Source Class Identification for DSLR and Compact Cameras

Yanmei Fang #,*1, Ahmet Emir Dirik #2, Xiaoxi Sun #, Nasir Memon #3

Dept. of Computer & Information Science
Polytechnic Institute of New York University, Brooklyn, NY, 11201, USA
1 yanmei@isis.poly.edu
2 emir@isis.poly.edu
3 memon@nyu.edu

* *School of Information Science & Technology*
Sun Yet-sen University, Guangzhou, 510275, China
1 fangym@mail.sysu.edu.cn

***Abstract*—The identification of image acquisition source is an important problem in digital image forensics. In this work, we focus on building a classifier to effectively distinguish between digital images taken from digital single lens reflex (DSLR) and compact cameras. Based on the architecture and the imaging features of DSLR and compact cameras, the images taken from different sources may have different statistical properties in both spatial and transform domains. In this work, we utilized wavelet coefficients and pixel noise statistics to model these two different source classes over 20 different digital cameras. The efficacy of the digital source class identifier, introduced in the paper, has been tested over 1000 high quality camera outputs and post-processed images (resized, re-compressed). Experimental analysis shows that the proposed method has good potential to distinguish DSLR and compact source classes.**

I. Introduction

Digital cameras are widely used in our daily lives. Point and shoot, compact cameras are easy to use and carry, due to their small weight and sizes. Digital single lens reflex (DSLRs) cameras are also getting popular very fast and being increasingly used by both professionals and ordinary users due to their falling costs, although they are bigger and heavier than the compact ones.

With the fast development of tools to manipulate multimedia data, the integrity of both content and acquisition device has become particularly important when images are used as critical evidence in journalism, reconnaissance, and law enforcement applications. So, multimedia forensics try to find some answers to image integrity and authenticity to guarantee the credibility of digital images. Such solutions would provide useful forensic information to law enforcement and intelligence agencies about which kind of camera is used to acquire an image [1], [2], [3], [4] or whether it is doctored or not.

For the source camera identification problem, several different methods have been proposed up to now. Recently, in [5], [6], the source identification problem was studied for a group of images taken with multiple cameras under controlled conditions with the same scene. In [2], the authors proposed a very successful method to identify individual imaging sensors utilizing the photo-response non-uniformity (PRNU) noise. Previously, Kurosawa [7] had proposed a unique camcorder identification method using defective pixels and dark currents of charge-coupled device (CCD) sensor. In [8], the source camera identification problem was studied for two different cameras utilizing complementary metal oxide semiconductor (CMOS) sensors. Authors reported that their method identifies the source cameras with high accuracy even for the images taken under very low and high lighting conditions.

Unlike individual camera model identification, there are a few works in the literature to distinguish image acquisition source classes. According to our best knowledge, this is the *first work* to identify DSLR and compact camera *classes* based on feature based classifiers. Available source camera identification methods utilizing PRNU noise [2], [3], color filter array (CFA) and demosaicing artifacts [4], [9] cannot be used in this problem.

Determining whether a given image is taken from a DSLR or a compact camera would help the forensic examiner very much since this information reduces the camera search space drastically. Even if the forensic examiner uses PRNU based camera identification method, he/she may deal with thousands or millions of images. Hence, testing the PRNU method on millions of images takes very long time. In such a case, the proposed camera class identifier can be used to reduce the search space and time significantly.

In this paper, we present a source camera identification scheme to distinguish digital SLR and compact camera classes. DSLR and compact cameras use different imaging sensors. For instance, DSLR cameras use larger sensors resulting sharper images with lessen noise levels. Differences in image quality and noise levels of DSLR and compact cameras can be detected with statistical analysis in spatial and transfer domains. Thus, here, we propose to extract some features from discrete wavelet transform (DWT) coefficients, noise residue, and image quality statistics to build up a classifier for identifying DSLR and compact cameras.

MMSP'09, October 5-7, 2009, Rio de Janeiro, Brazil.

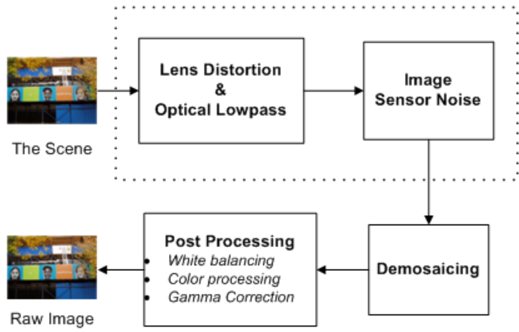

Fig. 1. The Digital Camera Imaging Pipeline.

