IMAGE AUTHENTICATION BY STATISTICAL ANALYSIS

Tong Qiao, Florent Retraint, Rémi Cogranne

ICD - LM2S - Université de Technologie de Troyes (UTT) - UMR STMR CNRS,
12 rue Marie Curie - CS 42060 - 10004 Troyes cedex - France.
E-mail : {tong.qiao , remi.cogranne , florent.retraint}@utt.fr

ABSTRACT

This paper investigates the discrimination between Photographic Images (PIM) and Computer Generated (CG) images. The proposed method exploits traces of Color Filter Array (CFA) interpolation, present in PIM images, together with the use of hypothesis testing theory. By using the Likelihood Ratio Test (LRT), the method proposed to distinguish PIM from CG images warrants a prescribed False Alarm Rate (FAR) and achieves the maximal detection power. Experimental results show the efficiency of the proposed methodology and the high robustness with respect to anti-forensic techniques.

Index Terms— CG, PIM, image forensics, hypothesis testing, linear parametric model, nuisance parameters.

1. INTRODUCTION AND CONTRIBUTIONS

Digital image forensics is a new technique for distinguishing a real image from a faked one. Different from digital image watermarking, image forensics judges images without imbedding any information previously for assisting authentication. To distinguish a photographic image (PIM) from a computer generated (CG) one is a research subfield of digital image forensics on which this paper focuses.

Driven by the pioneer work [1], most of the latest methods proposed exploits high order statistics and/or physical features to distinguish PIM from CG using supervised learning method (such as Support Vector Machine, SVM). Although those methods can achieve high detection accuracy, supervised statistical learning is time-consuming. In addition, several problems such as the robustness to training and testing set mismatch and the establishment of detection performance (false-alarm and missed-detection) remain open.

Note that PIM and CG images fundamentally differ as the formers are obtained from a complex imaging process, see Figure 1, while the latters are generated by a software, not by a camera. In [2], it is proposed to exploit the features of Color Filter Array (CFA) to detect CG images. Estimating the CFA pattern is also an effective approach to authenticate PIM images under some restrictive conditions, see [3].

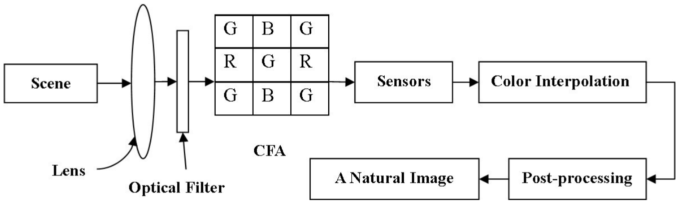


Fig. 1: Illustration of an imaging process in digital cameras.

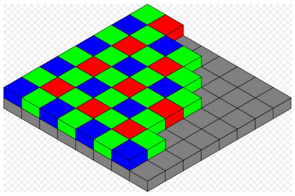


Fig. 2: Bayer Model.

Figure 1 illustrates the imaging process of digital still cameras. Photons radiating from an object go though the optical system. Then, the CFA filters the light spectrum so that each pixel records only one color channel (red, green, or blue) ; the two missing color channels are padded by color interpolation. Finally, a natural image is generated after several image post-processes such as white balance and gamma correction. In [4], the statistical features resulting from the imaging process are used for differentiating PIM from CG images.

Figure 2 shows the most commonly found CFA pattern referred to as Bayer. Based on the feature of Bayer model, PIM and CG images can be differentiated by the peak value in the frequency domain which is described specifically in Section 2 and in [5]. For a large number of PIM images the peak vanishes, hence this hardly permits the distinguishing from CG images. Thus, an improvement of the method proposed in [5] is needed.

This paper improves the Gallagher's method by two means. First, it is proposed to use the variance in the frequency domain on the assumption that the post-acquisition processes reduce the variance. Second, a linear parametric model is used to deal with nuisance parameters and based on the residual noise vector, a hypothesis testing model is established. Experimental results show the efficiency and the robustness compared to the algorithm proposed in [5].

This paper is organized as follows. Section 2 recalls the method proposed by Gallagher in [5]. Section 3 presents the proposed linear parametric model used to deal with nuisance parameters. The Likelihood Ratio Test (LRT) is established in Section 4. Numerical experiments are presented in Section 5 and, finally, Section 6 concludes this paper.

