IMAGE ACQUISITION FORENSICS: FORENSIC ANALYSIS TO IDENTIFY IMAGING SOURCE

Christine McKay, Ashwin Swaminathan, Hongmei Gou and Min Wu

ECE Department, University of Maryland, College Park

ABSTRACT

With widespread availability of digital images and easy-to-use image editing softwares, the origin and integrity of digital images has become a serious concern. This paper introduces the problem of image acquisition forensics and proposes a fusion of a set of signal processing features to identify the source of digital images. Our results show that the devices' color interpolation coefficients and noise statistics can jointly serve as good forensic features to help accurately trace the origin of the input image to its production process and to differentiate between images produced by cameras, cell phone cameras, scanners, and computer graphics. Further, the proposed features can also be extended to determining the brand and model of the device. Thus, the techniques introduced in this work provide a unified framework for image acquisition forensics.

Index Terms— Multimedia forensics, image acquisition forensics, color interpolation, noise statistics.

1. INTRODUCTION

Digital imaging technologies have seen tremendous growth in recent decades, and such digital imaging devices as digital cameras, scanners, cell phone cameras, video cameras, and camcorders have been used for a large number of day-to-day activities. Digital images captured by these imaging devices have been used in a number of applications from military, reconnaissance, and surveillance to consumer photography. When the image has security implications, or when the image's legitimacy is called into question, methods are needed to non-intrusively analyze the distinguishing features of the image in order to learn more about its origin. Knowledge of an image's acquisition source can be helpful in determining the authenticity of the image, as well as in determining who is responsible for creating the image.

In this work, we introduce *image acquisition forensics* as a new approach for forensic analysis aiming to determine the *device type* that was used to acquire the image in question. We present a feature based classification approach to facilitate image acquisition forensics, and show that the proposed methods provide a very high accuracy in differentiating between images from different sources such as cell phones cameras, standalone cameras, scanners, and computer-graphics. Further, the proposed methods can be extended to determine the brand and model of the device that was used to capture the image. To our best knowledge, this is the first work on image acquisition forensics that provides a unified framework to analyze images from four different sources to identify acquisition device type, and further extending such analysis to identify the device brand and model.

Related prior studies in multimedia forensics have had reasonable success in identifying an image by device brand or model, while assuming that the device acquisition type is known a priori. Features from such camera components as color interpolation and white balancing have been employed to differentiate between different brands and models of digital standalone cameras [1, 2], and cellphone cameras [3]. Noise features obtained by studying sensor imperfections have been used to identify camera model and set [4] and further extended to distinguish between images produced from different scanners [5, 6]. All these works inherently assume the knowledge of the image acquisition type. Recently, the work in [7] proposed a set of noise features for classifying images produced using different sensor types; however, their work is restricted to differentiating scanned and non-scanned camera images. Compared to the prior work, the methods proposed in this paper are more widely applicable. Our results with five different models of cell phones, five models of standalone cameras, and four models of scanners have demonstrated excellent forensic performance.

The rest of the paper is organized as follows. Section 2 provides a description of the image acquisition process in digital imaging devices and provides the basis of the proposed features discussed in Section 3. Simulation results are presented in Section 4, and the final conclusions are drawn in Section 5.

2. IMAGE ACQUISITION PROCESS IN DIGITAL IMAGING DEVICES

Fig. 1 shows the basic blocks in different types of image acquisition devices. The exact implementations vary among cell phone cameras, standalone cameras, and scanners, due to differences in color sensor type and alignment. Standalone digital cameras and scanners most commonly use charge coupled devices (CCD) to record the voltages generated by the light exposure corresponding to a particular color. Most cell phone cameras, on the other hand, use CMOS image sensors rather than CCDs for sampling the real-world scene as they are cheaper, faster, and consume much less power; however,

Email contact: {cemckay, ashwins, hmgou, minwu}@eng.umd.edu.

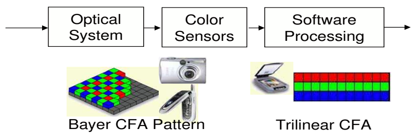

Fig. 1. Image acquisition process in digital imaging devices

CMOS sensors produce more noise than CCDs.

