USING SENSOR PATTERN NOISE FOR CAMERA MODEL IDENTIFICATION

Tomáš Filler, Jessica Fridrich, Miroslav Goljan

Dept. of Electrical and Computer Engineering, SUNY Binghamton, Binghamton, NY, USA
{tomas.filler,fridrich,mgoljan}@binghamton.edu

ABSTRACT

Sensor photo-response non-uniformity (PRNU) was introduced by Lukáš et al. [1] to solve the problem of digital camera sensor identification. The PRNU is the main component of a camera fingerprint that can reliably identify a *specific* camera. This fingerprint can be estimated from multiple images taken by the camera. In this paper, we demonstrate that the same fingerprint can be used for identification of camera brand and model. This is possible due to the fact that fingerprints estimated from images in the TIFF/JPEG format contain local structure due to various in-camera processing that can be detected by extracting a set of numerical features from the fingerprints and classifying them using pattern classification methods. We estimate and classify fingerprints for more than 4500 digital cameras spanning 8 different brands and 17 models. The average probability of correctly classified camera brand was 90.8%.

Index Terms— Digital forensic, camera model identification, pattern noise

1. INTRODUCTION

The trustworthiness of a digital image presented as silent witness in court has recently been questioned. This is mainly due to ease with which digital images can be manipulated using common image editing software. Often, the issue at question is the image origin. Proof that a given digital image was taken with a specific camera or a certain camera model can play a vital role whenever the digital object (e.g., image or video) is a result of a crime, such as in movie piracy cases.

Lukáš et al. [1] studied the problem of digital camera sensor identification using a sensor fingerprint based on photo-response non-uniformity (PRNU), which is a multiplicative noise that is unintentionally embedded by the digital camera into every image it takes. The authors proposed a method for estimating the fingerprint from a set of digital images and for detecting its presence in a specific image under investigation. Successful detection is indicative of the fact that the image was taken by the exact same camera. The methodology was further improved by Chen et al. [2].

Other authors proposed additional applications of this fingerprint to solve a variety of problems in digital forensic, including identification of cellular phones [3], digital camcoders [4], and scanners [5], [6]. The fingerprint can also be used for forensic classification of image origin [7] and for detection of digital forgeries in images [8].

In this paper, we show that the same fingerprint can be used to determine the camera model and brand, which is a problem previously investigated using other means [9, 10, 11]. We formulate the problem of brand/model detection as pattern classification, where each class corresponds to a different camera model. This work is motivated by the fact that the fingerprint obtained from images in the TIFF or JPEG format contains traces of in-camera processing, such as demosaicking or filtering. In fact, fingerprints are visually different across different brands due to presence of simple periodic patterns, which could be quantified and used as features.

The rest of this paper is structured as follows. In Section 2, we review a method for the fingerprint estimation from images. In Section 3, we describe the features computed from the estimated fingerprint that will be used for distinguishing camera models. The problem of collecting images for experiments is briefly mentioned in Section 4. The setup of experiments and their results are discussed in Section 5. Finally, the paper is concluded in Section 6.

2. CAMERA MODEL

The PRNU is a multiplicative noise caused by imperfections in the manufacturing process (e.g., slightly different pixel dimensions) and inhomogenities of sillicone. It has a stochastic nature and is unique to each sensor. Its high dimensionality and robustness to processing (it is a spread-spectrum signal) make it an ideal candidate for forensic aplications, such as camera identification. The PRNU must be estimated from raw sensor output. When the PRNU is estimated from images in typical viewable formats, such as TIFF or JPEG, the estimate (1), which we call the fingerprint, is already shaped by in-camera processing and thus carries information about

The work on this paper was supported by 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.

the camera brand or model.

The fingerprint is estimated using a minimum variance unbiased estimator derived from a simplified linearized model of sensor output in [2]

$$\hat{\mathbf{K}} = \frac{\sum_{i=1}^{m} \mathbf{W}_i \mathbf{I}_i}{\sum_{i=1}^{m} (\mathbf{I}_i)^2}, \tag{1}$$

where $\mathbf{I}_i, i = 1, \ldots, m$ stand for m images taken by the same camera, $\mathbf{W}_i$ is the noise residual of the i-th image, $\mathbf{W}_i = \mathbf{I}_i - \mathbf{I}_i^{(0)}$, $\mathbf{I}_i^{(0)}$ is $\mathbf{I}_i$ denoised using a denoising filter. We used a wavelet-based denoising filter as described in [1]. All operations among matrices are understood as element-wise. The number of images needed to obtain a good estimate of the fingerprint $\hat{\mathbf{K}}$ varies with the camera and the image content. In this paper, we used $m = 45$ images.

