Source Digital Camcorder Identification Using Sensor Photo Response Non-Uniformity

Mo Chen, Jessica Fridrich[*], Miroslav Goljan, Jan Lukáš
Department of Electrical and Computer Engineering
SUNY Binghamton, Binghamton, NY 13902–6000, USA

ABSTRACT

Photo-response non-uniformity (PRNU) of digital sensors was recently proposed [1] as a unique identification fingerprint for digital cameras. The PRNU extracted from a specific image can be used to link it to the digital camera that took the image. Because digital camcorders use the same imaging sensors, in this paper, we extend this technique for identification of digital camcorders from video clips. We also investigate the problem of determining whether two video clips came from the same camcorder and the problem of whether two differently transcoded versions of one movie came from the same camcorder. The identification technique is a joint estimation and detection procedure consisting of two steps: (1) estimation of PRNUs from video clips using the Maximum Likelihood Estimator and (2) detecting the presence of PRNU using normalized cross-correlation. We anticipate this technology to be an essential tool for fighting piracy of motion pictures. Experimental results demonstrate the reliability and generality of our approach.

Keywords: Video authentication, photo-response non-uniformity, camcorder identification, digital video forensics

1. INTRODUCTION

Digital video and digital TV continue to replace their analog counterparts in all aspects of human endeavor, including professional cinematography, home video, and surveillance cameras. With increasing bandwidth and decreasing price for storage and acquisition, sharing digital video over the Internet becomes increasingly more popular. Unfortunately, these advancements in technology also create problems with illegal copying and re-distribution. Digital camcorders are used by pirates in movie theaters to obtain copies of reasonable quality that are subsequently sold on a black market and transcoded to low bit-rates for illegal distribution over the Internet. This causes significant loss of revenues to the movie industry. Dan Glickman, Chairman and CEO of the Motion Picture Association, Inc. (MPAA) states in his Worldwide study of losses to the Film industry & international economies Due to piracy (available from `http://www.slyck.com/misc/mpaa_loss.`

---

[*] fridrich@binghamton.edu; phone +001 607 777 6177; fax +001 607 777 4464.

doc): "The film industry is a thriving economic engine that generates jobs and exports in countries all over the world. We are calling on governments internationally to continue to work with us in limiting the impact of piracy on local economies and the film industry. Movies are a valuable product and intellectual property must be respected." The soon-to-be-established consortium Movielabs is intended to provide funds to researchers working on camcorder detection and jamming.

Forensic methods capable of determining that two clips came from the same camcorder or that two transcoded versions of one movie have a common source will obviously help investigators draw connections between different entities or subjects and may become a crucial piece of evidence in prosecuting the pirates. Reliable, inexpensive, and fast identification of the source of digital video can also help the law enforcement with prosecution of child pornographers.

Previously, Kurosawa [2] proposed to use defective pixels and the dark current of CCD chips for camcorder identification. This approach is rather limited because dark current can only be extracted from dark frames. Another problem is that dark current is a relatively weak signal that does not survive video compression. Other recently proposed methods [3–5] might be used to identify camcorders from video clips by detecting traces of image processing unique to a specific camcorder model. Such methods, however, cannot distinguish between camcorders of the same model and thus have limited use in criminal cases.

In this paper, we adopt the techniques developed in [1] that identify individual *imaging sensors* using the photo-response non-uniformity noise. The PRNU is caused primarily by varying sensitivity of individual pixels to light due to inhomogeneity and impurities in silicon wafers and imperfections introduced by the sensor manufacturing process. The properties of the PRNU appear to be constant in time [1] and unique for each imaging sensor. Moreover, the PRNUs from different sensors are orthogonal (uncorrelated). The PRNU is not affected by light refraction on dust particles, optical surfaces, and optical zoom setting.

It is not possible to use the approach in [1] directly to identify a digital camcorder from a single video frame because the spatial resolution of the video is usually much smaller than for typical still images and each frame is highly compressed by complex compression systems (MPEG-x, H.26x, and their variants). In this paper, by taking advantage of the time resolution that is unique to video, we demonstrate that even at very low bit-rates and across various video formats, the PRNUs can be estimated and used to identify digital camcorders.

We start the description of the camcorder identification technique in Section 2 by introducing a simplified model of the imaging sensor output. Then, in Section 3 we describe the process for estimating the PRNU from a sequence of video frames. In Section 4, the source camcorder identification method based on normalized cross-correlation is described in detail and its performance tested in Section 5. Section 6 concludes the paper and outlines future research directions.

We reserve boldface font, e.g., $\mathbf{X}$ and $\mathbf{Y}$, for matrices with $\mathbf{X}[i, j]$ denoting the (i, j)-th element of $\mathbf{X}$. Everywhere in this paper unless specified otherwise, all operations among matrices, such as product, ratio, raising to a power,

etc., are *elementwise*. The dot product of matrices is $\mathbf{X} \odot \mathbf{Y} = \sum_{i=1}^{m} \sum_{j=1}^{n} \mathbf{X}[i,j]\mathbf{Y}[i,j]$ with $\|\mathbf{X}\| = \sqrt{\mathbf{X} \odot \mathbf{X}}$ being the norm of $\mathbf{X}$. The normalized correlation between $\mathbf{X}$ and $\mathbf{Y}$ is

$$corr(\mathbf{X},\mathbf{Y}) = \frac{(\mathbf{X}-\bar{\mathbf{X}}) \odot (\mathbf{Y}-\bar{\mathbf{Y}})}{\|\mathbf{X}-\bar{\mathbf{X}}\| \cdot \|\mathbf{Y}-\bar{\mathbf{Y}}\|}.$$

2. IMAGING SENSOR OUTPUT MODEL

The processing chain for the video signal in digital camcorders is quite complex and may vary greatly for camcorders from different manufacturers. It includes the quantization of the analog signal, white balance, demosaicking (color interpolation), color correction, gamma correction, filtering, and compression, for example into the VOB (MPEG 2) format. In this paper, we use a simplified model [6] that captures the most essential elements of typical in-camera processing. This enables us to develop a low-complexity camcorder identification procedure applicable to a wider spectrum of camcorders.

