

 Open access • Proceedings Article • DOI:10.1117/12.766824

Camera identification from printed images — Source link

Miroslav Goljan, Jessica Fridrich, Jan Lukás

Institutions: Binghamton University

Published on: 14 Feb 2008 - electronic imaging (International Society for Optics and Photonics)

Topics: Digital camera

Related papers:

- [Digital camera identification from sensor pattern noise](#)
- [Determining Image Origin and Integrity Using Sensor Noise](#)
- [Large scale test of sensor fingerprint camera identification](#)
- [Camera identification from cropped and scaled images](#)
- [Source Camera Identification Using Enhanced Sensor Pattern Noise](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/camera-identification-from-printed-images-s96k58ogja>

Camera Identification from Printed Images

Miroslav Goljan^{*}, Jessica Fridrich, and Jan Lukáš
Department of Electrical and Computer Engineering
SUNY Binghamton, Binghamton, NY 13902-6000

ABSTRACT

In this paper, we study the problem of identifying digital camera sensor from a printed picture. The sensor is identified by proving the presence of its Photo-Response Non-Uniformity (PRNU) in the scanned picture using camera ID methods robust to cropping and scaling. Two kinds of prints are studied. The first are postcard size (4" by 6") pictures obtained from common commercial printing labs. These prints are always cropped to some degree. In the proposed identification, a brute force search for the scaling ratio is deployed while the position of cropping is determined from the cross-correlation surface. Detection success mostly depends on the picture content and the quality of the PRNU estimate. Prints obtained using desktop printers form the second kind of pictures investigated in this paper. Their identification is complicated by complicated geometric distortion due to imperfections in paper feed. Removing this distortion is part of the identification procedure. From experiments, we determine the range of conditions under which reliable sensor identification is possible. The most influential factors in identifying the sensor from a printed picture are the accuracy of angular alignment when scanning, printing quality, paper quality, and size of the printed picture.

Keywords: Camera identification, Photo-Response Non-Uniformity, printed images identification.

1. INTRODUCTION

With the proliferation of digital imaging technology, the problem of establishing a link between an imaging hardware and a digital object, such as image or video, has tremendously increased on importance. This statement is confirmed by the recent surge in publications on this topic.¹⁻¹² In some situations, however, instead of a digital object only its analog form may be available. A valid question to ask is whether it is possible to identify a sensor from a printed picture rather than its digital form. An example of such situation is forged paper currency when a digital scan of a real banknote is subsequently printed on a printer.

The first camera identification technology was based on identifying defective pixels.⁶ These defects will likely show up in good quality printouts and thus form a very useful forensic entity for identifying sensors from printed pictures. Unfortunately, many cameras are able to correct for defective pixels (e.g., after proper firmware upgrade) or do not have any easily visible defects. The defective pixels can also be easily removed with retouching tools. Photo-response non-uniformity (PRNU) was recently proposed⁷ as a fingerprint for digital cameras or an electronic equivalent of “biometrics for sensors”. There are several important advantages of using this fingerprint for forensic purposes. The fingerprint is stable over time and under a wide range of physical conditions. It is also present in every picture (with the exception of completely dark images) independently of the camera optics, settings, or the scene content. The fingerprint is universal because both CCD and CMOS sensors found in today’s digital cameras exhibit PRNU. Most importantly, the fingerprint has large information content because it is a signal that contains a stochastic component due to material properties of silicone and the manufacturing process itself. Thus, it is unlikely that two sensors will possess similar fingerprints. Lastly, the fingerprint was proved to survive a wide range of common image processing operations, including lossy compression, filtering, and gamma adjustment.⁷

It is, however, a valid question whether the fingerprint survives the distortion due to digital-to-analog conversion because besides the distortion determined by the point-spread function of the hardware involved, we also lose synchronization, which is essential for successful detection of PRNU. It is the goal of this paper to experimentally

^{*} mgoljan@binghamton.edu; phone +001-607-777-5793; fax +001-607-777-4464.

investigate the range of conditions under which the presence of the camera fingerprint can be established from a printed picture. We carry out a series of targeted experiments to investigate this issue in more detail.

In the next section, we discuss various approaches that can be employed for our task under two different scenarios. In Section 3, we describe the detector used to identify pictures printed in common commercial printing labs. The experimental results on prints produced by 5 different digital cameras are in Section 4. Section 5 focuses on camera identification from pictures printed on a desktop ink-jet and laser printer. These experiments provide an insight as to which factors are the most influential for successful identification. Conclusions are drawn in Section 6.

2. SCENARIOS AND APPROACHES

The process of printing and scanning is quite complex because it involves several sources of distortion that vary considerably with the printer model and its settings, the paper and ink type, and the scanner. In the forensic application that we consider, the scanner is chosen by the analyst who has in possession a printed copy of a digital image. The printer, on the other hand, is unknown to the analyst even though some information may be deduced from the hard copy of the picture. The distortion due to the print and scan process can be roughly divided into two types: pixel value and geometric.¹³ An example of the pixel value distortion is gamma correction, halftoning, dot gain,¹⁴ print-to-print instability (horizontal banding), point-spread function, and the quantization noise. The geometrical distortion is a combination of cropping, rotation, scaling, and even some slight non-linearity (rubber-sheet deformation) due to uneven feed of the paper during printing. The last distortion concerns typical consumer printers more than commercial picture printing labs or print-in-minutes machines. The pixel value distortions are in general less of a problem than geometric distortions. The PRNU has been shown to well survive⁷ gamma correction and JPEG compression. The impact of halftoning becomes an issue when scanning at very high dpi. The ink spreading heavily depends on the paper quality and properties of the printing process itself. Both halftoning and the ink spreading influence the reliability of camera ID from printed images (as will be demonstrated in Section 4).