3. FEATURES

In this section, we describe the features that will be used in our pattern-classification approach to recognize camera brand and model from the fingerprint $\hat{\mathbf{K}}$. The features are designed to reflect differences in the CFA (color filter array), demosaicking algorithm, and the sensor signal transfer. Assuming the fingerprint was estimated separately in each color channel, we represent it as a three-dimensional array $\hat{\mathbf{K}} \in \mathbb{R}^{w \times h \times 3}$, where w and h are the width and height of the image in its native (highest) resolution.

Statistical moments

The first feature set is formed by the first 3 centralized sample statistical moments of the fingerprint $\hat{\mathbf{K}}$ in each color channel. This gives total of 9 features.

These features are influenced by the sensor PRNU. The estimated noise is approximately zero mean while its variance varies across different camera brands and models.

Cross-correlation

To capture local dependencies or periodicities among neighboring samples of $\hat{\mathbf{K}}$, we need to use higher-order statistical features. The local dependencies contain information about CFA, color interpolation, and processing. The color interpolation error exhibits periodicities mainly due to the periodic nature of the CFA. To describe this periodic structure, we calculate the normalized cross-correlation between color channels. For each color channel pair (C_1, C_2), $C_1, C_2 \in \{R, G, B\}$ and shift $\Delta_1 \in \{0, \ldots, 3\}$, $\Delta_2 \in \{0, \ldots, 3\}$, we calculate the normalized correlation, $\rho(\Delta_1, \Delta_2)$, between $C_1(i, j)$ and $C_2(i-\Delta_1, j-\Delta_2)$, where the normalized correlation between two matrices $\mathbf{A}, \mathbf{B}$ is defined in the usual way

$$\rho(A, B) = \frac{\sum_{i,j}(A_{i,j} - \bar{A})(B_{i,j} - \bar{B})}{\sqrt{\sum_{i,j}(A_{i,j} - \bar{A})^2 \sum_{i,j}(B_{i,j} - \bar{B})^2}}.$$

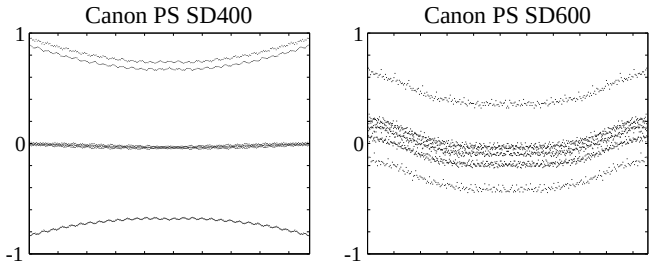

Fig. 1. Example of cyclic normalized cross-correlations for two camera models. Here we plot cyclic normalized cross-correlation of vector $\mathbf{x} \in \mathbb{R}^{h \times 1}$, where $\mathbf{x}_i$ was obtained as a mean of i-th row of red color channel of fingerprint $\hat{\mathbf{K}}$.

where $\bar{A}$, and $\bar{B}$ are sample means calculated from matrices $\mathbf{A}$ and $\mathbf{B}$. This results in $6 \times 4 \times 4 = 96$ numbers. Finally, we applied the principle component analysis (PCA) to obtain 4 features that we call principle components.

Block covariance

The most commonly used CFAs are obtained by periodically repeating blocks of 2×2 filters, which have direct impact on local dependencies among neighboring samples in $\hat{\mathbf{K}}$. We divide the whole matrix $\hat{\mathbf{K}}$ into disjoint squares of $k \times k$ pixels and consider each square as a sample of a random vector of length $3k^2$. From the whole image, we obtain $\frac{w \times h}{k^2}$ such sample vectors. Denoting their $3k^2 \times 3k^2$ covariance matrix as $\mathbf{C}$, we reshape it into a vector and compute its first 4 principal components. Using $k = 2$ and $k = 3$, we thus obtain $4 + 4 = 8$ features.

Linear-pattern cross-correlation

The last set of features is derived from the Linear Pattern as introduced in [2]. The PRNU is modeled as a collection of independent realizations of some random variable. Thus, the means of rows and columns should be zero. This is, however, not true for the estimated fingerprint. The row (or column) means thus contain useful information about in-camera processing.

We use this idea and calculate the vector $\mathbf{r} \in \mathbb{R}^{h \times 1}$, where the i-th element is the mean of the i-th row from the red channel of $\hat{\mathbf{K}}$ (we can use any color channel). This vector characterizes the systematic error of the processing algorithms in each row. To see the structure of the error, we calculate the cyclic normalized auto-correlation and plot it as a function of the shift. Figure 1 shows two examples of such plots for two Canon PowerShot cameras.

