

# The Game of Countering JPEG Anti-forensics Based on the Noise Level Estimation

Yunwen Jiang<sup>1</sup>, Hui Zeng<sup>1</sup>, Xiangui Kang<sup>1,2</sup>, Li Liu<sup>3</sup>

<sup>1</sup>School of Information Science and Technology, Sun Yat-sen University, Guangzhou, GD, China, 510006

<sup>2</sup>State Key Laboratory of Information Security (Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093), E-mail: isskxg@mail.sysu.edu.cn

<sup>3</sup>Marvell Semiconductor Inc., Santa Clara 95054, USA, liuli1@gmail.com

**Abstract**— It's well known that JPEG image compression can result in quantization artifacts and blocking artifacts. There are plenty of forensic techniques making use of image's compression fingerprints to verify digital images. However, when a forger exists, these methods are not reliable any more. One typical anti-forensic method is adding anti-forensic dither to DCT transform coefficients and erasing blocking artifacts to remove compression history. In this paper, we propose a new countering anti-forensic method based on estimating the noise added in the process of erasing blocking artifacts. The experimental results show that our method obtains an average detection accuracy of 98% on the UCID image database. Another advantage of our proposed method is that it has only one-dimensional feature and time-saving. Furthermore, we use the game theory to evaluate the performance of both sides, and identify the optimal strategies of both sides.

## I. INTRODUCTION

Nowadays, image editing softwares allow users to easily tamper an image without leaving any visual witness, which makes image forensics become a hot topic. JPEG compression detection is an important part of digital image forensics. By deciding whether an image is previously compressed, researchers can make a preliminary judgment of the authenticity of digital images, which can be used as the auxiliary decision-making of digital image forensics.

JPEG compression can be divided into four steps: color mode conversion and sampling, DCT transform, quantization, and entropy coding. JPEG compression starts by segmenting an input image into several nonoverlapping  $8 \times 8$  pixel blocks, then it uses the 2-D DCT to transform each block data into 64 DCT coefficients. In quantization step, each coefficient value is quantized by a parameter  $Q_{ij}$ . This procedure results in DCT coefficient quantization fingerprint and the tampering artifacts.

There are many works which aim at detecting the artifacts brought by JPEG compression [1 - 5] for digital image forensics. However, an anti-forensic method [6] can deceive detectors by first adding anti-forensic dither to DCT transform

coefficients to imitate the original uncompressed histograms and then erasing blocking artifacts via boundary blurring to remove the compression history.

To counter above anti-forensic method, in [7], the authors proposed a countering JPEG anti-forensic technique, which identifies image noisiness by re-compressing the forged image at different quality factors. [8] used the transition probability matrix of DCT coefficients to measure the possible modifications and identify the forged images from those original ones. [9] summarized several techniques in the cat-and-mouse game between digital image forensics and counter-forensics related to an image's JPEG compression history.

As only adding the anti-forensic dither cannot remove the blocking artifact, in [6], to erase the compression history, the forger must remove blocking artifact by first median filtering an image and then adding low-power white Gaussian noise to each of its pixel values. Both the window size of the median filter and the variance of the noise can affect the deblocking effect. Moreover, for a heavily compressed image, a larger median filter window size and greater noise variance to remove statistical traces of blocking artifacts are required.

In this paper, in order to detect the forged images, we employ noise level estimation [10] to estimate the noise added in the deblocking process. For a suspect image, we estimate the noise level of the image and compare it with a threshold to determine whether it is forged. Moreover, we imitate the interplay between forensics and anti-forensics as a extensive form game and evaluate the final results with the game theory.

The rest of the paper is organized as follows. Section II gives a brief review of the JPEG anti-forensic technique. Section III describes the process of applying the noise level estimation to counter anti-JPEG compression forensic and how to use game theory to evaluate this gaming process. Section IV shows the experimental results on the UCID image database [11]. The conclusion is made in Section V.

## II. ANTI-FORENSICS OF JPEG COMPRESSION

The segmentation of an input image into several nonoverlapping  $8 \times 8$  pixel block in JPEG compression result in pixel domain discontinuity which is called blocking artifact, especially in the boundaries of these segments. The following quantization and subband coding steps further introduce the quantization artifact. Fig. 1 shows the process how a digital

---

This work was supported by NSFC (Grant nos. 61070167, 61379155, U1135001), 973 Program (Grant no. 2011CB302204), the Research Fund for the Doctoral Program of Higher Education of China (Grant no. 20110171110042), National Science & Technology Pillar Program (Grant no. 2012BAK16B06) and NSF of Guangdong province (Grant no. s2013020012788).