2. OUTLINE OF GALLAGHER'S METHOD

In [5], Gallagher proposed to identify CG images by detecting the peak value of the image in the frequency domain. The specific algorithm is summarized below. Let us denote $\mathbf{I}(x,y,c)$ the pixels intensity of a given image with $c=\{r,g,b\}$ the color channel and (x,y) the pixel position. First, to avoid disturbances from low frequencies, $\mathbf{I}(x,y,g)$ is filtered by the following high-pass filter $\mathbf{H}(x,y)$.

$$\mathbf{H}(x,y)=\frac{1}{4}\begin{bmatrix}0 & 1 & 0\\ 1 & -1 & 1\\ 0 & 1 & 0\end{bmatrix}$$

Note that in Gallagher's Method [5] only the green channel is used because it carries more information, due to the specific Bayer's CFA pattern illustrated in Figure 2, but an extension to other channels is straightforward. Then, the mean of all diagonal values, from the filtered image, are calculated to obtain a vector denoted $\mathbf{d}=(d_1,\dots,d_N)^T$. Here n is the index number of the diagonal, N is the total number of diagonal and $\mathbf{M}^T$ is the transpose of matrix $\mathbf{M}$. Finally, it is proposed in [5] to use the frequential representation of $\mathbf{d}$, denoted $\mathbf{D}$ and formally defined as follows:

$$\mathbf{D}=|\mathcal{DFT}(\mathbf{d})| \text{ with } \mathbf{d}_n=N_n^{-1}\sum_{x+y=n}|\mathbf{H}*\mathbf{I}(x,y,g)| \quad (1)$$

$\forall n\in\{1,\dots,N\}$. Here $\mathcal{DFT}[\cdot]$ represents the calculation of DFT and N_n is the total number of pixels on the n_{th} diagonal.

Some examples of vectors $\mathbf{D}$, obtained from Gallagher's described in (1), are given in Figure 3. Roughly speaking, the very simple test proposed in [5] consists in declaring a given image as a PIM if a peak occurs at $\mathbf{D}(N/2)$. It is obvious that images 3 (a), 3 (b) and 3 (d) can be discriminated efficiently. On the opposite, image 3 (c) is likely to be detected as a CG because it has no peak. In fact, there is a large number of PIM images without peak, see Figure 3(g), and thus, that may be wrongly detected as CG images by Gallagher's method [5]. Consequently, for reliability and efficiency purposes, it is necessary to improve the detection scheme proposed in [5].

Note that the peak value is not the only characteristic that distinguishes PIM from CG images. For almost every PIM image, see Figure 3 (e) and 3 (g), the noise present in vector $\mathbf{D}$ has a much smaller variance than for CG images, see Figure 3.

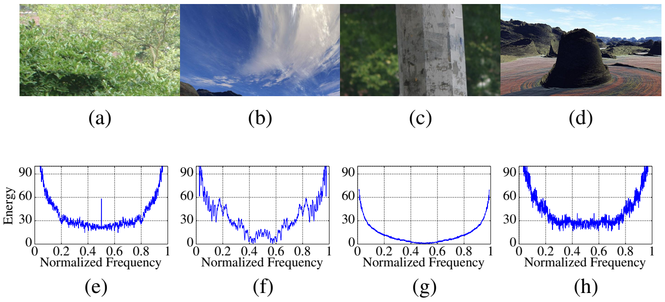


Fig. 3: PIM (a)-(c) and CG (b)-(d) images together with their diagonal mean spectrum, (e)-(g) and (f)-(h) respectively.

In the present paper it is proposed to use this property of noise variance to distinguish PIM from CG images.

3. DEALING WITH NUISANCE PARAMETERS

In this paper, a linear parametric model is proposed to deal with diagonal mean spectrum $\mathbf{D}$. To this end, $\mathbf{D}$ is splitted into K nonoverlapping vectors, denoted $\mathbf{y}_1,\dots,\mathbf{y}_K$, of m samples, see [6]. Let us define

$$\mathbf{y}_k\sim\mathcal{N}(\boldsymbol{\mu}_k,\sigma^2\mathbf{I}_m)=\boldsymbol{\mu}_k+\xi_k. \quad (2)$$

where $\boldsymbol{\mu}_k=(\mu_{k,1},\dots,\mu_{k,m})^T$ of expectations, ξ_k is the realization of a Gaussian vector with variance $\sigma^2\mathbf{I}_m$ and $\mathbf{I}_m$ the identity matrix of size $m\times m$. Hence, the problem of distinguishing PIM from CG images can be formulated as a choice between the following hypotheses:

