FORENSIC TECHNIQUES FOR CLASSIFYING SCANNER, COMPUTER GENERATED AND DIGITAL CAMERA IMAGES

Nitin Khanna[†], George T.-C. Chiu[‡], Jan P. Allebach[†], Edward J. Delp[†]

[†]School of Electrical and Computer Engineering
[‡]School of Mechanical Engineering
Purdue University
West Lafayette, Indiana

ABSTRACT

Digital images can be captured or generated by a variety of sources including digital cameras, scanners and computer graphics softwares. In many cases it is important to be able to determine the source of a digital image such as for criminal and forensic investigation. This paper presents methods for distinguishing between an image captured using a digital camera, a computer generated image and an image captured using a scanner. The method proposed here is based on the differences in the image generation processes used in these devices and is independent of the image content. The method is based on using features of the residual pattern noise that exist in images obtained from digital cameras and scanners. The residual noise present in computer generated images does not have structures similar to the pattern noise of cameras and scanners. The experiments show that a feature based approach using an SVM classifier gives high accuracy.

Index Terms— image forensics, digital camera, scanners, computer graphics, pattern noise.

1. INTRODUCTION

Advances in digital imaging technologies have led to the development of low-cost and high-resolution digital cameras and scanners. Digital images produced by various sources are widely used in a number of applications from medical imaging and law enforcement to banking and daily consumer use. There is also proliferation of software for generating as well as manipulating digital images. Forensic tools that help establish the origin, authenticity, and the chain of custody of digital images are essential for many applications [1].

There are various levels at which the image source identification problem can be addressed. One may want to find the particular device (digital camera or scanner) which generated the image or one might be interested in knowing only the make and model of the device. A number of robust methods have been proposed for source camera identification [2, 3, 4, 5, 1].

Poineering work in utilizing imaging sensor's pattern noise for source camera identification is presented in [3]. The identification is based on pixel nonuniformity noise for both CCD (Charged Coupled Device) and CMOS (Complementary Metal Oxide Semiconductor) sensors. The pattern noise is caused by several factors such as pixel non-uniformity, dust specks on the optics, optical interference, and dark currents [6]. The high frequency part of the pattern noise is estimated by subtracting a denoised version of the image from the original using a wavelet denoising filter [3]. A camera's reference pattern, estimated by averaging the noise patterns from multiple images, serves as an intrinsic signature of the camera. To identify the source camera, the noise pattern from an image is correlated with known reference patterns from a set of cameras [3].

There are similar approaches for source scanner identification using sensor noise. In [7], a direct extension of the camera identification algorithm [3] was used for source scanner identification. All the experiments have shown lower classification accuracy compared to similar methods for source camera identification. Further experiments show that one possible reason for the decline in performance is post-processing operations in the scanners such as denoising techniques including flat-fielding and heavy down-sampling [7].

Another approach for scanner model identification using sensor pattern noise described in [8] uses three sets of features for each scanned image. Experiments on a set of 26 images from 7 scanners give 90-96% average classification accuracy. In [9] pattern noise based source camera identification [3] was extended for scanner identification using a set of statistical features and a SVM classifier. Experiments show that 95% average classification accuacry is possible when scanning is done at native resolution of the scanner.

The techniques used for both source camera and scanner

---

identification are dependent upon having prior knowledge of the class of device (camera or scanner). If the image was generated by a digital camera, then the digital camera identification methods must be used. Similarly if the image was generated by a scanner, the scanner identification methods must be used.

In [10] a method for differentiating between computer generated (henceforth refered as CG) and photographic images based on wavelet statistics is presented. It has been shown that a model based on first and higher-order wavelet statistics reveals subtle but significant differences between CG images and photographic images. Motivated by the use of the pattern noise introduced during image acquisition as a unique characteristic of digital cameras [3], in [11] a method for distinguishing between digital camera images and computer generated images is proposed. This method is based on the observation that since the image sensor technology remains the same even if each individual camera has a unique noise pattern associated with it, pattern noise introduced by different digital cameras may have common properties and this common characteristic will not be present in computer generated images. This scheme can not be easily extended to include scanner generated images due to the registration problems in generating the scanner error reference pattern [7, 9].