Let $\mathbf{I}[i,j]$ be the signal in one color channel at pixel (i,j), $i, = 1, \ldots, m$, $j = 1, \ldots, n$, for a specific frame generated by the sensor before demosaicking is applied and $\mathbf{Y}[i,j]$ the incident light intensity at pixel (i,j). Dropping the pixel indices for better readability, the model of the sensor output is

$$\mathbf{I} = g^{\gamma} \cdot \left[(1+\mathbf{K})\mathbf{Y} + \Lambda + \Theta_s + \Theta_r\right]^{\gamma} + \Theta_q, \tag{1}$$

where g is the color channel gain, γ is the gamma correction factor (typically, $\gamma \approx 1/2.2$), $\mathbf{K}$ is a zero-mean multiplicative factor responsible for PRNU, and Λ, Θ_s, Θ_r, and Θ_q stand for the following noise sources – dark current, shot noise, read-out noise, and quantization (lossy compression) noise, respectively. We remind that all operations in (1) are element-wise. Because the dominant term in the square bracket in (1) is the light intensity $\mathbf{Y}$, we can factor it out and use Taylor expansion. Keeping only the first order terms, $(1+x)^{\gamma} \cong 1 + \gamma x$, we obtain from (1)

$$\mathbf{I} = \mathbf{I}^{(0)} + \gamma\mathbf{I}^{(0)}\mathbf{K} + \Theta, \tag{2}$$

where $\mathbf{I}^{(0)} = (g\mathbf{Y})^{\gamma}$ is the sensor output in the absence of noise or lossy compression (noise-free frame); Θ is a complex of independent noise components. As previously shown [1], the PRNU factor $\mathbf{K}$ can be used as a fingerprint that characterizes each imaging sensor and for identification and integrity verification [7].

3. PRNU ESTIMATION AND DETECTION

Camcorder identification can be formulated as a joint estimation and detection problem. It involves two major statistical signal processing procedures, which are (1) estimating the PRNUs from individual videos; (2) determining the common origin by establishing the presence of the same PRNUs. We first describe the details of estimating $\mathbf{K}$.

The first step is host signal rejection to improve the SNR between the signal of interest and observed data. We

suppress the influence of the noise-free frame $\mathbf{I}^{(0)}$ by subtracting from both sides of (2) an estimate $\hat{\mathbf{I}}^{(0)} = F(\mathbf{I})$ of $\mathbf{I}^{(0)}$ obtained using a denoising filter F

$$\mathbf{W} = \mathbf{I} - \hat{\mathbf{I}}^{(0)} = \gamma\hat{\mathbf{I}}^{(0)}\mathbf{K} + \mathbf{I}^{(0)} - \hat{\mathbf{I}}^{(0)} + \gamma(\mathbf{I}^{(0)} - \hat{\mathbf{I}}^{(0)})\mathbf{K} + \mathbf{\Theta}, \quad \text{or}$$
$$\mathbf{W} = \gamma\hat{\mathbf{I}}^{(0)}\mathbf{K} + \mathbf{\Xi}. \tag{3}$$

We use a wavelet based denoising filter [8] that extracts Gaussian noise $\mathbf{W} = \mathbf{I} - \hat{\mathbf{I}}^{(0)}$ of a given variance $\sigma_{\mathbf{W}}^2$. The term $\mathbf{\Xi}$ is a combination of $\mathbf{\Theta}$ with the additional distortion introduced by the denoising filter. Working with the noise residual significantly improves the SNR for our signal of interest $\hat{\mathbf{I}}^{(0)}\mathbf{K}$ and thus improves the reliability of the camcorder identification process.

Let us assume that we have a video clip consisting of N frames $\mathbf{I}_1, \ldots, \mathbf{I}_N$ from a given camcorder. From (3), we have for each frame index $k = 1, \ldots, N$

$$\frac{\mathbf{W}_k}{\gamma\hat{\mathbf{I}}_k^{(0)}} = \mathbf{K} + \frac{\mathbf{\Xi}_k}{\gamma\hat{\mathbf{I}}_k^{(0)}}, \quad \mathbf{W}_k = \mathbf{I}_k - \hat{\mathbf{I}}_k^{(0)}, \; \hat{\mathbf{I}}_k^{(0)} = F(\mathbf{I}_k). \tag{4}$$

Program streams, such as DVD and most videos transcoded for Internet, usually use variable bit rate coding (VBR) that compresses the video sequence as much as possible to a constant picture quality. Thus, the variance of $\mathbf{\Xi}_k$ in (4) should be approximately constant across the frames independently of their type (I/P/B frame, smooth-area frame/active-area frame, etc.). On the other hand, *transport streams*, such as DTV and broadcasting streams, use constant bit rate coding (CBR) that generates bit streams with constant bit rate but variable quality causing the variance of $\mathbf{\Xi}_k$ to be frame-dependent. In this case, adaptively adjusting the variance according to the quality of different type of frames or carefully selecting the frames might give us some gain in estimating the PRNU $\mathbf{K}$. We do not expect this gain, however, to be significant. Moreover, treating all frames equally by assuming that the variance of $\mathbf{\Xi}_k$ does not depend on the frame index k greatly simplifies the estimation.