To capture the three basic color components corresponding to red, green and blue, scanners typically employ a trilinear color filter array (CFA). Using the trilinear CFA along with the line-by-line acquisition mode enabled by the motion system, scanners can directly capture all the three color components of each raster line. On the contrary, digital cameras and cell phone cameras use a square CFA, such as the Bayer pattern as shown in Fig. 1, to capture the entire 2-D scene in one shot. Therefore, in standalone cameras and cell phone cameras, only one color is obtained for each pixel and the remaining two colors for any given pixel are estimated through an interpolation process, which in most cases are unique to each model of camera [1].

After color interpolation, the interpolated images go via a post-processing stage. This stage may include operations such as white-balancing, noise reduction, color correction, and JPEG compression. Standalone cameras either store the image in a proprietary raw format with no compression or perform JPEG compression with a quality factor close of 95 to 100 to minimize information loss; however, cell phone cameras often use lower quality by default, in order to keep file size smaller and conserve memory (the cell phone cameras we studied employed JPEG compression with quality factors ranging from 65-85). For this reason as well as the extra noise from the CMOS image sensors, picture quality in cell phone pictures is not as good as that of digital cameras.

3. FEATURE EXTRACTION

To create a successful identification scheme, one must first find sources of variation among different types of devices and between different models of a device. Then, these differences can be extracted and represented as unique features of each device which can be used for identifying the source of an unknown new image. In our work, we use the dissimilarities in the image acquisition process of the imaging devices to develop two groups of features, namely color interpolation coefficients and the noise features, and use these features for image acquisition forensics.

Color Interpolation Coefficients as Features: Most digital standalone cameras and cellphone cameras employ different types of color interpolation methods [1], and this difference can be exploited forensically to distinguish them from other types of imaging devices. To extract the color interpolation coefficients, we start by assuming the Bayer pattern for the CFA, which gives the locations of the pixels that are directly obtained from the sensor and those that are interpolated. We

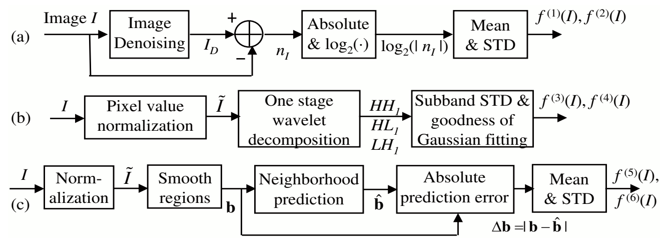

Fig. 2. Statistical noise feature extraction via (a) image denoising, (b) wavelet analysis, and (c) neighborhood prediction.

divide the image pixels into three different types of regions based on the local gradient values as (i) regions with significant horizontal gradient, (ii) regions with significant vertical gradient, and (iii) smooth regions. In each of these three regions, we approximate color interpolation to be linear, and represent the interpolated pixels as a weighted summation of its neighboring pixels that are directly obtained from the sensor. The set of equations obtained in this way is then solved to compute the interpolation coefficients [1].

Statistical Noise Feature Extraction: Noise occurs when photoelectrons are created in the imaging device. One example of measurable noise is dark signal non-uniformity, or variations between pixel voltage under conditions of no light. Photo response non-uniformity can be measured as the variations between pixel voltage under light with fixed intensity. While the imaging devices apply post-processing to compensate for and reduce noise in the image, some statistical properties of noise still remain unaltered depending on the specific nature of the sensors and filters used. In our work, we obtain noise features via three different types of analysis as shown in Fig. 2 and described below:

• *Features from image denoising:* We apply different types of denoising algorithms to the input image, and compute the mean and the standard deviation of the natural logarithm of the estimated noise magnitudes to obtain the first set of features. For our work, we apply four different denoising algorithms to an image: linear filtering with an averaging filter (filter size 3×3), linear filtering with a Gaussian filter (filter size 3×3 with standard deviation 0.5), median filtering (filter size 3×3), and Wiener adaptive image denoising (neighborhood sizes 3×3 and 5×5). These four denoising algorithms capture different statistical properties of the sensor noise giving us 30 different features [5].