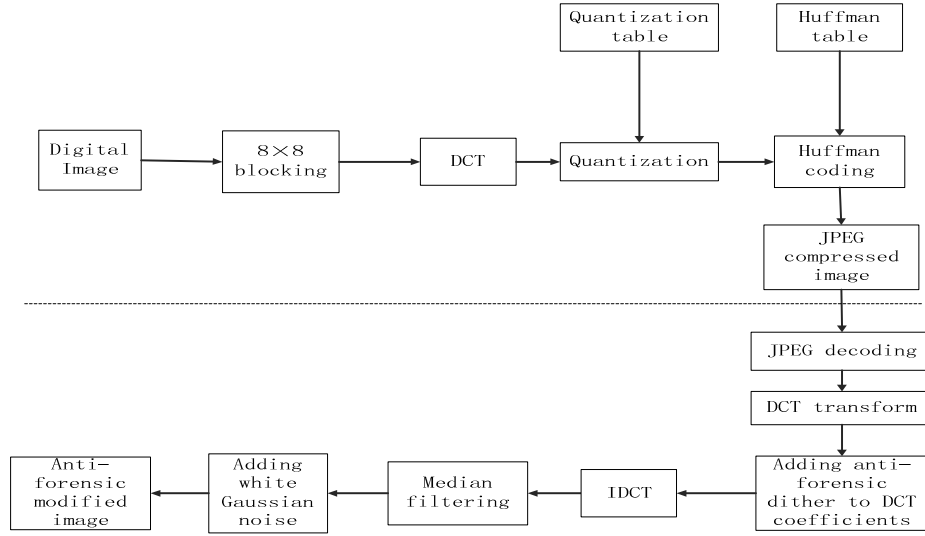


Fig. 1 JPEG compression process and its anti-forensic modified process.

TABLE I  
BLOCKING ARTIFACT DETECTION ACCURACY

Quality factor	JPEG compressed image	Anti-forensic modified image		
		$s=3, \sigma=3$	$s=3, \sigma=2$	$s=2, \sigma=2$
95	93.05%	51.38%	50.00%	50.04%
75	99.48%	51.64%	50.34%	51.20%
55	99.55%	52.95%	51.01%	56.61%

image is JPEG compressed and how it is anti-forensically modified using method [6]. The blocks above the dotted line are the JPEG compression process, and blocks below the dotted line shows the anti-forensic process.

For a JPEG compressed image, in order to remove the quantization artifact, i.e. make the distribution of its subband coefficient value match an original one, [6] added the anti-forensic dither to the DCT coefficient. The anti-forensic dither is described as follows.

The distribution of coefficient values within a particular AC DCT subband can be modeled as the Laplace distribution [12 - 14]. Let  $X$  be the DCT coefficient at the block position  $(i, j)$  for a uncompressed image,

$$P(X = x) = \frac{\lambda}{2} e^{-\lambda|x|} \quad (1)$$

where  $\lambda$  is a Laplacian parameter. According to this model and quantization rule, for a previously JPEG compressed image, AC coefficients of each subband are distributed as discrete Laplacian distribution. For the coefficients at the  $(i, j)$ -th position, if we use the quantization step  $Q_{i,j}$ , the distribution would be

$$P(Y = y) = \begin{cases} 1 - e^{-\frac{\lambda Q_{i,j}}{2}}, & \text{if } y=0 \\ e^{-\lambda|y|} \sinh(\frac{\lambda Q_{i,j}}{2}), & \text{if } y=kQ_{i,j} \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where  $Y = Q_{i,j} \text{round}(X / Q_{i,j})$ , and the parameter  $\lambda$  can be generated by maximum likelihood estimation. Let  $N = N_0 + N_1$  be the total number of observations of the current DCT subband,  $N_0$  represents the number of coefficients taking zero values, and  $N_1$  represents the number of nonzero coefficients, and  $S = \sum_{k=1}^N |y_k|$ . The model parameter  $\hat{\lambda}$  can be calculated using the equation

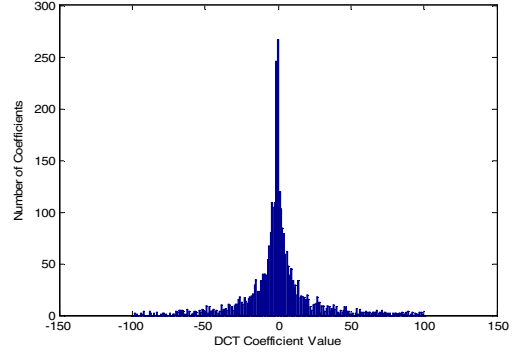
$$\hat{\lambda} = -\frac{2}{Q_{i,j}} \ln(\gamma) \quad (3)$$

where  $\gamma$  is defined as

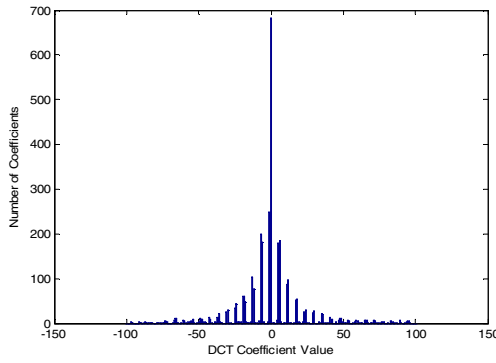
$$\gamma = \frac{-N_0 Q_{i,j}}{2N Q_{i,j} + 4S} + \frac{\sqrt{N_0^2 Q_{i,j}^2 - (2N_1 Q_{i,j} - 4S)(2N Q_{i,j} + 4S)}}{2N Q_{i,j} + 4S} \quad (4)$$



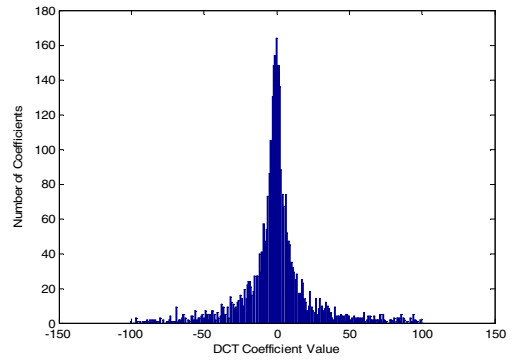
(a)



(b)



(c)



(d)

Fig. 2 An uncompressed image (a). (b) Histogram of coefficient values from the (2,2) DCT subband taken from (a), (c) the same image after JPEG compression, and (d) an anti-forensically modified copy of the JPEG compressed image.