By far the most difficult challenge to overcome is the geometric transformation. Sensor identification using PRNU is very sensitive to proper synchronization. If the image under investigation is misaligned with the fingerprint, direct detection will not succeed. The problem with geometrical transformations is well known to the watermarking community and is usually solved by a judicious watermark or template design by either making the watermark self-referencing,¹⁵ symmetrical,¹⁶ or by switching to a domain¹⁷ invariant to rotation, scaling, and translation. In our situation, however, the fingerprint (the “watermark”) is imposed on us and we have no control over its “design.” Thus, we cannot expect it to have any properties that would make the detection easier. Consequently, it is substantially more difficult to detect the fingerprint after geometrical processing than it is for a human-designed watermark.

To recover from a combination of change of scale and rotation, we can either

- 1) use the peaks in the Fourier transform due to printer halftone pattern as a template,
- 2) construct the detector in some domain invariant to rotation (and scaling), such as the Fourier-Mellin domain,
- 3) construct the detector directly in the spatial domain and use a brute force search for the parameters of the geometrical transform.

The first possibility is only applicable to laser printers because ink-jet printers do not exhibit the halftone pattern. For laser printers, however, this method can accurately estimate the rotation angle.¹⁴ It requires the image to be scanned at very high dpi (e.g., over 1200 dpi). After inverting the rotation, a brute force search for the scaling and cropping parameters could be applied.

The second possibility is appealing because the Fourier-Mellin transforms rotation and scaling to spatial shift. On the other hand, the F-M transform itself introduces additional distortion due to the necessity to resample the signals on a non-rectangular grid, which may be a substantial factor for our application since the SNR between the sensor fingerprint and the image is only -50dB or less.

In this paper, we opted for the third possibility due to its wide applicability, ease of implementation, and flexibility. Moreover, if the geometric transform is more complex and cannot be accurately approximated by a combination of rotation, scaling, and cropping, the first two methods might not provide satisfactory performance. In this case, the third approach can be applied on a block-by-block basis.

Two kinds of printed images will be considered in this work. The first kind represents pictures printed by professional quality digital minilabs (compact machines that print in minutes) or by a commercial printing lab (digital printer processor, digital photo center) where trained personnel typically makes prints within an hour. Such prints are usually not rotated or the rotation angle is so small that it can be neglected. The rotation due to scanning can also be largely eliminated by selecting a quality scanner and carefully controlling the scanning process. The scanning resolution will not, however, be synchronized with the original image pixel resolution. Cropping and scaling are thus the two main unknown factors that need to be detected and estimated. We have some control over the range of scaling ratios since the scanning resolution can be manually adjusted. The approach we have taken is to formulate the detection using hypothesis testing, derive a detector using the Generalized Likelihood Ratio Test (GLRT), and find the maximum of the test statistics using a brute force search described in Ref.¹⁸

The second kind of prints consists of full size images printed on a desktop laser or ink-jet printer. For simplicity, we assume that the images have a white border and thus no cropping is involved in this case. The desktop printers, however, introduce a non-linear distortion due to mechanical feed of paper in the printer that must be taken into consideration. The white border simplifies the search for the geometrical transformation.

3. THE CAMERA IDENTIFICATION ALGORITHM

In this section, we describe the algorithm we used for camera identification from prints produced by commercial printing labs. The experiments in Section 5 carried on prints from desktop printers were done with a different detector due to a different experimental setup.

3.1 PRNU estimation

The PRNU for each tested camera was estimated from the camera native resolution images of blue or cloudy sky using the maximum likelihood estimator.²² All images were taken at the highest available quality the camera allowed. With the exception of Canon G2 where the PRNU was estimated from 32 sky images in an uncompressed format (CRW), for the remaining cameras we used 40 images in the JPEG format. The estimate was also preprocessed using the column and row zero-meaning¹⁹ to remove any small systematic errors possibly caused by limited precision in color interpolation when color filter array is present in the camera. The original image dimensions for each camera are listed in Table 1.

Table 1. List of cameras tested, number of images used to estimate the PRNU, and original image dimensions $m \times n$.

	N_{PRNU}	$m \times n$
Fuji E550	40	2136 \times 2848
Canon G2	32	1704 \times 2272
Minolta G500	40	1944 \times 2592
Olympus C765	40	1712 \times 2288
Canon S2	40	1944 \times 2592

3.2 PRNU detection

We employ the brute force search algorithm for identification of the source camera from cropped and scaled images described in Ref.¹⁸ This section should be considered as an initial study investigating the feasibility and limitations of this approach on postcard-size commercial prints. Below, we only provide an abbreviated description of the detection algorithm, referring to the original publication¹⁸ for more details.