Assuming that for each pixel (i, j) the sequence $\mathbf{\Xi}_k[i, j]$ (in k) is WGN (white Gaussian noise) with variance σ^2, from (4) we can derive the MLE estimator of $\mathbf{K}$ given the measured data $\mathbf{W}_k/(\gamma\hat{\mathbf{I}}_k^{(0)})$ as

$$\gamma\hat{\mathbf{K}} = \frac{\sum_{k=1}^{N} \mathbf{W}_k\hat{\mathbf{I}}_k^{(0)}}{\sum_{k=1}^{N} (\hat{\mathbf{I}}_k^{(0)})^2}. \tag{5}$$

Because the observed data depends linearly on the unknown parameter, the MLE estimator is MVU (Minimum Variance Unbiased) and we obtain its variance from the Cramer-Rao Lower Bound

$$var(\hat{\mathbf{K}}) = \frac{\sigma^2}{\sum_{k=1}^{N} (\hat{\mathbf{I}}_k^{(0)})^2}. \tag{6}$$

Detailed derivation of (5) and (6) can be found in [9]. The estimator variance (6) provides us with some insight into how the estimation quality of $\mathbf{K}$ depends on the number and quality of video frames. We have the following two observations that we confirm experimentally in Section 5 through simulations.

(1) Under the same level of quality ($\text{var}(\Xi)$ is constant), the variance of the estimated PRNU is proportional to $1/N$. Thus, the estimation is more accurate when more video frames are used.

(2) On the other hand, if the total number of frames is fixed, videos of low quality will give us worse PRNU estimation than those of high quality because their quantization noise variance σ^2 is higher.

4. VIDEO SOURCE IDENTIFICATION USING CROSS-CORRELATION

In this section, we describe a method that can be used to decide whether two video-clips A and B were produced by the exact same camcorder. Let $\mathbf{K}_A$ and $\mathbf{K}_B$ be the PRNUs estimated from both clips. Because the PRNU is a unique signature of the camera, the task of origin identification is equivalent to discriminating $\hat{\mathbf{K}}_A$ from $\hat{\mathbf{K}}_B$. Due to estimation errors and varying quality and length of the video clips, the accuracy of the estimated PRNUs $\hat{\mathbf{K}}_A$ and $\hat{\mathbf{K}}_B$ might also vary. Moreover, there might be a translational shift (a, b) between $\hat{\mathbf{K}}_A$ and $\hat{\mathbf{K}}_B$, e.g., due to letterboxing. Hence, we capture the camcorder identification problem as simple binary hypothesis testing

$$H_0: \quad \hat{\mathbf{K}}_B[i,j] = \xi[i,j]$$

$$H_1: \quad \hat{\mathbf{K}}_B[i,j] = \hat{\mathbf{K}}_A[i-a, j-b] + \xi[i,j], \tag{7}$$

where ξ is a WGN with unknown variance. It is known that for this type of problem [10], the optimal detector is the normalized cross-correlation (NCC). In summary, to decide whether two estimated PRNUs $\hat{\mathbf{K}}_A$ and $\hat{\mathbf{K}}_B$ were obtained by the same camcorder, we first calculate the NCC between $\hat{\mathbf{K}}_A$ and $\hat{\mathbf{K}}_B$:

$$\mathbf{C}[u,v] = corr(\hat{\mathbf{K}}_A[i,j], \hat{\mathbf{K}}_B[i-u, j-v]). \tag{8}$$

Then, we examine the NCC surface $\mathbf{C}[u, v]$ and decide H_1 (e.g., both clips were taken by the same camcorder) by detecting the presence of a pronounced peak in $\mathbf{C}[u, v]$, which can be done using several different measures [11]. In this paper, we use the Peak to Correlation Energy (PCE) ($\mathcal{N}_{peak}$ is a small neighborhood of the peak)

$$PCE = \frac{\mathbf{C}[u_{peak}, v_{peak}]^2}{\frac{1}{mn - |\mathcal{N}_{peak}|} \sum_{(u,v) \notin \mathcal{N}_{peak}} \mathbf{C}[u,v]^2}. \tag{9}$$

4.1 Removing blockiness artifacts

Because PRNUs from two different sensors should be uncorrelated [1], if both clips are indeed from the same camcorder, we expect to see a sharp peak in $\mathbf{C}[u, v]$ (large PCE), otherwise $\mathbf{C}[u, v]$ will look like a low energy random noise. However, almost all camcorders use DPCM-Block DCT transform-type video coding, such as MPEG-x and H.26x. This creates (i) ringing artifacts at the frame boundaries caused by the padding required for frame dimensions not divisible by the block size and by operations such as motion estimation/compensation for out of frame movement; (ii) 16×16 blockiness artifacts inside the frame because most standard codecs are based on 16×16 macroblocks. These periodic pulse-like signals (see Fig. 1 (a)) propagate through the denoising filter into the estimated PRNUs and cause false correlations between otherwise uncorrelated PRNUs. Thus, they must be removed before calculating the NCC[1]. The boundary artifacts can be easily removed by cropping ~8 pixel

---

[1] Failure to remove the artifacts would result in substantially increased correlation between two unmatched

wide boundaries in the spatial domain. We remove the periodic pulse-like blockiness artifacts in the Fourier domain (see Fig. 1 (b)) by attenuating the Fourier coefficients at frequencies where most of the artifacts' energy is located. To illustrate how to locate the frequencies of these periodic pulse-like signals, let us consider the following one-dimensional periodic signal $x(n) = \delta(n-16m), \quad 0 \le n \le N-1$ whose DFT transform is $X(r)$

$$| X(r) |= \frac{\sin\left(\dfrac{2\pi}{N/16}\dfrac{k}{2}r\right)}{\sin\left(\dfrac{2\pi}{N/16}\dfrac{r}{2}\right)}, \tag{10}$$

where $k = \lfloor (N-1)/16 \rfloor$ and r is the DFT index. Equation (10) shows that the energy of $|X(r)|$ concentrates around frequencies of integer multiples of $N/16$. Therefore, setting $X(r) = 0$ for those frequencies and their neighborhood (3–6 times frequency resolution) effectively reduces the strength of the periodic signal. In our work, we used a similar idea to design an FFT domain filter to mitigate the deteriorating effect of blockiness on the NCC. Fig. 1(b) and (c) show the Fourier magnitude of the PRNU and the filtered PRNU. Since in practice the NCC is calculated in the Fourier domain, we can conveniently perform blockiness removal at the same time. Furthermore, we might remove other artifacts that manifest themselves as peaks in the Fourier domain, such as artifacts due to color filter array interpolation and other hardware or software operations [Section 7 in 9].

