ON RESAMPLING DETECTION IN RE-COMPRESSED IMAGES

Matthias Kirchner, Thomas Gloe

Technische Universität Dresden, Faculty of Computer Science, Institute of Systems Architecture
01062 Dresden, Germany

ABSTRACT

Resampling detection has become a standard tool in digital image forensics. This paper investigates the important case of resampling detection in re-compressed JPEG images. We show how blocking artifacts of the previous compression step can help to increase the otherwise drastically reduced detection performance in JPEG compressed images. We give a formulation on how affine transformations of JPEG compressed images affect state-of-the-art resampling detectors and derive a new efficient detection variant, which better suits this relevant detection scenario. The principal appropriateness of using JPEG pre-compression artifacts for the detection of resampling in re-compressed images is backed with experimental evidence on a large image set and for a variety of different JPEG qualities.

Index Terms— digital image forensics, tamper detection, resampling detection, re-compression artifacts

1. INTRODUCTION

Digital image forensics and tamper detection have become widely studied subjects in the multimedia security community [1, 2] and a broad fund of available techniques to distinguish original from manipulated images has evolved. As reported in the body of literature, most of the known forensic methods work pretty well under laboratory conditions. On the downside, however, there is definitely a gap to bridge between promising results in controllable test environments and 'real life' conditions. One of the largest issues for a lot of approaches is lossy compression, which is likely to smooth out subtle artifacts of previous manipulations. JPEG post-compression, probably the most common way to save images in a storage-friendly way, is at the same time a serious obstructor to a reliable detection of image manipulations.

Resampling detection [3–7], meanwhile a standard tool in image forensics, is a typical representative in this regard. While a reliable detection for almost arbitrary geometric transformations in uncompressed images has been reported, it is well-known that detection performance severely drops already for moderate JPEG compression [3, 6].

Two recent studies [8, 9] allude to the case of resampling of pre-compressed images, where JPEG artifacts from the first compression can actually help to increase the detection performance under a second (post-)compression. This interesting idea is of high practical relevance, since images from most digital cameras are already in JPEG format before they are further processed on a computer. Unfortunately, both author teams did not provide a thorough description or a quantitative experimental evaluation of these additional artifacts, which leaves the question about the detectability of resampling in pre-compressed images open.

This paper strives to fill this gap by applying our recent model of resampling artifacts in uncompressed images [5] to the case of resampling of pre-compressed images. We will start with a short review on 'classical' resampling detection in Sect. 2, before we give a description of resampling artifacts in pre-compressed images in Sect. 3. Section 4 presents an adapted detector, which is extensively tested on a large set of images in Sect. 5. The paper is concluded by Sect. 6 with remarks on future research and open problems.

2. RESAMPLING DETECTION IN BITMAP IMAGES

Using the notation of inverse mapping [10], the affine transformation of a digital image is to be formulated as resampling of discrete integer target coordinates $\chi = (\chi_1, \chi_2)$ to real-valued source coordinates $x = (x_1, x_2) = \mathbf{A}^{-1}\chi$ according to the 2×2 transformation matrix $\mathbf{A}$. Interpolation is the key to smooth and visually appealing image transformation and can be written as

$$s(\mathbf{A}^{-1}\chi) = \sum_{\chi' \in \mathbb{Z}^2} s(\chi')h(\mathbf{A}^{-1}\chi - \chi') \,, \qquad (1)$$

where s is the signal of interest, and h is the interpolation kernel. A virtually unavoidable side effect of typical interpolation algorithms is that they create periodic linear dependences between groups of neighboring samples. To the best of our knowledge, all resampling detectors exploit these artifacts by analyzing resampled signals in terms of their linear predictor residue [3, 5] or n-th order derivatives [4, 6, 7, 9].

2.1. Resampling Detection with Linear Predictor Residue

In the course of this paper, we shall focus on the predictor-based approaches, where, adhering to Popescu and Farid's

original scheme [3], a residue signal e is calculated as

$$e(\mathbf{A}^{-1}\boldsymbol{\chi}) = s(\mathbf{A}^{-1}\boldsymbol{\chi}) - \sum_{\boldsymbol{k}\in\{-K,...,K\}^2} \alpha_{\boldsymbol{k}}\, s\left(\mathbf{A}^{-1}(\boldsymbol{\chi}+\boldsymbol{k})\right), \quad (2)$$

with $\alpha_{0,0} := 0$ and K integer. The idea behind predicting a transformed pixel's intensity from its surrounding is that large absolute prediction errors indicate a minor degree of linear dependence and vice versa. A so-called *p-map* as a measure for the strength of linear dependence can be derived from the prediction error, which is modeled as a zero mean Gaussian random variable.[1] Previous resampling operations leave noticeable periodic pattern in the p-map, which result in distinct peaks in its Fourier spectrum. Typically, the detection is based on the existence of such peaks in the frequency domain.

