Steganalysis versus Splicing Detection

Yun Q. Shi[1], Chunhua Chen[1], Guorong Xuan[2], and Wei Su[1]

[1] New Jersey Institute of Technology, University Heights, Newark, NJ, USA 07102
[2] Tongji University, Shanghai, China 200092
{shi, cc86}@njit.edu

Abstract. Aiming at detecting secret information hidden in a given image using steganographic tools, steganalysis has been of interest since the end of 1990's. In particular, universal steganalysis, not limited to attacking a specific steganographic tool, is of extensive interests due to its practicality. Recently, splicing detection, another important area in the field of digital forensics has attracted increasing attention. Is there any relationship between steganalysis and splicing detection? Is it possible to apply universal steganalysis methodologies to splicing detection? In this paper, we address these intact and yet interesting questions. Our analysis and experiments have demonstrated that, on the one hand, steganography and splicing have different goals and strategies, hence, generally causing different statistical artifacts on images. However, on the other hand, both of them make the touched (stego or spliced) image different from the corresponding original (natural) image. Therefore, natural image model based on a set of carefully selected statistical features in the machine learning framework can be used for steganalysis and splicing detection. It is shown in this paper that those successful steganalytic schemes can make promising progress in splicing detection if applied appropriately. A more advanced natural image model developed from these state-of-the-art steganalysis methods is thereafter presented. Furthermore, a concrete implementation of the proposed model is tested on the Columbia Image Splicing Detection Evaluation Dataset, and the implementation has achieved an accuracy of 92%, indicating a significant advancement in splicing detection. Although images are the only medium discussed here, the principle discussed in this paper is expected applicable to other media as well.

Keywords: Steganography, steganalysis, splicing detection, tampering detection, digital forensics, natural image model.

1 Introduction

Covert communications emerged at the very beginning of our human history. From human hair to papyrus, many articles were ever used as carriers to pass secret messages. Nowadays, digital images have become an important channel to bear stego information.

People refer image steganography to as the art and science of *invisible* communication, which is to conceal the very existence of hidden messages in digital images. Some facts have motivated active research and abundant publications in the

field of image steganography. For example, popularly used today, especially on the Internet, images can convey a large amount of information. Moreover, the non-stationarity of images makes image steganography hard to break.

Accordingly, researchers have made efforts to develop image steganalysis schemes, which is to detect the very presence of hidden messages in a given image. According to the targeted steganographic tools, steganalysis methods may be classified into two categories, i.e., specific steganalysis and universal steganalysis [1]. Universal steganalysis aims to attack any steganographic tool, known or unknown in advance, while specific steganalysis is designed to detect some particular steganographic tool. Not limited to any specific steganographic tool, universal steganalysis usually tends to capture certain statistical artifacts introduced by steganographic embedding. Therefore, it may be feasible to apply universal steganalytic approaches to attack emerging steganographic tools.

Images are not only able to carry stego information, but also able to convey fake information. Sometimes people maliciously manipulate images to forge a scene that actually never exists to mislead the observers on purpose. Taking a look at Fig. 1(a), one can see that in this picture, John Kerry stands by Jane Fonda, who was speaking to Vietnam veterans at an anti-war rally in 1970's. Appearing at the beginning of 2004, this picture enraged some Vietnam veterans. It was substantiated later to be a composite of Fig. 1(b) and Fig. 1(c).

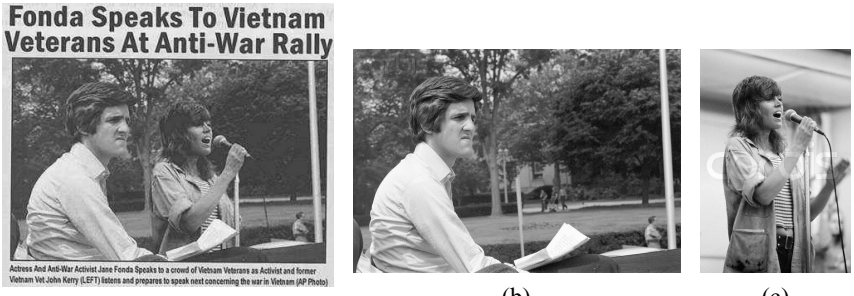


Fig. 1. An altered image and two original authentic images [2]

As its name implies, image splicing is a simple process of cropping and pasting regions from the same or different images to form another image without post-processing such as edge smoothing. Image splicing is one of the simple and commonly used image tampering schemes and often used as an initial and important step for image tampering. With modern image processing techniques, image splicing is hardly caught by the human visual system (HVS). Therefore, image splicing detection is of fundamental importance in image tampering detection and is urgently called for. People need to tell if a given image is spliced or not without any a priori knowledge. In other words, the splicing detection should be blind in nature.

Since both image steganalysis and splicing detection are critical for digital forensics and information assurance, in this paper, we take the lead to conduct a comparison study between universal steganalysis and splicing detection. Analytical research and extensive experiments have demonstrated that, those approaches used in

steganalysis can promisingly make progress in splicing detection applications if appropriately applied. This study leads to our conclusion that lessons learnt from steganalysis can shed light to splicing detection applications. Moreover, we raise and respond to some questions. Although the questions may have not been answered satisfactorily, our responses are expected to be thought provoking.