| (a) | (b) | (c) |
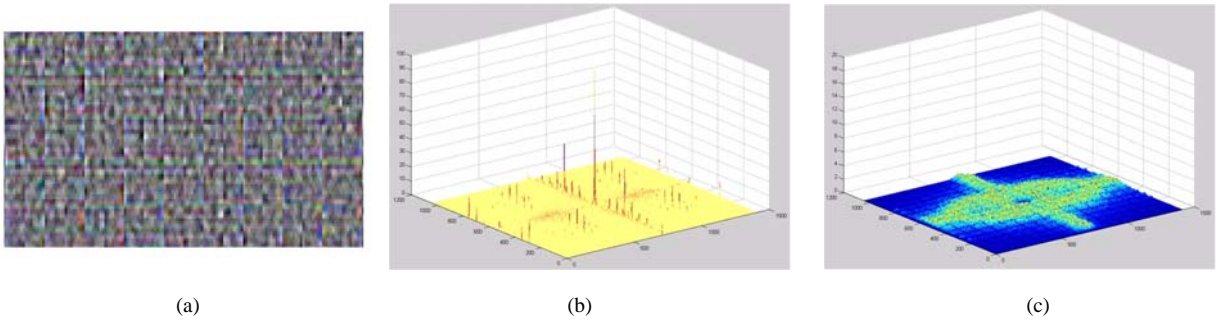
Figure 1. (a) Blockiness artifacts in a small magnified portion of the estimated PRNU; (b) Fourier magnitude of (a); (c) Fourier magnitude after removing the artifacts in the DFT domain.

5.   EXPERIMENTAL RESULTS

In this section, we present selected experiments to illustrate the effectiveness of the proposed approach in identifying the origin of video clips. Twenty-five consumer digital camcorders are used (20 SONY, 4 Hitachi, 1 Canon). The recording media was Mini-DV or DVD-RW and the sensor resolution varied from 0.68MP–4.1 MP. We selected three camcorders (one Canon DC40 and two camcorders of the same model SONY DCR-DVD105) and tested them against the remaining clips. We will address the two SONY camcorders as SONY DCR-1 and SONY DCR-2. With each camcorder, we prepared several high quality video clips (roughly 6 Mb/sec, DVD quality, resolution 536×720, frame rate 30 Hz, MPEG-2 VOB format) of various indoor and outdoor scenes. The clips contained brief periods of optical zooming in/out and panning. Some of the videos contained quickly

PRNUs, which would lead to an increased false acceptance rate.

moving objects (e.g., cars) while others had panned static scenes. All the camcorders had their Electronic Image Stabilization (EIS) and digital zooming turned off. All scenes were taped with the fully automatic settings.

The videos were also transcoded to low-bit rate formats, such as the MPEG-4 XviD format (~1Mbit/sec), the RealPlay format (~750 Kbit/sec), and the MPEG-4 DivX format (~450 Kbit/sec). These formats represent the most popular choices for distribution of video over the Internet today.
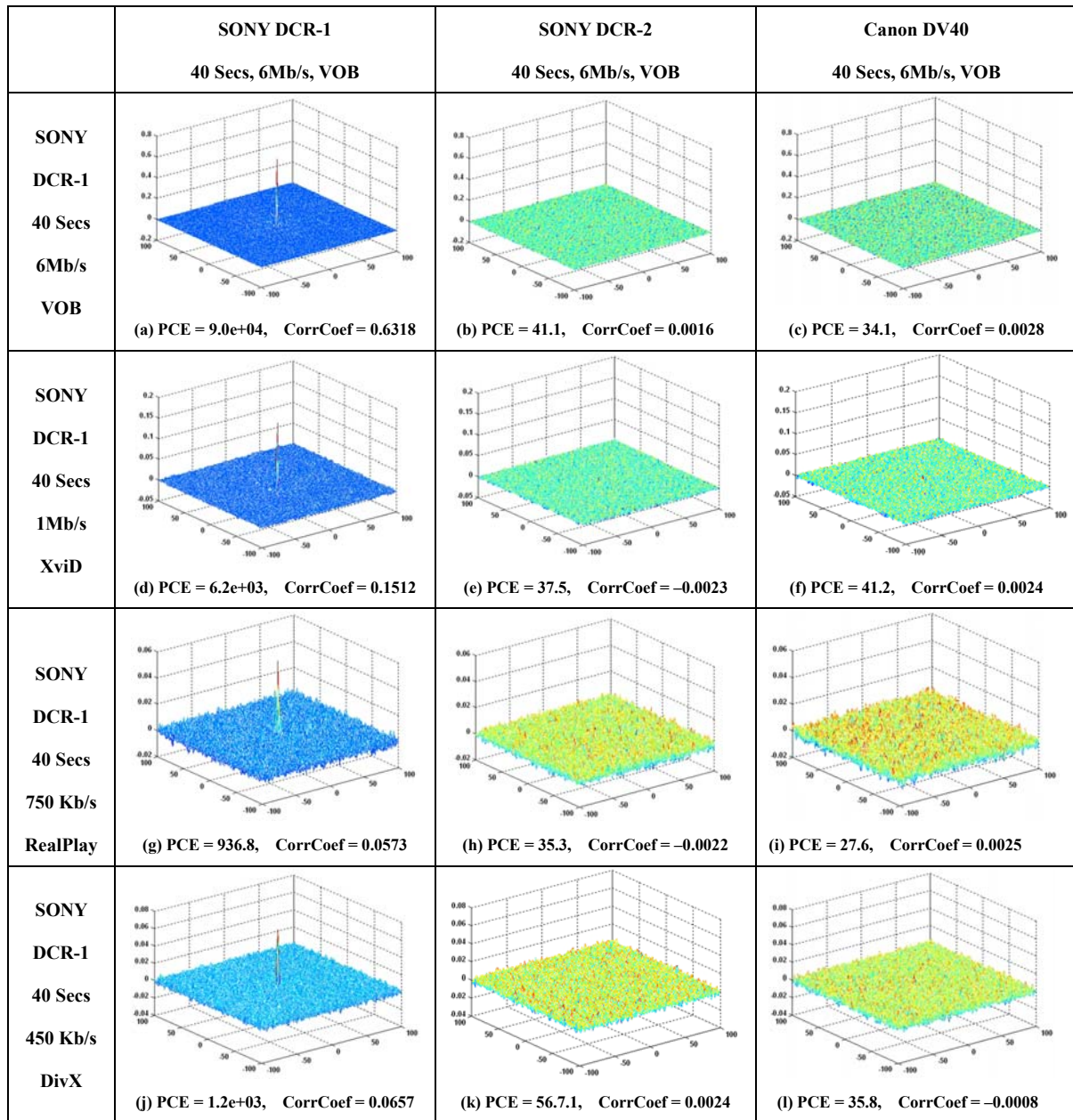
5.1 VOB, XviD, RealPlay, DivX vs. VOB

