Steganalysis of Adaptive JPEG Steganography Using 2D Gabor Filters

Xiaofeng Song*
State Key Laboratory of Mathematical Engineering and Advanced Computing
Zhengzhou 450001, China
xiaofengsong@sina.com

Fenlin Liu
Zhengzhou Information Science and Technology Institute
Zhengzhou 450001, China
liufenlin@sina.vip.com

Chunfang Yang†
Zhengzhou Information Science and Technology Institute
Zhengzhou 450001, China
chunfangyang@126.com

Xiangyang Luo‡
Zhengzhou Information Science and Technology Institute
Zhengzhou 450001, China
xiangyangluo@126.com

Yi Zhang
Zhengzhou Information Science and Technology Institute
Zhengzhou 450001, China
yizhang0125@163.com

ABSTRACT

Adaptive JPEG steganographic schemes are difficult to preserve the image texture features in all scales and orientations when the embedding changes are constrained to the complicated texture regions, then a steganalysis feature extraction method is proposed based on 2 dimensional (2D) Gabor filters. The 2D Gabor filters have certain optimal joint localization properties in the spatial domain and in the spatial frequency domain. They can describe the image texture features from different scales and orientations, therefore the changes of image statistical characteristics caused by steganography embedding can be captured more effectively. For the proposed feature extraction method, the decompressed JPEG image is filtered by 2D Gabor filters with different scales and orientations firstly. Then, the histogram features are extracted from all the filtered images.Lastly, the ensemble classifier is used to assemble the proposed steganalysis feature as well as the final steganalyzer. The experimental results show that the proposed steganalysis feature can achieve a competitive performance by comparing with the other steganalysis features when they are used for the detection performance of adaptive JPEG steganography such as UED, JUNIWARD and SI-UNIWARD.

Categories and Subject Descriptors

I.4.9 [**Computing Methodologies**]: Image Processing and Computer Vision—*Applications*

General Terms

Algorithms, Design, Security

Keywords

Adaptive steganography; JPEG; steganalysis; feature extraction; 2D Gabor filters

1. INTRODUCTION

JPEG is one of the most popular image formats on the internet, so the steganography and steganalysis techniques about JPEG image attract more attentions. For now, the steganography algorithms for JPEG image can be divided into two parts: non-adaptive steganography and adaptive steganography. The former include Jsteg, MB1 [22], MB2, Outguess [21], F5 [25], nsF5 [9], MME (modified matrix encoding) [16], PQ (perturbed quantization) [7], etc. The latter include PQt (texture-adaptive PQ), PQe (energy-adaptive PQ) [9], MOD (Model Optimized Distortion) [4], EBS (Entropy Block Steganography) [24], UED (Uniform Embedding Distortion) [11], JUNIWARD (JPEG Universal Wavelet Relative Distortion) [12], SI-UNIWARD (Side-informed Universal Wavelet Relative Distortion) [12] and so on. The main difference between non-adaptive and adaptive steganography algorithms is that the former do not consider the image content characteristics when the changed cover elements are selected while the latter constrain the embedding changes to the complicated image texture and edge regions which are difficult to model. For the above adaptive JPEG steganography algorithms, the frameworks of the steganographic schemes are similar. They all define a distortion function firstly and then the messages are embedded by encoding method. For example, as to PQt and PQe, the distortion functions are defined according to the texture and energy measure respectively, and then the given

*Xiaofeng Song is also a PhD candidates at Zhengzhou Information Science and Technology Institute .

†The corresponding author

‡Xiangyang Luo is also a researcher at State Key Laboratory of Information Assurance,Beijing 100072,China.

IH&MMSec'15, June 17–19, 2015, Portland, Oregon, USA.

http://dx.doi.org/10.1145/2756601.2756608.

messages are embedded by wet paper code [6]; as to MOD, UED, EBS, JUNIWARD and SI-UNIWARD, the different distortion functions are defined respectively and the messages are embedded while minimizing the distortion function by Syndrome-Trellis Codes (STCs) [5]. Compared with the non-adaptive JPEG steganography, the adaptive JPEG steganography can achieve better anti-detection abilities because the embedding changes are constrained to complicated image texture and edge regions.

For the detection of the adaptive JPEG steganography, some steganalysis methods have been proposed in recent years. In literature [18], for MOD steganography, the distortion function optimized to maximize security has been overtrained to an incomplete cover model, therefore the inter-block co-occurrences features beyond the optimized model are proposed for the detection performance. In literature [23], the principle of PQ steganography is analyzed and the enhanced histogram features are proposed to improve the detection performance for PQ, PQt and PQe. In literature [17], in order to capture embedding changes more comprehensively, a rich model of DCT coefficients in a JPEG files is proposed. In literature [12], the CC-JRM (Cartesian calibration JPEG rich model) features are combined with SRMQ1 (Spatial Rich Model with the single quantization q =1) [8] to detect adaptive JPEG steganography. In literature [13], by projecting neighboring residual samples onto a set of random vectors, the PSRM (projection spatial rich model) features are proposed. PSRM take the first-order statistic (histogram) of the projections as the feature instead of forming the co-occurrence matrix**.** In literature [14], the DCTR (Discrete Cosine Transform Residual) features which utilize 64 kernels of the discrete cosine transform are proposed. For DCTR, the decompressed JPEG image is convoluted with each DCT kernel firstly and then the first-order statistics of quantized noise residuals are obtained by subsampling residual images. The DCTR features can achieve better detection performance while preserve relative low feature dimensions. In addition, in literature [23, 17, 8, 13, 14], the final detection accuracy are all obtained by ensemble classifier [19] after features extraction.

From all above, it can be seen that the current steganalysis methods mainly depend on more effective features to improve the detection performance for adaptive JPEG steganography. However, the existing steganalysis features are not extracted from different scales and more orientations to capture the embedding changes. In fact, as we all know, adaptive JPEG steganography constrains the embedding changes to image texture and edge regions. Moreover, these steganography algorithms often define the embedding distortion in single scale, such as JUNIWARD steganography utilizes the wavelet decomposition coefficients in one scale and three different orientations to define distortion. Therefore, for the steganalyzer, if the image texture and edge features can be described accurately from different scales and orientations, then the statistical features extracted from the rich image representation can reflect the image changes more effectively and the detection performance can also be improved. Base on this idea, a steganalysis feature extraction method is proposed based on 2D Gabor filters. For the proposed method, the decompressed JPEG image is decomposed by the 2D Gabor filters [2] with different scales and orientations, and then the steganalysis features are extracted from the image filtering coefficients. The 2D Gabor filter acts as a local band-pass filter with certain optimal joint localization properties in the spatial domain and in the spatial frequency domain. It can describe the image texture and edge features effectively. In contrast to DCTR features which utilize 64 DCT kernels for image filtering, the 2D Gabor filters can capture the embedding changes from more scales and orientations, so the steganalysis feature extracted from image filtering coefficients got by 2D Gabor filters might be more effective for the detection performances of adaptive JPEG steganography. The usage of 2D Gabor filters is the main differences between DCTR features and the proposed steganalysis feature in this paper.