Let \mathbf{Z} be the scanned image (represented as an array of integers $M \times N \times 3$) and let $\hat{\mathbf{K}}$ be the maximum likelihood estimate of the camera PRNU from images taken with the camera in its native resolution. We represent $\hat{\mathbf{K}}$ with an array of $m \times n \times 3$ real numbers. Using a denoising filter F , we obtain the noise residual of the scan $\mathbf{W} = \mathbf{Z} - F(\mathbf{Z})$ in each color channel. Prior to running the detection algorithm, the three color channels for both \mathbf{Z} and $\hat{\mathbf{K}}$ are combined to form a grayscale signal using the usual RGB to grayscale conversion.

The task of detecting the PRNU in the noise residual can be formulated as the two-channel problem

$$\begin{aligned} H_0: \mathbf{K} &\neq \mathbf{K}' & \hat{\mathbf{K}} &= \mathbf{K} + \boldsymbol{\Xi}_1 \\ H_1: \mathbf{K} &= \mathbf{K}' & \mathbf{W} &= T_r \circ T_s(\mathbf{I}\mathbf{K}' + \boldsymbol{\Xi}_2) \end{aligned}$$

where T_r and T_s are the operations of scaling by factor r and shifting by vector \mathbf{s} , respectively, and \mathbf{I} is the original digital image that was printed. The statistics for a simplified GLRT for this problem derived in Ref.¹⁸ is

$$\max_{i,j,r} \text{NCC}[i, j, r],$$

where $\text{NCC}[i, j, r]$ is the normalized cross-correlation between $\hat{\mathbf{K}}$ and $T_{1/r}(\mathbf{WZ})$

$$\text{NCC}[i, j, r] = \frac{\sum_{k=1}^m \sum_{l=1}^n T_{1/r}(\mathbf{WZ})[k, l] \hat{\mathbf{K}}[k+i, l+j]}{\|T_{1/r}(\mathbf{WZ})\| \cdot \|\hat{\mathbf{K}}\|}, \quad i=1, 2, \dots, m, \quad j=1, 2, \dots, n. \quad (1)$$

For each scaling ratio r , the maximum is taken over (i, j) that lie within the $(m-M+1) \times (n-N+1)$ region of the NCC corresponding to all possible croppings (i.e., no ‘‘circular’’ cropping is allowed). The search for the scaling ratio starts by first resizing the scan so that it just fits the PRNU (scaling ratio r_0) down to $r_{R-1} = r_0/2$ (see Fig. 1)

$$r_0 > r_2 > \dots > r_{R-1}, \quad r_0 = \min\{m/M, n/N\}, \quad r_{R-1} = r_0/2, \quad 1/r_{i+1} - 1/r_i = 0.005.$$

The signals are appropriately padded with zeros before calculating the NCC. The norm in the denominator of (1) is the L_2 norm when considering the matrices as vectors. Finally, note that (1) can be quickly evaluated using the fast Fourier transform.

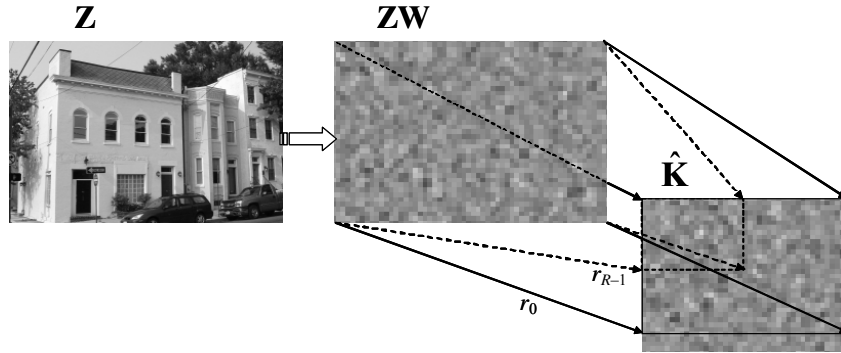


Fig. 1. The scaling ratio search range starts at r_0 and proceeds down to r_{R-1} .

Denoting the location of the maximum value of $\text{NCC}[i, j, r]$ over (i, j) (for a fixed r) as $(s_1, s_2) = \mathbf{s}_{\text{peak}}$, the statistical significance of the peak is evaluated using the ‘‘Peak to Correlation Energy’’ ratio²⁰ (PCE)

$$\text{PCE} = \frac{\text{NCC}[\mathbf{s}_{\text{peak}}]^2}{\frac{1}{mn - |\mathcal{A}|} \sum_{\mathbf{s} \in \mathcal{A}} \text{NCC}[\mathbf{s}]^2}, \quad (2)$$

where \mathcal{A} is a small square area (15×15 pixels) around the peak. If the maximum PCE over the search range is above a certain threshold, the image is deemed to be taken by the camera, otherwise the camera is not determined.