To hide the compression evidence, [6] introduced dither into the AC coefficients to approximately restore the histogram of each subband, by

$$Z = Y + D \quad (5)$$

where  $Z$  is the anti-forensically modified coefficient and  $D$  is the additive dither. For the coefficient  $Y$  of zeros value at the  $(i, j)$ -th position, the noise distribution is given by

$$P(D = d | Y = 0) = \begin{cases} \frac{1}{c_0} e^{-\hat{\lambda}|d|}, & \text{if } \frac{-Q_{i,j}}{2} \leq d \leq \frac{Q_{i,j}}{2} \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

where  $c_0 = (2/\hat{\lambda})(1 - e^{-\hat{\lambda}Q_{i,j}/2})$ . The distribution of the anti-forensic dither added to nonzero quantized DCT coefficients is

$$P(D = d | Y = q_k) = \begin{cases} \frac{1}{c_1} e^{-\text{sgn}(q_k)\hat{\lambda}(d + \text{sgn}(q_k)(Q_{i,j}/2))}, & \text{if } \frac{-Q_{i,j}}{2} \leq d \leq \frac{Q_{i,j}}{2} \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

where  $c_1 = (1/\hat{\lambda})(1 - e^{-\hat{\lambda}Q_{i,j}})$ .

Fig. 2 shows the histogram of coefficient values in the (2,2) DCT subband in an uncompressed image along with the corresponding coefficient value histograms from the JPEG compressed and anti-forensically modified image. From Fig. 2, we can see that for an image which is JPEG compressed but anti-forensically modified, its distribution of AC coefficient values also obeys the Laplace distribution. It is shown that the quantization fingerprint is removed via adding anti-forensic dither. The experimental results [6] show that no evidence of quantization [4] is presented after adding the forensics dither.

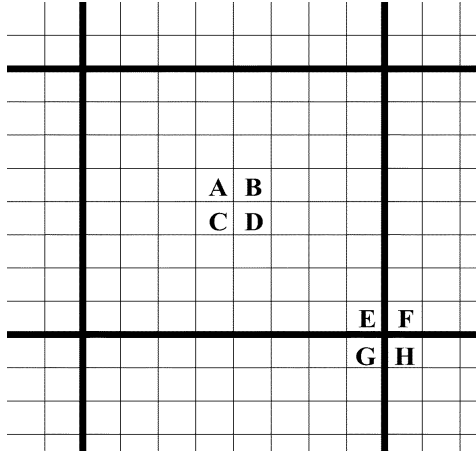


Fig. 3 For each block the numbers  $Z'(i, j) = |A + D - B - C|$  and  $Z''(i, j) = |E + H - F - G|$  are computed, i.e., involving same pixel pattern but spanning or not multiple blocks.

Even though adding anti-forensic dither can remove the evidence of quantization, it cannot remove blocking artifacts. Difference within a block and spanning across a block boundary are illustrated in Fig. 3 [4]. For each block  $(i, j)$ , we compute

$$Z'(i, j) = |A + D - B - C| \quad Z''(i, j) = |E + H - F - G| \quad (8)$$

where  $A$  through  $H$  are the values of the pixels in the positions depicted in Fig. 3. Then blocking artifacts are detected by examining the difference between two histograms:

$$K = \sum_n |h_1(n) - h_2(n)| \quad (9)$$

where  $h_1$  is the histogram of  $Z'$  each image block, and  $h_2$  is the histogram of  $Z''$ . Fig. 4(a) and Fig. 4(b) show the two histograms for an uncompressed image and an JPEG compressed image. It is observed that the difference ( $K$ ) between  $h_1$  and  $h_2$  is very small for an uncompressed image, whereas  $K$  is remarkable large for a JPEG compressed image. Hence, the statistic  $K$  can be used to determine whether an image is JPEG compressed. That is, if  $K$  is greater than the threshold  $t_1$ , then the image is classified as a JPEG compressed one. Moreover, as described in [6],  $h_1(1) > h_1(0)$  and  $h_2(1) > h_2(0)$  shall present in an uncompressed image. Whereas in a JPEG compressed image,  $h_1(1) > h_1(0)$  and  $h_2(1) > h_2(0)$  may not meet or not meet at the same time.

To remove blocking artifacts, the authors in [6] remove the difference between  $h_1$  and  $h_2$  by median filtering (as shown in Fig. 4(c)). However, median filtering can not remove the blocking artifacts completely because of  $h_1(1) < h_1(0)$  and  $h_2(1) < h_2(0)$ . Hence, a zero mean Gaussian random noise is added after median filtering. Fig. 4(d) shows that adding noise after median filtering can achieve the deblocking purpose.