The rest of this paper is organized as follows. A general comparison between image steganalysis and splicing detection, and the concept of natural image model are contained in Section 2. Machine learning framework is presented in Section 3. In Section 4, some successful universal steganalysis approaches are reviewed and the natural image models they used are applied to splicing detection, and their performances are reported. In Section 5, a newly developed natural image model and one of its concrete versions designed for the available Columbia Image Splicing Detection Evaluation Dataset [3] are given to boost splicing detection capability. Further discussion on steganalysis and splicing detection with respect to natural image model is given in Section 6 and conclusion is drawn in Section 7.

2 Natural Image Model for Steganalysis and Splicing Detection

In this section, we first study the difference and then the similarity between image steganalysis and splicing detection. At the end, we prompt a question for discussion.

Obviously, steganalysis and splicing detection have different motivations and objectives. Specifically, steganalysis is to deter the secret communication, while splicing detection is to authenticate a given image by determining if it has been spliced. To look further, it is necessary to study their counterparts, i.e., steganography and splicing.

Steganography encodes information bits (possibly encrypted) and then embeds the bits into the cover image. Splicing is to replace one or more parts of a host image with fragment(s) from the same host image or other source images. Therefore, the statistical artifacts left with steganography are likely different from those caused by splicing.

Because steganography is to transmit secret message to the receiver, it is essential to hide the message in a digital image without drawing suspicion. Therefore, it is important for steganography to reduce the difference between a cover image and the stego image. Furthermore, steganography must ensure the embedded message "readable" to the receiver. However, the receiver of information conveyed by spliced image is assumed to be human eyes instead of a machine, and therefore, it is critical to make the host image and the cut-and-pasted image fragment(s) look accordant. In other words, it is essential to make sure a spliced image perceived by human eyes is just like a non-spliced (natural) image, though the image content has been changed. That is, once the HVS observes that a spliced image is not spliced, this specific splicing operation is considered to be successful.

To resist steganalytic attacks, steganography often embeds data in a cover image as widely as possible, while splicing just touches part(s) of the host image. In this sense, steganography is more global while splicing is more local in nature. Consequently, the relative change between a cover image and its stego image is small. Differently, splicing generally changes the content of a host image. Therefore, the relative change

between a host image and its spliced version is larger, more dramatic, and more local though.

From information assurance point of view, both image steganography and splicing are to disguise something: they try to make their touched images look like intact. In other words, they try to impress observers that these stego images or spliced images are natural images.

Since these stego images and spliced images are ever touched, the steganographic embedding and splicing operation shall cause disturbance on the smoothness, regularity, continuity, consistency, and/or periodicity of the images. Therefore, they do cause some statistical artifacts and are thus detectable using certain statistical attack approaches. As shown in the next a few sections, if we have a well designed natural image model, which can separate stego images (with hidden data and therefore unnatural) and spliced images (with inconsistent image fragment(s) and therefore unnatural) from natural images, both steganography and splicing are detectable by machine learning schemes.

When analyzing the steganographic artifacts and splicing artifacts, we may treat the noise (defined as the difference between a stego image and its cover image or a spliced image and its authentic host image) as additive noise. This assumption does not always hold. It, however, will simplify our analysis.

How to measure the strength of change brought to the cover image (for steganography) or host image (for splicing) is an interesting question. We may have subjective measurement and objective measurement. The HVS generally provide subjective measurement. For steganography, we may use BPP (bits per pixel) and/or MSE (mean square error) (or PSNR, peak signal to noise ratio) as objective measures. For splicing, it is hard to measure the change strength. MSE (or PSNR) may be an objective measure candidate. Other possible measurements include the ratio of the perimeter of the cut-and-pasted image fragment(s) to that of the host image and the ratio of the area of the cut-and-pasted image fragment(s) to that of the host image. More study in this regard seems necessary. Covert communications emerged at the very beginning of our human history. From human hair to papyrus, many articles were ever used as carriers to pass secret messages. Nowadays, digital images have become an important channel to bear stego information.

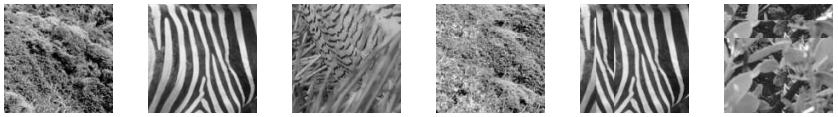
3 Machine Learning for Steganalysis and Splicing Detection

As mentioned above, a well designed natural image model may separate stego images or spliced images from natural images. An image model consists of a set of features, or a feature vector, characterizing a given image. With a dataset comprising both natural images and non-natural (stego or spliced) images, universal steganalysis or splicing detection can be carried out under the machine learning framework. The dataset and classifier used in our experimental study are described in this section.