Fig. 2 shows the structure of the whole camera ID algorithm. The search continues three more loops after finding the maximum PCE . This ensures (with high probability) that the global maximum has been located. A detailed search around the maximum with the step corresponding to one pixel change in the larger dimension further improves the accuracy of the estimate of the geometrical transformation. A detailed pseudo-code for the whole algorithm can be found in Ref.¹⁸

If the camera identification is positive, the location of the maximum correlation in the cross-correlation surface determines the position of the cropped image within the original. This includes both the cropping that could have occurred before or after printing and, possibly, the cropping during scanning. The scaling ratio found during this process is of lesser interest since it depends on the scanning resolution.

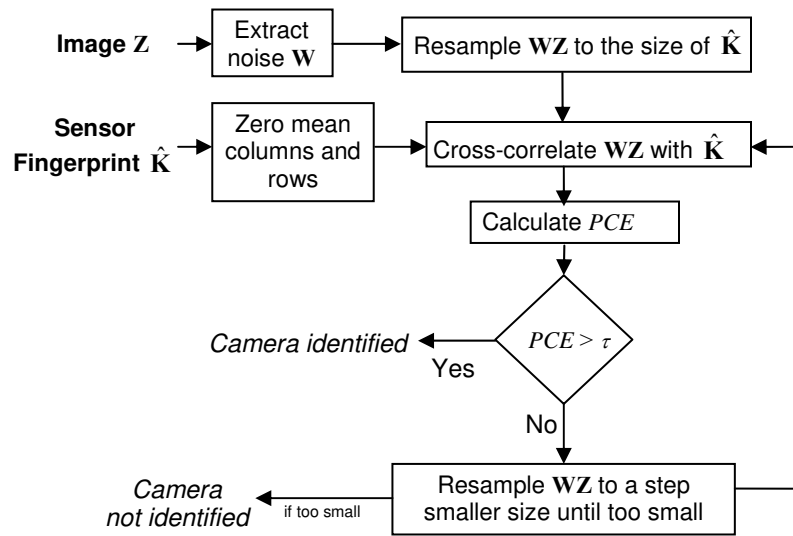


Fig. 2. Simplified block diagram of camera identification from a cropped and arbitrarily scaled image.

3.3 Error analysis

Modeling the NCC (with the exception of the area \mathcal{A}) as a Gaussian signal (see the discussion on plausibility of this assumption in¹⁸), the probability of obtaining $PCE \geq \tau$ under the hypothesis that the image was not taken by the camera is

$$P_{FA} = 1 - \left(1 - Q(\sqrt{\tau})\right)^{(m-M+1)(n-N+1)k_{\max}}, \quad (3)$$

where $Q(x) = 1 - F(x)$ with $F(x)$ denoting the cumulative density function of a normal random variable $N(0,1)$, and k_{\max} is the maximum number of scaling ratio samples the search runs through. Solving (3) for τ allows us to calculate the threshold for PCE , given the false acceptance rate P_{FA} ,

$$\tau = \left[Q^{-1} \left(1 - (1 - P_{FA})^{\frac{1}{(m-M+1)(n-N+1)k_{\max}}} \right) \right]^2. \quad (4)$$

3.4 Factors influencing reliability of identification

The following factors influence the reliability of camera identification the most:

- 1) quality of the PRNU estimate,
- 2) scanning resolution,
- 3) rotation angle between the scanned image and the (unknown) original,
- 4) severity of non-linear distortion caused by printing and by scanning,
- 5) amount of saturation in the image,
- 6) image content.

In general, we should use as many images for PRNU estimation as possible. As explained above, in our experiments with commercially printed images, we used 32–40 images of blue (cloudy) sky in the best format the camera allowed.

Intuitively, the higher the scanning resolution the better the noise extraction will work thus making the correlation between \mathbf{WZ} and $\hat{\mathbf{K}}$ larger. However, it could be detrimental to have the scanned image much larger than the PRNU because \mathbf{WZ} is then down-sampled to the size of the PRNU or smaller. Experiments with 30 prints from 4 megapixel Canon G2 camera and scanning resolutions 200, 300, and 600 dpi suggest that the resolution should be adjusted to produce images of approximately the same size as the PRNU. For post-card size prints, 300 dpi was generally the best overall choice, although 200 dpi was not, surprisingly, much worse and sometimes even better. On the other hand, 600 dpi lead to an overall lower PCE and required smaller step than 0.005% in the search for the scaling ratio to avoid missing narrower peaks. The other possibility to up-sample $\hat{\mathbf{K}}$ instead of down-sampling \mathbf{WZ} , which leads to increased computational complexity and false alarms due to the larger exponent in (3).

The rotation of the scanned image due to imprecise placement of the print on the scanner glass has an important impact on the reliability of the camera identification process. To know exactly how precise we have to be in this regard, we conducted the following experiment and repeated it for two printed images from Olympus C765 one from Canon G2. The printed picture was placed on the scanner and rotated by 0.1 degree each time it was scanned. Two rotational degrees were thus covered by 21 scans. The first and the last such scanned images misaligned by 2 degrees are shown in Fig. 3. Each scan was tested for the presence of the correct PRNU. Plotting the largest PCE for each rotated picture in Fig. 4, we see that camera identification is successful when the rotation angle is smaller than ± 0.1 degree for Olympus C765 and about ± 0.2 degree for Canon G2. This quite a narrow margin indicates that it may be worth to include a small search for this angle to improve the identification reliability. For the Olympus, the cropped area was correctly detected for all four rotations #11–14. For illustration, in Fig. 5 we show the detected cropped area for rotations #12 and #13.