Hence, the deblocking algorithm can be summarized as follows:

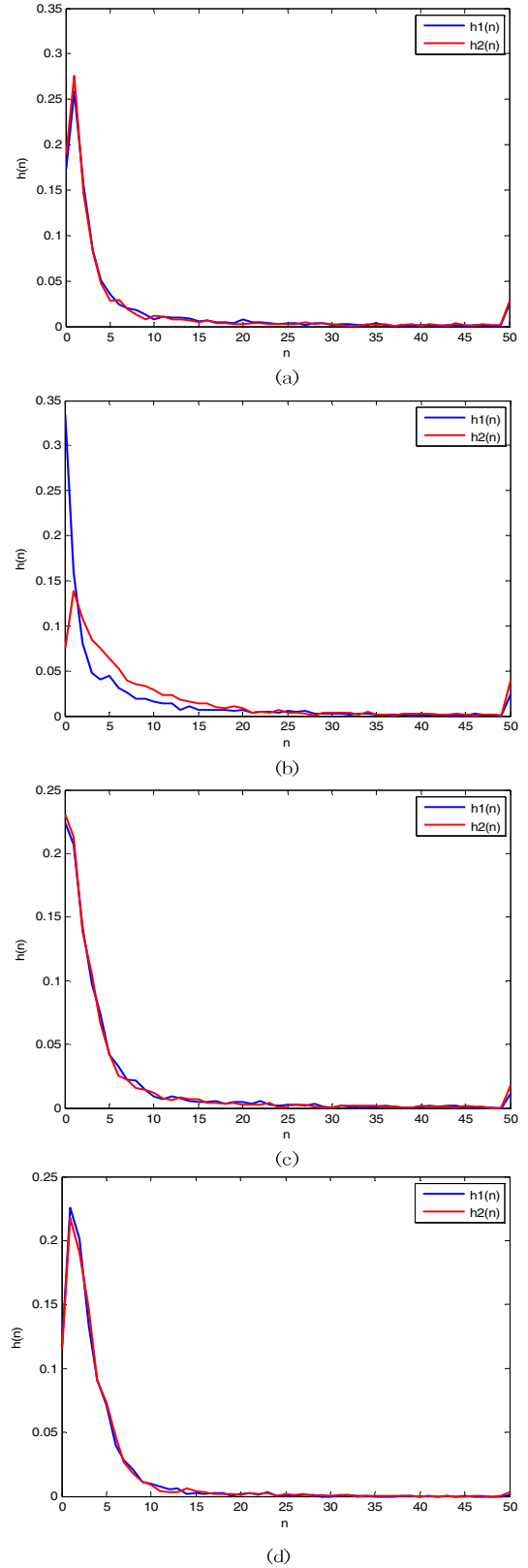


Fig. 4  $h_1$  and  $h_2$  obtained from (a) an uncompressed image, (b) the same image after JPEG compression using a quality factor of 75, (c) deblocking with median filtering, (d) deblocking with median filtering and adding Gaussian noise.

$$v_{i,j} = \text{med}_s(u_{i,j}) + n_{i,j} \quad (10)$$

where  $u_{i,j}$  represents the pixel value at location  $(i, j)$  in an unmodified image, and  $v_{i,j}$  denotes its deblocked counterpart,  $\text{med}_s(\cdot)$  denotes a two-dimensional median filter with a square window of size  $s$  pixels, and  $n_{i,j}$  is a zero mean Gaussian random noise with variance  $\sigma^2$ . By adjusting the window size  $s$  and variance  $\sigma^2$ , better anti-forensic results can be achieved. Table I shows the results of blocking artifact detection accuracy from our experiments. For FPR (False positive rates) varies in a given interval, we choose the optimal threshold to obtain the Accuracy Rate = (TPR + TNR) / 2. TPR is true positive rates and TNR is true negative rates. The results show that the above method can remove the blocking artifact effectively.

### III. COUNTERING TECHNIQUE BASED ON THE NOISE LEVEL ESTIMATION

#### A. Image Noise Estimation

Recently, several methods related to noise level estimation have been proposed. [15] used the changes in kurtosis values of images to estimate the noise level. In [10], the authors estimated noise level from the selected weak textured patches using PCA. Fig. 5 shows the two methods' performance on the same image Lena. It can be observed that the method in [10] is more accurate than the method in [15] in most cases.

In this paper, we choose the method from [10] to estimate the noise level.

In [10], the noise level estimation can be summarized as the following major steps:

1. Decomposing the test image into overlapping patches. The default patch size is  $7 \times 7$  pixels.
2. Estimating an initial noise level  $\sigma_e$  from the covariance matrix as (11).

$$\sigma_e^2 = \lambda_{\min}(\sum_{y'} \cdot) \quad (11)$$

where  $\sum_{y'}$  is the covariance matrix of the selected patches and  $\lambda_{\min}(\sum_{y'})$  is the minimum eigenvalue of  $\sum_{y'}$ .

3. Selecting the weak textured patches from the test image using a threshold that varies with  $\sigma_e$ .

4. Estimating a new noise level  $\sigma_e$  using the selected patches. The process of step 3 and 4 is iterated until  $\sigma_e$  is stable.

In this paper, we use  $\sigma_e$  to denote the estimated noise level of a test image.