2.2. Effects of JPEG Post-compression

Results documented in the literature show that resampling with low-order interpolation kernels is reliably detectable for almost arbitrary transformation parameters. However, it has been mentioned from the very beginning [3, 4] that subsequent JPEG compression diminishes detection performance mainly for two reasons. First, the lossy compression generally tends to blur the output signal. Second, JPEG compression is based on a 8×8 block-based processing of the image, which can lead to sharp transitions between neighboring pixels at the borders of two consecutive blocks. Both effects are not without consequences to the prediction error. While subtle periodic traces in resampled images are smoothed out due to smaller prediction errors within a block, *new* periodic artifacts originating from systematically increased prediction errors at the block borders are introduced. As a result, the magnitude of resampling peaks in JPEG images is generally lower compared to uncompressed images. Further, the detector must ignore those peaks that stem from block-based lossy compression. An immanent side-effect is the detector's blindness to any resampling peak superimposed with a JPEG peak.

3. RESAMPLING ARTIFACTS IN PRE-COMPRESSED IMAGES

Research on resampling detection so far was mainly concerned with the analysis of resampling of never-compressed bitmap images (and possibly subsequent compression). Here, JPEG artifacts are associated with negative influence on the detection performance. In the case of pre-compressed images, however, JPEG artifacts of the first compression actually give rise to the formation of additional indicative traces [8, 9].

For a description of resampling artifacts in pre-compressed images we will adhere to our recent work [5] on the analysis of periodic patterns in the prediction error of resampled images. Modeling the prediction error as a zero-mean random

variable, we used the variance $\mathrm{Var}[e(\boldsymbol{\chi})]$ as a simple model of the p-map of a wide-sense stationary signal and showed how resampling introduces the known periodicities.

In the following, denote the (continuous) Fourier transformation of $\mathrm{Var}[e(\boldsymbol{x})]$ and $\mathrm{Var}[e(\mathbf{A}\boldsymbol{x})]$ by V and $V^{(\mathbf{A})}$, respectively. The two spectra are interrelated by [10, p. 53]

$$V^{(\mathbf{A})}(\boldsymbol{f}) = \mathcal{F}(\mathrm{Var}[e(\mathbf{A}\boldsymbol{x})]) = \frac{1}{\det \mathbf{A}} V\left(\boldsymbol{f}^{(\mathbf{A})}\right), \quad (3)$$

with $\boldsymbol{f}^{(\mathbf{A})}$ being the 'affine transformed' frequency,

$$\boldsymbol{f}^{(\mathbf{A})} = (\mathbf{A}')^{-1}\boldsymbol{f}. \quad (4)$$

Taking the discrete sampling coordinates $\boldsymbol{\chi} \in \mathbb{Z}^2$ into account, the spectrum of $\mathrm{Var}[e(\boldsymbol{\chi})]$ is given by its discrete Fourier transformation (DFT) $\tilde{V}^{(\mathbf{A})}$,

$$\tilde{V}^{(\mathbf{A})} = \sum_{\forall \mathbf{n}\in\mathbb{Z}^2} V\left(\boldsymbol{f}^{(\mathbf{A})} - \boldsymbol{n}\right). \quad (5)$$

Note, that we have dropped the scalar term from (3) for notational convenience. From the Nyquist theorem, we know that the baseband of $\tilde{V}^{(\mathbf{A})}$ is given by $\left|\boldsymbol{f}^{(\mathbf{A})}\right| \leq (0.5, 0.5)$, where $|\boldsymbol{f}| = (|f_1|, |f_2|)$ denotes element-wise absolute values. Thus, whenever $|(\mathbf{A}')^{-1}\boldsymbol{f}| > (0.5, 0.5)$, alias frequencies from high frequency bands will be mapped to baseband frequencies $\boldsymbol{b}^{(\mathbf{A})}$,

$$\left|\boldsymbol{b}^{(\mathbf{A})}\right| = \left|\boldsymbol{f}^{(\mathbf{A})} - \left[\boldsymbol{f}^{(\mathbf{A})}\right]\right|, \quad (6)$$

with $[\boldsymbol{f}] = ([f_1], [f_2])$ denoting element-wise rounding.