The rest of this paper is organized as follows. In section 2, the image representation using 2D Gabor filters is introduced; in section 3, the changes of image filtering coefficients are analyzed after steganography embedding; in section 4, the proposed steganalysis feature extraction method is given and analyzed in details, and then the parameters setting are also discussed; in section 5, the proposed steganalysis feature are compared with the other steganalysis features for adaptive JPEG steganography, in section 6, the conclusion is drawn.

2. IMAGE REPRESENTATION USING 2D GABOR FILTERS

The Gabor transform belongs to a short time Fourier transform, and it adds the Gaussian window to Fourier transform in order to realize the local analysis in spatial domain and spatial frequency domain. In literature [2], Daugman proposes the 2D Gabor filter theory based on Gabor transform and points out an important property of the family of 2D Gabor filters is their achievement of the theoretical lower bound of joint uncertainty in the two conjoint domains of visual space and spatial frequency variables. In other words, 2D Gabor filter can achieve optimal joint localization properties in the spatial domain and in the spatial frequency domain. In literature [3], the 2D Gabor filter has been used for the distortion definition of adaptive JPEG steganography, and the better anti-detection ability is achieved than db8 wavelet. In the following, the 2D Gabor filter is introduced.

When the 2D Gabor filters are used for image processing and analysis, the image should be filtered by 2D Gabor filters firstly, and then the feature extraction, edge detection, denoising and so on processing or analysis can be performed. The 2D Gabor filtering for image is that an input image $I(x, y)$ is convolved with a 2-D Gabor function $g(x, y)$ to obtain a Gabor feature image $u(x, y)$ as follows:

$$u(x,y) = \iint_{\Omega} I(\xi,\eta) g(x-\xi, y-\eta) d\xi d\eta \qquad (1)$$

where, $(x, y) \in \Omega$, Ω denotes the set of image points.

In this paper, the 2D Gabor function $g(x, y)$ in equation (1) use the following family of Gabor functions [10], it is a product of a Gaussian and a cosine function.

$$g_{\lambda,\theta,\phi}(x,y) = e^{-\left(\left(x'^2+\gamma^2 y'^2\right)/2\sigma^2\right)} \cos\left(2\pi\frac{x'}{\lambda}+\phi\right) \qquad (2)$$

where, $x' = x\cos\theta + y\sin\theta$, $y' = -x\sin\theta + y\cos\theta$, $\sigma = 0.56\lambda$, $\gamma = 0.5$. In equation (2), σ represents the scale parameter, the small σ means high spatial resolution, the image filtering coefficients reflect local properties in fine scale,

while the large σ means low spatial resolution, the coefficients reflect local properties in coarse scale. The other parameters in equation (2) can be explained as follows: θ specifies the orientations of 2D Gabor filters, λ denotes the wavelength of the cosine factor, γ is the spatial aspect ratio and specifies the ellipticity of Gaussian factor, ϕ specifies the phase offset of the cosine factor ($\phi = 0, \pi$ correspond to symmetric "centre-on" functions, while $\phi = -\pi/2,\ \pi/2$ correspond to anti-symmetric functions). In addition, in order to capture the embedding changes, all the 2D Gabor filters are made zero mean by subtracting the kernel mean from all its elements to form high-pass filter.

In Figure 1, the filtered images generated by convolving Lena image[1] with 2D Gabor filters are given. For the top filtered images,the corresponding parameters are set as $\sigma = 0.5,\ \phi = \pi/2,\ \theta = \{0, \pi/4, \pi/2, 3\pi/4\}$; as to the bottom filtered images, only the scale parameter σ is different. From the filtered image shown in Figure 1, it can be seen that the image local properties in different scales and orientations can be captured effectively. For real applications, the image should be filtered by 2D Gabor filters with more scales and orientations to capture rich texture and edge information. From Figure 1, it also can be found that the image local properties such as texture, edge are more obvious when the scale parameter σ is relative small and the spatial resolution is relative high, for example, as to the top filtered images, the texture and edge is highlighted. As to the bottom filtered images, the scale parameter σ is relative large, the filtered image reflect the image local properties in relative low resolution, so the texture and edge are relative obscure.

Adaptive JPEG steganography constrains the embedding changes to complex image texture and edge regions. The image filtering by 2D Gabor filters with different scales and orientations can capture the image texture and edge features accurately. So, the steganalysis feature extracted from image 2D Gabor filtering coefficients would be more effective.

3. EMBEDDING CHNAGES OF IMAGE FILTERING COEFFICIENTS

Adaptive JPEG steganography embeds the messages by modifying the DCT coefficients. The modifications will cause the changes of image filtering coefficients in different scales and orientations. In this section, we will take a look how the image filtering coefficients are affected by modifying one DCT coefficient of JPEG image.

Before the JPEG image is filtered by 2D Gabor filters, the JPEG file should be decompressed to the spatial domain. In order to avoid any loss of information, the JPEG image should be decompressed without quantizing the pixel values to {0,1,...,255}. Let us suppose the decompressed JPEG image is denoted as $\mathbf{I}'$, then the filtered image $\mathbf{U}^{s,l}=\mathbf{I}'*\mathbf{G}^{s,l}$, $\mathbf{G}^{s,l}$ specifies the 8×8 2D Gabor filter in s scale and l orientation, '*' denotes a convolution without padding. Furthermore, suppose $\mathbf{B}^{(i,j)}$ denotes a 8×8 DCT basis pattern, $\mathbf{B}^{(i,j)} = \left(B_{mn}^{(i,j)}\right), 0 \le m, n \le 7,\ 0 \le i, j \le 7,$

$$B_{mn}^{(i,j)} = \frac{w_i w_j}{4} \cos\frac{\pi i(2m+1)}{16} \cos\frac{\pi j(2n+1)}{16} \qquad (3)$$

where, $w_0 = 1/\sqrt{2},\ w_i = 1\ (i > 0)$.