#### B. Countering Technique

With the algorithm adopted from [10], we show how to apply the noise level estimation method to counter the anti-forensic method [6]. As described in section II, in order to forge an image, the forger must add the Gaussian noise to erase the blocking artifacts. That is, forged images may

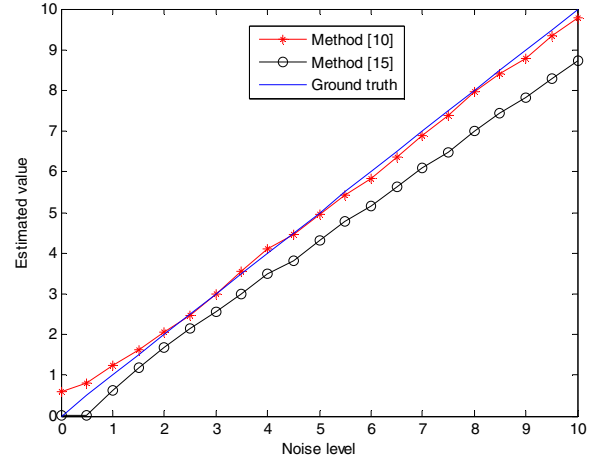


Fig. 5 The performance of noise level estimation.

present a higher noise level than original images. Therefore, estimating the noise level of the image can be an effective way of detecting forgery.

Above assumption is proved by the experimental results showed in the Fig. 6(a), which is in line with our conjecture. The blue “+” represents the original image of TIFF format, and green “\*” represents the images after anti-forensics ( $s = 2$ ,  $\sigma = 2$ ) [6]. It can be seen that most of the original images' noise level are lower than 0.5. As the noise level is very sensitive to JPEG compression, when images are JPEG compressed, the noise level decreases to nearly zero. The forged images are first JPEG compressed, and then added low-power white Gaussian noise. Hence, the estimated noise level is near to  $\sigma$ . Furthermore, since noise level estimation is related to the texture and size of an image, complex textured images can have a higher noise level. The larger size image can have more accurate detection rate than that of small size images. Fig. 6(b) shows the noise level of another group images which are in TIFF format ( $1024 \times 768$ ). It is observed that the noise level estimation is more accurate when the image size is larger.

#### C. Game Theoretic Evaluation

In this section, we apply the game theory to analyze the interplay between forensics and anti-forensics. To avoid the analysis of the game being too overwhelming, we assume that the forger adopts the anti-forensic method proposed in [6] and can only adjust the added noise variance  $\sigma$ . We also assume that the investigator adopts our proposed countermeasure.

As described in Section II, to cover the tampering trace and not introduce new artifacts, the forger will choose reasonable variance of the noise.

For the investigator side, the reasonable strategy is to perform blocking detection first and then detect the noise level.

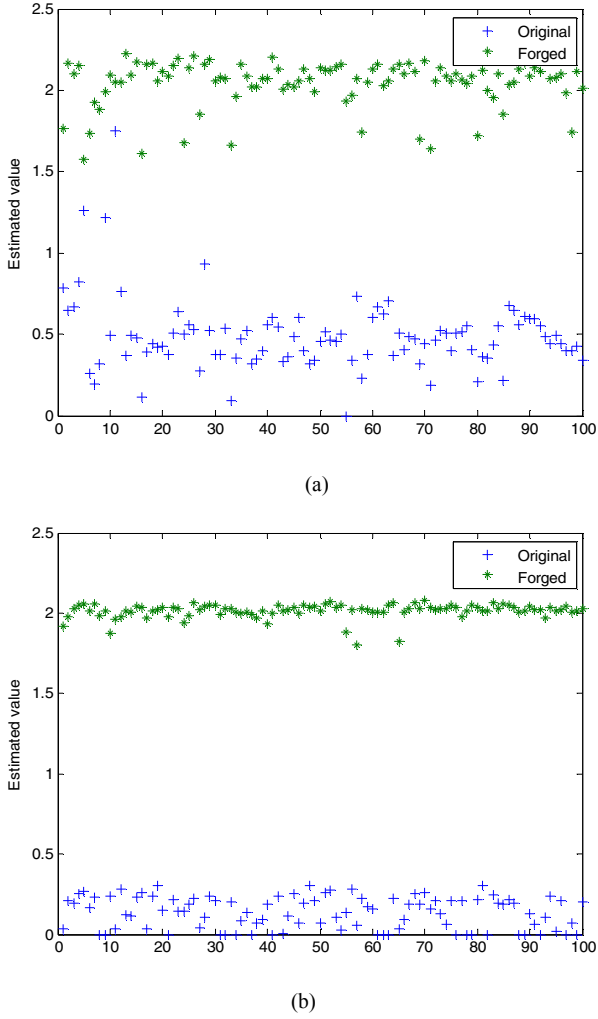


Fig. 6 (a): the noise level estimation in which images come from UCID with the resolution of  $512 \times 384$ .  
(b): the noise level estimation in images with the resolution of  $1024 \times 768$ .

1) Blocking detection. The forger perform blocking detection  $\delta_b$  based on the statistics  $K$ ,  $h_1(1) - h_1(0)$  and  $h_2(1) - h_2(0)$ . For an uncompressed image,  $K$  is relatively small and  $h_1(1) > h_1(0)$  and  $h_2(1) > h_2(0)$ . For a JPEG compressed image,  $K$  is larger. Moreover,  $h_1(1) > h_1(0)$  and  $h_2(1) > h_2(0)$  may not meet or not meet at the same time. Hence for an test image  $I$ , the blocking detection in the first stage can be expressed as a hypothesis testing problem.