3.1 Image Database for Splicing Detection

The Columbia Image Splicing Detection Evaluation Dataset [3] is used in our

experimental work by courtesy of digital video and multimedia lab (DVMM), Columbia University. This data set is created by DVMM for benchmarking the blind passive image splicing detection algorithms. There are 933 authentic and 912 spliced images in this data set, each of which is of the same size 128×128. Some authentic images and spliced images from [3] are shown in Fig. 2. This image dataset is open for downloading. For more information about [3], readers are referred to the technical report [4].



(a) (b) (c) (d) (e) (f)

Fig. 2. Some sample authentic images ((a), (b), and (c)) and spliced images ((d), (e), and (f)) used in this experimental work (source: [3])

3.2 Classifier, Classification, and Result Analysis

From machine learning point of view, both universal steganalysis and universal splicing detection consist of the following two parts: feature extraction, training and using trained classifier for testing.

In supervised machine learning, the support vector machine (SVM) has been widely used as the classifier. The SVM codes of Matlab are downloaded from [5]. Specifically, we use the radial basis function (RBF) kernel in our experimental works.

In each experiment, randomly selected 5/6 of the authentic images and 5/6 of the spliced images are used to train a SVM classifier. The remaining 1/6 of these images are used to test the trained classifier. The receiver operating characteristic (ROC) curve is thereafter obtained to demonstrate the trained classifier's performance. Two numerical methods can also be used to show the classifier's performance. One is to calculate the area under the ROC curve (AUC). Readers are referred to [6] for more information of ROC and AUC. Another is to obtain true negative (TN) rate, true positive (TP) rate, and accuracy of the trained classifier. To reduce the effect caused by variation incurred in selection of training/testing images, we individually conduct each random experiment 20 times and report the arithmetic average of these random tests.

4 Applying Natural Image Models Created in Universal Steganalysis to Splicing Detection

Since splicing detection emerges later than steganalysis, it is expected that splicing detection can "borrow" something from steganalysis. In this section, four natural image models, i.e., feature vectors used in four state-of-the-art universal steganalysis schemes are introduced. We then apply these models to the Dataset [3] and use the trained SVM as classifier to detect spliced images. Experimental results are reported.

4.1 Review of Some State-of-the-Art Universal Steganalysis Methods

In [7], the first four order statistical moments of wavelet coefficients and their prediction errors of nine wavelet high frequency subbands are used to form a 72-dimensional (72-D) feature vector for steganalysis. The block diagram of this image model (thereafter Hyu and Farid's) is shown in Fig. 3.

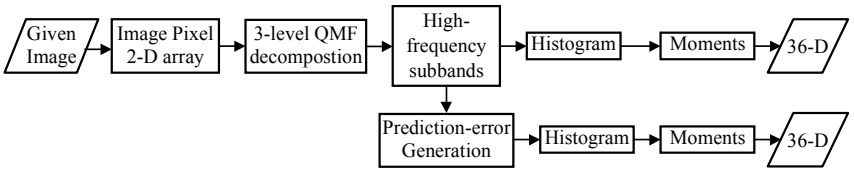


Fig. 3. Feature generation block diagram of Hyu and Farid's method [7]

A steganalysis scheme utilizing statistical moments of characteristic functions of the test image, its prediction-error image, and all of their wavelet subbands is proposed in [8], where a 78-D feature vector is used as the image model. The block diagram of this image model (thereafter Shi et al.'s) is given in Fig. 4.

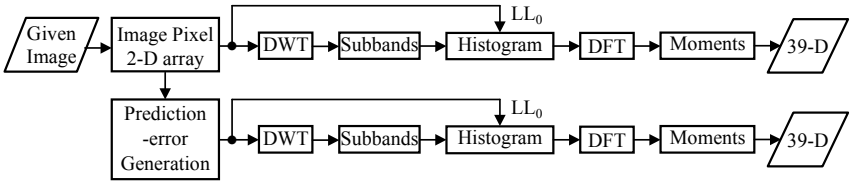


Fig. 4. Feature generation block diagram of Shi et al.'s method [8]

In [9], a steganalysis system based on 2-D Markov chain of thresholded prediction-error image is proposed. Image pixels are predicted with their neighboring pixels, and the prediction-error image is generated by subtracting the prediction value from the pixel value and then thresholded with a predefined threshold (set as 4). The empirical transition matrices of Markov chain along the horizontal, vertical, and main diagonal directions serve as features. The block diagram of this image model (thereafter Zou et al.'s) is shown in Fig. 5. Please note that those parts within dot lines in Fig. 5 are not used in [9]. Instead, these parts are used in [10].