[1]Lena image [EB/OL]. http://en.wikipedia.org/wiki/lenna.

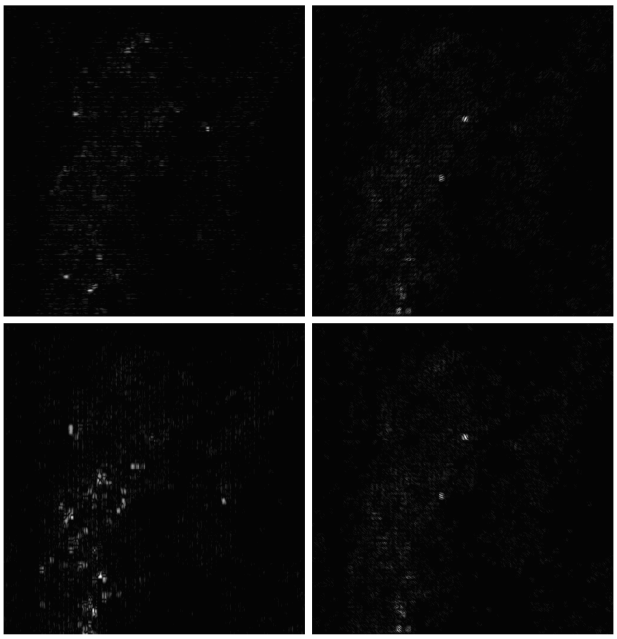

Figure 2: The absolute difference values between image filtering coefficients of Lena image and the corresponding stego image, from left to right, top to bottom, $\theta = 0, \pi/4, \pi/2, 3\pi/4$ respectively

Then , the modification of DCT coefficient in mode (i, j) of 8×8 DCT block will affect all 8×8 pixels in the corresponding block, and an entire 15×15 neighborhood of values in $\mathbf{U}^{s,l}$. The values will be modified by "unit response" [14] expressed in equation (4).

$$\mathbf{R}^{(i,j)\,(s,l)} = \mathbf{B}^{(i,j)} \otimes \mathbf{G}^{s.l} \qquad (4)$$

where, $\otimes$ denotes the full cross-correlation.

In Figure 2, the absolute difference values between image filtering coefficients of Lena and the corresponding stego image are shown. The parameters of 2D Gabor filters are set as $\sigma = 1,\ \phi = \pi/2,\ \theta = \{0, \pi/4, \pi/2, 3\pi/4\}$. The payload of stego image is 0.4 bit per non-zero AC DCT coefficient (bpac). From Figure 2, it can be seen that the image filtering coefficients are changed after steganography embedding.

4. STEGANALYSIS FEATURE EXTRACTION BASED ON 2D GABOR FILTERS

In this section, the steganalysis feature extraction method based on 2D Gabor filters is described in details firstly, and then the parameters setting are discussed for feature extraction.

4.1 Feature extraction method

When the steganalysis feature is extracted, the decompressed JPEG image is convolved with 2D Gabor filters to get the filtered images firstly. However, the statistical feature is not extracted directly from the filtered images in different scales and orientations. As we known, the JPEG image can be divided into 8×8 DCT blocks. According to the 64 DCT modes in 8×8DCT block, the each filtered image can be subsampled by step 8 to form 64 subimages. The

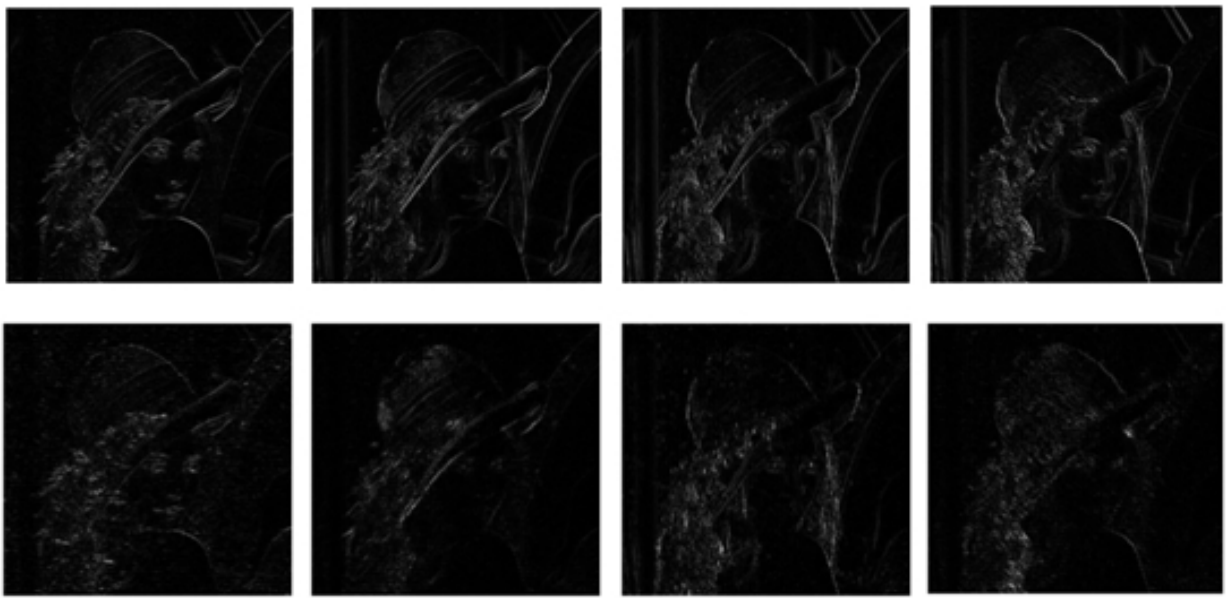

Figure 1: The filtered image obtained by convolving Lena image with different 2D Gabor filters (orientations are 0^0,45^0, 90^0, 135^0), the Scale parameter σ equals 0.5 for the top filtered images and it equals 1 for the bottom.

feature of each filtered image is formed by combing the features of 64 corresponding subimages. Lastly, the features of all filtered image are joined to obtain the steganalysis feature.