$\delta_b(I) = H_0^{(1)} : I$  is an uncompressed image, i.e.  $I$  is original.

$\delta_b(I) = H_1^{(1)} : I$  is a JPEG compressed image.

where upper index "(1)" denotes the first stage. The acceptance region of  $H_0^{(1)}$  is:

$$I: K < t_1 \text{ and } h_1(1) > h_1(0) \text{ and } h_2(1) > h_2(0) \quad (12)$$

where  $t_1$  is chosen according to the false alarm rate  $P_{fa}^{(1)}$ .  $P_{fa}^{(1)}$  is defined as:

$$P_{fa}^{(1)} = P(\delta_b(I) = H_1^{(1)} | I \text{ is uncompressed}) \quad (13)$$

2) Noise level detection. The investigator estimates the noise level of the images that accepted in  $H_0^{(1)}$ . As described in Section III.B, if  $I$  is forged, the noise level  $\sigma_e$  is higher with respect to the case that  $I$  is originally uncompressed. Thus, the noise level detection can be expressed as a hypothesis testing problem.

$H_0^{(2)} : I$  is an originally uncompressed image, i.e.  $I$  is original.

$H_1^{(2)} : I$  is a forged image. In other words,  $I$  is JPEG compressed and anti-forensically modified, i.e.  $I$  is forged. where upper index "(2)" denotes the second stage. The acceptance region of  $H_0^{(2)}$  is:

$$I: \sigma_e(I) < t_2 \quad (14)$$

where  $t_2$  is chosen according to the false alarm rate  $P_{fa}^{(2)}$ .  $P_{fa}^{(2)}$  is defined as:

$$P_{fa}^{(2)} = P(\sigma_e(I) > t_2 | I \text{ is original}) \quad (15)$$

The total detection rate of the blocking detection and the noise level detection can be denoted as:

$$P_d = P((\delta_b(I) = H_1^{(1)} \cup \sigma_e(I) > t_2) | I \text{ is forged}) \quad (16)$$

and the total probability of false alarm rate is defined as:

$$P_{fa} = P_{fa}^{(1)} + P_{fa}^{(2)} \quad (17)$$

For a given  $P_{fa}$ , the investigator would make a tradeoff between  $P_{fa}^{(1)}$  and  $P_{fa}^{(2)}$  to seek an optimal  $P_d$ . Obviously, the optimal  $P_{fa}^{(1)}$  would vary with  $\sigma$ . From the forger's side, there is a tradeoff in choosing the added noise variance  $\sigma$ . The optimal  $\sigma$  would vary with  $P_{fa}^{(1)}$ .

Based on analysis above, we model this interplay between the investigator and the forger as a zero-sum game. The set of strategies that the investigator can use is  $P_{fa}^{(1)} \in [0, P_{fa}]$ , and the set of strategies that the forger can use is  $\sigma \in [0, \sigma_m]$ , where  $\sigma = \sigma_m$  corresponds to a too strong noise which would introduce significant visual distortions. The utility of the investigator is denoted as:

$$U_1(P_{fa}^{(1)}, \sigma) = P_d(P_{fa}^{(1)}, \sigma) \quad (18)$$

and the utility of the forger is denoted as:

$$U_2(P_{fa}^{(1)}, \sigma) = -P_d(P_{fa}^{(1)}, \sigma) \quad (19)$$

#### IV. EXPERIMENTAL RESULTS

##### A. The Performance of Counter Anti-forensics

In order to test the algorithm described in Section III, we performed experiments on 1338 images from the Uncompressed Color Image Database (UCID) [11]. All the images are converted to 8-bit gray-scale images. The 1338 uncompressed images are first JPEG compressed at a given quality factor. Then anti-forensic method [6] is applied to remove DCT coefficient quantization fingerprints from the JPEG compressed images by adding anti-forensic dither to the DCT coefficients. Afterwards, each image is deblocked with the algorithm proposed in [6]. During the deblocking process, the window size of the median filter  $s$  and standard  $\sigma$  variance of the noise are chosen according to [6]. Finally, the noise levels of the forged images and the original images are estimated using the method described in Section III. After computing all the images' noise level, we choose a threshold to classify the two group images. If an image's noise level is lower than the threshold, this image is classified as an original one. Otherwise it will be regarded as a forged image. Fig. 7 is the ROC (receiver operating characteristic) curves for different quality factors of 55, 75 and 95 respectively, with parameters ( $s = 3, \sigma = 3$ ). Fig. 7 shows that the TPR increases with the quality factor. Table II shows the detection accuracy with different quality factors and different parameters. From Table II, it can be observed that our proposed countering anti-forensic method can achieve an average detection accuracy of 98%. Furthermore, as described in [6], lower quality factor needs larger  $\sigma$  to remove the blocking artifacts, which would be detected by the proposed method more easily.

The performance between our proposed method and other state-of-the-art methods are compared and the results are shown in Table III. All results are obtained with the median filter window size  $s = 3$  and noise standard variance  $\sigma = 2$ . Here,  $D$  represents the dimension and  $QF$  represents the quality factor. From Table III, it can be observed that the detection accuracy of the method [7] is less than 80%. While the detection accuracy of the method [8] and our proposed method are over 99%. The advantage of our proposed method is that only one-dimensional feature set is used and it is of low computational complexity, whereas detection method in [8] based on a feature set of 100-dimensional that is much more time-consuming.