This paper is organized as follows. The image features used in source class identification are introduced and analyzed in Section 2. Experimental results and identification performance for authentic and post-processed images are given in Section 3. Finally, the conclusion of this work is drawn in Section 4.

II. Image Features

A. DSLR and Compact Cameras

DSLRs are often preferred by professional still photographers as they allow an accurate preview of framing close to the moment of exposure, and they also allow the user to choose from a variety of interchangeable lenses. Many professionals also prefer DSLRs for their larger sensors compared to most compact digital ones. These large sensors allow for similar depths of field and picture angle to film formats. Besides, they yield better image quality high ISO performance, and lessen noise levels.

However, compact digital cameras are less expensive and more convenient to use when compared with DSLR. Benefits of compact digital cameras include easier to use and, for the most part, easier to learn.

According to the digital camera imaging pipeline shown in Fig.1, the major differences, between DSLR and compact camera image, are caused by lenses, optical filters, and particularly sensors (size and noise), shown with the dot box in Fig.1.

B. Wavelet Coefficient Features

DSLR cameras have larger sensors compared to compact cameras, leading to low sensor noise. In other words, they produce less noisy and sharp images. Therefore, we propose to use statistical noise features of digital images to discriminate direct camera outputs of different classes. The underlying idea of our approach is that different image sensors and lenses affect the image noise statistics in a specific way. If such effects can be extracted, they can be used in source camera identification.

It is known that wavelet coefficient statistics are useful in modelling image quality and pixel noise statistics [10]. So, in this work, we perform wavelet decomposition to images and, then, extract statistical features from sub-band coefficients. The image decomposition employed here is based on separable QMFs [11], [12]. The QMF bank is a multirate digital filter bank. QMF decomposition is better than more traditional wavelets, e.g., Haar or Daubechies; because, unlike other wavelets, QMFs minimize spatial aliasing within the decomposition sub-bands. On the other hand, QMFs do not afford perfect reconstruction of the original image though reconstruction errors can be minimized with a careful filter design [13]. The QMFs are separable, and comprised of a pair of one-dimensional low-pass and high-pass filters, e.g., $l(\cdot)$ and $h(\cdot)$. The first level of decomposition includes a vertical, horizontal and diagonal sub-band. It is generated by convolving the gray channel of the image, $I(x, y)$. The filters are as follows:

$$L_1(x, y) = I(x, y) * l(x) * l(y) \tag{1}$$

$$V_1(x, y) = I(x, y) * h(x) * l(y) \tag{2}$$

$$H_1(x, y) = I(x, y) * l(x) * h(y) \tag{3}$$

$$D_1(x, y) = I(x, y) * h(x) * h(y) \tag{4}$$

Where, $*$ is the convolution operator. L_1 is the low-pass sub-band, which is down-sampled by a factor of two filtered in the same way as above, to yield $V_i(x, y)$, $H_i(x, y)$, $D_i(x, y)$, $i = 2, 3$. So, we get a three-scale QMF decomposition.

The first component of the feature set is the higher order wavelet sub-band coefficient statistics, HOW(36), and the second component is composed of estimated error statistics as defined in [11], called HOW(72) in this work. In this paper, the wavelet decomposition is only applied in green channel.

The statistical model is composed of the mean(μ), variance(σ^2), skewness(ξ) and kurtosis(κ) of the sub-band coefficients and estimated errors, calculated as follows:

$$f_1 = \mu = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} f(i, j) \tag{5}$$

$$f_2 = \sigma^2 = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} \left(f(i, j) - \mu\right)^2 \tag{6}$$

$$f_3 = \xi = \frac{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f(i, j) - \mu\right)^3}{\left[\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f(i, j) - \mu\right)^2\right]^{\frac{3}{2}}} \tag{7}$$

$$f_4 = \kappa = \frac{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f(i, j) - \mu\right)^4}{\left[\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f(i, j) - \mu\right)^2\right]^{2}} - 3 \tag{8}$$