$$\begin{cases}\mathcal{H}_0=\left\{\mathbf{y}_k\sim\mathcal{N}(\boldsymbol{\mu}_k,\sigma_0^2\mathbf{I}_m),\forall k=(1,...,K),\sigma_0\leq\sigma\right\}\\ \mathcal{H}_1=\left\{\mathbf{y}_k\sim\mathcal{N}(\boldsymbol{\mu}_k,\sigma_1^2\mathbf{I}_m),\forall k=(1,...,K),\sigma_1>\sigma\right\}\end{cases} \quad (3)$$

where σ_0^2 and σ_1^2 respectively represent the variance under each hypothesis $\mathcal{H}_0=\{$the image is PIM$\}$ and $\mathcal{H}_1=\{$the image is CG$\}$ and σ is the threshold. Obviously, the expectation $\boldsymbol{\mu}_k$ is the nuisance parameter without any interest to distinguish PIM from CG images.

Furthermore, $\mathbf{y}_k$ can be described with the following linear parametric model

$$\boldsymbol{\mu}_k=\mathbf{A}\mathbf{x}_k. \quad (4)$$

where $\mathbf{A}$ is a known full rank matrix of size $m\times n$, with $m>n$, and $\mathbf{x}_k$ is a $n\times 1$ vector of parameters describing the expectation of $\mathbf{y}_k$.

The idea of using such a linear parametric model is that it allows an easy elimination of nuisance parameter $\boldsymbol{\mu}_k$ which can be used in a hypothesis test using invariance theory [7, chap.6]. To apply this theory, let us define $C(\mathbf{A})$ the column space spanned by $\mathbf{A}$, with $\dim(C(\mathbf{A}))=\text{rank}(A)=n$ and $C(\mathbf{A})^\perp$ its orthogonal complement, sometimes referred

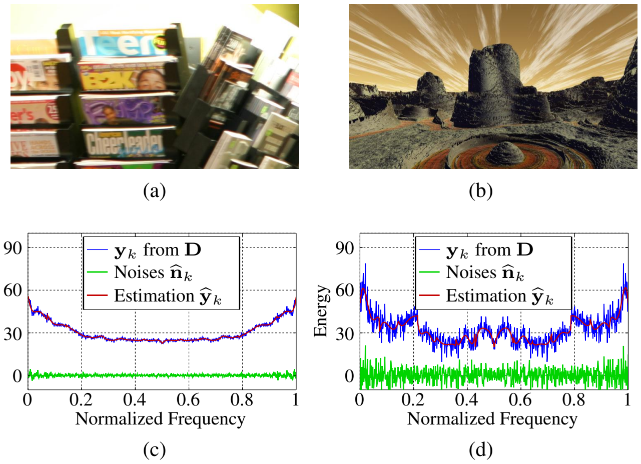


Fig. 4: PIM (a) and CG (b) images with their diagonal mean spectrum, estimated expectation and residual noises (c)-(d).

to as the "parity space", with $\dim\left(C\left(\mathbf{A}\right)^{\perp}\right) = m - n$. The projection of observation vector $\mathbf{y}_k$ onto the parity space is obtained by $\mathbf{n}_k = \mathbf{W}\mathbf{y}_k$ where the matrix $\mathbf{W}$ verifies, among others, the following useful properties:

$$\mathbf{W}\mathbf{A} = \mathbf{0} \;\text{ and }\; \mathbf{W}\mathbf{W}^T = \mathbf{I}_{m-n}. \qquad (5)$$

Hence, by using the definitions of hypotheses (3), the projection of observation vector $\mathbf{y}_k$ onto the parity space $C\left(\mathbf{A}\right)^{\perp}$ yields: $\mathbf{n}_k = \mathbf{W}\mathbf{y}_k = \mathbf{W}\xi_k \sim \mathcal{N}(0, \sigma_i \mathbf{I}_{m-n})$ with $i = \{0, 1\}$ depending on the hypothesis.

