Steganalysis using Regional Correlation and Second-order Markov Features

Su Jinyang, Zeng Xianting and Wang Lei

College of information and engineering, China Jiliang university, Hangzhou, 310018 mico@cjlu.edu.cn

Abstract

A blind steganalysis based on regional correlation and second-order Markov transition probability matrix is proposed for JPEG images. By analyzing the region correlation of JPEG image, the Markov transition probability matrix is used to capture the correlation of intra-block and inter-block DCT coefficients. In addition, the calibrated features are extracted from the calibrated images. The difference features between the original and calibrated images are used for training and classifying. Three different image libraries are used to detect the five kinds of typical JPEG steganography schemes. Experimental results show that, in comparison to the several effective steganalysis, the proposed scheme improves the detection accuracy on some JPEG-based steganography schemes, including outguess, Steghide and MB1.

Keywords: Steganography, Steganalysis, Markov transition probability matrix, regional correlation

1. Introduction

Steganography, which is sometimes referred to as data hiding, is used to conceal secret message into the cover images imperceptibly. Opposite to the steganography, steganalysis focuses on detecting the presence of the hidden data. In general, steganalytic techniques can be divided into two categories — targeted approaches and blind steganalysis[1-5]. It is likely that for a specific steganography the targeted approaches would provide more accurate and reliable results than blind steganalysis, while in practice, blind steganalysis is more important, because there is no need to develop a new targeted approach each time when a new steganography appears.

Shi *et al.*, [2] proposed a Markov process based approach to effective attacking JPEG steganography. In their scheme, difference JPEG matrix along horizontal, vertical, and diagonal directions are used to enhance changes caused by JPEG steganography. Markov process is applied to modeling these difference JPEG matrices so as to utilize the second order statistics for steganalysis. By setting the threshold to construct 324 dimensional feature vectors, which can be more effective for classification.

Subsequently, by introducing inter-block correlation of DCT coefficients, Chen *et al.*, [3] improved the method of Shi *et al.*, [2]. In Chen's scheme, they calculated block coefficient difference matrix of the Markov transition probability matrix, and finally constructed 486 dimensional feature vectors.

By extending the 23 DCT feature set, Pevny *et al.*, [4] applied calibration to the Markov features. The resulting feature sets are merged, producing a 274-dimensional feature vector. The new feature set is then used to construct a Support Vector Machine multi-classifier capable of assigning stego images to six popular steganographic algorithms. The new feature set provides significantly more reliable results.

Huang *et al.*, [6] proposed an improved calibration-based universal JPEG steganalysis. By combining the microscopic and macroscopic calibrations, they calibrated the local and global distribution of the quantized BDCT coefficients of the test image. All features are generated from the difference signal between the quantized BDCT coefficients of the test image and its corresponding microscopic calibrated image, or from the difference between the signal extracted from test image and its corresponding macroscopic calibrated image. The extracted features will be more effective for classification.

Li *et al.*, [7] proposed a regional correlation based blind dection of JPEG image steganography. To estimate the cover image, they use microscopic and macroscopic calibrations.

Fridrich *et al.*, [8] presented a method for detection of steganographic methods that embed in the spatial domain by adding a low-amplitude independent stego signal, an example of which is least significant bit (LSB) matching. First, arguments are provided for modeling the differences between adjacent pixels using first order and second-order Markov chains. Subsets of sample transition probability matrices are then used as features for a steganalyzer implemented by support vector machines.

In this paper, we have extended Fridrich *et al.*'s algorithm [8]. By scanning image along the horizontal, vertical, Zigzag and Hibert directions, we can obtain the absolute values of DCT coefficients of original image and scanning images. Then, we can compute the difference values of DCT adjacent coefficients along horizontal, vertical, diagonal and back diagonal directions. Next, we can get the difference matrices of Markov transition probability and the same calculation will be applied to calibration image. Finally we can extract 686 features. Support vector machine is used as classifier. Experimental results have shown that, the proposed scheme has a good detection effect.