In this test, we investigated whether it is possible to correctly identify the source camera from videos that were transcoded to 4 different formats and bit-rates. We first estimated the PRNUs from a 40-second randomly selected video segment from SONY DCR-1 clips in the VOB format and from its three transcoded formats, Xvid, RealPlay, and DivX, obtaining thus four SONY DCR-1 PRNUs of varying quality. Then, we calculated the NCC with the PRNUs from a different 40-second SONY DCR-1 video clip in the VOB format and 24 PRNUs from 24 40-second video clips from all the other camcorders, also in the VOB format. For the SONY DCR-1, SONY DCR-2, and Canon DC40 camcorders, we show the NCC surface and the PCE in a pictorial form in Fig. 2. The results for the remaining 22 camcorders are summarized in the table below the figure. In the same manner, two 40-second randomly selected SONY DCR-2 clips and Canon DC40 clips were randomly chosen and tested against all the PRNUs from the 25 camcorders (obtained from VOBs). The results are shown in the same format in Fig. 2 (b) and Fig. 2(c). The figures reveal the reliability of the proposed identification approach for all four bit rates and also support observation (1) from Section 3 that with the same number of frames, the quality of the estimated PRNUs decreases as the video quality decreases (measured by the bit rate). The degradation of the estimated PRNUs is the reason for deterioration of the NCC surface (and the decrease in PCE and correlation coefficient). Regardless of the video format, the PCE and the correlation coefficients obtained for the matched case are by several orders of magnitude larger than for the unmatched case.

5.2 Xvid vs. Xvid for clips of different length

In the second experiment, we estimated two PRNUs from two 40-second SONY DCR-2 video clips of different scenes in the XviD-format and calculated the NCC between them. Then, we repeated the same process but increased the length of the clips to 80 seconds and 120 seconds. The resulting NCCs are shown in Fig. 3, which verifies observation (2) made in Section 3: with a constant video quality, the PRNU estimation improves with the increased number of frames.

5.3 Low bit-rate experiment

The third experiment we carried out targeted identification of "Internet-quality" clips with low resolution and very low bit-rate. We took two clips, one using SONY DCR-1 and one with Canon DC40 at LP resolution of 264×352 pixels and then transcoded both clips to 150kb/sec. in the RMVB format. Then we tested both clips for the presence of a PRNU estimated from four 2.5min VOB clips from SONY DCR-1. The NCC surfaces and PCEs are shown in Figure 4. The identification is again possible and improves with the length of the clip.

	SONY DCR-1 40 Secs, 6Mb/s, VOB	SONY DCR-2 40 Secs, 6Mb/s, VOB	Canon DV40 40 Secs, 6Mb/s, VOB
SONY DCR-1 40 Secs 6Mb/s VOB	(a) PCE = 9.0e+04, CorrCoef = 0.6318	(b) PCE = 41.1, CorrCoef = 0.0016	(c) PCE = 34.1, CorrCoef = 0.0028
SONY DCR-1 40 Secs 1Mb/s XviD	(d) PCE = 6.2e+03, CorrCoef = 0.1512	(e) PCE = 37.5, CorrCoef = –0.0023	(f) PCE = 41.2, CorrCoef = 0.0024
SONY DCR-1 40 Secs 750 Kb/s RealPlay	(g) PCE = 936.8, CorrCoef = 0.0573	(h) PCE = 35.3, CorrCoef = –0.0022	(i) PCE = 27.6, CorrCoef = 0.0025
SONY DCR-1 40 Secs 450 Kb/s DivX	(j) PCE = 1.2e+03, CorrCoef = 0.0657	(k) PCE = 56.7.1, CorrCoef = 0.0024	(l) PCE = 35.8, CorrCoef = –0.0008

Other 22 camcorders		SONY DCR-1 40 Secs, 6Mb/s VOB		SONY DCR-1 40 Secs, 1Mb/s XviD		SONY DCR-1 40 Secs, 750 Kb/s RP		SONY DCR-1 40 Secs, 450 Kb/s DivX	
		CorrCoef	PCE	CorrCoef	PCE	CorrCoef	PCE	CorrCoef	PCE
Statistics	Min	–0.0041	28.0	–0.0044	28.2	–0.0053	26.8	–0.0035	25.5
	Max	0.0084	89.0	0.0045	90.3	0.0046	73.9	0.0050	156.9
	Median	–0.0004	43.2	–0.0005	37.3	0.0012	32.5	0.0007	38.7

Figure 2 (a). NCC of PRNUs of 4 differently transcoded versions of a SONY DCR-1 clip with PRNUs estimated from 25 camcorders in the VOB format.