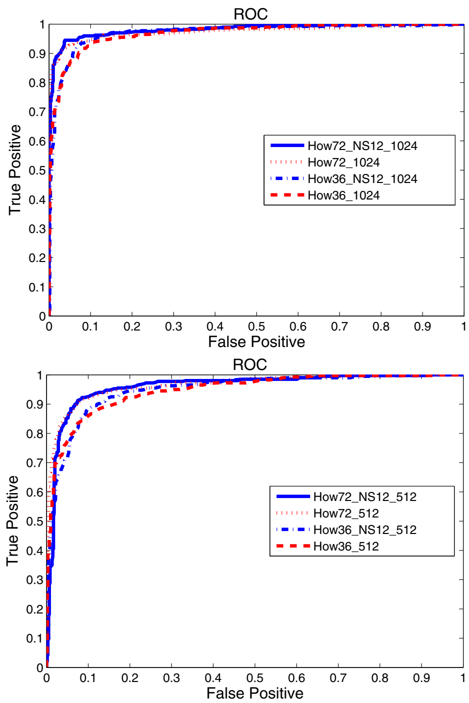

Fig. 2. Receiver Operating Characteristics. From up to down: (a) Image size 1024×1024; (b) Image size 512×512.

C. Image Noise Features

The image sensor noise is very useful for distinguishing digital camera sources. For example, PRNU noise, is used successfully to distinguish unique source camera devices [3].

In this work, the noise features will be extracted through an image denoising algorithm. Here, we utilized several image denoising algorithms to measure sensor noise statistics. Specifically, to capture the different aspects of sensor noise, we apply three different denoising algorithms. These denoising methods utilize separable 2-D DWT, real 2-D dual-tree DWT, and complex 2-D dual-tree DWT [14].

Using these three denoising methods, we obtain three denoised versions of the input image and corresponding noise residues. For each noise residue, we measured 4 first order statistics as given in formula (5)-(8) in intensity channel, and obtained totally $3 * 4 = 12$ features.

III. Experimental Results

The experiments in this study were conducted based on 20 different camera models, including 8 DSLR and 12 compact cameras of different models, as shown in Table I. For each camera, 100 images were taken, resulting 800 images for DSLR and 1200 images for compact camera classes. Utilizing the features introduced in the previous section, several support vector machine classifiers (SVM) [15] were built up. For benchmarking, image quality metrics (IQM), introduced in [16], were also used in the experiments. To train the classifiers, 50% of the images were used in training phase and the rest were used in testing.

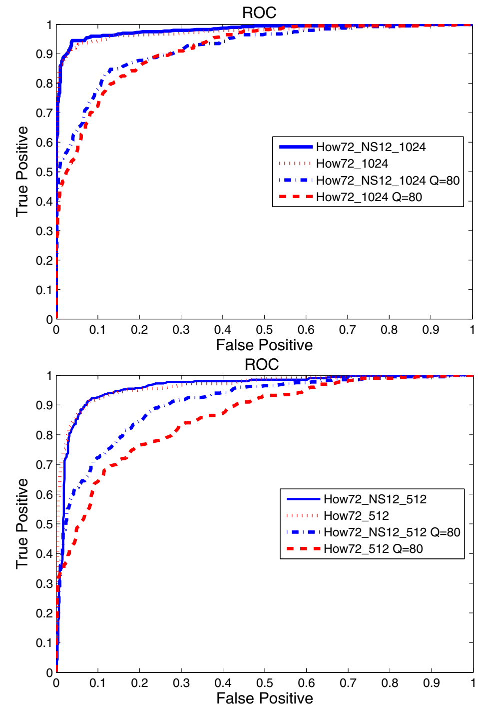

Fig. 3. Performance under image compression manipulation, $Q = 80$. From up to down: (a) Image size 1024×1024; (b) Image size 512×512.

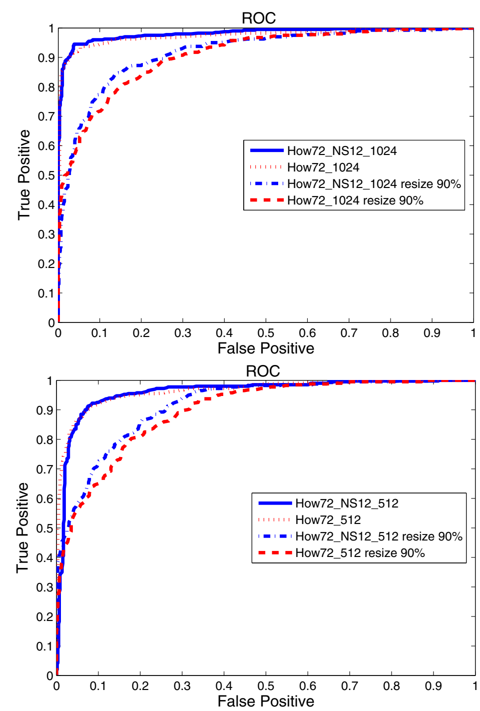

Fig. 4. Performance under image resizing manipulation, $\alpha = 90\%$. From up to down: (a) Image size 1024×1024; (b) Image size 512×512.