Both the printing and scanning processes suffer from small non-linear distortions caused mostly by tolerances in moving parts and mechanisms. Since the scanner is our choice, we can test it and thus at least minimize the contribution to the geometrical distortion due to scanning.

Saturation in bright parts of the scanned image effectively wipes out all noise that may be present on the print. Whenever such situation occurs, one should adjust the scanner contrast/brightness appropriately to avoid saturation. If the printed image already contains large saturated areas, the chance of successful camera identification from the image sharply diminishes.



Fig. 3. Two images misaligned by 2 degrees (camera Olympus C765).

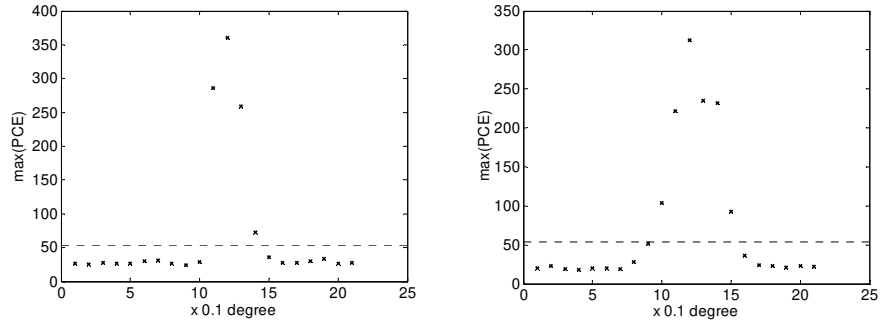


Fig. 4. Maximum PCE obtained after picture rotation during scanning. Camera Olympus C765 (left) ($\tau = 53.927$) and Canon G2 (right) ($\tau = 53.861$). The dotted line marks the threshold τ obtained from (4) by setting $P_{FA} = 10^{-5}$.



Fig. 5. Detected cropping for rotated image #12 (left), $PCE = 360.35$ and image #13 (right), $PCE = 259.14$, for Olympus C765. The gray border is the area that was cropped out.

4. CAMERA IDENTIFICATION FROM PRINTS PRODUCED BY COMMERCIAL PRINTING LABS

We conducted an experiment with prints from five digital cameras printed at a Wal-Mart Photo Center. Total of 20 pictures (all in the JPEG format) from each camera were developed and printed on a 4"×6" matte photo paper. Prior to printing, each image was cropped, often in both dimensions. The cropping was performed at the printing center after uploading the images to the computer. All pictures were scanned on an inexpensive HP Photosmart 2610 scanner at 300 dpi; automatic contrast adjustment was on. All 100 scans underwent camera identification testing for all five cameras. The search method for scaling and cropping parameters explained in the previous section was applied with the (valid) assumption that none of the printed images were cropped from the original more than 50% in either dimension.

Because forensic applications require very low false alarm rate, we report the results by showing successful detections for the decision threshold τ adjusted to guarantee the probability of false alarms less than 10^{-5} . We believe that this reporting is generally more appropriate in forensic setting than producing confusion matrices.

Table 2. Camera identification test for printed and scanned images – the part in which the camera matches the print. Gray background highlights the cases where the camera identification failed (PCE below the threshold).

Im #	Fuji E550		Canon G2		Minolta G500		Olympus C765		Canon S2		
	PCE orig.	PCE scan	PCE orig.	PCE scan	PCE orig.	PCE scan	PCE orig.	PCE scan	PCE orig.	PCE scan	PCE rot.
1	13854	569	11775	156	9540	65	2194	30	11071	47	89
2	15803	421	13597	267	15920	123	2645	33	16608	125	214
3	1622	73	12203	401	31789	355	9623	121	22624	26	117
4	1116	32	14077	386	20479	260	4632	50	20595	28	98
5	8544	90	13401	174	44013	149	5604	91	25633	29	55
6	1552	139	6261	267	6703	29	2979	29	24796	89	139
7	2309	31	8125	67	26080	355	3022	28	30264	113	256
8	931	30	8274	182	14370	251	599	29	36071	143	286
9	4440	88	5425	81	21421	293	2071	24	34821	57	206
10	7947	80	21250	173	17913	27	8603	105	56625	51	342
11	26666	168	19655	712	8399	43	25198	234	40674	29	245
12	19039	53	15376	374	14772	108	12167	115	27422	33	171
13	3616	38	21290	459	33592	181	11545	32	9210	35	55
14	23991	390	12869	121	7225	172	9579	103	22525	123	286
15	17155	221	12989	160	16628	138	12260	133	9398	40	106
16	41457	763	16956	90	7888	26	7323	102	38827	62	261
17	3788	30	25014	751	14315	64	12534	303	22186	73	149
18	15 dz	119	16326	821	10726	54	6349	123	25655	33	124
19	36887	382	22864	645	5495	119	14811	205	49951	145	347
20	4412	35	17011	658	5204	79	17632	182	5659	29	48
τ		54.8		53.9		54.5		53.9		54.4	58.2
ω		45.1		41.8		39.4		46.8		42.1	