	SONY DCR-1 **40 Secs, 6Mb/s, VOB**	**SONY DCR-2** **40 Secs, 6Mb/s, VOB**	**Canon DV40** **40 Secs, 6Mb/s, VOB**
SONY **DCR-2** **40 Secs** **6Mb/s** **VOB**	**(a) PCE = 38.0, CorrCoef = 0.0028**	**(b) PCE = 8.2e+04, CorrCoef = 0.6464**	**(c) PCE = 38.8, CorrCoef = 0.0058**
SONY **DCR-2** **40 Secs** **1Mb/s** **XviD**	**(d) PCE = 32.8, CorrCoef = 0.0006**	**(e) PCE = 1.0e+04, CorrCoef = 0.2038**	**(f) PCE = 34.8, CorrCoef = 0.0006**
SONY **DCR-2** **40 Secs** **750 Kb/s** **RealPlay**	**(g) PCE = 28.1, CorrCoef = 0.0005**	**(h) PCE = 2.1e+03, CorrCoef = 0.0867**	**(i) PCE = 28.0, CorrCoef = 0.0013**
SONY **DCR-2** **40 Secs** **450 Kb/s** **DivX**	**(j) PCE = 37.2, CorrCoef = –0.0016**	**(k) PCE = 1.9e+03, CorrCoef = 0.0871**	**(l) PCE = 33.6, CorrCoef = –0.0018**

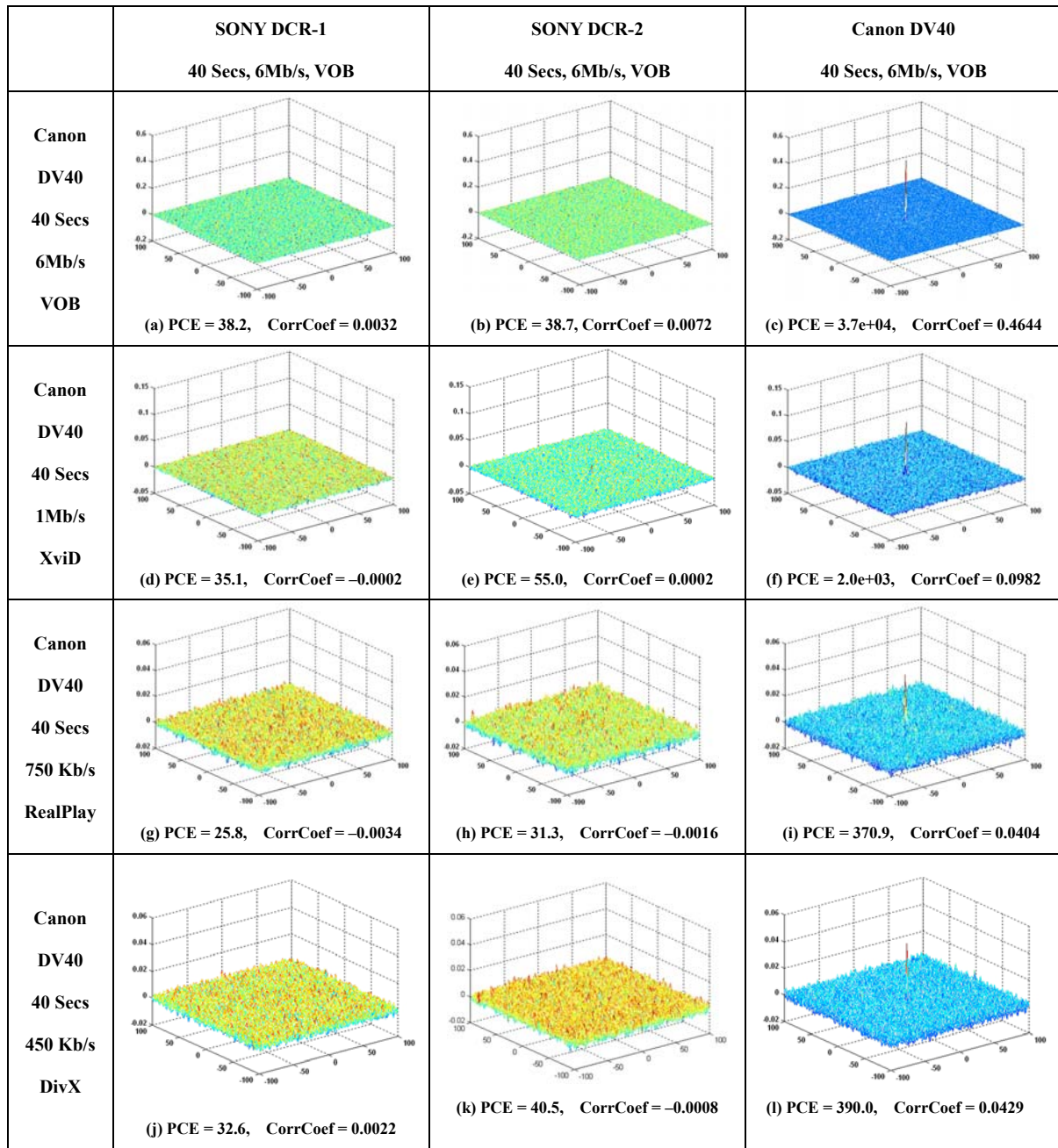| **Other 22** **camcorders**		SONY DCR-2 40 Secs, 6Mb/s VOB		SONY DCR-2 40 Secs, 1Mb/s XviD		SONY DCR-2 40 Secs, 750 Kb/s RP		SONY DCR-2 40 Secs, 450 Kb/s, DivX	
		CorrCoef	PCE	CorrCoef	PCE	CorrCoef	PCE	CorrCoef	PCE
Statistics	Min	–0.0070	22.6	–0.0058	29.5	–0.0060	28.4	–0.0044	27.6
	Max	0.0059	73.7	0.0051	100.8	0.0059	116.5	0.0065	61.6
	Median	–0.0022	43.1	–0.0015	38.5	–0.0005	39.3	0.0005	35.8