The flaw diagram of the proposed feature extraction method is shown in Figure 3 and the detailed extraction procedures are described as follows:

Step1: The JPEG image is decompressed to spatial domain without quantizing the pixel values to {0,1,...,255} to avoid any loss of information.

Step2: The 2D Gabor filter bank is generated and the filter bank includes 2D Gabor filters with different scales and orientations.

Step3: The decompressed JPEG image is convolved with each 8×8 2D Gabor filter $\mathbf{G}^{s,l}$, the filtered image $\mathbf{U}^{s,l}$ is operated as the following:

(1) According to the 64 DCT modes $(a,b)(0 \leq a \leq 7, 0 \leq b \leq 7)$ in 8×8 DCT block, the filtered image $\mathbf{U}^{s,l}$ is subsampled by step 8 to get 64 subimages $\mathbf{U}^{s,l}_{a,b}$ (as shown in Figure 4);

(2) For each subimage $\mathbf{U}^{s,l}_{a,b}$, the histogram feature is extracted by equation (5),

$$\mathbf{h}^{s,l}_{a,b}(x) = \frac{1}{\left|\mathbf{U}^{s,l}_{a,b}\right|} \sum_{u \in \mathbf{U}^{s,l}_{a,b}} \left[Q_T\left(|u|/q\right) = x\right] \tag{5}$$

where, Q_T is a quantizer with integer centroids $\{0, 1, \cdots, T\}$, q denotes the quantization step, and $[P]$ is the Iverson bracket equal to 0 when the statement P is false and 1 when P is true.

(3) According to the method in literature [14], all the histogram features of 64 subimages $\mathbf{U}^{s,l}_{a,b}$ is merged and combined to obtain the histogram features $\mathbf{h}^{s,l}$ of the filtered image $\mathbf{U}^{s,l}$.

Step4: For the filtered image generated by 2D Gabor filters with the same scale parameter σ, the corresponding histogram features are merged according to the symmetrical orientations. For example, suppose the orientation parameter $\theta = \{0, \pi/8, 2\pi/8, \cdots, 6\pi/8, 7\pi/8\}$, then the histogram features of the filtered image with $\theta = \pi/8$,$7\pi/8$, $\theta = 2\pi/8$,$6\pi/8$ and so on should be merged by averaging.

Step5: All the merged histogram features are combined to form the final steganglysis feature.

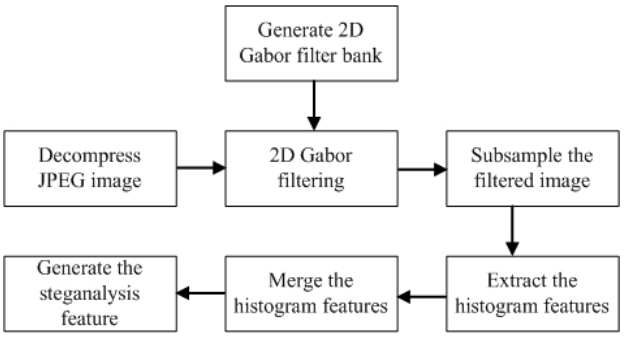

Figure 3: The flaw diagram of the proposed feature extraction method

From the above descriptions, it can be seen that the proposed feature extraction method mainly includes three parts: the generation of 2D Gabor filter bank, the histogram features extraction, the merging of histogram features. In the following, the further descriptions are given for these three parts.

(1) Generation of 2D Gabor filter bank

For the generation of 2D Gabor filter bank with different scales and orientations, suppose the number of scales equals L (the scale parameter σ has L different values), the number of orientations equals S, the parameter ϕ is set to 0 and $\pi/2$, so the number of the 2D Gabor filters is $2 \cdot L \cdot S$. For example, if L=4 and $S = 32$, then the number of the 2D Gabor filters

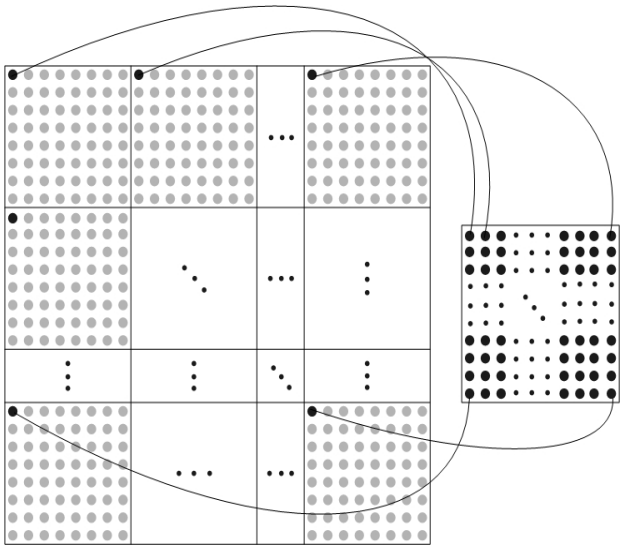

Figure 4: The procedure of the filtered image subsampling

is 256. The scale parameter L and orientation parameter S are both important for the extraction of steganalysis feature. In this paper, the two parameters are set according to the experiences.

(2) Histogram features extraction

For the histogram features extraction, the decompressed JPEG image is convolved with each 2D Gabor filter, and then the filtered image is subsampled and the histogram features are extracted. The features extraction by subsamping the filtered image into 64 subimages according to the 64 DCT modes in 8×8 DCT block can enhance the diversity and effectiveness of the steganalysis features, the similar idea has been used for the local histogram features extraction in literature [20, 14]. In addition, the histogram features extracted from the 64 subimages can be merged [14] to reduce the feature dimensionality according to the affects for filtered image which are caused by modifying the DCT coefficients in different DCT modes. For each filtered image, the 64 histogram features with $T+1$ dimensions can be got when the cut threshold is set to T, then these histogram features can be merged to one histogram feature with $25\times(T+1)$ dimensions. Furthermore, the histogram feature with $2\cdot L\cdot S\cdot 25\cdot(T+1)$ dimension can be obtained by all the $2\cdot L\cdot S$ 2D Gabor filters.

(3) Merging of histogram features

When the histogram feature of the decompressed JPEG image is obtained, the dimensions of the histogram feature can be reduced further by merging the features extracted by 2D Gabor filters with symmetrical direction and same scale. For example, the histogram features of the filtered image with $\theta=\pi/8$,$7\pi/8$, $\theta=2\pi/8$,$6\pi/8$and so on should be merged by averaging. This merging operation has been used in many feature extraction method [17, 8, 13] because images have similar statistical characteristics in symmetrical orientations and same scale. After the histogram features are merged by symmetrical orientations, the dimensions of the final feature are $2\cdot L\cdot(S/2+1)\cdot 25\cdot(T+1)$. In the experiments, it can be found that the detection performance of the merging features will be better.