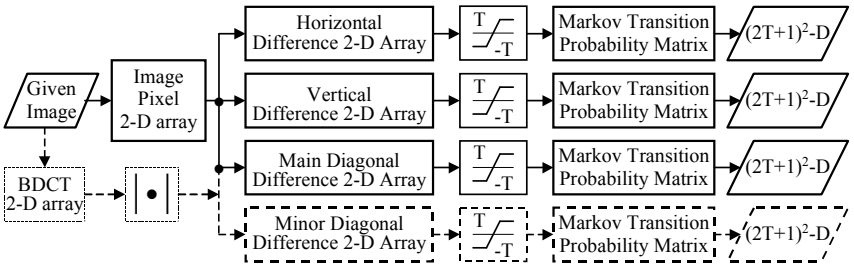
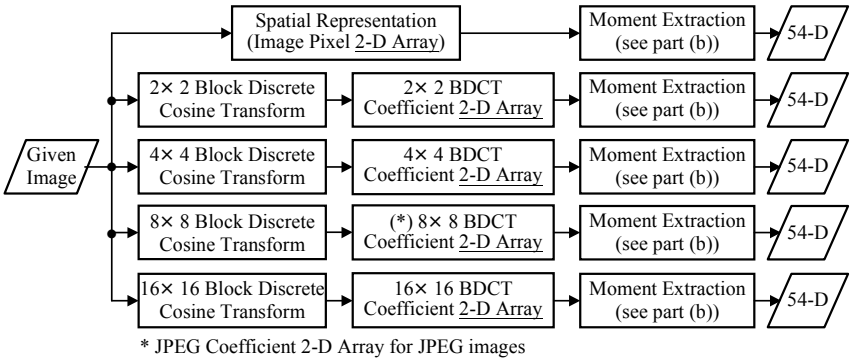
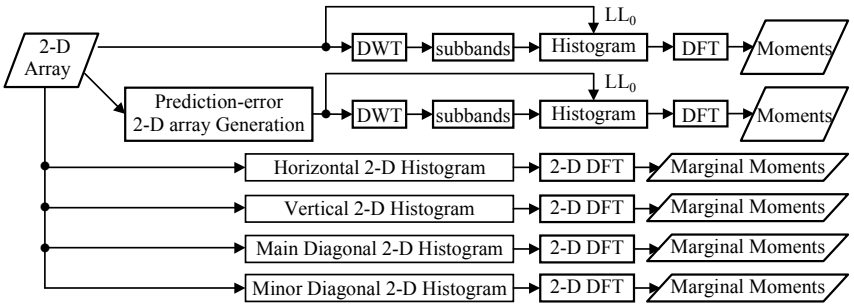


Fig. 5. Feature generation block diagram of Zou et al.'s method

Another effective universal steganalyzer is proposed in [11], which combines statistical moments of 1-D and 2-D characteristic functions extracted from the image pixel 2-D array and the multi-size block discrete cosine transform (MBDCT) 2-D arrays. This scheme greatly improves the capability of attacking steganographic methods applied to texture images, which has been shown to be a tough task [11]. In addition, this scheme can also be used as an effective universal steganalyzer for non-texture images. The block diagram of this method (thereafter Chen et al.'s) is shown in Fig. 6.



(a)



(b)

Fig. 6. Feature generation block diagram of Chen et al.'s method. (a).General block diagram; (b).Moment generation block diagram

4.2 Applying These Natural Image Models to Splicing Detection

Although image steganography and image splicing result in different statistical artifacts, both of them cause disturbances on the correlation between image pixels. Since both artifacts deviate from natural image model, both steganography and

splicing can be detected by using natural image model and machine learning framework. Therefore it is possible to apply natural image model established in steganalysis to splicing detection applications via machine learning methodology.

To evaluate the performances of those four natural image models when applied to splicing detection, we directly use these methods to extract features from images in Dataset [3], then train the SVM classified, and finally use the trained SVM classifier to detect spliced images. Experimental results are provided in Table 1.

It is observed from Table 1 that all these four steganalysis approaches can be used to detect spliced images with quite good performances, while the method using moments generated from both 1-D and 2-D characteristic functions and MBDCT [11], and the method using Markov based features [9] perform even better. Thus we have shown that the natural image models established in these steganalysis approaches can be straightforwardly applied to detection of spliced images.

Table 1. Detecting spliced images using four natural image models established in universal steganalysis (standard deviation in parentheses)

	Hyu and Farid's	Shi et al.'s	Zou et al.'s	Chen et al.'s
TN Rate	69.42% (4.34%)	76.07% (2.41%)	75.10% (3.47%)	86.32% (3.02%)
TP Rate	78.22% (4.14%)	75.59% (3.85%)	77.43% (4.24%)	87.83% (2.95%)
Accuracy	73.78% (2.01%)	75.83% (2.36%)	76.25% (3.20%)	87.07% (2.28%)
AUC	0.7892 (0.0206)	0.8162 (0.0247)	0.8232 (0.0244)	0.9292 (0.0170)

5 An Advanced Natural Image Model to Boost Splicing Detection Capability

Steganalysis methods presented in Section 4 can be summarized as follows. An image model (i.e., a feature vector) is used to represent an image in the statistic sense. This model describes characteristics of different image classes, i.e., non-stego images and stego images. A classifier is then used to separate these classes. These natural image models have demonstrated their promising capability on splicing detection.

In this section, we present a novel natural image model to separate spliced images from natural images. In this model, two types of features are used: moments (of characteristic functions) based features and Markov process based features. This model further combines features derived from the image pixel 2-D array and features derived from the multi-size block discrete cosine transform (MBDCT) coefficient 2-D arrays.