3.1. Resampling Peaks

The formation of 'classical' resampling peaks in the p-map's spectrum can be modeled by noting that the predictor residue of a wide-sense stationary signal have periodic variance with a period of 1 [5],

$$\mathrm{Var}[e(x_1, x_2)] = \mathrm{Var}[e(x_1 + 1, x_2)] =$$
$$\mathrm{Var}[e(x_1, x_2 + 1)] = \mathrm{Var}[e(x_1 + 1, x_2 + 1)]. \quad (7)$$

Consequently, we expect V to have distinct peaks at integer frequencies $\boldsymbol{p} \in \mathbb{Z}^2$. Under an affine transformation $\mathbf{A}$, these peaks are mapped to the baseband according to (6) to form the characteristic baseband resampling peaks.

3.2. Shifted JPEG Peaks

As mentioned in Sect. 2.2, JPEG compression generally introduces its own characteristic artifacts to a signal and hence to the prediction error. In terms of our model of the p-map, a block-wise quantization will typically lower the variance of the prediction error of pixels within a block and increase it for pixels near the block borders. As a result, V will have additional distinct peaks at frequencies $(k/8, l/8)$, $(k, l) \in \mathbb{Z}^2$.

---

[1] See [3, 5] for a description of the computation of the p-map.

In case of a geometric transformation, these JPEG peaks underlie the affine transformation mapping in (6), similar to the previously mentioned resampling peaks. It also follows from our intuition that resampling affects the shape of the 8×8 block structure and hence shifted versions of the JPEG peaks occur in the p-map's spectrum, $\tilde{V}^{(\mathbf{A})}$. As a result, the spectrum of a pre-compressed resampled image's p-map exhibits a mixture of resampling and shifted JPEG peaks. The latter however, can be even more pronounced. This is not surprising when we think of the visibility of JPEG artifacts (depending on the JPEG quality) compared to the subtle periodicity due to interpolation.

4. DETECTION PROCEDURE

Knowing that resampling yields periodic artifacts in the p-map that are visible as distinct peaks in the p-map's Fourier spectrum, resampling detectors base their decision on some kind of a peak detector. While Popescu and Farid's method employs an exhaustive search over a large set of pre-computed synthetic p-maps [3], we recently proposed to calculate the maximum gradient in the cumulative periodogram to test for the existence of strong peaks more efficiently [5].

Both detectors, however, have weaknesses when pre-compression comes into play. In order to fully exploit the additional artifacts, Popescu and Farid's detector [3] would need a complete second set of synthetic p-maps, which does not seem to be very practical. The periodogram-based detector [5] has a different flaw: The existence of additional peaks decreases the maximum gradient in the cumulative periodogram (which, by definition, is normalized to the total energy in the spectrum).

Additionally, every peak detector will be sensitive to peaks due to JPEG post-compression. Consequently, these peaks must be ignored when searching for resampling peaks [4]. However, it is important to note that it is not enough to simply zero-out all spectral components at frequencies $(k/8, l/8)$, since the JPEG peaks are typically spread over a certain range around their theoretical position.

We will now outline an adapted detection procedure that is not only capable of detecting typical resampling peaks, but is also able to exploit the existence of shifted JPEG peaks. The basis for predictor-based resampling detection is the p-map and its Fourier transformation [3, 5]. In the following denote p as the p-map and P as the magnitude of its Fourier spectrum, $P = |\mathcal{F}(p)|$. Before peaks are evaluated, we employ the following pre-processing:

Normalization To attenuate the in general strong low-frequency components due to image content, the spectrum P is first normalized to its median filtered version P_{median},

$$P_n = P/P_{\mathrm{median}} . \tag{8}$$

Peak finding In order to find interesting peaks in the spectrum, we process P_n with a maximum filter of size W and zero-out all non-maximum frequency components:

$$P_m(\boldsymbol{f}) = \begin{cases} P_n(\boldsymbol{f}) & \text{if } P_n(\boldsymbol{f}) = \max_{\boldsymbol{w} \in \{-W,\dots,W\}^2} P_n(\boldsymbol{f} + \boldsymbol{w}) \\ 0 & \text{else.} \end{cases} \tag{9}$$

Remove JPEG peaks of post-compression In the case of JPEG images, we remove all remaining frequency components at frequencies $(k/8, l/8)$, $(k, l) \in \{-4, \dots, 4\}^2$.