Note that the use of projection matrix $\mathbf{W}$ can be replaced by using the Maximum Likelihood Estimation (MLE). By rejecting the nuisance parameter, let us define the estimation of $\mathbf{n}_k$ as

$$\widehat{\mathbf{n}}_k = \mathbf{y}_k - \widehat{\mathbf{y}}_k = \mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{y}_k \;\text{ with }\; \mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{I}_m - \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T \qquad (6)$$

In fact, a straightforward calculation, using the properties (5), shows that :

$$\|\,\mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{y}_k\|_2^2 = \mathbf{y}_k^{\mathrm{T}}\mathbf{W}^{\mathrm{T}}\mathbf{W}\mathbf{W}^{\mathrm{T}}\mathbf{W}\mathbf{y}_k = \|\,\mathbf{W}\mathbf{y}_k\|_2^2.$$

For clarity, in the present paper the matrix $\mathbf{W}$ is used in all calculus while matrix $\mathbf{P}_{\mathbf{A}}^{\perp}$ is used for illustrations and figures, see Figure 4, to keep the same number of observations.

Note that in the present paper, the chosen linear parametric model is an algebraic polynomial of degree $n-1$; this yields the following matrix $\mathbf{A}$:

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & \dots & \dots & 1 & 1 \\ 1 & 2 & 4 & \dots & \dots & 2^{n-2} & 2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 1 & m & m^2 & \dots & \dots & m^{n-2} & m^{n-1} \end{pmatrix}$$

In addition, as detailed in Section 5, it has been chosen to remove the few first and last samples from vector $\mathbf{D}$ as well as few samples around the $N/2$.

4. LIKELIHOOD RATIO TEST PERFORMANCES

By using a linear parametric model, as described in the methodology presented in Section 3, the problem of detecting PIM and CG images can be formulated as follows:

$$\begin{cases} \mathcal{H}_0 = \{\mathbf{n}_k \sim \mathcal{N}(\mathbf{0}, \sigma_0^2 \mathbf{I}_{m-n}), \forall k = (1, \dots, K), \sigma_0 \leq \sigma^\star\} \\ \mathcal{H}_1 = \{\mathbf{n}_k \sim \mathcal{N}(\mathbf{0}, \sigma_1^2 \mathbf{I}_{m-n}), \forall k = (1, \dots, K), \sigma_1 > \sigma^\star\} \end{cases} \qquad (7)$$

For solving statistical detection problem such as (7), it follows from the Neyman-Pearson lemma [7, Theorem 3.2.1] that the Likelihood Ratio Test (LRT) is optimal in the sense described below. For definition, let

$$\mathcal{K}_\alpha = \left\{ \delta : \sup_{\sigma_0^2 \leq \sigma^\star} \mathbb{P}_0[\delta(\mathbf{D}) = \mathcal{H}_1] \leq \alpha \right\} \qquad (8)$$

be the class of tests, solving problem (7), with an upper-bounded false-alarm probability α. Here $\mathbb{P}_j[\cdot]$ is the probability under $\mathcal{H}_j, j \in \{0, 1\}$. Among all the tests in $\mathcal{K}_\alpha$ the LRT is the most powerful test, it maximizes the detection power

$$\beta_\delta = \mathbb{P}_1[\delta(\mathbf{D}) = \mathcal{H}_1]. \qquad (9)$$

From the statistical independence of vectors $\mathbf{y}_k$, the LRT is given by the following decision rule:

$$\delta(\mathbf{D}) = \begin{cases} \mathcal{H}_0 \text{ if } \Lambda(\mathbf{D}) = \sum_{k=1}^{K} \Lambda(\mathbf{y}_k) \leq \tau_\alpha \\ \mathcal{H}_1 \text{ if } \Lambda(\mathbf{D}) = \sum_{k=1}^{K} \Lambda(\mathbf{y}_k) > \tau_\alpha \end{cases} \qquad (10)$$

where the decision threshold τ_α is the solution of equation $\sup_{\sigma_0^2 \leq \sigma^\star} \mathbb{P}_0[\Lambda(\mathbf{D}) > \tau_\alpha] = \alpha$ to guarantee that $\delta(\mathbf{D}) \in \mathcal{K}_\alpha$. From the model of tested hypotheses, a straightforward calculation shows that the Likelihood Ratio (LR) $\Lambda(\mathbf{y}_k)$ is given by