4.2 Parameters setting

For the feature extraction method proposed in this paper, all the parameters θ,σ,q,T should be set before feature extraction. In the following, the detection performances of different parameter settings are discussed. In the experiments of this section, 10000 grayscale images from BossBase1.01 [1] are converted into JPEG image with quality factor 75, and then stego images are generated by JUNIWARD steganography which is a state-of-the-art hiding method in JPEG domain. The detectors are trained as binary classifiers implemented using ensemble classifier proposed in literature [19].The E_{OOB}(out-of-bag estimate of the testing error) is used to evaluate the detection performance of the proposed steganalysis feature.

(1) Direction parameter θ

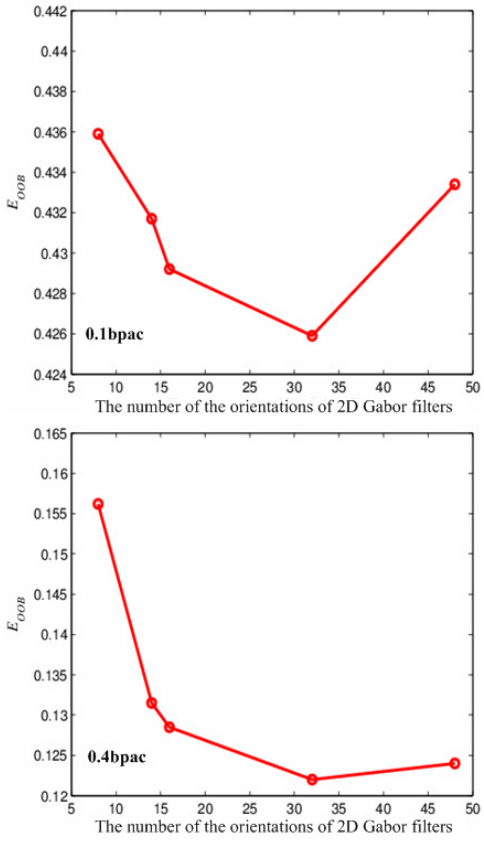

Figure 5: The effect of the number of orientations of 2D Gabor filters on detection accuracy(quality factor (QF) is 75).

By the 2D Gabor filters with different orientations, the changes of the statistical features of JPEG image can be captured more effectively for steganography embedding. In Figure 5, for JUNIWARD with paylaod 0.1bpac and 0.4bpac,the detection performance are presented when the orientations of 2D Gabor filters equal 8, 14, 16, 32, 48 respectively. The other parameters for the detection performance are set as $\sigma=1$, $q=6$, and $T=4$. In addition, in Figure 5, from top to bottom, the payloads of JUNIWARD steganography are 0.1bpac and 0.4bpac respectively.

From Figure 5, it can be seen that the detection accuracy of the proposed steganalysis feature will be better when

more 2D Gabor filters with different orientations are utilized. At the same time, it should be noticed that the feature dimension will be increased with more orientations. The increase in feature dimension will lead to more time and space consumptions; on the other hand, more training samples are need for classifier and this will affect the detection accuracy when the training samples are limited.

(2) Scale parameter σ

With the 2D Gabor filters with different scales, the changes of JPEG image statistical features can be captured from more scales. In Table 1, the detection performances are given for JUNIWARD with quality factor 75 when the scale parameter σ of 2D Gabor filters is set as 0.5, 0.75, 1, and 1.25 respectively. The orientations of 2D Gabor filters equal 32 for each scale and the threshold T equals 4. The Quantization step q is set to 2, 4, 6, 8 respectively for different scales. The detection performances of combinatorial feature which is formed by combining the histogram features got in different scales are also given in Table 1.

Table 1: Detection error E_{OOB} for JUNIWARD by 2D Gabor filters with different scales.

Payload	Scale parameter σ				
	0.5	0.75	1	1.25	combine
0.1bpac	0.4300	0.4277	0.4273	0.4350	**0.4158**
0.2bpac	0.3373	0.3233	0.3259	0.3438	**0.2974**
0.3bpac	0.2376	0.2202	0.2233	0.2504	**0.1847**
0.4bpac	0.1626	0.1326	0.1237	0.1553	**0.0956**

From Table 1, it can be seen that the detection error of the features extracted in different scales are also different and the combinatorial feature can achieve better detection accuracy. This is because that the features extracted from different scales reflect the changes of JPEG image statistical characteristics in different scales after steganography embedding. The combination of these features enhances the diversity and effectiveness of the steganalysis feature, thus the detection accuracy is improved.

(3) Quantization step q

For the steganalysis feature extraction, the quantization for the image filtering coefficients can make the feature more sensitive to embedding changes at spatial discontinuities in the image [8]. Therefore, the image filtering coefficients are quantized to improve the detection accuracy in the proposed feature extraction method. In Figure 6, the effects of the quantization step q on detection accuracy are shown for JUNIWARD steganography at 0.4bpac payload with quality factor 75 and 95. The other parameters of 2D Gabor filters are set as $\sigma = 1$, the orientations equal 32, the threshold $T = 4$.

From Figure 6, it can be seen that the detection accuracy will be changeable with different quantization step q and appropriate value of q can improve the detection accuracy. In addition, for the same quality factor, we notice that the quantization step q for optimal detection accuracy is larger when the scale parameters σ is relative large.

(4) Threshold T

The threshold T is also a parameter for feature extraction. For the histogram feature, the feature dimension will be increased when the larger threshold T is set. In Table 2, the effects of the threshold T on detection accuracy are given for quality factor 75. The scale parameter σ is set as 0.5, 0.75, 1, 1.25 respectively, the corresponding quantization steps q are 2, 4, 6, 8, and the orientations equal 32 for all scales.

Table 2: The effect of the threshold T on detection accuracy.

T	3	4	5	6
E_{OOB}	0.0999	0.0956	0.0996	0.1004

From Table 2, it can be seen that the detection performances are insensitive to threshold T, so the threshold T is set to 4 in this paper.