Emphasize strong peaks In order to give more weight to strong peaks we finally perform a gamma correction step,

$$P_\gamma = \max(P_m) \cdot \left(\frac{P_m}{\max(P_m)} \right)^\gamma . \tag{10}$$

The so processed spectrum is now fed through a simple peak detector. Due to symmetry in the Fourier domain we assume that for basic affine transformations like scaling, rotation or shearing, in theory, there will be at least two and at most four resampling or shifted JPEG peaks with the same distance from the DC frequency. Writing $P_\gamma(r; \phi)$ as the spectrum P_γ indexed by polar coordinates $(r; \phi)$, only the four maximum peaks on each radius r are evaluated by summing up their magnitudes. Denote Σ_r as the sum corresponding to radius r, we then take the ratio δ

$$\delta = \max_{r > r_t} \Sigma_r \Big/ \operatorname*{median}_{r > r_t} \Sigma_r \tag{11}$$

as decision criterion. Note that all frequencies within radius r_t are ignored in order to preclude a false detection due to low frequency "noise" stemming from the image content. We expect δ to be larger for transformed images than for originals, since resampling and/or shifted JPEG peaks cause a few predominant outliers, which however do not significantly influence the median value. Thus, for a given threshold T, whenever $\delta > T$, the image under investigation is flagged as resampled.

5. EXPERIMENTAL RESULTS

For a quantitative evaluation of the detectability of resampling in JPEG compressed images, we test our detector with a subset of the 'Dresden Image Database' [11]. For full control over JPEG compression qualities, we randomly chose each 100 original RAW images from a Nikon D70 and Panasonic DMC-FZ50 digital camera and converted them into uncompressed TIFF images using Adobe Lightroom. We further cropped a 1024×1024 pixel region and used these 200 images as test database.

The test setup includes a wide range of scaling factors and considers two different scenarios. The *baseline detection results* report the detection of resampling in the absence of pre-compression. This is basically a reproduction of already known results [3, 5]. The *re-compression scenario* investigates the detection of resampling in images that had been
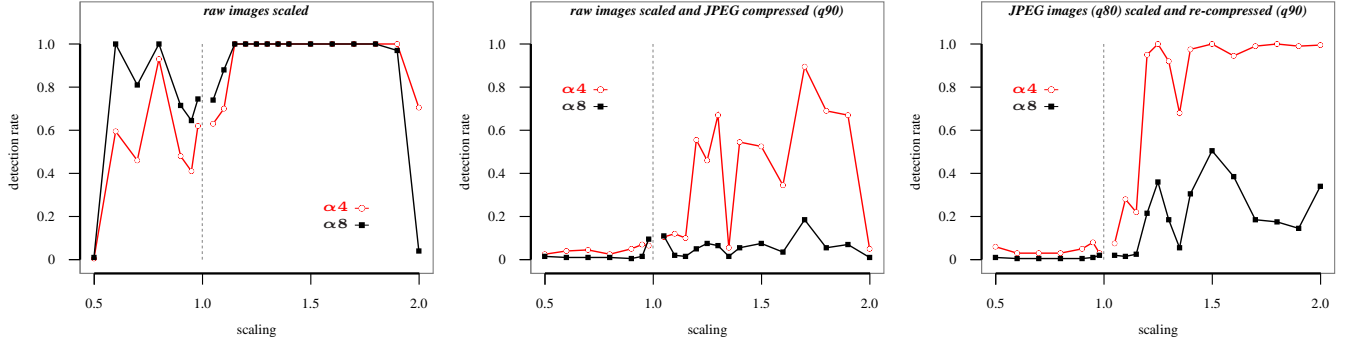
Fig. 1. Detection results for filters $\alpha 8$ and $\alpha 4$. (left) scaled raw images, (middle) scaled and JPEG compressed images, and (right) scaled and re-compressed images. post-compression quality 90; pre-compression quality 80. FAR $< 1\,\%$.

JPEG compressed before the transformation. All tests were conducted using using ImageMagick's `convert` with linear interpolation and varying JPEG pre- and post-compression qualities in the range of $\{40, 50, 70, 75, 80, 90, 95, 98, 100\}$ where appropriate. The overall test set accumulates to more than 400 000 analyzed images. Due to space constraints, however, we are only able to report a small subset of results, which we deem particularly insightful.