$$\Lambda(\mathbf{y}_k) = \|\mathbf{n}_k\|_2^2$$

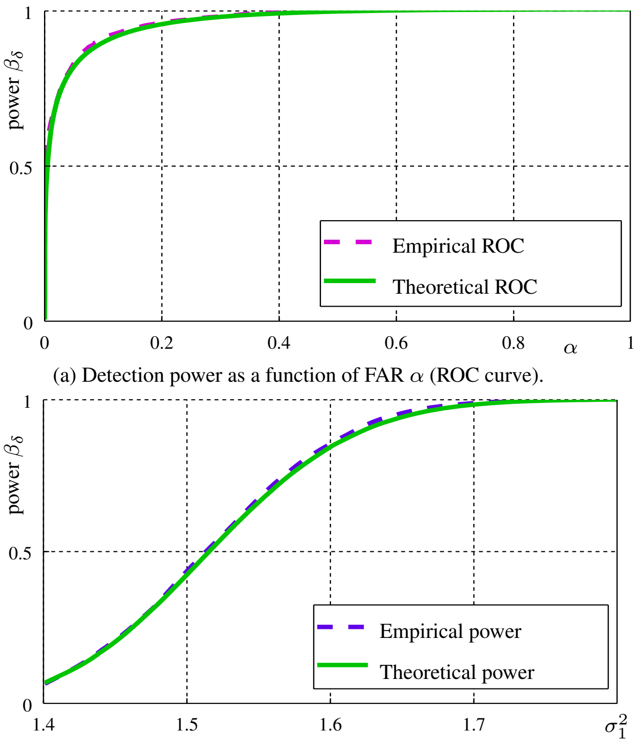
Finally, from the statistical distribution of noise residuals $\mathbf{n}_k$, and from the properties of Gaussian random variables, one immediately obtains that under hypothesis $\mathcal{H}_i, i = \{0; 1\}$:

$$\begin{aligned} & \tfrac{1}{\sigma_i^2}\|\mathbf{n}_k\|_2^2 \sim \chi^2_{m-n} \\ \Leftrightarrow \quad & \Lambda(\mathbf{y}_k) = \|\mathbf{n}_k\|_2^2 \sim \Gamma(\tfrac{m-n}{2}, 2\sigma_i^2) \end{aligned}$$

where $\Gamma(k, \theta)$ represents the Gamma distribution with a shape parameter k and a scale parameter θ. Subsequently, it follows from the stability under summation of Gamma random variables, that the statistical distribution of the LR $\Lambda(\mathbf{D})$ is given under hypothesis $\mathcal{H}_i, i = \{0; 1\}$ by:

$$\Lambda(\mathbf{D}) \sim \Gamma\left(\frac{K(m-n)}{2}, 2\sigma_i^2\right) \qquad (11)$$

It is thus immediate to establish the statistical properties of the proposed test (11) which are given in the following theorems; for clarity, $\mathbf{F}_\Gamma(\cdot)$ and $\mathbf{F}_\Gamma^{-1}(\cdot)$ represent the Gamma cumulative distribution function and its inverse respectively.



(a) Detection power as a function of FAR α (ROC curve).

(b) Detection power as a function of the variance σ_1^2.

Fig. 5: Comparison between theoretically established and empirically obtained performance of the proposed test (10).

Theorem 1. *Assume that the model hypothesis (3) holds, then for any* $\alpha \in (0; 1)$ *the decision threshold:*

$$\tau_\alpha = F_\Gamma^{-1}\left(1-\alpha; \frac{K(m-n)}{2}, 2\sigma_0^2\right) \qquad (12)$$

guarantees that the LRT δ *(10) is in the class* $\mathcal{K}_\alpha$*.*

Theorem 2. *Assume that the model hypothesis (3) holds, for any decision threshold* $\tau_\alpha \in \mathbb{R}$*, the power function associated with the test* δ *(10) is given by*

$$\beta_\delta = 1 - F_\Gamma\left(\tau_\alpha; \frac{K(m-n)}{2}, 2\sigma_1^2\right) \qquad (13)$$