Denoting the auto-correlation vector as $\mathbf{x} \in \mathbb{R}^{h \times 1}$, we define vectors $\mathbf{x}^1, \ldots, \mathbf{x}^8$ as $\mathbf{x}^j = (\mathbf{x}_j, \mathbf{x}_{j+8}, \mathbf{x}_{j+16}, \ldots) \in \mathbb{R}^{h/8 \times 1}$. We further calculate the mean of each vector $\mathbf{x}^j$ and consider the first 4 principal components from this 8 element vector as 4 features. Visual inspection of the plots of $\mathbf{x}$ for different camera models (see Figure 1) prompted us to define

		C	F	K	M	N	O	P	S
Canon	C	93.47	*	*	*	3.46	*	*	*
Fujifilm	F	2.5	93.39	*	*	*	*	*	*
Kodak	K	*	*	96.45	*	*	*	*	*
Minolta	M	*	*	*	79.79	18.88	*	*	*
Nikon	N	2.35	*	*	4.35	86.16	5.12	*	*
Olympus	O	*	*	*	*	7.11	87.29	*	*
Panasonic	P	*	*	*	*	*	*	94.44	*
Sony	S	*	*	*	*	*	*	*	95.97

Fig. 2. Confusion matrix calculated for camera brands as an average over 8 experiments. The symbol $*$ represents values smaller than 2%.

3 more features. Considering each vector $\mathbf{x}^j$ as a curve (plot of the vector $\mathbf{x}^j$), we compute the sample mean and variance of the vector $(\mathbf{x}^{min}+\mathbf{x}^{max})/2$, where $\mathbf{x}^{min}$ and $\mathbf{x}^{max}$ are the vectors with minimal and maximal sample means. Finally, by calculating the area between $\mathbf{x}^{min}$ and $\mathbf{x}^{max}$, we obtain 7 features.

4. DATA COLLECTION

In order to test the ability of the PRNU based fingerprint to capture the camera brand or model, we need a large database of images coming from cameras of various brands and models. Also, it is very important to have multiple physically different cameras for each brand and model to avoid overtraining to a *specific* camera rather than a *class* of cameras. Simulating different cameras by dividing images from one camera to disjoint clusters brings the obvious danger of overtraining, because the estimated fingerprints would be similar.

The image sharing portal, www.flickr.com, was used as our image source. From this portal, we downloaded full-resolution images about which we assumed that they were not subjected to further geometrical processing. The camera model and brand was extracted from the EXIF header. We also made the reasonable assumption that all images posted by the same user using the same camera model were all taken by the exact same camera. These assumptions allowed us to use these images for estimation of the camera fingerprint. Only landscape oriented images were used in our experiments to avoid ambiguity in $\pm 90^\circ$ rotation.

Table 1 shows the list of available camera models along with the number of different cameras (different fingerprints).

5. EXPERIMENTAL RESULTS

We now describe the details of our experiment. For each user, we downloaded 45 randomly chosen images from which we estimated the fingerprint $\hat{\mathbf{K}}$. We made an effort to find as many different users on Flickr as possible for each camera model. In our experiment, we only used those camera models for which we could obtain at least 100 different users. Having at least 45 images for each user, it was not possible to find approximately the same number of users for given camera model. The number of different cameras per model thus varied from 101 to 650.

Camera model	# of cameras	Sensor size (MPix)
Canon PowerShot S3 IS	465	6.0
Canon PowerShot SD400	647	5.0
Canon PowerShot SD600	213	6.0
Fujifilm FinePix A345	140	4.0
Kodak CX7300	150	3.2
Kodak Z740 Zoom	245	5.0
Minolta DiMAGE XT	117	3.1
Nikon Coolpix 3200	352	3.1
Nikon Coolpix 4300	262	3.9
Nikon Coolpix 4600	394	3.9
Olympus C350 Zoom	101	3.1
Olympus Stylus 300	320	3.1
Panasonic DMC-FX01	246	6.0
Panasonic DMC-FX7	119	4.9
Panasonic DMC-FZ7	241	6.0
Sony DSC-P200	283	7.1
Sony DSC-W50	270	6.0

Table 1. List of available camera models.

We used the classical voting system and a set of $\frac{17\times 16}{2} = 136$ binary classifiers to perform the multi-classification. The following approach was used to train the binary classifiers. First, we randomly selected 70 estimated fingerprints for training and left the rest for testing. The features that require calculating PCA were evaluated on the union of both training datasets. Thus, the principal components were extracted separately for each binary classifier. Using this approach, we obtained 28 features that were specific to the binary classifier. To avoid possible overtraining, we reduced the number of features using BAHSIC feature selection method proposed by Song et al. [12] from 28 to 5. This allowed us to choose different features for different binary classifiers. Finally, the SVM classifier [13] with the RBF kernel was used for classification.

Experimental results obtained by averaging confusion matrices over 8 trials are presented in Figure 2 and in Figure 3.