• *Features from wavelet analysis:* We apply wavelet analysis to the image to measure the statistical properties of noise in the frequency domain. As a first step, we normalize the image to unit energy as $\tilde{I}(i,j) = \frac{I(i,j)}{\left(\frac{1}{MN}\sum_{k=1}^{M}\sum_{l=1}^{N} I(k,l)^2\right)^{\frac{1}{2}}}$, where $I(.,.)$ and $\tilde{I}(.,.)$ represent the input image and the normalized image, respectively. We then perform a 2-D one level wavelet decomposition of $\tilde{I}$ to get the LH_1, HL_1, and HH_1 subbands. The mean, μ, and the variance, σ^2, of the subband coefficients are obtained and the variance forms the second set of features [5]. We quantify the goodness of fitting this Gaussian distribution to the distribution of the wavelet subband coefficients to obtain additional features. Let $p(y)$ and $q(y)$ denote the probability density functions of the Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$ and the distribution of the subband wavelet coefficients, respectively. We quantify the goodness of Gaussian fitting by measuring the distance between $p(y)$ and $q(y)$ as $f^{(4)}(I) = \sum_i |p(y_i) - q(y_i)|\Delta y$, where i is the index of the histogram bins, Δy is the length of the bin, and $p(y_i)$ and $q(y_i)$ denote the histogram values at the bin centers. This gives a total of 18 features.

• *Features from neighborhood prediction:* Lastly, we employ neighborhood prediction and measure the error in the prediction of neighboring pixels in smooth regions. This is based on the observation that the image acquisition noise in the smooth regions of the image results in non-trivial prediction errors which can provide forensic evidence about the origin of the image. Given an image I, we first normalize it to obtain $\tilde{I}$ as before and identify its smooth regions by examining the local image gradient values. The pixels in the chosen smooth regions are then expressed as a weighted summation of its neighborhood values to obtain a set of linear equations. More specifically, we predict each pixel value b_i in a given region using a linear model on its eight surrounding neighbors $\{a_{i,1} - a_{i,8}\}$: $\hat{b}_i = \sum_{k=1}^{8} x_k a_{i,k}$. The set of linear equations thus obtained are solved under the non-negativity constraints $x_k \geq 0$ to obtain x_k. The absolute prediction errors are then obtained as $\Delta b_i = |\hat{b}_i - b_i|$. The mean and the variance of Δb_i over all the pixels in the smooth regions form our third set of 12 statistical noise features [5].

4. SIMULATION RESULTS AND DISCUSSIONS

This section presents the experimental results by applying the proposed features for image acquisition forensics to identify both the device type and the brand/model of the device that was used to acquire the image. For our study, we use 100 images from each of the four scanner models (Epson Perfection 2450 photo, Acer Prisa Acerscan, Canon CanoScan D1250U2F, and Microtek ScanMaker 3600), five different cell phone cameras models (Nokia 6102, Motorola V550, Samsung c417, Sony Ericsson W810, and Audiovox CDM-8910), and five standalone cameras models (Canon Powershot A75, Canon Powershot S410, FujiFilm Finepix S3000, Casio QV-UX2000, and Minolta DiMage F100). A separate set of 100 computer graphics (CG) images were obtained from [8]. The sample images were taken in completely random conditions, without any controlled setup in order to simulate the challenging non-intrusive testing conditions. The image dataset thus simulates real-world data in terms of lighting, color, texture, and subject. The color interpolation coefficients and the noise features were estimated from each of the 1500 images in our database and employed for subsequent studies.