##### B. The Nash Equilibrium Performance

As described in Section III, when applying the game theory to analyze the process, from the experimental results, it can be observed that if the attacker knows the existence of our method, their optimal strategy is adding less noise after median filtering. For the forger side, the problem is choosing the reasonable strength of the noise which helps them to avoid the detection of blocking artifact and leave less other detectable traces. This restricts the forger in the process of anti-forensics. For the investigator side, the strategy is how to allocate the FPR between the blocking detection and the noise level detection.

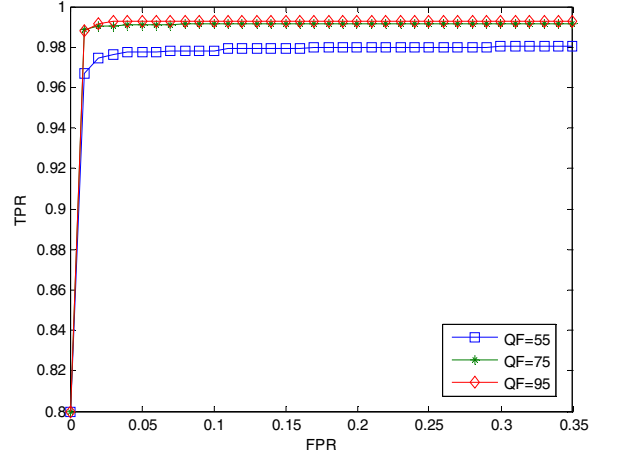


Fig. 7 ROC curves with proposed method for different quality factor

TABLE II  
DETECTION ACCURACY WITH THE PROPOSED METHOD.

Quality factor	$s=3, \sigma=3$	$s=3, \sigma=2$	$s=2, \sigma=2$
95	99.48%	99.14%	98.95%
85	99.48%	99.29%	97.05%
75	99.48%	99.33%	98.99%
65	99.48%	99.36%	98.77%
55	99.48%	99.36%	97.91%

TABLE III  
DETECTION ACCURACY

	$D$	$QF=95$	$QF=85$	$QF=75$	$QF=65$
Method[7]	1	71.26%	75.90%	79.79%	79.94%
Method[8]	100	99.40%	99.61%	99.70%	99.82%
Proposed	1	99.14%	99.29%	99.33%	99.36%

We find the Nash equilibrium strategies by solving the following equation [16]:

$$(P_{fa}^{(1)*}, \sigma^*) = \arg \max_{P_{fa}^{(1)}} \min_{\sigma} U_1(P_{fa}^{(1)}, \sigma) \quad (20)$$

Fig. 8 shows the performance when the total probability of false alarm rate  $P_{fa}$  is 0.2 and JPEG quality factor is 50. The x-axis represents the false alarm rate  $P_{fa}^{(1)}$  that is used in the blocking detection, and the y-axis is the noise standard variance  $\sigma$ . The Nash equilibrium  $(P_{fa}^{(1)*}, \sigma^*)$  is (0.02, 0.4) and the total detection rate  $P_d$  is 86.7% under this setting.  $P_{fa}^{(1)} = 0.02$  denotes that the optimal strategy of the investigator is choosing  $P_{fa}^{(1)} = 0.02$ , and  $\sigma^* = 0.4$  means that choosing  $\sigma = 0.4$  can achieve the purpose for the forger. In this experiment, we found that the optimal noise strength ranges between  $\sigma = 0.4$  and  $\sigma = 0.8$ . Fig. 9 shows the NEROC (Nash equilibrium