Matching pairs. Table 2 contains the maximum PCE for each matching pair of image and the camera (column “PCE scan”). Because the threshold τ depends on both the image aspect ratio and the PRNU dimensions, it had to be calculated for every image separately. However, because the differences in the aspect ratio were small overall, the threshold came out essentially the same for every set of 20 images (the row τ). The column “PCE orig.” corresponds to the case when the original (not printed) digital image was identified instead of the scan. It reflects how reliably would the camera be identified if the image was not processed outside of the camera, i.e., if it did not go through the process of cropping, printing, and scanning. “PCE orig.” thus provides information about the influence of image content. As a rule of thumb, the larger the PCE for the original image is, the more likely can the camera be identified from the printed image as well.

Note that all pictures from Canon G2 were successfully identified. Seven pictures from Fuji E550 slipped through the camera identification due to their darker, textured, or saturated content. The same can be said about 6 out of 8 prints from Olympus 765 that were not identified. Notice that image #18 from Fuji E550 was correctly matched to the camera despite the fact that the original was not! This is because the image was acquired using digital zoom (dz), which we confirmed by applying to the original image the camera ID method robust to scaling and cropping¹⁸.

The higher missed detection rate for Canon S2 cannot be explained by a worse quality PRNU estimate or “more difficult” image content because the PCEs for their original counterparts are mostly very large (tens of thousands). These prints were ordered a few weeks later than the other tested prints and were all rotated by about 0.15 degree. Therefore, for these prints, we included a small search for 7 rotation angles $-0.15, -0.1, -0.05, 0, 0.05, 0.1,$ and 0.15 degree. The enlarged search required increasing the threshold τ by 3.8 to 58.2 to keep the same false alarm rate. After incorporating the search for the rotation, only 3 images were not detected (image #5, #13, and #20) at our false alarm level. The last column in Table 2 shows the maximum PCE among the 7 rotations.

Overall, the main reason for many missed detections is the presence of non-linear distortion. A good but simple model for some typical cases of this distortion may help to overcome this problem.

Non-matching pairs. The maximum PCE obtained by testing each PRNU against all 4×20 images from the other four cameras is included as ω in the last row of Table 2. As this maximum is below the threshold τ , we experienced no false alarm in our experiments.

5. PICTURES PRINTED ON A DESKTOP INK-JET AND LASER PRINTERS

Although desktop printers typically exhibit larger point-spread function than professional printing devices, by far the most influential factor complicating camera ID from pictures printed on a desktop printer is the geometrical distortion. It typically exhibits nonlinearities introduced by the mechanical parts that transport the paper under the printing head. Because inverting the non-linear distortion can be quite difficult, we simplify the task by assuming that the image has not been cropped (e.g., a small white border around the picture is intentionally left). Assuming that the deformation is not severe, after locating the image corners, we can resynchronize the image with the camera PRNU using an affine map and then apply the correlation detector (1).

Starting with the scanned image \mathbf{Z}_s , the following steps form the camera identification method.

1. Locate the corners of the scanned image and determine an affine transformation T that maps its corners to the corners of $\hat{\mathbf{K}}$, the estimated PRNU.
2. Using T , resample the scanned image, \mathbf{Z}_s , to the size of $\hat{\mathbf{K}}$ to obtain \mathbf{Z} .
3. Combine the noise residuals from all three color channels of \mathbf{Z}_s to a grayscale signal. Transform it using the affine transform T . The transformed signal, denoted as \mathbf{W} , has the same dimensions as $\hat{\mathbf{K}}$ and \mathbf{Z} .
4. Evaluate the normalized cross-correlation, $\text{NCC}[i, j]$ (1), between $\hat{\mathbf{K}}$ and $\mathbf{Z}\mathbf{W}$.
5. Find the peak of $\text{NCC}[i, j]$ in a small (10×10) area around the origin $(i, j) = (0, 0)$ and compute the PCE (2).
6. Set $P_{FA} = 10^{-5}$. For this false alarm, we need $PCE \geq \tau$, where

$$\tau = \left[Q^{-1} \left(1 - (1 - P_{FA})^{10^{-2}} \right) \right]^2 = 27.033.$$

The affine transform is the simplest one that maps the corners of the printed and scanned image to the corners of the original image. Other pixels may not be mapped precisely to those in the original image. (The corner pixels, too, can be off a fraction of the pixel size.) This is why the correlation peak for the matching camera can be shifted by a few pixels from the origin (0, 0). To stay on the safe side, we search for the maximum correlation in a 10×10 area around the origin in the cross-correlation surface.

4.1 Desktop printing tests

In tests #1–3, we used 10 images (see Fig. 6) from the Canon G2 camera. A different set of 10 images taken using the Olympus C765 camera were used in tests #4 and #5. The printed images were initially stored as high quality JPEGs (90% standard quality). As before, the HP Photosmart 2610 was the printing (ink-jet) and scanning device. The automatic contrast adjustment for printing was switched off to prevent a low color depth.