Identifying Image Acquisition Device: For our study, 100 images from each device type (cell phone camera, standalone camera, and scanner) were selected with an equal number from each model, and all CG images were used, to create four classes of 100 images each. A randomly chosen set of 99 images from each class were used in training the SVM classifier, and the remaining image was used in testing to obtain the *leave-one-out* performance. The experiment was repeated 100 times with different set of training images and the average confusion matrix is shown in Table 1. Here, the $(i, j)^{th}$ element of the matrix corresponds to the fraction of images from source type$-i$ classified as belonging to source type$-j$. The main diagonal elements give the percentage of correct identification. From the results in Table 1, we find that overall identification accuracy is 93.75%, suggesting that the proposed features are good for identifying the source type.

Table 1. Confusion matrix for device type identification

Device	Phone camera	Standalone Camera	Scanner	CG
Phone camera	93%	2%	0%	5%
Standalone	1%	98%	1%	0%
Scanner	1%	3%	94%	2%
CG	4%	2%	4%	90%

Identifying Device Brand/Model: Once an image's source device has been determined, further analysis can be performed using the same set of features to identify the particular brand or model of the device that was used to capture the image. In this subsection, we present the results for cell phone camera identification in detail and similar results are also obtained for scanner identification and camera identification.

• *Cell phone camera identification:* Finding the type of cell phone camera from its output images poses additional challenges, compared to standalone cameras and scanners, due to their lower image resolution, noisier image sensors, and a higher rate of default JPEG compression. In our results with cell phone cameras, we found that using interpolation coefficients alone, rather than a combination of interpolation coefficients and noise features, produced higher accuracies. This result for cell phone cameras is expected because most cell phone camera brands/models employ different algorithms for color interpolation; and therefore, these coefficients alone provide tell-tale evidence to distinguish images from different brands/models. For our experiments with cell phone cameras, we used a randomly chosen 90 random images for training and the remaining 10 for testing, and the corresponding results are shown in Table 2. We find from the table that the average identification accuracy is close to 97.7% for five models, and this is significantly better than existing techniques that produce average accuracies close to 92% over four camera models from two different camera brands [3].

We test the robustness of the proposed system for post-processing operations such as JPEG compression. To generate data, we compress the original cell phone camera images under different JPEG quality factors from 60 to 100. The color interpolation coefficients are then obtained from the compressed images and used as features for classifica-

Table 2. Confusion matrix for cell phone camera identification

Cell Phone	Nokia	Motorola	Samsung	Sony	Audiovox
Nokia	95.8%	0.4%	0%	3.8%	0%
Motorola	2.8%	97.2%	0%	0%	0%
Samsung	1.2%	0%	97.8%	0.2%	0.8%
Sony	2.4%	0%	0%	97.6%	0%
Audiovox	0%	0%	0%	0%	100%

tion. A randomly chosen 90 images were used in training the classifier and the remaining 10 were used in testing. The experiment was repeated 100 times and the average accuracies under different JPEG quality factor are shown in Fig. 3. The figure shows that as the JPEG quality factor decreases, the identification accuracy decreases as expected. Further, an accuracy close to 91% is achieved even when the images are compressed with a quality factor of 60.

We compare the performance of the proposed features for cell phone camera identification with the higher order statistical features introduced in [9]. In our experiments with [9], we employ the same set of cell phone camera images (with 90 for training and 10 for testing) and examine the identification accuracies as a function of JPEG quality factors. The performance, averaged over 100 iterations, is shown alongside in Fig. 3. The results suggest that the proposed features perform at least 12% better in identifying the cell phone brand/model, establishing the goodness of the proposed features.

• *Scanner and standalone camera identification:* For scanner identification, we found that using a combination of interpolation coefficients and noise feature parameters gave the best results. 100 images from each of the four models of scanners were used, with 90 random images used for training and the remaining 10 used for testing. The overall identification accuracy for scanner brand was 96.2%. Further, the identification results were found to be robust to moderate levels of post-processing operations such as JPEG compression, image sharpening, gamma correction, and contrast enhancement.

Our results with standalone digital cameras suggests that the best results for camera identification were obtained using just the color interpolation coefficients, rather than using a combination of interpolation coefficients and noise features. In our experiments, we used a randomly selected set of 90 images from each of the five camera brands for training the SVM classifier and tested it with the remaining 10 images. The experiments were repeated over 100 times and we obtained an overall average accuracy of 94.3% for identifying the correct camera brand. These results establish the goodness of the proposed features for image acquisition forensics.