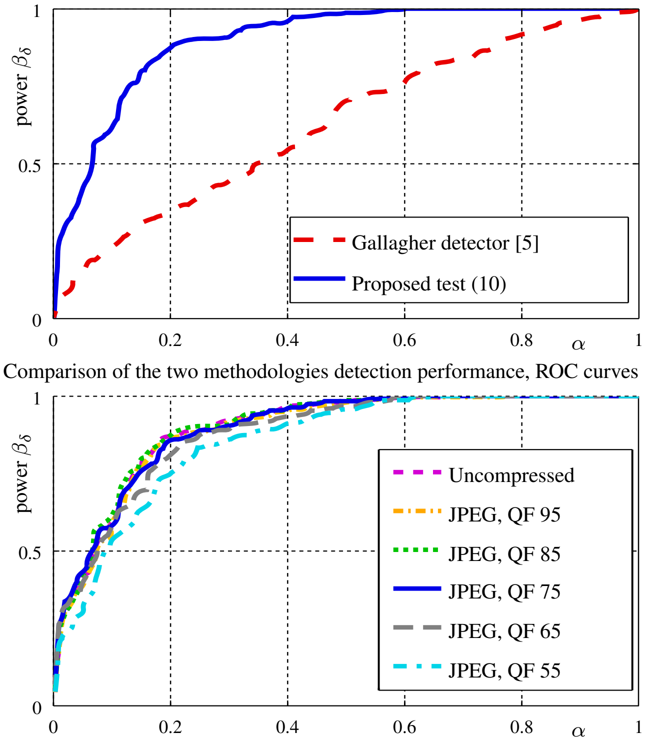
5. NUMERICAL RESULTS

To verify the sharpness of the theoretically established results, a Monte-Carlo simulation is performed. Prior to our experiments, it is proposed to use an image database containing 300 PIM images (with 150 images from Nikon D70 and 150 images from Canon 10D) from Columbia's ADVENT dataset [8] and 300 CG images downloaded from www.pandromeda.com. All these 600 images are cropped to reduce their size to 700×400 pixels and finally saved as Jpeg format with the quality factor 85.

The parametric linear model (2) is defined by a polynomial order $n-1 = 4$ and the size of vector $\mathbf{y}_k$ is set to $m = 64$ samples. Note that to avoid dealing with different variance, possibly non-uniform, the first, last and middle samples are excluded from analysis. To this end, it is proposed in practice not to consider the first, $\mathbf{y}_1$, the last, $\mathbf{y}_K$, and the two middle vectors, $\mathbf{y}_{K/2}$ and , $\mathbf{y}_{K/2+1}$. From the remaining sample, the variance of $\mathbf{D}$ is calculated using $\mathbf{n}_k$ in each image. Two sets containing 10000 vectors of 768 samples are randomly generated with zero mean and variance $\sigma_0^2 = 1.39$, to simulate residual noises from PIM images, or hypothesis $\mathcal{H}_0$, and with variance $\sigma_1^2 = 1.80$ to simulate CG images, hypothesis $\mathcal{H}_1$. The detection performances obtained with the proposed test are illustrated in Figure 5 (a); the Receiver Operating Characteristic (ROC), that is the detection power β_δ as a function of false-alarm probability α, of both empirical and theoretically established results (13) are compared. Similarly, Figure 5 (b) shows a comparison between empirical and theoretical detection power as a function samples variance, σ_1^2, under alternative hypothesis $\mathcal{H}_1$. The numerical results presented in Figure 5 (b) are obtained with a false-alarm probability set to $\alpha = 0.05$. From Figure 5 (a) and 5 (b) it is obvious that the empirical detection powers are almost identical to the theoretically established ones (13); this shows the sharpness and the relevance of theoretical findings.

To emphasize the improvement of the proposed test compared to [5], Figure 6(a) presents the detection performance, as ROC curves, of those detectors. It can be noted that, for instance, for $\alpha = 0.2$, the proposed test power is above 0.85 while it is below 0.35 using Gallagher's method. In addition, when the detected images with the low image quality, the proposed test preserves a high detection performance. This is emphasized in Figure 6(b), which presents ROC curves of the proposed test for uncompressed images and compressed images with Jpeg standard and quality factors ranging from 55 to 95.