For each image under investigation the p-map was calculated from a 512×512 region of the corresponding luminance channel using our fast linear filtering approach [5] with the filter masks

$$\boldsymbol{\alpha 8} = \begin{bmatrix} -0.25 & 0.50 & -0.25 \\ 0.50 & 0 & 0.50 \\ -0.25 & 0.50 & -0.25 \end{bmatrix}, \; \boldsymbol{\alpha 4} = \begin{bmatrix} 0 & 0.25 & 0 \\ 0.25 & 0 & 0.25 \\ 0 & 0.25 & 0 \end{bmatrix},$$

as $\qquad p(\mathbf{A}^{-1}\boldsymbol{\chi}) = \exp\left(-|e(\mathbf{A}^{-1}\boldsymbol{\chi})|^2\right) .$ $\qquad$ (12)

Filter $\boldsymbol{\alpha 8}$ was already proposed in [5], where it gave good results for uncompressed images. The right mask $\boldsymbol{\alpha 4}$ is more related to derivative-based resampling detectors [4, 6], however capable of detecting signal changes in two dimensions.

For a given false acceptance rate (FAR), the detection threshold T was determined for each post-compression quality by applying the detector to all corresponding original images in the database. The peak detector employed a 7×7 median filter for the normalization of the spectrum and was operated with $\gamma = 4$, $W = 4$. For uncompressed images we set $r_t = 1/16$, whereas for JPEG images $r_t = 1/8$ was chosen.[2]

5.1. Baseline Detection Results

To demonstrate the generally good performance of resampling detection in uncompressed images, the left graph in Fig. 1 reports detection rates at FAR $< 1\,\%$ for varying scaling factors. Each point in the plot corresponds to 200 resam-

	post-compression quality									
	40	**50**	**60**	**70**	**75**	**80**	**90**	**95**	**98**	**100**
40	70.0	83.5	94.5	99.5	99.5	99.0	100.0	100.0	100.0	99.5
50	8.5	61.0	82.5	98.0	99.0	99.5	99.5	99.0	99.5	99.5
60	2.0	11.5	63.0	96.0	100.0	100.0	100.0	100.0	100.0	100.0
70	1.5	2.0	7.0	86.0	98.0	99.5	100.0	100.0	100.0	100.0
75	4.0	3.0	2.5	55.0	94.0	95.5	100.0	100.0	100.0	100.0
80	3.0	1.5	1.5	24.0	71.0	91.5	100.0	100.0	100.0	100.0
90	2.5	2.5	1.5	5.5	10.0	25.5	99.5	100.0	100.0	100.0
95	3.0	2.5	2.0	7.0	11.0	11.5	87.5	100.0	100.0	100.0
98	2.5	3.0	2.5	5.5	9.0	10.0	56.0	99.5	100.0	100.0
100	2.0	3.0	3.0	5.5	8.5	10.0	58.0	99.0	100.0	100.0
*	2.5	2.0	2.5	5.0	9.0	10.5	52.5	98.5	100.0	100.0

(pre-compression quality labels the rows on the left)

Table 1. Detection rates [%] for varying combinations of pre- and post-compression quality. 150 % upscaling; FAR $< 1\,\%$; filter $\boldsymbol{\alpha 4}$. The bottom row corresponds to uncompressed input images.

pled images. Upscaling is generally better detectable than downscaling.[3]

After JPEG compression, the detection rates considerably drop both for up- and downscaling (cf. middle graph in Fig. 1 for JPEG quality 90). Interestingly, here filter $\boldsymbol{\alpha 4}$ gives better results throughout the whole range of scaling factors. In general, we found $\boldsymbol{\alpha 4}$ to be superior for stronger compression, while $\boldsymbol{\alpha 8}$ is preferable for high quality images. The reason for this effect is subject to future research.

5.2. Re-compression Results

From the rightmost graph in Fig. 1, we can see that resampling detection in re-compressed images benefits from blocking artifacts of the previous compression. While the post-compression quality is the same as in the baseline scenario (middle graph), upscaling is much better detectable since the input images were JPEG compressed with quality 80 before. It is obvious, that the pre-compression JPEG quality is a very influential factor on the detection performance after resampling and post-compression. The stronger the blocking

---

[2]Due to lossy compression, JPEG images tend to have a stronger low-frequency part.