5.1 General Framework of This Natural Image Model

This natural image model is shown in Fig. 7. That is, we consider the spatial representation of the given test image (i.e., the image pixel 2-D array, or referred to as image 2-D array for short in this paper) and extract statistical moments of characteristic functions and Markov transition probabilities from this image 2-D

array. Furthermore, we apply block discrete cosine transform (BDCT) with a set of block sizes to the test image, resulting in a set of BDCT coefficient 2-D arrays (i.e., MBDCT coefficient 2-D arrays, or MBDCT 2-D arrays for short in this paper). From these MBDCT 2-D arrays, we also extract statistical moments of characteristic functions and Markov transition probabilities as features.

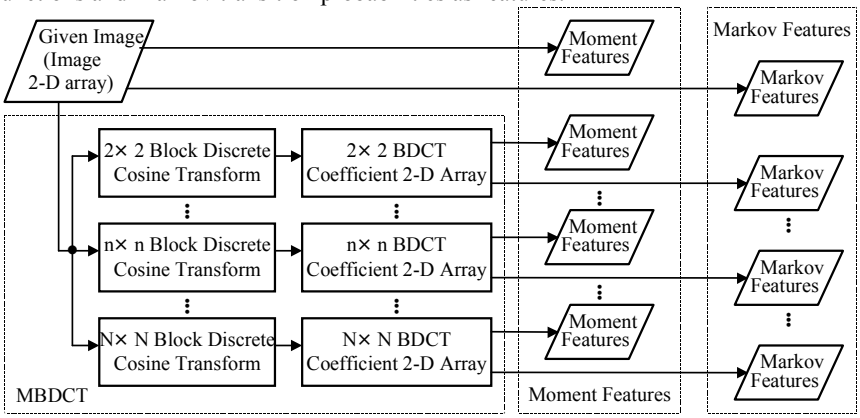


Fig. 7. A natural image model

5.2 Multi-size Block Discrete Cosine Transform (MBDCT) 2-D Arrays

As in [11], BDCT with a set of block sizes is used in this novel natural image model. This is to utilize the complementary decorrelation capabilities contributed by BDCT's with various block sizes. The splicing procedure changes the frequency distribution of the host images. Coefficients of the BDCT's can reflect these changes. It is noted that the pattern in which the correlation changes is various and complicated due to different host images (contents), different cut-and-pasted image fragments (contents and shapes), and different possible positions of those image fragments. Therefore, one cannot expect to catch this kind of changes effectively by using only one single-block-size BDCT. With various block sizes, the MBDCT coefficients can perceive the changes of frequency distribution in a variety of ways and hence the spliced images can be distinguished from natural images with features extracted from these MBDCT 2-D arrays.

The application of an n×n BDCT is described as follows. Firstly, the given image is divided into non-overlapping n×n blocks. Then, 2-D discrete cosine transform (DCT) is applied to each block independently. Finally, we obtain a 2-D array consisting of all the BDCT coefficients of all these blocks.

With each individual block size, we obtain one BDCT 2-D array. From the image 2-D array or each BDCT 2-D array, we can generate corresponding features.

5.3 Moment Based Features

As shown in Fig. 6(b), the moment feature extraction procedure is similar to that described in [11]. That is, moment based features consist of moments of 1-D

characteristic functions (discrete Fourier transform (DFT) of the first-order histograms) of the image 2-D array or MBDCT 2-D array, its prediction-error 2-D array, and all of their wavelet subbands, and marginal moments of the 2-D characteristic functions (2-D DFT of the second-order histograms) of the image 2-D array or MBDCT 2-D array.

Wavelet analysis, prediction-error, characteristic function, and 2-D histogram are key points of moment based features. Wavelet analysis has been widely used in digital image processing applications owing to its superior multi-resolution and space-frequency analytical capability. Wavelet transform is suitable to catch transient or localized changes in spatial and frequency domains and hence good for splicing detection, which has demonstrated its efficiency in steganalysis applications [7, 8, 10]. A prediction-error 2-D array is used to reduce the influence caused by diversity of image contents and to simultaneously enhance the statistical artifacts introduced by splicing. Moments of characteristic functions have been shown more effective than moments of histograms [8]. By measuring the intensity change of pixels (or coefficients) with respect to their neighbors, 2-D histograms can reflect the statistical effects of splicing artifacts more efficiently than 1-D histograms, which consider one pixel (or coefficient) at a time and thus do not reflect the intensity/position relationship among neighboring pixels.

5.4 Markov Process Based Features

Moment based features are a kind of measures reflecting the statistical changes caused by image splicing. Introduced in this section, Markov based features are another kind of measures able to reflect those statistical changes. Combining these two kinds of features, this novel natural image model can be enhanced to detect spliced images more effectively.

The Markov feature extraction procedure is similar to that described in [9]. At first, we form difference 2-D arrays (difference arrays for short) from the given image and/or coefficient 2-D array. These difference 2-D arrays are modeled by Markov process and then the transition probability matrix is calculated for each difference array. The entries of all the transition probability matrices are utilized as features to build up another part of the natural image model. In addition, a thresholding technique is developed to greatly reduce the dimensionality of the transition probability matrices, and hence the dimensionality of feature vectors, thus making the computational complexity manageable. The general block diagram of Markov feature extraction is shown in Fig. 5, in which four difference arrays are formed from an image 2-D array or BDCT 2-D array (rounded magnitude). This procedure is also similar to that used in [10].

