1:

Average co-occurrence matrix defines the average of four GLCMs that θ is 0°, 45° 90° and 135°, the distance is d.

$$C_d = (C_{d0} + C_{d1} + C_{d2} + C_{d4}) / 4 \tag{4}$$

Average co-occurrence matrix reflects the change of the pixels in four directions. Comparing with co-occurrence matrix in a single direction, Average co-occurrence matrix can reflect the distribution of pixels in the smooth area better, and show the correlation between the pixels better.

GLCM is composed by the elements in N diagonals, as shown in Fig.2. The i th diagonal is denoted as D_i, $i \in (-255, 255)$ and i is the Difference-value of the pixel pairs. Because the GLCM elements above and below main diagonal are symmetrical, therefore i can be considered only $i \in (0, 255)$.

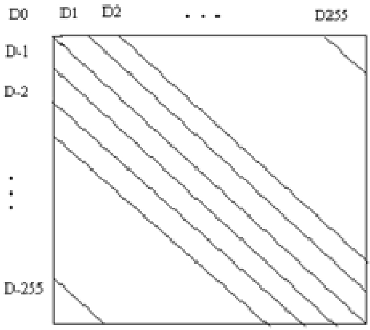

Fig. 2. Elements in the diagonal of the grayscale co-occurrence matrix

Def 2:

Sum feature of average co-occurrence matrix defines the sum of all the elements in all the GLCM's diagonals:

$$G_i = \sum D_i \;,\; i = 0,1,2,...,255 \tag{5}$$

In eq(5), i denotes the absolute difference between two pixels, when i is relatively small, the two pixel values are close, the two pixels locate in the smooth area of the image, the correlation between them is strong; As the value of i increases, the two pixels locates in the edge area of the image, the correlation is weak. The change due to data embedding in the edge area of the image is not significant. And if the image has complex textures, the edge information may cause interference to steganalysis. So combine eq (2), eq (5) is changed to eq (6):

$$G_i = \sum D_i \;,\; i = 0,1,2,3 \tag{6}$$

If we use the Sum Features of the Average Co-occurrence Matrix to detect LSB matching, when the embedding rate is 100%, the detection accuracy is close to 100%. However, when the embedding rate is lower than 20%, the detection accuracy is only about 70%.

3.3 Image Estimation

Wavelet transform has low entropy, multi-resolution, decorrelation and multi-choice of wavelet basis and has significant superiority in the image denoising. Image noise energy generally concentrates in the high frequency part of the signal. The noisy signal is transformed to the frequency domain, and then the contradiction between the protection of local details and the suppression of noise become obvious. Wavelet transform has good time-frequency localization property, which may solve the contradiction above. In image processing, there exists a lot of denoising methods using wavelet transform, such as wavelet threshold denoising method, hard threshold denoising method, and the corresponding soft threshold denoising method, the wavelet energy filtering is the improvement of wavelet threshold denoising method, which uses the wavelet coefficients energy features to revise the wavelet coefficients, and filter out the noise, thus achieve the purpose of image restoration. The adaptive wavelet energy filter recovery method proposed here is different from the general method in which the wavelet coefficients shrinkage is done pixel by pixel, its wavelet coefficients shrinkage depends on the energy of all pixels in the neighborhood, and thus it has better adaptability.

Wavelet transform has different decomposition scales. There still exists some redundancy among the decomposition scales. A natural image usually has similar wavelet transform coefficients with the same resolution scales. The recovery method of adaptive wavelet energy filtering gives the wavelet energy shrinkage function and the recovery function.

Wavelet energy shrinkage function:

$$S_{m,n} = \frac{1}{N^2}\sum_{m=1}^{N}\sum_{n=1}^{N} d_{m,n}^2 \tag{7}$$

N is the filtering window size, $d_{m,n}$ are elements in the window function.

Recovery function:

$$\hat{d}_{ij} = \begin{cases} d_{ij}(1-\alpha\times\frac{\lambda^2}{S_{ij}}) & S_{ij} > \beta\times\lambda^2 \\ d_{ij} & other \end{cases} \tag{8}$$

$$\lambda^2 = 4\times\sigma^2 \log\sqrt{N} \tag{9}$$

In eq (7), d_{ij} is the center element of the filtering window. Select odd number for the filtering window, and alpha, beta for undetermined coefficient. In eq (9), σ denotes noise variance. For a test image, σ is unknown and needs to be estimated from the image. In general, the energy of the noise after wavelet transform is mainly in HH frequency band. Here, use the wavelet coefficients of HH band after 1 level wavelet decomposition to estimate $\sigma = median(abs(d))/0.6745$.

The steps of image estimation:

Step 1: Calculate the first level wavelet decomposition of the test image, extract HH sub-band coefficients, and then calculate the local variance of HH sub-band coefficient, according to $\sigma = median(abs(d))/0.6745$;

Step 2: Calculate the second level wavelet decomposition of the test image, 6 high-frequency sub-bands are calculated ; use the energy shrinkage function and the recovery function to modify wavelet coefficients of three high-frequency sub-band in each level respectively, then another 6 high-frequency sub-bands are calculated;

Step 3: Merge the wavelet coefficients of the 12 sub-bands and reconstruct the image.