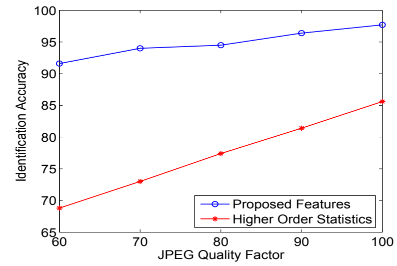

Fig. 3. Robustness to JPEG compression for cell phone camera identification

5. CONCLUSIONS

In this work, we have introduced a unified approach for image acquisition forensics to identify both the type of image acquisition device and the brand/model of the device. We have proposed to jointly employ color interpolation coefficients and noise statistics as features for forensic analysis. We show that the combined set of features can provide tell-tale clues and help accurately trace the origin of the input image to its production process and help differentiate between camera, cell phone camera, scanner, and computer-graphics images. Detailed simulation results with five cell phone cameras, four scanners, five digital cameras, and computer graphics images suggest that the proposed techniques are very effective giving an overall accuracy around 93.75%. The proposed techniques can also be extended to identify the brand and model of such imaging devices as cell phone cameras with an accuracy close to 97.7%, scanners with an accuracy of 96.2%, and standalone digital cameras with over 94.3% accuracy. Further, the features introduced in this work are also robust to post-processing operations such as moderate JPEG compression, demonstrating their effectiveness for image acquisition forensics. Overall, the proposed technique provides a promising unified framework to establish the origin of digital images with broad forensics applications.

6. REFERENCES

[1] A. Swaminathan, M. Wu, and K. J. R. Liu, "Non Intrusive Component Forensics of Visual Sensors Using Output Images," *IEEE Trans. of Information Forensics and Security*, vol. 2, no. 1, pp. 91-106, Mar. 2007.

[2] S. Bayram, H. T. Sencar, and N. Memon, "Improvements on source camera-model identification based on CFA interpolation," *Proc. of the WG 11.9 Intl. Conference on Digital Forensics*, Orlando, FL, Jan. 2006.

[3] M-J. Tsai, C-L. Lai, and J. Liu, "Camera/Mobile Phone Source Identification for Digital Forensics," *Proc. of the ICASSP*, vol. 2, pp. 221–224, Honolulu, HI, Apr. 2007.

[4] J. Lukas, J. Fridrich, and M. Goljan, "Determining Digital Image Origin Using Sensor Imperfections," *SPIE Conf. on Security, Stego., and Watermarking of Mult. Contents VII*, vol. 5681, pp. 249-260, Jan. 2005.

[5] H. Gou, A. Swaminathan, and M. Wu, "Robust Scanner Identification based on Noise Features," *SPIE Conf. on Security, Stego., and Watermarking of Multimedia Contents IX*, Jan. 2007.

[6] N. Khanna, A. K. Mikkilineni, G. T. C. Chiu, J. P. Allebach, and E. J. Delp, "Scanner Identification Using Sensor Pattern Noise," *SPIE Conf. on Security, Stego., and Watermarking of Mult. Contents IX*, Feb. 2007.

[7] N. Khanna, A. K. Mikkilineni, G. T. C. Chiu, J. P. Allebach, and E. J. Delp, "Forensic Classification of Imaging Sensor Types," *SPIE Conf. on Security, Stego., and Wmk of Mult. Contents IX*, San Jose, CA, Feb. 2007.

[8] T-T. Ng, S-F. Chang, J. Hsu, and M. Pepeljugoski, "Columbia Photographic Images and Photorealistic Computer Graphics Dataset," *ADVENT Technical Report 205-2004-5*, Columbia University, Feb. 2005.

[9] S. Lyu and H. Farid, "Detecting Hidden Messages Using Higher-Order Statistics and Support Vector Machines," *5th Int. Workshop on Information Hiding*, Netherlands, 2002.