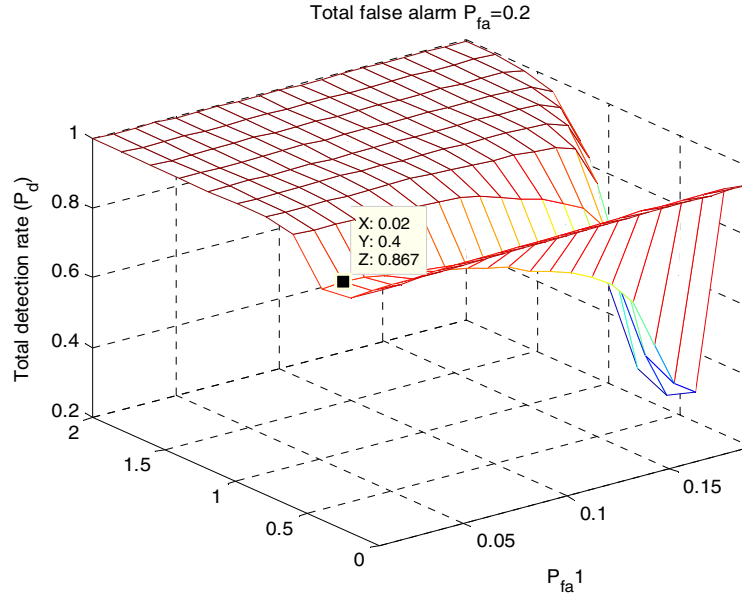


Fig. 8 The detection performance when  $P_{fa} = 0.2$

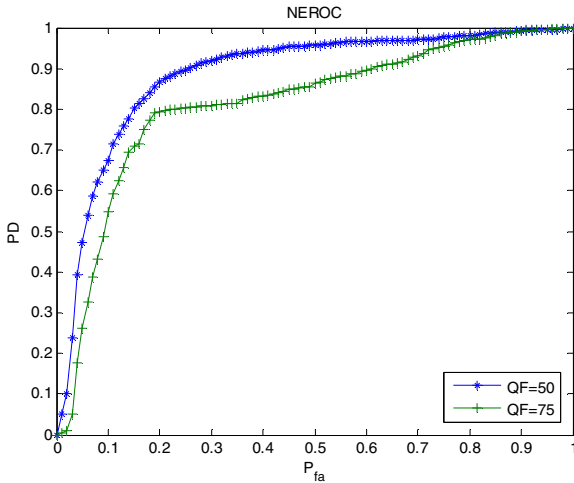


Fig. 9 NERO C curves for different quality factors and with  $s=3$

ROC) [16] for different quality factors. It can be observed that even though the forger can choose the optimal noise strength, the performance of our proposed method is showed in Fig. 9.

## V. CONCLUSIONS

Though anti-forensic technique can remove JPEG compression trace, it also introduces other detectable artifacts. In this paper, we proposed a new countering anti-JPEG compression method based on noise level estimation. The

experimental results show that our method is effective over a range of quality factors. Another advantage of the proposed method is that it only uses one-dimensional feature set thus is time-saving. Moreover, we apply the game theory to analyze the forensics and anti-forensic process and achieve the NERO C curves for different quality factors.

In this paper, we limit our analysis to the interplay between a specific anti-forensics and its countermeasures. We assume that the forger adopts the anti-forensic method proposed in [6] and the forensic investigator adopts our proposed method. However, the game theory model would be suitable for analyzing similar interplays between forensics and anti-forensics.

## REFERENCES

- [1] J. Lukáš, J. Fridrich, "Estimation of primary quantization matrix in double compressed JPEG images," in *Proc. of SPIE 6819, Security, Forensics, Steganography, and Watermarking of Multimedia Contents X*, pp. 681911, February 26, 2008.
- [2] T. Pevny, J. Fridrich, "Detection of double-compression in JPEG images for applications in steganography," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 2, pp. 247-258, June 2008.
- [3] W. Luo, Z. Qu, J. Huang, G. Qiu, "A novel method for detecting cropped and recompressed image block," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing 2007 (ICASSP 2007)*, pp. II-217-II-220.
- [4] Z. Fan, R. L. de Queiroz "Identification of bitmap compression history: JPEG detection and quantizer estimation," *IEEE Transactions on Image Processing*, vol.12, no. 2, pp. 230-235, Feb. 2003.



- [5] D. Fu, Y. Q. Shi, W. Su, "A generalized Benford's law for JPEG coefficients and its applications in image forensics," In *Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6505, Security, Steganography, and Watermarking of Multimedia Contents*, pp. 65051L-1~65051L-11, Jan. 2007, San Jose, CA, USA.
- [6] M. C. Stamm, K. J. R. Liu, "Anti-forensics of digital image compression," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 1050-1065, Sept. 2011.
- [7] G. Valenzise, V. Nobile, M. Tagliasacchi, et al. "Countering JPEG anti-forensics," in *Proc. 18th. IEEE International Conference on Image Processing 2011 (ICIP2011)*, pp. 1949-1952, IEEE, 2011.
- [8] H. Li, W. Luo, J. Huang, "Countering anti-JPEG compression forensics," in *Proc. 19th. IEEE International Conference on Image Processing 2012 (ICIP2012)*, pp. 241-244, IEEE, 2012.
- [9] S. Y. Lai, R. Böhme, "Countering counter-forensics: The case of JPEG compression," in *Proc. Information Hiding 2011*, pp. 285-298, Springer Berlin Heidelberg, 2011.
- [10] X. Liu, M. Tanaka, M. Okutomi. "Noise level estimation using weak textured patches of a single noisy image," in *Proc. 19th. IEEE International Conference on Image Processing 2012 (ICIP2012)*, pp. 665-668, IEEE, 2012.
- [11] G. Schaefer and M. Stich, "UCID-An uncompressed color image database," in *Proc. SPIE 5307, Storage and Retrieval Methods and Applicat. for Multimedia*, Jan. 2004, pp. 472-480.
- [12] E Y Lam, J W Goodman, "A mathematical analysis of the DCT coefficient distributions for images," *IEEE Transactions on Image Processing*, vol. 9, no. 10, pp. 1661-1666, 2000.
- [13] L. Liu and X. Zhuang, "A novel square root rate control algorithm for H. 264/AVC encoding," In *Proc. of IEEE International Conference on Multimedia and Expo 2009( ICME 2009)*, pp. 814-817, 2009.
- [14] L. Liu, X. Zhuang, Z. He, and Y. Sun. "H. 264/AVC rate control with enhanced rate-quantisation model and bit allocation," *IET image processing*, vol. 5, no. 7, pp. 619-629, 2011.
- [15] D Zoran, Y. Weiss "Scale invariance and noise in natural images," *IEEE 12th International Conference on Computer Vision 2009*, pp. 2209-2216, IEEE, 2009.
- [16] M. C. Stamm, W. S. Lin, K. J. R. Liu "Forensics vs. anti-forensics: A decision and game theoretic framework," *IEEE International Conference on Acoustics, Speech and Signal Processing 2012 (ICASSP2012)*, pp. 1749-1752, IEEE, 2012.