Figure 2 (b). NCC of PRNUs of 4 differently transcoded versions of a SONY DCR-2 clip with PRNUs estimated from 25 camcorders in the VOB format.
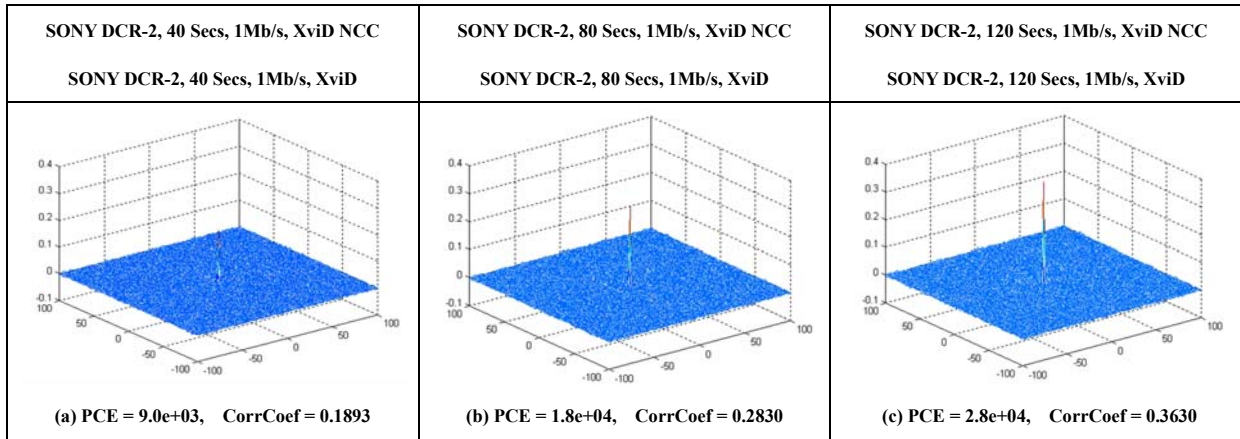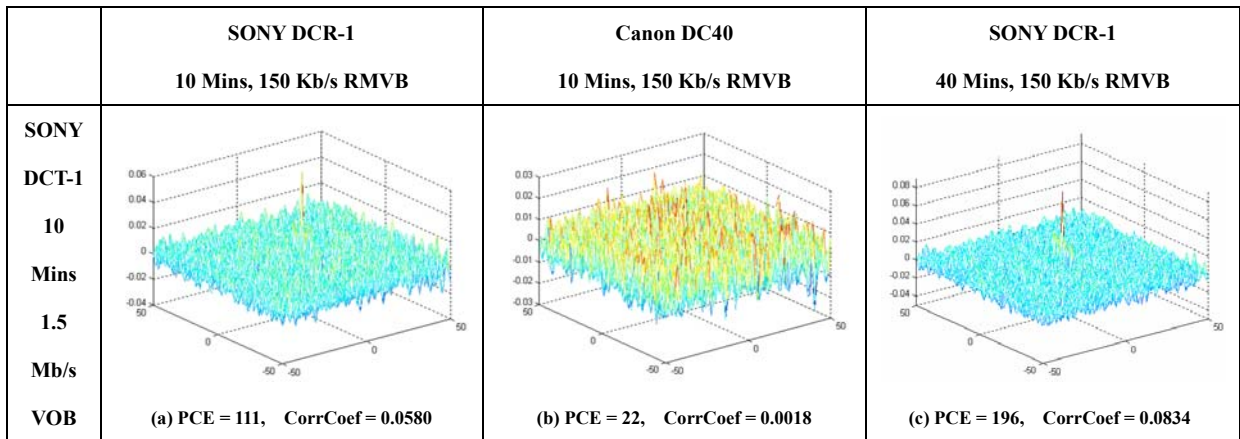
	SONY DCR-1 40 Secs, 6Mb/s, VOB	SONY DCR-2 40 Secs, 6Mb/s, VOB	Canon DV40 40 Secs, 6Mb/s, VOB
Canon DV40 40 Secs 6Mb/s VOB	(a) PCE = 38.2, CorrCoef = 0.0032	(b) PCE = 38.7, CorrCoef = 0.0072	(c) PCE = 3.7e+04, CorrCoef = 0.4644
Canon DV40 40 Secs 1Mb/s XviD	(d) PCE = 35.1, CorrCoef = –0.0002	(e) PCE = 55.0, CorrCoef = 0.0002	(f) PCE = 2.0e+03, CorrCoef = 0.0982
Canon DV40 40 Secs 750 Kb/s RealPlay	(g) PCE = 25.8, CorrCoef = –0.0034	(h) PCE = 31.3, CorrCoef = –0.0016	(i) PCE = 370.9, CorrCoef = 0.0404
Canon DV40 40 Secs 450 Kb/s DivX	(j) PCE = 32.6, CorrCoef = 0.0022	(k) PCE = 40.5, CorrCoef = –0.0008	(l) PCE = 390.0, CorrCoef = 0.0429