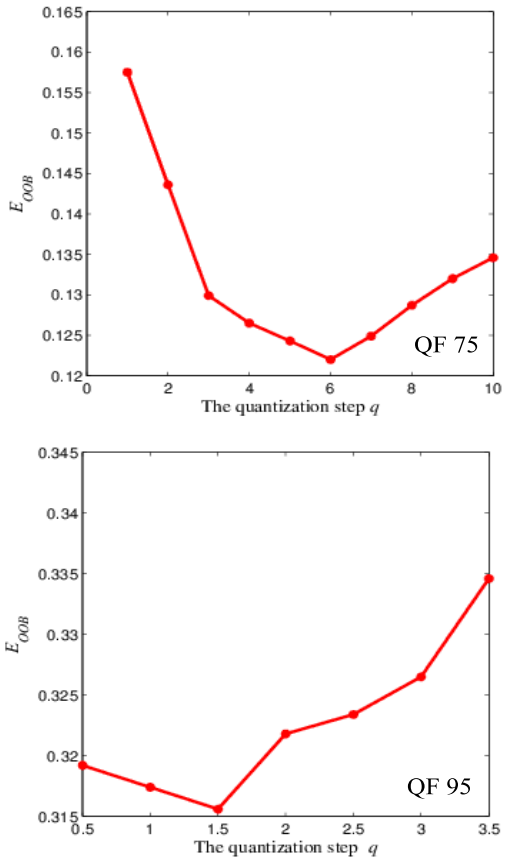

Figure 6: The effect of the quantization step q on detection accuracy.

5. EXPERIMENTAL RESULTS AND ANALYSIS

This section is organized as the following. In section 5.1, the image database and experiment setup are introduced; in section 5.2, by comparing to CC-JRM and DCTR, the proposed feature are evaluated according to the detection performance for adaptive JPEG steganography such as UED, JUNIWARD and SI-UNIWARD; in section 5.3, the detection performances of the combinatorial feature steganalysis are presented and the complement among different features are analyzed.

5.1 Image Database and Experiment setup

In the experiments, the image database is BossBase-1.01 containing 10000 grayscale 512×512 images with PGM format. For UED and JUNIWARD steganography, all these grayscale images are converted into JPEG image with quality factor 75 and 95 respectively, and then the corresponding stego images are generated with payload 0.05, 0.1, 0.2, 0.3, 0.4, 0.5bpac. For SI-UNIWARD steganography, the original grayscale images are used as precover images, and then the corresponding stego images are generated with payload from 0.05bpac to 0.5bpac when the grayscale images are compressed to JPEG image with quality factor 75 and 95. So, for each steganography algorithm and quality factor, we have one group cover images and six group stego images, one group cover images and one group corresponding stego images are used as image samples for one payload.

For the steganalysis feature extraction, the parameters are set as: scales parameter $\sigma = 0.5,\ 0.75,\ 1,\ 1.25$ respectively, the number of the orientations of 2D Gabor filters equal 32 for each scale, the threshold T is set to 4, the quantization step q is set to 2, 4, 6, 8 for different scales with σ in ascending order when quality factor is 75, the q is set to 0.5, 1, 1.5, 2 for quality factor is 95. Lastly, the steganalysis feature is obtained with 17000 dimensions. In all experiments, ensemble classifier [19] is used for the training and testing. The proportion of training set to test set is set to 5:5 and the E_{OOB} is used to evaluate the detection performance of the steganalysis feature. The detection accuracy is the average value of ten duplicate experiments.

5.2 Comparison to prior art

In this section, the proposed steganalysis feature is compared with CC-JRM and DCTR, the detection errors E_{OOB} are given for UED, JUNIWARD and SI-UNIWARD with different payloads.

In Figure 7, the detection errors E_{OOB} of the three different steganalysis features are presented for UED with seven payloads and two quality factors. From Figure 7, it can be seen that the proposed steganalysis feature based on 2D Gabor filters can achieve competitive detection performance by comparing with the other steganalysis features. For example, in contrast to DCTR, the testing error E_{OOB} can be improved by 5.27% when payload is 0.2bpac and quality factor is 75, the improvement is 12.47% by comparison with CC-JRM. When quality factor is 95, the improvement is 3.16% and 6.88% respectively for payload 0.2bpac. This is because that the 2D Gabor filters with different scales and orientations can capture the embedding changes more effectively.

In Figure 8, the detection errors E_{OOB} of the three steganalysis features are given for JUNIWARD with different payloads. From Figure 8, it can be seen that the proposed steganalysis feature can achieve the best detection performance for JUNIWARD with different payloads and quality factors. For example, when quality factor is 75, in contrast to DCTR, the testing error E_{OOB} can be improved by 4.05% when payload is 0.2bpac, the improvement is 12.07% by comparison with CC-JRM; when quality factor is 95, the improvements are 2.44% and 4.78% respectively for payload 0.2bpac.

In Figure9, the detection errors E_{OOB} of the three steganalysis features are given for SI-UNIWARD with different payloads. From Figure 9, it can be seen that the proposed steganalysis feature can achieve the best detection performance when quality factor is 75. For example, when payload is 0.2bpac and quality factor is 75, in contrast to DCTR, the testing error E_{OOB} can be improved by 1.1%, the improvement is 2.1% by comparison with CC-JRM. However, the detection performance of DCTR is more accurate when quality factor is 95.

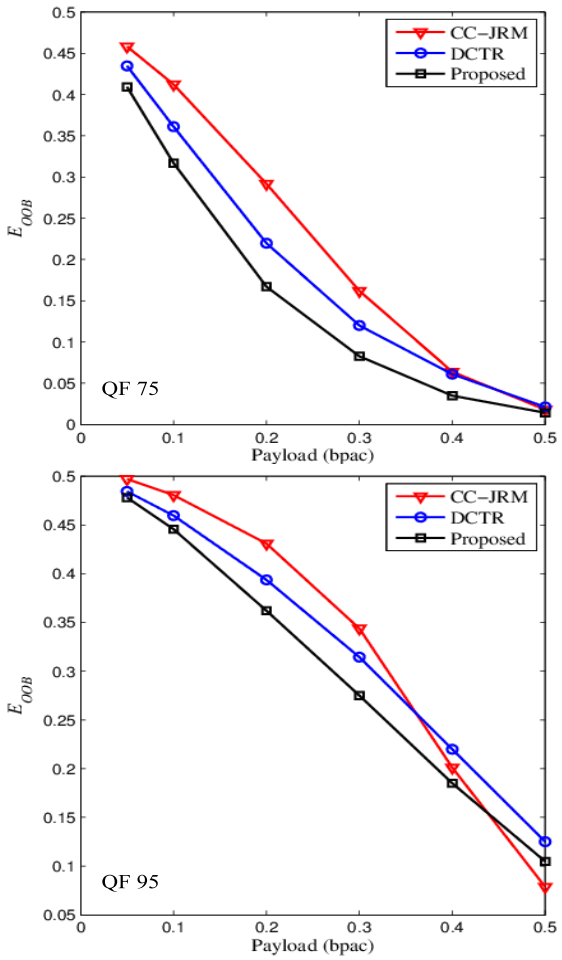

Figure 7: The detection error E_{OOB} for UED for quality factor 75 and 95 when steganalyzed with different features.