In [12], a novel technique for classification of images based on their sources, scanned and non-scanned images, is presented. A SVM classifier is used with appropriate features of the sensor pattern noise. For distinguishing images scanned at native resolution of the scanner from those captured using a digital camera, an average classification accuracy of greater than 95% is obtained.

In this paper we will extend the above feature vector based method [12] for classifying images captured using digital cameras, computer generated images and images captured using scanners.

2. FEATURE EXTRACTION

Both digital cameras and scanners work on similar principles in terms of their imaging pipeline. However, digital cameras use a two dimensional sensor array while most scanners use a one dimensional linear array. In the case of flatbed scanners, the same linear array is translated to generate the entire image. It is expected to find periodic correlation between rows of the fixed component of the sensor noise of a scanned image. There is no reason to find a similar periodic correlation between columns of the sensor noise of a scanned image. Neither the rows nor the columns of the fixed component of the sensor noise of an image generated by a digital camera are expected to exhibit such periodicity. This difference can be used as a basis for discriminating between the two image source classes. Further, due to the fundamental differences in the image generation process, the residual noise in computer generated images may not have properties similar to those of

images from the other two classes. Thus, for distinguishing images we develope an approach similar to that of [12] with suitable modifications in the features.

Let I denote the input image of size $M \times N$ pixels (M rows and N columns) and I_{noise} be the noise corresponding to the image. Let $I_{denoised}$ be the result of applying the denoising filter [13] on I. Then as in [3],

$$I_{noise} = I - I_{denoised} \qquad (1)$$

To save computation time, only the green channel is used for denoising and feature extraction. Let $\widetilde{I}^r_{noise}$ and $\widetilde{I}^c_{noise}$ denote the average of all the rows and the columns of the noise (I_{noise}) (Equation 2, 3).

$$\widetilde{I}^r_{noise}(1,j) = \frac{1}{M} \sum_{i=1}^{M} I_{noise}(i,j); \ \ 1 \leq j \leq N \qquad (2)$$

$$\widetilde{I}^c_{noise}(i,1) = \frac{1}{N} \sum_{j=1}^{N} I_{noise}(i,j); \ \ 1 \leq i \leq M \qquad (3)$$

Further, let $\rho_{row}(i)$ denote the value of correlation between the average of all the rows ($\widetilde{I}^r_{noise}$) and the ith row of the noise (I_{noise}) (Equation 4). (Similarly $\rho_{col}(j)$, Equation 5).

$$\rho_{row}(i) = \mathbf{C}(\widetilde{I}^r_{noise}, I_{noise}(i,.)) \qquad (4)$$

$$\rho_{col}(j) = \mathbf{C}(\widetilde{I}^c_{noise}, I_{noise}(.,j)) \qquad (5)$$

Where $\mathbf{C}(X,Y)$ is the normalized correlation between two vectors X and Y. For scanned images, ρ_{row} is expected to have higher values than ρ_{col} since there is a periodicity between rows of the fixed component of the sensor noise of a scanned image. The mean, standard deviation, skewness and kurtosis of ρ_{row} and ρ_{col} are the first eight features, extracted from each input image. The standard deviation, skewness and kurtosis of $\widetilde{I}^r_{noise}$ and $\widetilde{I}^c_{noise}$ correspond to features 9 through 14. The last feature for every input image is given by the following:

$$\mathbf{f}_{15} = \left(1 - \frac{\frac{1}{N} \sum_{j=1}^{N} \rho_{col}(j)}{\frac{1}{M} \sum_{i=1}^{M} \rho_{row}(i)} \right) * 100 \qquad (6)$$

A total of 15 features is obtained for each image. These features capture the essential properties of the image which are useful for discriminating between different image sources. Note that for extracting these statistical features, we need not to know the scan direction, that is, whether the image was scanned as portrait or as landscape. This is because the average of ρ_{row} is always higher than the average of ρ_{col} and so if needed, we can just rotate the image (rotating noise will be sufficient) before estimating the feature vector.