By simply predicting an image pixel or a BDCT coefficient using its immediate neighbor, it is expected that the disturbances caused by splicing can be emphasized by the prediction-error, i.e., the difference between an element and its neighbor in an image or BDCT 2-D array.

For a given image or BDCT 2-D array (each of its element is the magnitude of a coefficient rounded into an integer), we can form four difference arrays, i.e., the horizontal, vertical, main diagonal, and minor diagonal difference 2-D arrays $F_h(u,v)$, $F_v(u,v)$, $F_d(u,v)$, and $F_m(u,v)$. The former three are called horizontal,

vertical, and diagonal prediction-error image in [9]. The minor diagonal difference 2-D array is given by

$$F_m(i,j) = x(i,j+1) - x(i+1,j), \tag{1}$$

where $i \in [0, S_i - 2],\ j \in [0, S_j - 2]$, S_i, S_j denote the 2-D array's dimensions in the horizontal direction and vertical direction, respectively, and $x(i,j)$ is either an image pixel value or a rounded magnitude of BDCT coefficient.

Proverbially, there exists correlation between pixels/coefficients in an image/coefficient 2-D array and the difference arrays. Therefore, we can use Markov process to model these difference arrays. According to the theory of random process, a transition probability matrix can be used to characterize these Markov processes. In our natural image model, we use the so-called one-step transition probability matrix to characterize those difference arrays [12].

To further reduce computational complexity, we resort to a thresholding technique, i.e., the value of an element in a difference array is represented by T or –T respectively if it is either larger than T or smaller than -T, resulting in a transition probability matrix of (2T+1) × (2T+1), the elements of which are given by

$$p\{F_h(i,j) = n \mid F_h(i+1,j) = m\} = \frac{\sum_{j=0}^{S_j-2}\sum_{i=0}^{S_i-2}\delta(F_h(i+1,j) = m, F_h(i,j) = n)}{\sum_{j=0}^{S_j-2}\sum_{i=0}^{S_i-2}\delta(F_h(i+1,j) = m)}, \tag{2}$$

$$p\{F_v(i,j) = n \mid F_v(i,j+1) = m\} = \frac{\sum_{j=0}^{S_j-2}\sum_{i=0}^{S_i-2}\delta(F_v(i,j+1) = m, F_v(i,j) = n)}{\sum_{j=0}^{S_j-2}\sum_{i=0}^{S_i-2}\delta(F_v(i,j+1) = m)}, \tag{3}$$

$$p\{F_d(i,j) = n \mid F_d(i+1,j+1) = m\} = \frac{\sum_{j=0}^{S_j-2}\sum_{i=0}^{S_i-2}\delta(F_d(i+1,j+1) = m, F_d(i,j) = n)}{\sum_{j=0}^{S_j-2}\sum_{i=0}^{S_i-2}\delta(F_d(i+1,j+1) = m)}, \tag{4}$$

$$p\{F_m(i+1,j) = n \mid F_m(i,j+1) = m\} = \frac{\sum_{j=0}^{S_j-2}\sum_{i=0}^{S_i-2}\delta(F_m(i,j+1) = m, F_m(i+1,j) = n)}{\sum_{j=0}^{S_j-2}\sum_{i=0}^{S_i-2}\delta(F_m(i,j+1) = m)}, \tag{5}$$

where $m, n \in \{-T, -T+1, \cdots, 0, \cdots, T\}$, and

$$\delta(A = m, B = n) = \begin{cases} 1, & if\ A = m\ and\ B = n \\ 0, & Otherwise \end{cases}. \tag{6}$$

Note that all the elements of the transition probability matrix are used as features to form our natural image model.

5.5 A Concrete Implementation on Image Dataset [3]

Thus far in this section, we have presented the general framework of the novel natural image model. Here, we provide a concrete implementation of this model on image dataset [3] with only 933 authentic images and 912 spliced images, each of which has

a dimension of 128×128. These facts have determined this concrete implementation. The most obvious constraint is that the feature vector's dimensionality is bonded by the limited number of images in the dataset [3].

In this implementation, the BDCT block sizes are selected as 2×2, 4×4, and 8×8, partially because this choice (power of 2) is of computational benefits in implementing BDCT. As a result, for each given test image, we have the image pixel 2-D array, and 2×2, 4×4, and 8×8 BDCT 2-D arrays in this concrete implementation.

Haar wavelet is used in our implementation due to its simplicity in implementation. Furthermore, to balance splicing detection capability and computational complexity, we only conduct one-level wavelet decomposition and thus have five subbands (the image 2-D array is considered as the LL_0 subband). Compared to three-level DWT decomposition, which results in 13 subbands, the number of subbands reduces by 62%.