Furthermore, it should be noticed that the dimensions of the proposed steganalysis feature can be reduced 8000 when scale parameter σ is still set to four different values and orientations equal 14 for each scale. In this case, the feature dimension of the proposed steganalysis feature is same to DCTR feature. In the experiments, the detection performances of the proposed steganalysis feature with 8000 dimensions are also evaluated. The experimental results show the proposed steganalysis feature with 8000 dimensions can also achieve competitive detection accuracy by comparison with other steganalysis features. For example, for UED and JUNIWARD with payload 0.2bpac, the improvements are 4.2% and 2.14% respectively for JPEG image with quality factor 75 in contrast to DCTR; when quality factor is 95, the improvements are 2.15% and 1.91%. For SI-UNIWARD, the detection performances of the proposed steganalysis feature

with 8000 dimensions are comparative to DCTR when quality factor is 75; however, it is still inferior to DCTR when quality factor is 95.

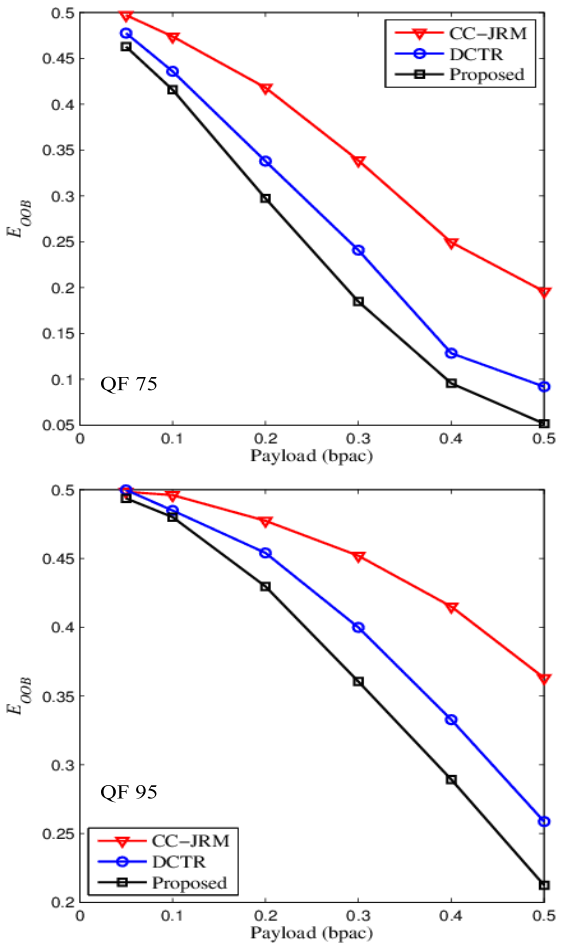

Figure 8: The detection error E_{OOB} for JUNIWARD for quality factor 75 and 95 when steganalyzed with different features.

5.3 Detection accuracy of the combinatorial features

In this section, the detection performances of the combinatorial features which are formed by combining different steganalysis features are compared for JUNIWARD with payload 0.4bpac. In Table 3 and Table 4, the detection error $E_{OOB}$is given for quality factors 75 and 95.

From Table 3 and Table 4, it can be seen that the combination of different steganalysis features only can improve the detection performance slightly. It is possible that the complements among these steganalysis features are not strong.

6. CONCLUSION

In this paper, a steganalysis feature extraction method based on 2D Gabor filters is proposed for adaptive JPEG steganography which often constrains the embedding changes to complicated texture and edge regions. Firstly, the definition and advantages of 2D Gabor filters are introduced. Secondly, the proposed feature extraction method based on 2D

Table 3: The detection error E_{OOB} of the combinatorial features for QF 75.

Proposed	DCTR	CC-JRM	E_{OOB}	Dim.
•			0.0956	17000
	•		0.1284	8000
		•	0.2494	22510
•	•		0.0860	25000
•		•	0.0929	39510
•	•	•	0.0872	47510

Table 4: The detection error E_{OOB} of the combinatorial features for QF 95

Proposed	DCTR	CC-JRM	E_{OOB}	Dim.
•			0.2893	17000
	•		0.3328	8000
		•	0.4151	22510
•	•		0.2739	25000
•		•	0.2846	39510
•	•	•	0.2759	47510

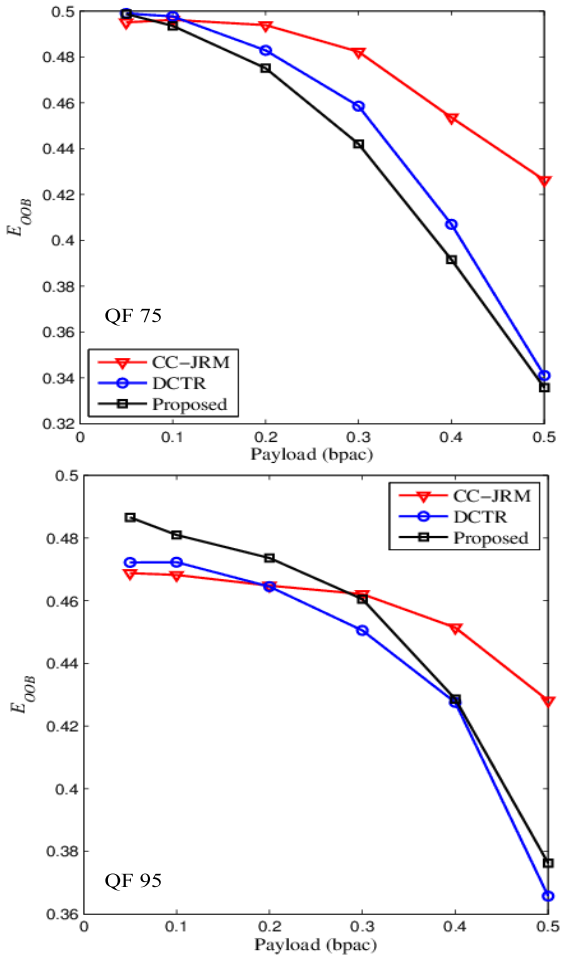

Figure 9: The detection error E_{OOB} for SI-UNIWARD for quality factor 75 and 95 when steganalyzed with different features.