Other 22 camcorders		Canon DV40 40 Secs, 6Mb/s VOB		Canon DV40 40 Secs, 1Mb/s XviD		Canon DV40 40 Secs, 750 Kb/s RP		Canon DV40 40 Secs, 450 Kb/s DivX	
		CorrCoef	PCE	CorrCoef	PCE	CorrCoef	PCE	CorrCoef	PCE
Statistics	Min	–0.0058	26.8	–0.0033	28.7	–0.0045	32.5	–0.0041	26.0
	Max	0.0111	121.1	0.0080	122.3	0.0050	94.5	0.0039	82.2
	Median	0.0021	56	–0.0012	38.4	–0.0013	32.4	–0.0005	39.9

Figure 2 (c). NCC of PRNUs of 4 differently transcoded versions of a Canon clip with PRNUs estimated from 25 camcorders in the VOB format.

SONY DCR-2, 40 Secs, 1Mb/s, XviD NCC	SONY DCR-2, 80 Secs, 1Mb/s, XviD NCC	SONY DCR-2, 120 Secs, 1Mb/s, XviD NCC
SONY DCR-2, 40 Secs, 1Mb/s, XviD	SONY DCR-2, 80 Secs, 1Mb/s, XviD	SONY DCR-2, 120 Secs, 1Mb/s, XviD

(a) PCE = 9.0e+03, CorrCoef = 0.1893	(b) PCE = 1.8e+04, CorrCoef = 0.2830	(c) PCE = 2.8e+04, CorrCoef = 0.3630

Figure 3. NCCs of PRNUs from different SONY DCR-2 XviD-format video clips with the length 40, 80, and 120 seconds.

	SONY DCR-1	Canon DC40	SONY DCR-1
	10 Mins, 150 Kb/s RMVB	10 Mins, 150 Kb/s RMVB	40 Mins, 150 Kb/s RMVB
SONY DCT-1 10 Mins 1.5 Mb/s VOB			

	(a) PCE = 111, CorrCoef = 0.0580	(b) PCE = 22, CorrCoef = 0.0018	(c) PCE = 196, CorrCoef = 0.0834

Figure 4. NCC surface and PCE coefficient for two low-resolution, low bit-rate clips from SONY DCR-1 and Canon DC40 with PRNU estimated from a 10 minute VOB clip from SONY DCR-1. 4a) is for a 10 minute clip from SONY DCR-1 and 4b) for 40 minute clip.

6. CONCLUSIONS

We present a new approach to the problem of digital camcorder identification from digital videos. The identification is based on the imaging sensor photo-response non-uniformity (PRNU), which is a unique fingerprint of imaging sensors. The proposed method can verify whether two video clips came from the same camcorder. First, the PRNU is estimated from both clips using the Maximum Likelihood Estimator. Then the PRNUs are filtered to remove the blockiness artifacts due to lossy compression. Finally, they are processed using the normalized cross-correlation. The Peak to Correlation Energy coefficient is used to establish the common origin of both PRNUs. Experiments with 25 camcorders show that only 40 seconds of video is sufficient for a very reliable decision from clips encoded as low as 450kb/sec. With decreasing video quality (larger compression) and decreasing spatial resolution, the length of the video clip necessary for reliable decision must be increased. For "Internet quality" videos in LP resolution (264×352) and 150 kb/sec. bit-rate, we obtained good results with clips of length 10 minutes.

ACKNOWLEGEMENT

REFERENCES

[1] J. Lukáš, J. Fridrich, and M. Goljan, "Digital Camera Identification from Sensor Noise," *IEEE Transactions on Information Security and Forensics* **1**(2): 205–214, June 2006.

[2] 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.

[3] M. Kharrazi, H.T. Sencar, and N.M. Memon, "Blind Source Camera Identification," *Proc. ICIP' 04*, Singapore, October 24–27, 2004.

[4] A.C. Propescu and H. Farid, "Statistical Tools for Digital Forensic," in J. Fridrich (ed.): *6th International Workshop on Information Hiding*, LNCS vol. 3200, Springer-Verlag, Berlin-Heidelberg, New York, pp. 128–147, 2004.

[5] A. Swaminathan, M. Wu, and K.J.R. Liu, "Non-intrusive Forensic Analysis of Visual Sensors Using Output Images," *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP' 06)*, May 2006.

[6] G.E. Healey and R. Kondepudy, "Radiometric CCD Camera Calibration and Noise Estimation," *IEEE Trans. Image Process* **16**(3): 267–276, March 1994.

[7] J. Lukáš, J. Fridrich, and M. Goljan, "Detecting Digital Image Forgeries Using Sensor Pattern Noise," *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VIII*, vol. 6072, San Jose, CA, January 16–19, pp. 0Y1–0Y11, 2006.

[8] M.K. Mihcak, I. Kozintsev, and K. Ramchandran, "Spatially Adaptive Statistical Modeling of Wavelet Image Coefficients and its Application to Denoising," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 6, Phoenix, AZ, March, 1999, pp. 3253–3256.

[9] M. Chen, J. Fridrich, and M. Goljan, "Digital Imaging Sensor Identification (Further Study)," *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, San Jose, CA, January 28–February 1, 2007.

[10] C.R. Holt, "Two-Channel Likelihood Detectors for Arbitrary Linear Channel Distortion," *IEEE Transactions on Acoustics, Speech, and Signal Processing* **ASSP-35**(3): 267–273, March 1987.

[11] B.V.K. Vijaya Kumar and L. Hassebrook, "Performance Measures for Correlation Filters," *Appl. Opt.* **29**(20): 2997–3006, July 1990.