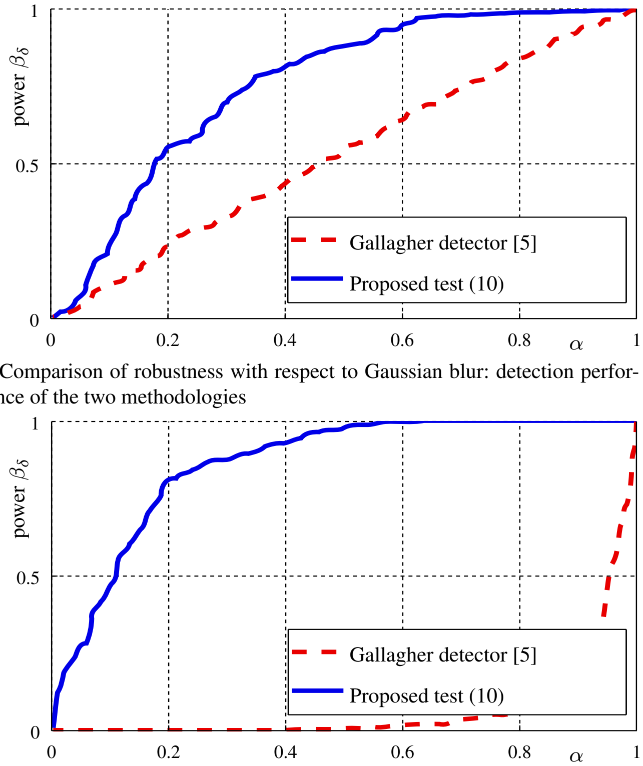
Finally it is proposed to study the detection performance of the proposed test in the context of anti-forensics considered in this paper. First, it is proposed to apply a Gaussian blur to both PIM and CG on the assumption that such a blurring process should largely reduce the noise variance. In addition, PIM images are subjected to deterministic degradations, which are well modeled as blurring processes, during its acquisition. Figure 7(a) shows the empirically obtained detection performance, of both the proposed test and the method proposed in [5], after image blurring. Second, knowing the image acquisition pipeline, it is reasonable to assume that one may try to simulate the CFA interpolation in CG images. Hence, it is proposed in this paper to apply the well known bi-linear demosaicing filter on all the CG images. The empirically obtained results are shown in Figure 7(b); these particularly highlight that the detection method proposed in [5] performs poorly. In fact since most of the PIM images have a small peak, or no peak at all, the simulation of CFA interpolation artificially creates a periodic pattern which results in a peak in CG diagonal mean spectrum. The method proposed



(a) Comparison of the two methodologies detection performance, ROC curves

(b) Detection performance for different image quality factor (QF) of JPEG compression

Fig. 6: Illustration of the proposed test (10) performance for real PIM and CG images.



(a) Comparison of robustness with respect to Gaussian blur: detection performance of the two methodologies

(b) Comparison of robustness with respect to simulated CFA interpolation

Fig. 7: Comparison of detection performance in the presence of anti-forensic process.

in [5] hence easily classifies CG images as PIM images as soon as the CFA interpolation process is simulated. The results from Figure 6 emphasizes the lack of the robustness of the detection method proposed in [5] and, on the opposite, highlights the efficiency as well as the good robustness of the proposed statistical test (10).

6. CONCLUSION

In this paper, we describe an approach of distinguishing between PIM and CG images based on statistical decision theory. A linear parametric model is developed to deal with nuisance parameters. By using the residual noise $\mathbf{n}_k$ representing the property of each detected image, hypothesis testing model is exploited to categorize two kinds of images. The method proposed in this paper overcomes the detector proposed in [5] and improves the detection accuracy. Moreover, experimental results also emphasize that the proposed method has a good robustness with respect to basic anti-forensic techniques.

7. REFERENCES

[1] S. Lyu and H. Farid, "How realistic is photorealistic?," *Signal Processing, IEEE Transactions on*, vol. 53, no. 2, pp. 845–850, 2005.

[2] S. Bayram *& al.*, "Source camera identification based on cfa interpolation," in *IEEE Intl. Conf. on Image Processing*, 2005.

[3] H. Cao and A.C Kot, "Accurate detection of demosaicing regularity for digital image forensics," *Information Forensics and Security, IEEE Transactions on*, vol. 4, no. 4, pp. 899–910, 2009.

[4] A. Swaminathan *& al.*, "Digital image forensics via intrinsic fingerprints," *Information Forensics and Security, IEEE Transactions on*, vol. 3, no. 1, pp. 101–117, 2008.

[5] A.C. Gallagher and T. Chen, "Image authentication by detecting traces of demosaicing," in *Proc. of IEEE CVPR conf.*, 2008, pp. 1–8.

[6] R. Cogranne and F. Retraint, "An asymptotically uniformly most powerful test for LSB matching detection," *Information Forensics and Security, IEEE Transactions on*, vol. 8, no. 3, pp. 464–476, 2013.

[7] E. Lehman and J. Romano, "Testing statistical hypotheses," in *Second Edition*. Springer, 2005.

[8] T.T. Ng *& al.*, "Columbia photographic images and photorealistic computer graphics dataset," *Columbia University Technical Report*, 2005.