Table 1. Image Sources Used in Experiments

Image class	Devices used
Digital Camera	Canon PowerShot SD200, Nikon Coolpix 4100, Nikon Coolpix 7600
Computer Generated	www.3dlinks.com, www.irtc.org, www.raph.com, www.digitalrepose.com, www.maxon.net, www.realsoft.com
Flatbed Scanners	Epson Perfection 4490 Photo, HP ScanJet 6300c-1, HP ScanJet 6300c-2, HP ScanJet 8250, Mustek 1200 III EP, Visioneer OneTouch 7300, Canon LiDe 25, Canon Lide 70, OpticSlim 2420, Visioneer OneTouch 7100, Mustek ScanExpress A3

3. EXPERIMENTS AND RESULTS

Table 1 shows the sources of different classes of digital images used in our experiments. Some of the scanners have CCD sensor while others have CIS sensor. Computer generated images include images from number of different methods such as 3ds max, Maya, Softimage and Lightwave. From each of the 11 scanners 108 images were scanned at 200dpi resolution and stored in TIFF format (1024×768 pixels). Approximately 300 images were captured from each of the three cameras at 1024×768 resolution and stored in the best quality JPEG format supported by each camera. Computer generated images, in JPEG format, were downloaded from publicaly available websites listed in Table 1. For computer generated images of varying sizes, a central 1024×768 or smaller block is used for feature extraction depending upon the size of the image. In total there are approximately 1000 images from each of the three source classes. The `LIBSVM` package [14] is used in this study for the SVM classifier. A radial basis function is chosen as the kernel function and a grid search is performed to select the best parameters for the kernel. Unless stated otherwise, randomly chosen 80% of the images are used for training the classifier and rest of the images are used for testing. This training and testing is repeated multiple times to obtain the final average classification results.
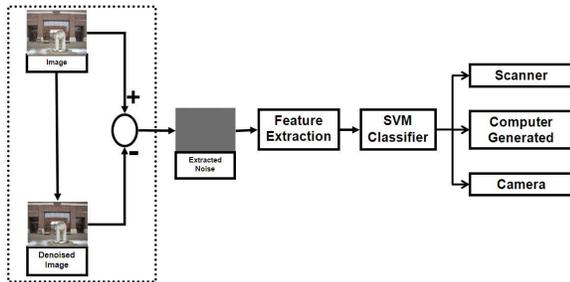
Fig. 1. Image Source Classification

The experiment is shown in Figure 1. In first set of experiments three separate SVM classifiers are designed for distinguishing between three possible pairs of image source classes: scanner, computer generated and camera. Tables 2, 3 and 4 show the confusion matrices for classifying these pair of classes. The average classification accuracy for distinguishing Scanner images from CG images is 97.6%. The average classification accuracy for distinguishing CG images

from Camera images is 91.5%. While the average classification accuracy for distinguishing Camera images from Scanner images is 89.4%, the lowest among three pairs.

Table 2. Confusion Matrix for Scanner vs. CG

		Predicted	
		Scanner	CG
Actual	Scanner	98.2	1.8
	CG	3.1	96.9

Table 3. Confusion Matrix for CG vs. Camera

		Predicted	
		CG	Camera
Actual	CG	88.3	11.6
	Camera	5.2	94.8

Table 4. Confusion Matrix for Camera vs. Scanner

		Predicted	
		Camera	Scanner
Actual	Camera	89.5	10.5
	Scanner	10.7	89.3

Table 5. Confusion Matrix for Scanner, CG and Camera

		Predicted		
		Scanner	CG	Camera
Actual	Scanner	85.3	1.0	13.7
	CG	1.7	88.3	10.0
	Camera	11.9	3.9	84.2

The average classification accuracy for classifying images from all three classes is 85.9%. Corresponding confusion matrix is shown in Table 5. Thus, by training an SVM classifier on the 15 dimensional feature vectors from each image, using randomly chosen 800 images from each class for training and separate 200 images for testing, proposed method is able to give average classification accuracy of 85.9%.