Gabor filters is given in details, and then the parameters setting are also discussed. Lastly, the detection performance of the proposed steganalysis feature is evaluated by comparing with CC-JRM and DCTR features. From the experimental results, it can be seen that the detection performance can be improved effectively. For the proposed steganalysis feature extraction method, the generation of 2D Gabor filters is very important, in the future, the effect of 2D Gabor filter construction for steganalysis of adaptive JPEG steganography will be studied to improve the detection accuracy further. In addition, in literature [15], PHARM feature has been proposed for steganalysis of JPEG steganography and can achieve competitive detection performance. So, we will compare the proposed feature with PHARM feature in the future research.

7. ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (No. 61379151, 61272489, 61302159, and 61401512), the Excellent Youth Foundation of Henan Province of China (No.144100510001), the National Cryptography Development Fund of China (No.MMJJ201301005) and the Foundation of Science and Technology on Information Assurance Laboratory (No.KJ-14-108).

8. REFERENCES

[1] P. Bas, T. Filler, and T. Pevnỳ. ąś break our steganographic systemąś: The ins and outs of organizing boss. In *Information Hiding*, pages 59–70. Springer, 2011.

[2] J. G. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *JOSA A*, 2(7):1160–1169, 1985.

[3] T. Denemark, V. Sedighi, V. Holub, R. Cogranne, and J. Fridrich. Selection-channel-aware rich model for steganalysis of digital images. In *IEEE Workshop on Information Forensic and Security, Atlanta, GA*, 2014.

[4] T. Filler and J. Fridrich. Design of adaptive steganographic schemes for digital images. In *IS&T/SPIE Electronic Imaging*, pages 78800F–78800F. International Society for Optics and Photonics, 2011.

[5] T. Filler, J. Judas, and J. Fridrich. Minimizing additive distortion in steganography using syndrome-trellis codes. *Information Forensics and Security, IEEE Transactions on*, 6(3):920–935, 2011.

[6] J. Fridrich, M. Goljan, P. Lisonek, and D. Soukal. Writing on wet paper. *Signal Processing, IEEE Transactions on*, 53(10):3923–3935, 2005.

[7] J. Fridrich, M. Goljan, and D. Soukal. Perturbed quantization steganography. *Multimedia Systems*, 11(2):98–107, 2005.

[8] J. Fridrich and J. Kodovsky. Rich models for steganalysis of digital images. *Information Forensics and Security, IEEE Transactions on*, 7(3):868–882, 2012.

[9] J. Fridrich, T. Pevnỳ, and J. Kodovskỳ. Statistically undetectable jpeg steganography: dead ends challenges, and opportunities. In *Proceedings of the 9th workshop on Multimedia & security*, pages 3–14. ACM, 2007.

[10] S. E. Grigorescu, N. Petkov, and P. Kruizinga. Comparison of texture features based on gabor filters. *Image Processing, IEEE Transactions on*, 11(10):1160–1167, 2002.

[11] L. Guo, J. Ni, and Y.-Q. Shi. An efficient jpeg steganographic scheme using uniform embedding. In *WIFS*, pages 169–174, 2012.

[12] V. Holub and J. Fridrich. Digital image steganography using universal distortion. In *Proceedings of the first ACM workshop on Information hiding and multimedia security*, pages 59–68. ACM, 2013.

[13] V. Holub and J. Fridrich. Random projections of residuals for digital image steganalysis. *Information Forensics and Security, IEEE Transactions on*, 8(12):1996–2006, 2013.

[14] V. Holub and J. Fridrich. Low complexity features for jpeg steganalysis using undecimated dct. *Information Forensics and Security, IEEE Transactions on*, 10(2):219–228, 2015.

[15] V. Holub and J. Fridrich. Phase-aware projection model for steganalysis of jpeg images. *Proc. SPIE, Electronic Imaging, Media Watermarking, Security, and Forensics XVII, to appear, San Francisco, CA*, 2015.

[16] Y. Kim, Z. Duric, and D. Richards. Modified matrix encoding technique for minimal distortion steganography. In *Information hiding*, pages 314–327. Springer, 2007.

[17] J. Kodovskỳ and J. Fridrich. Steganalysis of jpeg images using rich models. In *IS&T/SPIE Electronic Imaging*, pages 83030A–83030A. International Society for Optics and Photonics, 2012.

[18] J. Kodovsky, J. Fridrich, and V. Holub. On dangers of overtraining steganography to incomplete cover model. In *Proceedings of the thirteenth ACM multimedia workshop on Multimedia and security*, pages 69–76. ACM, 2011.

[19] J. Kodovsky, J. Fridrich, and V. Holub. Ensemble classifiers for steganalysis of digital media. *Information Forensics and Security, IEEE Transactions on*, 7(2):432–444, 2012.

[20] T. Pevny and J. Fridrich. Merging markov and dct features for multi-class jpeg steganalysis. In *Electronic Imaging 2007*, pages 650503–650503. International Society for Optics and Photonics, 2007.

[21] N. Provos. Defending against statistical steganalysis. In *Usenix Security Symposium*, volume 10, pages 323–336, 2001.

[22] P. Sallee. Model-based steganography. In *Digital watermarking*, pages 154–167. Springer, 2004.

[23] X. Song, F. Liu, X. Luo, J. Lu, and Y. Zhang. Steganalysis of perturbed quantization steganography based on the enhanced histogram features. *Multimedia Tools and Applications*, pages 1–27, 2014.

[24] C. Wang and J. Ni. An efficient jpeg steganographic scheme based on the block entropy of dct coefficients. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 1785–1788. IEEE, 2012.

[25] A. Westfeld. F5ąła steganographic algorithm. In *Information hiding*, pages 289–302. Springer, 2001.