6. CONCLUSION

In this paper we show that the camera brand and model can be determined from the PRNU based camera fingerprint originally proposed for identification of a specific camera. The approach is based on classification of features derived from the fingerprint. We report an average probability of correctly classified camera brand at 90.8%. We would like to point out that the presented results were obtained by using a large

		C1	C2	C3	F1	K1	K2	M1	N1	N2	N3	O1	O2	P1	P2	P3	S1	S2
Canon PS S3	C1	69.4	∗	22.3	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗
Canon PS SD400	C2	∗	95.0	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗
Canon PS SD600	C3	22.6	∗	65.2	∗	∗	∗	∗	∗	∗	5.8	∗	∗	∗	∗	∗	∗	∗
Fuji Finepix A345	F1	∗	∗	∗	93.4	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗
Kodak CX7300	K1	∗	∗	∗	∗	94.6	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗
Kodak Z740 Zoom	K2	∗	∗	∗	∗	∗	98.2	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗
Minolta Dimage XT	M1	∗	∗	∗	∗	∗	∗	79.8	10.6	∗	8.2	∗	∗	∗	∗	∗	∗	∗
Nikon Coolpix 3200	N1	∗	∗	∗	∗	∗	∗	6.8	82.2	∗	5.6	∗	∗	∗	∗	∗	∗	∗
Nikon Coolpix 4300	N2	∗	∗	∗	∗	∗	∗	∗	∗	85.5	3.9	3.9	∗	∗	∗	∗	∗	∗
Nikon Coolpix 4600	N3	∗	∗	∗	∗	∗	∗	5.8	8.3	∗	71.2	8.2	∗	∗	∗	∗	∗	∗
Olympus C350z	O1	∗	∗	∗	3.2	∗	∗	∗	3.2	∗	8.4	77.8	∗	∗	∗	∗	∗	∗
Olympus S300	O2	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	96.2	∗	∗	∗	∗	∗
Panasonic DMC-FX01	P1	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	90.9	∗	∗	∗	∗
Panasonic DMC-FX7	P2	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	95.9	∗	∗	∗
Panasonic DMC-FZ7	P3	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	90.3	∗	∗
Sony DSC-P200	S1	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	93.4	∗
Sony DSC-W50	S2	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	∗	96.1

Fig. 3. Confusion matrix obtained as an average over 8 experiments. The symbol ∗ represents values smaller than 3%.

number of different physical cameras to avoid the danger of overtraining to a cluster of specific cameras.

This work should be viewed as tool for the camera brand/model classification that complements existing approaches based on other principles, such as [9, 10, 11].

7. REFERENCES

[1] J. Lukáš, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 205–214, June 2006.

[2] M. Chen, J. Fridrich, and M. Goljan, "Digital imaging sensor identification (further study)," *in Delp III and Wong, [14]*.

[3] O. Çeliktutan, I. Avcibas, and B. Sankur, "Blind identification of cellular phone cameras," *in Delp III and Wong, [14]*.

[4] M. Chen, J. Fridrich, and M. Goljan, "Source digital camcorder identification using ccd photo response nonuniformity," *in Delp III and Wong, [14]*.

[5] T. Gloe, E. Franz, and A. Winkler, "Forensics for flatbed scanners," *in Delp III and Wong, [14]*.

[6] H. Gou, A. Swaminathan, and M. Wu, "Robust scanner identification based on noise features," *in Delp III and Wong, [14]*.

[7] N. Khanna, A. K. Mikkilineni, G. T. C. Chiu, J. P. Allebach, and E. J. Delp III, "Forensic classification of imaging sensor types," *in Delp III and Wong, [14]*.

[8] M. Chen, J. Fridrich, and M. Goljan, "Imaging sensor noise as digital x-ray for revealing forgeries," in *Proc. of 9th Information Hiding Workshop, Saint Malo, France*, June 2007.

[9] M. Kharrazi, H. T. Sencar, and N. Memon, "Blind source camera identification," in *Proc. ICIP*, October 2004, vol. 1, pp. 709–712.

[10] S. Bayram, H. T. Sencar, N. Memon, and I. Avcibas, "Source camera identification based on CFA interpolation," in *Proc. ICIP*, September 2005, vol. 3, pp. 69–72.

[11] A. Swaminathan, M. Wu, and K.J.R. Liu, "Non-intrusive component forensics of visual sensors using output images," in *IEEE Transactions on Information Forensics and Security*, March 2007, vol. 2, pp. 91–106.

[12] L. Song, A. J. Smola, A. Gretton, K. M. Borgwardt, and J. Bedo, "Supervised feature selection via dependence estimation," in *ICML*, June 2007, pp. 823–830.

[13] Chih-Chung Chang and Chih-Jen Lin, *LIBSVM: a library for support vector machines*, 2001, Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[14] E. J. Delp III and P. W. Wong, *eds., Security, Steganography, and Watermarking of Multimedia Contents IX, Proceedings of SPIE 6505, (San Jose, California, USA), SPIE and IS&T, January 28 – February 1 2007*.