In the tests below, we chose several different settings for the print quality, print size, and paper type to better map out the conditions under which successful camera identification can be achieved. We set the scanning resolution to 1200 dpi, which worked overall the best. We were able to obtain good results also when scanning at 600 dpi. The main reason for the small difference is that higher scanning resolution allows for a better precision in locating the printed image corners. Anticipating more difficult detection, for this experiment we estimated the PRNU $\hat{\mathbf{K}}$ for each camera from 300 images. In the first 5 tests, the printed image size was 3.75"×5". The missed detections (PCE below 27.033)

are indicated by gray background in the result tables below. The p -values show the probability of observing the value of PCE or higher purely by chance.

Test #1. Setup: HP Glossy Photopaper, normal print quality. Gray indicates missed detections at $P_{FA} = 10^{-5}$.

Image	1	2	3	4	5	6	7	8	9	10
<i>PCE</i>	37.24	37.71	28.81	28.17	17.68	33.52	36.10	16.51	38.58	38.58
<i>p</i> -value	5.22×10^{-8}	4.10×10^{-8}	3.99×10^{-6}	5.56×10^{-6}	0.0013	3.53×10^{-7}	9.37×10^{-8}	0.0024	2.63×10^{-8}	2.63×10^{-8}

Test #2. Setup: HP Glossy Photopaper, best print quality. Gray indicates missed detections at $P_{FA} = 10^{-5}$.

Image	1	2	3	4	5	6	7	8	9	10
<i>PCE</i>	231.58	95.09	190.64	115.46	62.88	190.65	297.81	62.76	215.01	273.41
<i>p</i> -value	$< 10^{-13}$	$< 10^{-13}$	$< 10^{-13}$	$< 10^{-13}$	1.11×10^{-13}	$< 10^{-13}$	$< 10^{-13}$	1.22×10^{-13}	$< 10^{-13}$	$< 10^{-13}$

Comparing Test #1 and #2, the print quality substantially influenced the identification success rate.

Test #3. Setup: HP Normal ink-jet paper, best print quality.

The identification failed for all 10 images. The paper quality thus plays an important role. We were able, however, to successfully identify the camera from normal quality prints when the size of the printed image was doubled.

The next two tests repeat the first two tests for 10 (different) images from the Olympus C765 camera.

Test #4. Setup: HP Glossy Photopaper, normal print quality. Gray indicates missed detections at $P_{FA} = 10^{-5}$.

Image	1	2	3	4	5	6	7	8	9	10
<i>PCE</i>	18.50	17.08	15.81	17.40	16.14	37.86	17.76	16.41	31.14	29.72
<i>p</i> -value	0.00085	0.00179	0.0035	0.00151	0.00294	3.8×10^{-8}	0.00125	0.00255	1.2×10^{-6}	2.5×10^{-6}

Test #5. Setup: HP Glossy Photopaper, best print quality. Gray indicates missed detections at $P_{FA} = 10^{-5}$.

Image	1	2	3	4	5	6	7	8	9	10
<i>PCE</i>	58.57	172.68	53.17	83.90	193.14	131.48	79.32	224.72	173.96	131.23
<i>p</i> -value	9.77×10^{-13}	$< 10^{-13}$	1.53×10^{-13}	$< 10^{-13}$	$< 10^{-13}$	$< 10^{-13}$	$< 10^{-13}$	$< 10^{-13}$	$< 10^{-13}$	$< 10^{-13}$

Tests #4 and #5 indicate that the printing quality also plays a very important role. While many prints obtained through normal printing quality could not be detected, the best printing quality produced images that were reliably identified.

In the next batch of tests, we increased the size of the printed image from $3.75" \times 5"$ to $7.5" \times 10"$ and again ran similar tests to see the influence of the paper quality and printing settings. We only summarize the results of these experiments in words. We were able to successfully detect the PRNU from the large printouts even when printed on regular ink-jet paper with the best quality printing setting.

Besides the tests with an ink-jet printer, we also performed limited number of tests with HP Color Laser Jet 4600dtn. To our surprise, even though the quality of the laser printouts was visibly much worse compared to the output from the ink-jet (the images contained pronounced horizontal banding), we were successful in extracting the sensor PRNU (sufficiently correlated to our PRNU estimate) in large printouts ($7.5" \times 10"$) at the best printing quality even when printed on a regular paper.



Fig. 6. Ten test images from the 4 megapixel Canon G2 camera.

5. CONCLUSIONS

We present a method for identifying digital camera from printed pictures. The method itself is based on our previous work on camera identification. The presented results suggest that camera signature in the form of PRNU in most cases survives high quality printing and scanning. The most influential factors are the printing quality, namely the lack of any severe geometrical distortion, physical size of the printed picture, quality of the paper, and the image content. Reliable identification of source camera from professional (in store) postcard size (4" × 6") prints is often possible with very low false alarm probability. To obtain successful camera ID from postcard size of 3.75" × 5" pictures printed on a desktop ink-jet printer, it was necessary to print full size images on high quality photopaper at the best printing quality. When printed on regular quality paper, the physical size of the pictures had to be increased to 7.5" × 10" to obtain positive identification. This was true for printouts coming from both the ink-jet and laser jet printers.