The rest of this paper is organized as follows. The proposed scheme is presented in Section II. Experimental results are illustrated in Section III, and conclusions are drawn in Section IV.

2. Feature Generation

In this section, we shall present the new feature set for steganalysis of JPEG steganography. First four scanning patterns are defined, and a merged Markov transition probability matrix on inter-block and intra-block is obtained, and then the matrix is calibrated using local and global calibrations respectively to get our feature set for further classification.

2.1. Image Scanning Patterns

To begin with, a image is divided into a number of non-overlapping blocks sized 8×8 each. Then, scan each block along row, column, Zigzag and Hibert directions respectively, as shown in Figure 1.





Figure 1. Four Scanning Patterns

By scanning the image along different paths as shown above, we can obtain the spatial correlation of the image better.

2.2. Intra-block Features

After the image reordering, as mentioned above, we can extract Intra-block features. Assume that the size of an image is $M \times N$, and let F denote the matrix of DCT coefficient absolute value, as shown in Figure 2.



Figure 2. DCT-coefficient Difference Matrices

The differences of adjacent coefficients in horizontal, vertical, diagonal and back diagonal directions can be obtained by

where $1 \le u \le M - 1$, $1 \le v \le N - 1$, and F_h , F_v , F_d and F_m are sized of (M-1)×(N-1).

The Markov stochastic process can be used to describe the difference matrix. And based on the random process theory, the transition probability matrix can describe the stochastic process well. As stated in ref [8], second-order Markov transition probability matrix can be given by

Where

Because the dimension of feature vectors extracted from the DCT coefficients is very much, in order to reduce the computational complexity, we first calculate all elements of M_h , M_v , M_d and M_m , and then we take their average value as feature vector, as shown in formula (10).

$$F_1^{2n} \stackrel{1}{=} [\mathcal{M}_n + \mathcal{M}_n + \mathcal{M}_n + \mathcal{M}_n] \quad (10)$$

In addition, set the threshold T=3, and choose all the elements in the Markov transition probability matrix that satisfy (i, j) \in ([-3,3] \times [-3,3]) as features, whose number is (2T+1) 3=343 in total. Finally, compute the transition probability matrix from the original image and the row, column, Zigzag and Hibert scanning paths respectively, and calculate their average value as the intra-block feature vector, which is given by

$$F_{inver} = \frac{1}{5} [F_{iv} + F_{iv} + F_{iv} + F_{iv} + F_{iv}] \quad (11)$$

2.3. Inter-block Features





As shown in Figure 3, rearrange DCT coefficients in all 8×8 blocks and form new matrices in which all are of the same AC coefficients, leading to 63 matrices (DC coefficient is not considered), denoted by F1~F63. For each matrix Fi (i \in [1, 63]), according to the formulas (1) - (4), we calculate the 4 difference matrices Fih, Fiv, Fid and Fim from horizontal, vertical, diagonal and back diagonal direction respectively. And by employing the formulas (5) ~ (8), we can calculate the transition probability matrices of Mih, Miv, Mid and Mim, and calculate the average value for all the transition probability matrix. In order to avoid the excessive number of features, we set the threshold T=3, the calculation formula is as follows

$$F_{2}^{\mathcal{H}} = \sum_{\mu} \mathcal{M}_{\mu} \mathcal{M} \mathcal{M}_{\mu} \mathcal{$$

Finally, compute the transition probability matrices from the original image and the row, column, Zigzag and Hibert scanning paths respectively, and calculate their average value as the inter-block feature vector, which is given by

$$\mathbf{F}_{\mathbf{n}\mathbf{k}\sigma} = \begin{bmatrix} \mathbf{I}_{\mathbf{2}\sigma} + \mathbf{I}_{\mathbf{2}} + \mathbf{I}_{\mathbf{2}} + \mathbf{I}_{\mathbf{2}} + \mathbf{I}_{\mathbf{2}} \end{bmatrix}$$
(13)

2.4. Image Calibration

Image calibration is used to accurate the obtained statistics. We can get the calibrated JPEG image from the given one by cropping and recompressing as following.