For each wavelet subband, we compute the first three order moments. When calculating marginal moments from the second-order histograms, we only use horizontal and vertical 2-D histograms and obtain the first three order marginal moments. As a result, 42 moment features are obtained from each given image 2-D array or BDCT 2-D array. Since we have one image pixel 2-D array and three derived BDCT 2-D arrays, 168 moment features for each given image are formed in this specific implementation.

To limit the dimensionality of Markov features, in this specific implementation, we only use two difference arrays (i.e., horizontal difference array and vertical difference 2-D array) and we only generate these two types of difference arrays from the 8×8 BDCT 2-D array of the given image. With threshold $T = 4$, a transition probability matrix of dimensionality $9 \times 9 = 81$ is formed for each direction and the derived Markov features are of dimensionality 162 consequently.

To sum up, 330 features are used to represent this novel natural image model.

5.6 Experimental Results with this Concrete Implementation

The averaged ROC curve of 20 experiments applying our proposed natural image model to image dataset [3] is given in Fig. 8, where the ROC curves of experiments reported in Section 4.2 are also included. This concrete implementation of the novel approach has achieved TN rate 91.52% (2.19%), TP rate 92.86% (1.72%), accuracy 92.18% (1.30%), and averaged AUC 0.9537 (0.0112). Compared to the accuracies 72%, 80%, and 82% achieved by [14, 15, and 16] over the same image dataset [3], this novel natural image model has made a significant advancement in splicing detection.

5.7 Detecting Real Images

In Section 1, we have given an altered image and its associated originals (three images in total). We used the trained classifier resulted in the 20 random experiments mentioned in Section 5.6 to test these three images. The 20 test results are shown in Table 2. That is, among these 60 image-tests, 57 provide correct classification. These

results are rather encouraging. No doubt, however, much more efforts need to be made for image splicing/tampering detection research in the future.

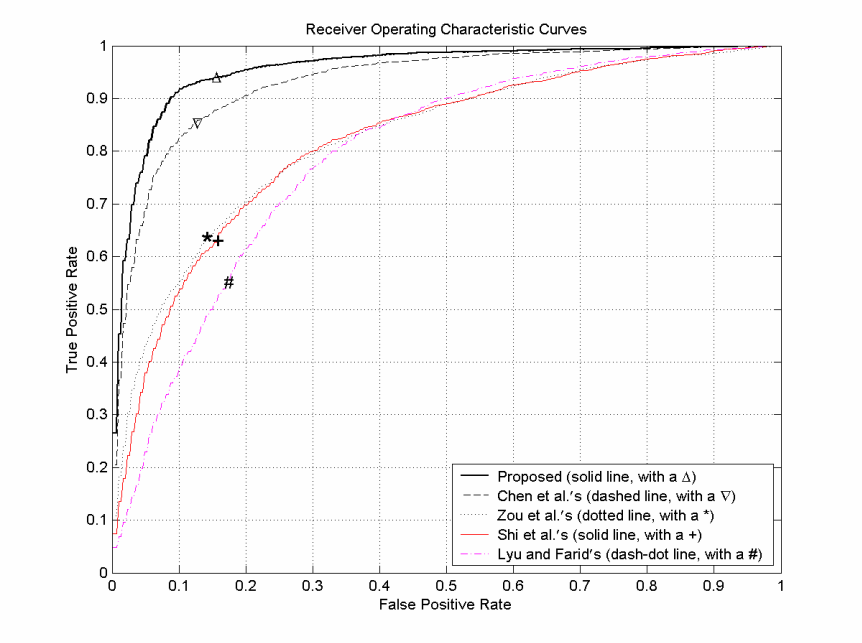


Fig. 8. The ROC curves of applying the natural image model to [3]

Table 2. Test results on real images (√— correct, × — wrong)

Test Sequence	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Fig. 1(a)	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Fig. 1(b)	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Fig. 1(c)	√	×	√	√	√	√	√	×	√	√	√	√	√	√	√	√	√	√	×	√

6 Discussion

We have presented a novel splicing detection scheme in Section 5. This scheme constructs a natural image model derived from some effective steganalysis approaches to distinguish spliced images from natural authentic images.

We can describe the relationship of steganalysis and splicing detection using Fig. 9. With a well designed natural image model, stego images and spliced images can be separated from natural images in the feature space. Although stego images and spliced images occupy different regions in the high dimensional feature space, both of them are separable from natural images. In other words, if the image model is advanced

enough, both image steganalysis and splicing detection can be successfully performed using the advanced natural image model provided the corresponding and sufficient image datasets are available.

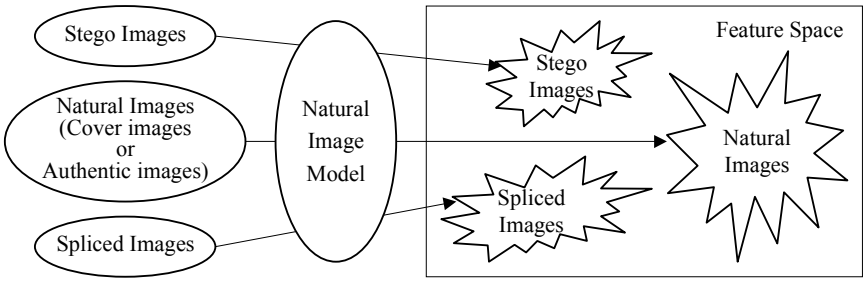


Fig. 9. Mapping images to feature space

As we discussed in this paper, splicing detection can learn lessons from steganalysis. Can steganalysis learn some lesson from splicing detection? The answer seems to be positive. It is expected that both steganalysis and splicing detection can learn from each other and advance each other.