ACKNOWLEDGEMENT

The work on this paper was supported by the AFOSR grant number FA9550-06-1-0046. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation there on. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of Air Force Research Laboratory, or the U.S. Government. Special thanks belong to Lisa Carlsson for posing her question about possible survival of the camera fingerprint in printed pictures.

REFERENCES

1. S. Bayram, H.T. Sencar, and N. Memon: "Source Camera Identification Based on CFA Interpolation." *Proc. ICIP'05*, Genoa, Italy, September 2005.
2. A. Swaminathan, M. Wu, and K.J.R. Liu: "Non-intrusive Forensic Analysis of Visual Sensors Using Output Images." *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP'06)*, May 2006.

3. A. Swaminathan, M. Wu, and K.J.R. Liu: "Image Authentication via Intrinsic Fingerprints." *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. 6505, January 29–February 1, San Jose, CA, pp. 1J–1K, 2007.
4. M. Kharrazi, H.T. Sencar, and N. Memon: "Blind Source Camera Identification," *Proc. ICIP'04*, Singapore, October 24–27, 2004.
5. K. Kurosawa, K. Kuroki, and N. Saitoh: "CCD Fingerprint Method – Identification of a Video Camera from Videotaped Images." *Proc. ICIP'99*, Kobe, Japan, pp. 537–540, October 1999.
6. Z. Geradts, J. Bijhold, M. Kieft, K. Kurosawa, K. Kuroki, and N. Saitoh: "Methods for Identification of Images Acquired with Digital Cameras." *Proc. of SPIE, Enabling Technologies for Law Enforcement and Security*, vol. 4232, pp. 505–512, February 2001.
7. J. Lukáš, J. Fridrich, and M. Goljan: "Digital Camera Identification from Sensor Pattern Noise." *IEEE Transactions on Information Security and Forensics*, vol. 1(2), pp. 205–214, June 2006.
8. H. Gou, A. Swaminathan, and M. Wu: "Robust Scanner Identification Based on Noise Features." *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. 6505, January 29–February 1, San Jose, CA, pp. 0S–0T, 2007.
9. N. Khanna, A.K. Mikkilineni, G.T.C Chiu, J.P. Allebach, and E.J. Delp III: "Forensic Classification of Imaging Sensor Types." *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. 6505, January 29–February 1, San Jose, CA, pp. 0U–0V, 2007.
10. B. Sankur, O. Celiktutan, and I. Avcibas: "Blind Identification of Cell Phone Cameras." *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. 6505, January 29–February 1, San Jose, CA, pp. 1H–1I, 2007.
11. T. Gloe, E. Franz, and A. Winkler: "Forensics for Flatbed Scanners." *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. 6505, January 29–February 1, San Jose, CA, pp. 1I–1J, 2007.
12. N. Khanna, A.K. Mikkilineni, G.T.C. Chiu, J.P. Allebach, and E.J. Delp III: "Scanner Identification Using Sensor Pattern Noise." *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. 6505, January 29–February 1, San Jose, CA, pp. 1K–1L, 2007.
13. C.Y. Lin and S.F. Chang: "Distortion Modeling and Invariant Extraction for Digital Image Print-and-Scan Process," presented at the Int. Symp. Multimedia Information Processing, December 1999.
14. K. Solanki, U Madhow, B. S. Manjunath, S. Chandrasekaran, and I. El-Khalil: "Print and Scan Resilient Data Hiding in Images," *IEEE Transactions on Information Security and Forensics*, vol. 1(4), pp. 464–478, December 2006.
15. C.W. Honsinger and S.J. Daly: "Method for detecting rotation and magnification in images," United States Patent 5,835,639, Kodak, 1998.
16. N. Nikolaidis and I. Pitas: "Circularly Symmetric Watermark Embedding in 2-D DFT Domain," In: *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP99)*, vol. 6, pp. 3469–3472, Phoenix, AZ, March 15–19, 1999.
17. J.J.K. Ó Ruanaidh and T. Pun: "Rotation, Scale and Translation Invariant Spread Spectrum Digital Image Watermarking," *Signal Processing*, vol. 66(5), pp. 303–317, May 1998.
18. M. Goljan and J. Fridrich, "Camera Identification from Scaled and Cropped Images," *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents X*, San Jose, CA, 2008.
19. J. Fridrich, M. Chen, M. Goljan, and J. Lukáš: "Digital Imaging Sensor Identification (Further Study)," *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. 6505, San Jose, CA, January 28–February 2, pp. 0P–0Q, 2007.
20. B.V.K. Vijaya Kuma and L. Hassebrook: "Performance Measures for Correlation Filters," *Appl. Opt.* 29, 2997–3006, 1990.
21. S.M. Kay, *Fundamentals of Statistical Signal Processing*, Volume II, Detection theory, Prentice Hall, 1998.
22. J. Fridrich, M. Chen, M. Goljan, "Source Digital Camcorder Identification Using Sensor Photo-Response Non-Uniformity", *Proc. of SPIE Electronic Imaging, Photonics West*, January 2007, pp. 1G–1H.