1. Decompress the given JPEG image J1 into the spatial domain to get B1.

2. Crop B1 by 4 pixels in each of horizontal and vertical direction to obtain B2.

3. Recompress B2 with the same quantization table as J1 and generate the calibrated JPEG image J2.

One can think that the cropped stego image is perceptually similar to the cover image.

Assume that the feature extracted from the test image is denoted as F_o and feature extracted from the calibration image is denoted as F_c . We take the difference between F_o and F_c as statistics, *i.e.*, $F=F_o-F_c$. Finally, we obtain 686 features for our classification.

3. Experimental Results

In this section, experimental results are used to evaluate the performance of our method. The experimental results shown in Tables I - III are the arithmetic average of 20 times of tests. TPR and TNR represent the true positive rate and the true negative rate respectively, and AR is the final accuracy rate, which is calculated by AR = (TPR+TNR)/2. By comparison, our scheme shows a good detection accuracy on stego images which are embedded with five popular steganographic algorithms. And our scheme presents a good ability to construct a universal steganalyzer to detect an "unknown" steganography.

3.1. Image Database and JPEG Steganography Tools

We use three image libraries, *i.e.*, CorelDraw, Greenspun and UCID.v2, which consist of various types of raw color images. To evaluate the performance of our scheme, we focus on attacking five widely used JPEG steganography, *i.e.*, Jphide, Outguess, Steghide, MB1 and MB2, over the image libraries. We use bpnc (bits per non-zero DCT coefficients) to represent the embedding rate, *i.e.*, a ratio between the length of hidden message and the non-zero DCT coefficients in a JPEG image. In our

experiments, we use three kinds of embedding rate, *i.e.* 0.05, 0.10 and 0.20 bpnc for each steganography, and we use LibSVM [9] to construct the classifiers.

3.2. Experimental Results

Tables I-III are the experimental results on three kinds of image libraries respectively.

We compare the performance of our scheme with three recent universal algorithms Huang324 [6], Fridrich274 [8] and Chen486 [3]. For each steganography, we select 500 cover images and the corresponding 500 stego images randomly as the training set, and the rest images are used for testing.

As we can observe, the proposed scheme can provide a good accuracy rate. Among the five types of typical JPEG steganography, compared with three effective universal methods, our experiments have indicated that, for Outguess, Steghide, MB1, our approach has a better accuracy rate. An exceptional steganography is Jphide, Fridrich274 algorithm is better than ours in two kinds of image libraries.

	1	Huang324			Fr	Fridrich274			Chen486			the proposed		
	opne	TNR	TPR	AR	TNR	TPR	AR	TNR	TPR	AR	TNR	TPR	AR	
Jphide	0.05	81.24	31.31	56.27	88.73	49.25	68.99	43.11	57.45	50.28	95.74	45.22	70.48	
	0.10	79.55	37.49	58.52	87.70	62.12	74.91	47.38	59.10	53.24	93.60	57.87	75.74	
	0.20	90.60	85.13	87.87	93.39	90.24	91.82	65.02	81.87	73.45	95.73	86.94	91.33	
Outguess	0.05	93.75	93.60	93.68	93.53	94.58	94.05	94.12	91.42	92.77	97.91	96.05	96.98	
	0.10	100	99.93	99.96	100	99.96	99.98	98.24	99.10	98.67	99.99	99.73	99.86	
	0.20	100	100	100	100	99.96	99.98	99.85	100	99.93	100	100	100	
Steghide	0.05	93.30	94.16	93.73	93.26	95.21	94.23	96.70	94.46	95.58	99.13	99.57	99.35	
	0.10	99.48	99.74	99.61	99.40	99.74	99.57	99.29	99.74	99.51	100	99.91	99.96	
	0.20	99.40	99.74	99.57	100	100	100	99.63	100	99.81	100	100	100	
MB1	0.05	79.55	82.73	81.14	82.62	80.90	81.76	84.31	29.96	57.13	85.99	89.68	87.84	
	0.10	96.48	95.54	96.01	96.52	96.25	96.39	85.69	83.90	84.79	99.05	97.97	98.51	
	0.20	99.59	99.89	99.74	100	99.44	99.72	98.84	99.81	99.33	100	99.61	99.81	
	0.05	81.69	83.78	82.73	83.82	84.19	84.01	73.45	53.82	63.63	87.07	89.96	88.52	
MB2	0.10	96.89	96.37	96.63	98.09	97.60	97.85	88.61	90.60	89.61	99.10	98.92	99.01	
	0.20	99.63	99.93	99.78	100	99.78	99.89	99.63	99.81	99.72	100	99.67	99.84	