In order to validate the similarity between the estimation image after filtering and the original cover image, we use reconstruction bit error rate (BER), BER denotes the rate of the number of the pixels that the original cover image and corresponding estimation image have the same value and the total number of image pixels. Filtering the cover image and stego image respectively, the cover image BER is about 0.9 or so, the stego image BER is about 0.7 or so. Although estimation images of stego image and cover image are quite different, the difference highlights the difference between stego images and cover images. So extracting the sum features of the average GLCM from the estimation image can improve the steganalysis performance.

3.4 Feature Extraction

This section shows the detail of the feature extraction method and the flow diagram of our steganalysis approach, as shown in Fig.4.

Step 1: Use adaptive wavelet energy filter to get the estimation image from test image;

Step 2: Calculate the average GLCM C_1 , C_2 of the test image, according to eq (5); extract sum features $G_i, i = 0,1,2,3$ of C_1 , C_2 respectively, get 8 features;

Step 3: Calculate the average GLCM C_1^* , C_2^* of the estimate image, according to eq (5); extract sum features $G_i, i = 0,1,2,3$ of C_1^* , C_2^* respectively, get 8 features;

Step 4: Combine features from step3 and step4, get 16 average GLCM sum features.

Step 5: Use FISHER linear classifier for classification.

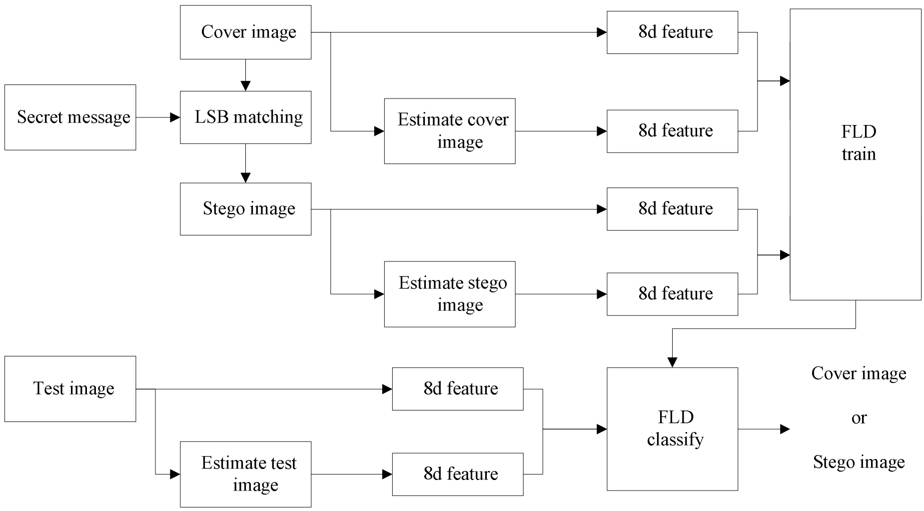

Fig. 3. Diagram of LSB matching steganalysis

4 Experimental Results

In order to verify the performance of the proposed method, we test out algorithm on the following database. The database consists of 485 uncompressed TIFF format images with the resolution of 1440 x960. All the images are converted into gray image. We generate stego images with the LSB Matching algorithm for different embedding rates of 100%, 75%, 50%, 25%, 15%, 10%. We randomly choose 200 images as the training set from the cover and the stego image database respectively, and the others for testing.

To compare the performance with other steganalysis methods, the 16 features using estimation image, the 8 features without using estimation image, WAM 27 features and improved HCF features are respectively classified by FISHER classifier. For different embedding rates and different false alarm rates, experimental results can be shown in ROC curve. Fig.4 shows detection accuracy of different methods when the embedding rate is 100% .Table 1 presents the details of the results for different embedding rates and different features (false alarm rate is 0.1).

Table 1. Average detection accuracy with the fisher classifier

Embedding rate %	16d features	8d features	WAM	Improved HCF
100	99.65%	99.30%	99.47%	96.90%
75	99.30%	98.42%	98.25%	96.84%
50	98.95%	97.72%	96.14%	81.05%
25	95.09%	91.40%	90.70%	62.46%
15	86.72%	76.49%	74.61%	60.42%
10	82.36%	73.51%	70.46%	58.42%