TABLE I
DSLR & COMPACT CAMERAS USED IN THE EXPERIMENTS

Compact Camera Model	n	SLR Camera	n
Konica Minolta Dimage Z3	100	Canon EOS Digital Rebel XT	100
Casio EX-Z850	100	Canon EOS 10D	100
Canon PowerShot A80	100	Canon EOS 30D	100
Canon PowerShot A70	100	Canon EOS 350D	100
SONY DSC-S90	100	Canon EOS Kiss X2 (450D)	100
HP 635 Digital Camera	100	Nikon D40	100
Panasonic DMC-TZ1	100	Nikon D50	100
Olympus FE230/X790	100	Nikon D70	100
Sony Cybershot	100		
Canon PowerShot S1 IS	100		
Sony H1	100		
Sony P150	100		

TABLE II
FEATURES VS CLASSIFICATION ACCURACY (t REFERS TO FEATURE EXTRACTION TIME PER IMAGE)

	Size of 1024×1024				Size of 512×512			
Features	ACC	TP	FP	t(sec)	ACC	TP	FP	t(sec)
Bior(36)	86.80	84.8	11.2	2.8	82.56	77.8	13.8	1.1
IQM(22)	-	-	-	-	80.13	72.3	14.2	147
NoiseStats(NS)(12)	80.84	68.8	10.5	19.1	78.77	63.0	10.3	6.1
HOW(36) +IQM (22)	91.23	88.5	6.3	157	89.47	87.0	8.2	149
HOW(72)+IQM(22)	93.77	91.8	4.4	154	91.36	88.3	6.0	152
HOW(36)	91.32	88.7	7.0	4.7	87.60	82.8	8.3	2.5
HOW(36)+NS(12)	91.56	89.0	6.0	23.8	89.04	85.8	8.2	8.6
HOW(72)	94.46	92.5	3.8	5.4	91.16	89.5	7.2	2.7
HOW(72)+NS(12)	94.50	92.0	3.5	24.5	91.76	88.8	5.8	9.6

TABLE III
FEATURES VS CLASSIFICATION ACCURACY AGAINST IMAGE PROCESSING

	Size of 1024×1024		Size of 512×512	
Manipulation	HOW(72)	HOW(72)+NS(12)	HOW(72)	HOW(72)+NS(12)
Original	94.46	94.50	91.16	91.76
Q=80	82.30	81.90	75.30	78.90
Q=60	68.20	68.00	67.20	64.00
Resize 0.90	84.90	83.60	80.20	82.50
Resize 0.50	77.00	84.60	73.20	73.60

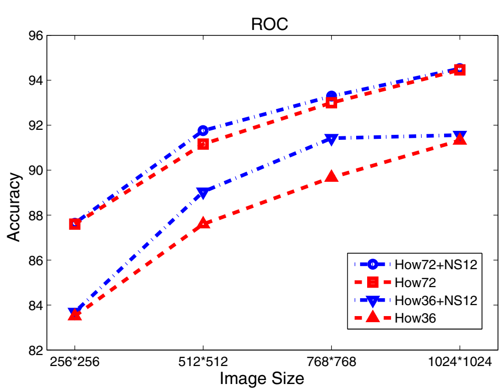

Fig. 5. Experimental results: Accuracy (%) vs. image size

The accuracy of each classifier are given in Table II. In 4th and 8th columns, feature extraction times for different feature sets are also presented in seconds. For every feature set shown in the table, we had trained and tested classifiers 10 times and calculated maximum accuracy, and corresponding false positive (FP) and true positive (TP) rates by averaging the results of 10 classifiers. The experiments were taken on Intel Pentium 3.40 GHz with 2GB RAM.

We can see from the results that the higher order wavelet (HOW) statistics based on QMFs have an important role in distinguishing DSLR and compact images. It is seen that the performance of QMFs-based statistics are superior than the Bior-based wavelet statistics. This is due to QMFs, as a multirate digital filter bank can minimize spatial aliasing in decomposition. Moreover, IQM features in [16] take a longer computing time to extract out and their performance is not satisfactory. Hence, we choose HOW and HOW+NS features as considerable solutions for source class identification problem. As a result, the contribution of features HOW(36) is more remarkable than noise-based features, e.g., NS(12).