Table I. Detection Accuracy in CorelDraw Image Library (in the unit of %)

Table II. Detection Accuracy in Orcenspun inage Eistary (in the unit or /

	1	Huang324		Fi	Fridrich274			Chen486			the proposed		
	opnc	TNR	TPR	AR	TNR	TPR	AR	TNR	TPR	AR	TNR	TPR	AR
Jphide	0.05	87.98	24.92	56.45	82.45	47.81	65.13	34.96	65.49	50.22	93.17	35.42	64.29
	0.10	89.39	27.85	58.62	82.57	62.57	72.57	27.50	78.96	53.23	91.35	42.54	66.95
	0.20	92.35	86.73	89.54	92.49	89.53	91.01	83.20	77.91	80.56	94.39	82.76	87.42
Outguess	0.05	93.49	91.95	92.72	92.75	92.72	92.73	89.94	85.84	87.89	94.85	93.58	94.21
	0.10	99.80	99.90	99.85	99.78	99.90	99.84	98.19	98.71	98.45	99.72	99.53	99.62
	0.20	99.80	99.90	99.85	100	100	100	99.88	100	99.94	100	100	100
Steghide	0.05	93.12	93.54	93.33	93.54	94.41	93.98	95.33	92.57	93.95	97.78	97.97	97.87
	0.10	99.50	99.50	99.50	99.88	99.53	99.70	99.50	99.18	99.34	99.85	99.85	99.85
	0.20	100	100	100	99.80	99.90	99.85	99.98	99.98	99.98	100	100	100
MB1	0.05	76.32	80.55	78.43	80.80	79.11	79.95	76.20	44.32	60.26	82.45	80.41	81.43
	0.10	94.01	93.54	93.78	95.45	93.89	94.67	82.01	91.33	86.67	97.36	94.78	96.07
	0.20	99.90	99.30	99.60	99.83	98.98	99.40	98.41	99.70	99.06	99.96	98.90	99.43

	0.05	78.83	82.11	80.47	83.45	82.06	82.76	76.67	54.56	65.61	82.23	82.62	82.42
MB2	0.10	95.13	94.46	94.80	97.27	94.81	96.04	84.47	95.23	89.85	97.91	95.49	96.70
	0.20	100	99.35	99.68	100	99.55	99.78	99.18	99.95	99.57	99,95	99.14	99.55