The efficacy of the proposed method is also tested on images that have been JPEG compressed. An average classi-

fication accuracy of 79.8% is obtained when all the scanned images are saved as JPEG (Q=90) before feature extraction for classifier training and testing. Corresponding confusion matrix is shown in Table 6. This slight decrease in performance is as expected since the pattern noise degrades with JPEG compression and down-sampling, as observed for source camera identification [3] and source scanner identification [7]. Whenever a camera or scanner is used to capture images at resolution lower than maximum resolution supported by the device, generally down-sampling is done in the device driver, for example scanning at 200dpi from a 4800dpi scanner. Further experiments by varying the size of training dataset show that average classification accuracy remains close to 80% even when only 40% images (400 images from each source class) are used for training the classifier.

Table 6. Confusion Matrix for Classifying JPEG Compressed Images

		Predicted		
		Scanner	CG	Camera
Actual	Scanner	86.4	1.7	11.9
	CG	11.8	70.6	17.6
	Camera	13.4	4.2	82.4

4. CONCLUSION

In this paper we investigated the use of the sensor pattern noise for classifying digital images based on their sources. Selection of proper features is the key to achieve accurate results. The scheme presented here utilizes statistical properties of the residual noise and the difference in the geometry of the imaging sensors and demonstrates promising results. Future work will include, tests on images that have undergone various post-processing operations.

5. REFERENCES

[1] N. Khanna, A. K. Mikkilineni, A. F. Martone, G. N. Ali, G. T.-C. Chiu, J. P. Allebach, and E. J. Delp, "A survey of forensic characterization methods for physical devices," *Digital Investigation*, vol. 3, pp. 17–28, 2006.

[2] Z. J. Geradts, J. Bijhold, M. Kieft, K. Kurosawa, K. Kuroki, and N. Saitoh, "Methods for identification of images acquired with digital cameras," in *Enabling Technologies for Law Enforcement and Security*. 2001, vol. 4232, pp. 505–512, SPIE Press.

[3] J. Lukas, 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.

[4] A.C. Popescu and H. Farid, "Exposing digital forgeries in color filter array interpolated images," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3948–3959, 2005.

[5] S. Bayram, H.T. Sencar, N. Memon, and I. Avcibas, "Source camera identification based on cfa interpolation," in *Proceedins of the IEEE International Conference on Image Processing*, 2005, pp. 69–72.

[6] G. C. Holst, *CCD Arrays, Cameras, and Displays, Second Edition*, JCD Publishing & SPIE Press, USA, 1998.

[7] T. Gloe, E. Franz, and A. Winkler, "Forensics for flatbed scanners," in *Proceedings of the SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents IX*. 2007, vol. 6505, p. 65051I, SPIE.

[8] H. Gou, A. Swaminathan, and M. Wu, "Robust scanner identification based on noise features," in *Proceedings of the SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents IX*. 2007, vol. 6505, p. 65050S, SPIE.

[9] N. Khanna, A. K. Mikkilineni, G. T. C. Chiu, J. P. Allebach, and E. J. Delp, "Scanner identification using sensor pattern noise," in *Proceedings of the SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents IX*. 2007, vol. 6505, p. 65051K, SPIE.

[10] S. Lyu and H. Farid, "How realistic is photorealistic?," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 845–850, 2005.

[11] S. Dehnie, T. H. Sencar, and N. D. Memon, "Digital image forensics for identifying computer generated and digital camera images," in *Proceedings of the International Conference on Image Processing*, Atlanta, Georgia, USA, October 2006, pp. 2313–2316.

[12] N. Khanna, A. K. Mikkilineni, G. T. C. Chiu, J. P. Allebach, and E. J. Delp, "Forensic classification of imaging sensor types," in *Proceedings of the SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents IX*. 2007, vol. 6505, p. 65050U, SPIE.

[13] A. Foi, V. Katkovnik, K. Egiazarian, and J. Astola, "A novel local polynomial estimator based on directional multiscale optimizations," in *Proceedings of the 6th IMA Int. Conf. Math. in Signal Processing*, 2004, vol. 5685, pp. 79–82.

[14] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," 2001.