Fig. 2 and Fig. 3 show the receiver operating characteristics (ROC) of different composite features for different image sizes, i.e., 1024×1024 and 512×512 and qualities (JPEG Q100 and Q80). Different sized images here were obtained by cropping the authentic images from the center. It is seen from the figures that HOW+NS features provide relatively good results. More details about Fig. 2 and Fig. 3 are given in Table III. Fig. 4 shows the robustness of the forensic scheme to image resizing 90%. The results of robustness to 50% resizing is given in Table III. Fig. 5 also shows the performance of the presented scheme for different image dimensions (obtained by cropping authentic images from their centers). It is seen that the larger the image is, the better performance we obtain.

IV. CONCLUSION

In this paper, we introduced a forensic scheme to distinguish between DSLR and compact images. Since DSLR and compact cameras use different type of sensors and lenses, their camera output quality in terms of sharpness and ISO sensitivity differs significantly. Such differences also affect the sensor noise levels and can be detected through wavelet decomposition and noise analysis. Thus, in this work, a source camera class identification scheme for DSLR and compact cameras is proposed based on machine learning classifiers utilizing statistical features of wavelet sub-bands and noise residues. The proposed scheme is also compared with image quality metrics to evaluate its performance. The experimental results show that the proposed forensic scheme has a potential to identify DSLR and compact images even they are re-compressed or down-sampled with 50%.

ACKNOWLEDGMENT

The authors would like to thank Sevinc Bayram for helpful discussions and the anonymous reviewers for their useful suggestions.

References

[1] A. E. Dirik, H. T. Sencar, and N. Memon, "Digital single lens reflex camera identification from traces of sensor dust," *IEEE Trans. on Information Forensics and Security*, vol. 3, pp. 539–552, 2008.

[2] M. Chen, J. Fridrich, M. Goljan, and J. Lukac, "Source digital camcorder identification using sensor photo response non-uniformity," *Proc. of SPIE Electronic Imaging*, vol. 6505, pp. 1G–1H, 2007.

[3] J. Lukac, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Security and Forensics*, vol. 1, pp. 205–214, 2006.

[4] A. Swaminathan, M. Wu, and K. J. R. Liu, "Digital image forensics via intrinsic fingerprints," *IEEE Tran. on Information Forensics and Security*, vol. 3, pp. 101–117, 2008.

[5] M. Kharrazi, H. Sencar, and N. Memon, "Blind source camera identification," *Proc. ICIP, International Conference on Image Processing*, vol. 1, pp. 709–712, 2004.

[6] S. Bayram, H. T. Sencar, N. Memon, and I. Avcibas, "Source camera identification based on cfa interpolation," *Proc. ICIP, International Conference on Image Processing*, vol. 3, pp. 69–72, 2005.

[7] K. Kurosawa, K. Duroki, and N. Saitoh, "Ccd fingerprint method - identification of a video camera from videotaped images," *Proc. of ICIP, International Conference on Image Processing*, vol. 1, pp. 537–540, 1999.

[8] X. Xue, S. Zhang, and Y. Guo, "An image identification system of seal with fingerprint based on cmos image sensor," *Proc. of SPIE*, vol. 6027(3)A, pp. 1–8, 2006.

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

[10] H. Gou, A. Swaminathan, and M. Wu, "Noise features for image tampering detection and steganalysis," *Proc. ICIP, International Conference on Image Processing*, vol. 6, pp. 97–100, 2007.

[11] S. Lku and H. Farid, "Steganalysis using higher-order image statistics," *IEEE Tran. on Information Forensics and Security*, vol. 1, pp. 111–119, 2006.

[12] H. Farid and S. Lyu, "Higher-order statistics and their applications to digital forensics," *Proc. of IEEE Workshop on Statistics Analysis in Computer Vision*, vol. 8, pp. 1–8, 2003.

[13] J. Portilla, V. Strela, and M. Wainwright, "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Tran. on Image Processing*, vol. 12, pp. 1338–1351, 2003.

[14] I. W. Selesnick, "The double-density dual-tree dwt," *IEEE Trans. on Signal Processing*, vol. 52, pp. 1304–1313, 2004.

[15] C. C. Chang and C. J. Lin, "Libsvm – a library for support vector machines," *http://www.csie.ntu.edu.tw/ cjlin/libsvm*, 2001.

[16] I. Avcibas, B. Sankur, and K. Sayood, "Statistical evaluation of image quality measures," *J. Electron. Imaging*, vol. 112, pp. 202–223, 2002.