	bpnc	Huang324			Fridrich274			Chen486			the proposed		
		TNR	TPR	AR	TNR	TPR	AR	TNR	TPR	AR	TNR	TPR	AR
Jphide	0.05	82.63	39.82	61.23	87.85	46.92	67.38	35.31	65.62	50.46	92.54	37.55	65.04
	0.10	76.53	49.45	62.99	85.10	62.71	73.90	21.64	81.55	51.60	91.19	43.93	67.56
	0.20	87.80	85.41	86.61	88.39	83.75	86.07	80.30	76.53	78.42	90.85	83.75	87.30
Outguess	0.05	86.23	87.39	86.81	86.35	86.14	86.24	86.31	82.72	84.51	91.32	87.34	89.33
	0.10	96.61	96.49	96.55	96.49	96.28	96.39	96.95	97.10	97.03	98.91	97.54	98.22
	0.20	100	100	100	99.97	99.91	99.94	99.82	99.88	99.85	99.99	99.80	99.89
Steghide	0.05	91.69	91.18	91.43	90.85	91.84	91.35	92.88	92.29	92.59	96.09	95.70	95.89
	0.10	99.16	99.22	99.19	98.86	98.83	98.85	98.62	99.13	98.88	99.55	99.11	99.33
	0.20	99.97	100	99.99	100	99.85	99.93	99.91	100	99.96	99.99	99.93	99.96
MB1	0.05	72.84	73.02	72.93	74.50	68.46	71.48	82.90	28.97	55.93	81.19	69.68	75.44
	0.10	89.84	89.69	89.76	91.06	88.31	89.69	75.19	86.58	80.88	94.39	91.14	92.76
	0.20	99.88	99.37	99.63	99.52	98.09	98.80	97.79	98.95	98.37	99.72	99.03	99.38
	0.05	75.10	80.48	77.79	79.79	72.62	76.20	73.93	57.67	65.80	82.88	71.84	77.36
MB2	0.10	93.57	93.39	93.48	94.50	91.06	92.78	81.35	94.50	87.92	95.78	91.79	93.78
	0.20	99.88	99.64	99.76	99.82	99.37	99.60	98.95	99.73	99.34	99.84	99.34	99.59

Table III. Detection Accuracy in UCID.v2 Image Library (in the unit of %)

4. Conclusion

We have proposed a blind approach to detect JPEG steganography. By utilizing the row, column, Zigzag and Hibert scanning paths and calculating the second-order Markov probability matrix, we have further improved the detection accuracy. Among the five kinds of typical JPEG steganography, compared with three effective universal methods, experiments have indicated that, for Outguess2, Steghide, MB1, our approach has a better accuracy rate, and an exceptional steganography is Jphide, Fridrich274 algorithm is better than ours in two kinds of image libraries.

Acknowledgements

This work was supported in part by the natural science foundation of Zhejiang Province (Grant No. Y1110450, Grant No. LQ12F03016).

References

- Z. Li, J. Chen, X.-n. Jiang, X.-t. Zeng and X.-z. Pan, "Blind JPEG steganalysis based on multi-domain feature", Journal of Zhejiang University (Engineering Science), vol. 45, no. 9, (2011), pp. 1528-1538 (in Chinese).
- [2] Q. Yun, S. C. Chen and W. Chen, "A Markov Process Based Approach to Effective Attacking JPEG Steganography", Proceedings of the IEEE international Conference on Multimedia & Expo 2005. Amstenlam, Netberlands: [s.n.], (2005), pp. 269-272.
- [3] C. Chen and Y. Q. Shi, "JPEG Image Steganalysis Utilizing both Intrablock and Interblock Correlations", IEEE international Symposium on Circuits and Systems. Seattle, Washington, USA: IEEE, (2008), pp. 3029-3032.
- [4] T. Pevny and J. Fridrich, "Merging Markov and DCT features for multi-class JPEG steganalysis", Proceedings of the 8th information Hiding Workshop. Alexandria, VA, USA: Springer-Verlag, (2007), pp. 249-264.

- [5] J. Kodovsky and J. Fridrich, "Calibration Revisited", Proceeding of the 11th ACM Multimedia Security Workshop, (2009), pp. 63-74.
- [6] F. Huang and J. Huang, "Calibration based universal JPEG steganalysis", Science in china series F: Information sciences, vol. 52, no. 2, (2009), pp. 260-268.
- [7] K. Li and X.-j. Ping, "Regional Correlation Based Blind Dection of JPEG Image Steganography", Journal of Information Engineering University, vol. 13, no. 3, (**2012**), pp. 306-311 (in Chinese).
- [8] T. Pevny, P. Bas and J. Fridrich, "Steganalysis by subtractive pixel adjacency matrix", IEEE Transaction on Information Forensics and Security, vol. 5, no. 2, (2010), pp. 215-224.
- [9] C. C. Chang and C. J. Lin, "LIBSVM: A library for Support Vector Machines", http://www.csie.ntu.edu.tw/~cjlin/libsvm.