[3]The doubling of image size is problematic since the resampling peaks are superimposed with CFA peaks [12], which are ignored by the detector.

artifacts were, the more they will remain in the processed images. Further, it is advantageous when the pre-compression quality is below that of re-compression because strong JPEG compression causes blurred blocks. This effect can also be inferred from Table 1 that reports detection results (using $\alpha 4$) for varying pre- and post-compression qualities at a fixed scaling rate of 150 %. Observe how low pre-compression as well as high post-compression qualities increase the ability to detect resampling. In general, resampling of re-compressed images is better detectable than resampling of solely post-compressed images. Similar results have been be obtained for all 21 investigated scaling factors. The strength of pre-compression, which is necessary for a successful resampling detection, however generally depends on the resampling parameters. Compared to upscaling, a detection of downscaling requires a lower JPEG quality in the first compression step. For instance, we found that for a post-compression quality of 90 (cf. Fig. 1), a pre-compression quality below 50 is required to obtain detection rates comparable to the baseline results.

Overall, we can conclude from the reported results that it is important to keep in mind that, for re-compressed images, a low JPEG quality not necessarily means a bad detection performance *per se*.

6. CONCLUDING REMARKS

In this paper, we have investigated the detection of resampling in re-compressed images. While lossy compression so far has mostly been discussed in terms of negative effects on the detectability of geometric transformations, we have shown how blocking artifacts in re-compressed images can actually help to increase detection performance. Based on linear predictor residue resampling detection [3, 5], the main contributions of this paper lie in a description of how affine transformations of pre-compressed images affect the detector output (i. e. the p-map's Fourier spectrum) and the presentation of a suitable detection variant. Experimental results confirm that resampling detection in JPEG images is not by definition a lost cause.

As to the limitations, we have to note that our detector is no exception with regard to downscaling as the most problematic type of geometric transformation. It remains an open research question if (and how) a more reliable detection of downscaling is possible.

Since the results presented in this paper are limited to the case of scaling, our future research will include an examination of additional affine transformations. We further believe that it is worth studying how the size of the analyzed image (region) influences the ability to detect resampling. First experimental tests indicate that a peak detector can better distinguish between shifted and original JPEG peaks (that have to be ignored) the larger the spectral resolution becomes.

Finally, it is worth mentioning that the idea of using shifted peaks for the detection of resampling is not limited to JPEG peaks. The same effect can be observed for peaks stemming from color filter array (CFA) interpolation inside the camera [12]. This probably allows a reliable detection of very slight transformations, which has not always been possible so far.

7. REFERENCES

[1] H.T. Sencar and N. Memon, "Overview of state-of-the-art in digital image forensics," 2008, Mimeo.

[2] H. Farid, "Image forgery detection," *IEEE Signal Processing Magazine*, vol. 26, no. 2, pp. 16–25, 2009.

[3] A.C. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of re-sampling," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 758–767, 2005.

[4] A.C. Gallagher, "Detection of linear and cubic interpolation in JPEG compressed images," in *Second Canadian Conference on Computer and Robot Vision*, 2005, pp. 65–72.

[5] M. Kirchner, "Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue," in *ACM Multimedia and Security Workshop (ACM MM&Sec)*, 2008, pp. 11–20.

[6] B. Mahdian and S. Saic, "Blind authentication using periodic properties of interpolation," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 3, pp. 529–538, 2008.

[7] A. Suwendi and J.P. Allebach, "Nearest-neighbor and bilinear resampling factor estimation to detect blockiness or blurriness of an image," *Journal of Electronic Imaging*, vol. 17, no. 2, 023005, 2008.

[8] M.-C. Poilpré, P. Perrot, and H. Talbot, "Image tampering detection using Bayer interpolation and JPEG compression," in *e-Forensics 2008*, 2008.

[9] W. Wei, S. Wang, and Z. Tang, "Estimation of rescaling factor and detection of image splicing," in *IEEE International Conference on Communication Technology (ICCT)*, 2008, pp. 676–679.

[10] B. Jähne, *Digital Image Processing*, Springer-Verlag, Berlin, Heidelberg, 6. edition, 2005.

[11] T. Gloe and R. Böhme, "The 'Dresden Image Database' for benchmarking digital image forensics," *submitted to ACM SAC 2010*.

[12] A.C. Popescu and H. Farid, "Exposing digital forgeries in color filter array interpolated images," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3948–3959, 2005.