7 Conclusion

In this paper, we have conducted a comparison study between steganalysis and splicing detection and proposed an advanced natural image model for splicing detection. We can summarize our study as follows.

1) Although different in target and application, image steganography and splicing do have some common aspects. One of such aspects is that both of them cause the touched images to deviate from natural images. The generated statistical artifacts are different though. Hence, steganography and splicing can be detected by natural image model (a set of natural image features).
2) Our experimental works have shown that applying the natural image models established in several state-of-the-art steganalysis schemes to splicing detection with a well-designed dataset for splicing detection has resulted in promising splicing detection capability.
3) Furthermore, a novel natural image model, developed from the state-of-the-art steganalysis schemes, has been presented and applied to splicing detection. Its performance has demonstrated significant advancement in splicing detection.
4) Lessons learnt from steganalysis can be applied to splicing detection and it is expected that steganalysis can learn something from splicing detection as well.

Acknowledgments. Credits for the use of the Columbia Image Splicing Detection Evaluation Dataset are given to the DVMM Laboratory of Columbia University, CalPhotos Digital Library, and the photographers, whose names are listed in [4].

References

1. Kharrazi, M., Sencar, H. T., and Memon, N.: Image Steganography: Concepts and Practice. Lecture Note Series, Institute for Mathematical Sciences, National University of Singapore (2004)
2. http://www.snopes.com/photos/politics/kerry2.asp
3. Columbia DVMM Research Lab: Columbia Image Splicing Detection Evaluation Dataset, http://www.ee.columbia.edu/ln/dvmm/downloads/AuthSplicedDataSet/AuthSplicedDataSet.htm (2004)
4. http://www.ee.columbia.edu/dvmm/publications/04/TR_splicingDataSet_ttng.pdf
5. Chang, C. C. and Lin, C. J.: LIBSVM: A Library for Support Vector Machines. http://www.csie.ntu.edu.tw/cjlin/libsvm
6. Fawcett, T.: Roc Graphs: Notes and Practical Considerations for Researchers. http://home.comcast.net/~tom.fawcett/public _html/papers/ROC101.pdf
7. Lyu, S. and Farid, H.: Detecting Hidden Messages Using Higher-order Statistics and Support Vector Machines. Proc. of 5th *International Workshop on Information Hiding*, Noordwijkerhout, The Netherlands (2002)
8. Shi, Y. Q., Xuan, G., Zou, D., Gao, J., Yang, C., Zhang, Z., Chai, P., Chen, W., and Chen, C.: Steganalysis Based on Moments of Characteristic Functions Using Wavelet Decomposition, Prediction-error Image, and Neural Network. *International Conference on Multimedia and Expo*, Amsterdam, the Netherlands (2005)
9. Zou, D., Shi, Y. Q., Su, W., and Xuan, G.: Steganalysis Based on Markov Model of Thresholded Prediction-Error Image. *International Conference on Multimedia and Expo*, Toronto, ON, Canada (2006)
10. Shi, Y. Q., Chen C., and Chen, W.: A Markov Process Based Approach to Effective Attacking JEPG Steganography. *Information Hiding Workshop* (*IH06),* Old Town Alexandria, VA, USA (2006)
11. Chen, C., Shi, Y. Q., and Xuan, G.: Steganalyzing Texture images. *International Conference on Image Processing*, St. Antonio, TX, USA (2007, accepted)
12. Leon-Garcia, A.: *Probability and Random Processes for Electrical Engineering*, 2nd Edition, Addison-Wesley Publishing Company (1994)
13. Vapnik, V. N.: *The Nature of Statistical Learning Theory*, 2nd Edition, Springer (1999)
14. Ng, T.-T., Chang, S.-F., and Sun, Q.: Blind Detection of Photomontage Using Higher Order Statistics. *IEEE International Symposium on Circuits and Systems*, Vancouver, BC, Canada (2004)
15. Fu, D., Shi, Y. Q., and Su, W.: Detection of Image Splicing Based on Hilbert-Huang Transform and Moments of Characteristic Functions with Wavelet Decomposition. In: Shi, Y. Q., Jeon, B. (eds.): Digital Watermarking, Proceeding of 5th *International Workshop on Digital Watermarking*, Jeju Island, Korea (2006)
16. Chen, W., Shi, Y. Q., and Su, W.: Image Splicing Detection Using 2-D Phase Congruency and Statistical Moments of Characteristic Function. In: Delp, E. J., Wong, P.W. (eds.): *Security, Steganography and Watermarking of Multimedia Contents IX*, Proceeding. *of SPIE*, San Jose, CA, USA (2007)