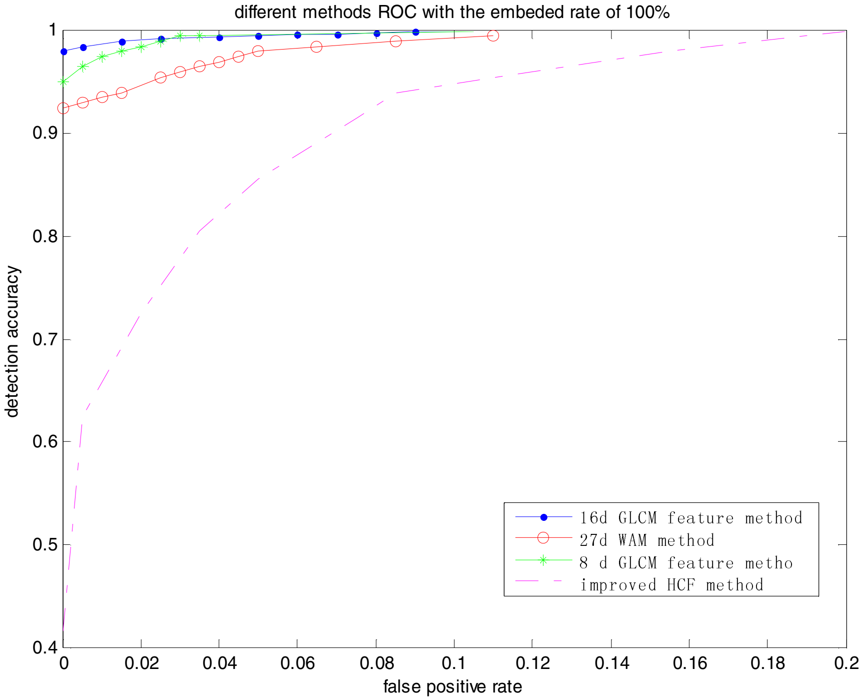

Fig. 4. ROC of different methods with embedding rate 100%

It is clear from Fig.4 and Table 1 that the detection accuracy of our method outperforms the other methods in detecting the LSB matching method for grayscale images with the same false alarm rate and the features using estimation image do better than the 8 features without using estimation image.

Experimental results show that when the embedding rate is lower than 25%, the detection accuracy is increased above 20% compared with the improved HCF feature, is increased about 10% compared with the feature without using estimation image, and is increased about 3% compared with WAM feature.

As shown, the detection accuracy of our proposed method is better than other methods.

5 Conclusion

This article starts from the correlation in the space domain of image, takes image as a local area stationary source, is based on characteristic of LSB matching, and proposes a steganalysis method based on the sum features of average co-occurrence matrix using estimation image. Experimental results show that the detection accuracy of this method is about 100% with the embedding rate of 100%; even in low embedding rate, the detection accuracy reaches 80%. Continuously improving the accuracy of image estimation is a new direction for steganalysis.

Acknowledgments. This work is supported by the National Natural Science Foundation of China (Grant No. 60971095) and China Postdoctoral Science Foundation (NO.20110490343).

References

1. Harmsen, J., Pearlman, W.: Steganalysis of additive noise modelable information hiding. In: Delp, E., Wong, P.W. (eds.) Proceedings SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents V, Santa Clara, California, USA, vol. 5020, pp. 131–142 (2003)
2. Ker, A.D.: Steganalysis of LSB matching in grayscale images. IEEE Signal Processing Letters 12(6), 441–444 (2005)
3. Zhang, T., Li, W.X., Zhang, Y., et al.: Steganalysis of LSB matching based on statistical modeling of pixel difference distributions. Information Sciences 180(23), 4685–4694 (2010)
4. Goljan, M., Fridrich, J., Holotyak, T.: New Blind Steganalysis and its Implications. In: Delp, E., Wong, P.W. (eds.) Proceedings SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VIII, San Jose, California, USA, vol. 6072, pp. 1–13 (2006)
5. Sullivan, K., Madhow, U., Chandrasekaran, S.: Steganalysis of spread spectrum data hiding exploiting cover memory. In: Delp, E., Wong, P.W. (eds.) Proceedings SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VII, San Jose, California, USA, vol. 5681, pp. 38–46 (2005)
6. Xuan, G., Shi, Y.Q., Huang, C., Fu, D., Zhu, X., Chai, P., Gao, J.: Steganalysis using high-dimensional features derived from co-occurrence matrix and class-wise non-principal components analysis (CNPCA). In: Shi, Y.Q., Jeon, B. (eds.) IWDW 2006. LNCS, vol. 4283, pp. 49–60. Springer, Heidelberg (2006)
7. Fridrich, J.: Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes. In: Fridrich, J. (ed.) IH 2004. LNCS, vol. 3200, pp. 67–81. Springer, Heidelberg (2004)
8. Pevný, T., Fridrich, J.: Towards multi-class blind steganalyzer for JPEG images. In: Barni, M., Cox, I., Kalker, T., Kim, H.-J. (eds.) IWDW 2005. LNCS, vol. 3710, pp. 39–53. Springer, Heidelberg (2005)
9. Wong, G.X., Ping, X.J., Xu, M.K., et al.: Steganalysis Based on Neighbor Image Pixels Correlation. Journal of Information Engineering University 8(1), 56–58 (2007)
10. Deng, Q.L., Lin, J.J.: Image steganalysis based on co-occurrence matrix. Microcomputer Information 12(1), 6–8 (2009)