Survey of Scanner and Printer Forensics at Purdue University

Nitin Khanna, Aravind K. Mikkilineni,
George T.-C. Chiu, Jan P. Allebach, and Edward J. Delp

Purdue University, West Lafayette IN 47907, USA *

Abstract. This paper describes methods for forensic characterization of scanners and printers. This is important in verifying the trust and authenticity of data and the device that created it. An overview of current forensic methods, along with current improvements of these methods is presented. Near-perfect identification of source scanner and printer is shown to be possible using these techniques.

1   Introduction

Advances in digital imaging technologies have led to the development of low-cost and high-resolution digital cameras and scanners, both of which are becoming ubiquitous. Digital images generated by various sources are widely used in a number of applications from medical imaging and law enforcement to banking and daily consumer use. The increasing functionality of image editing software allows even an amateur to easily manipulate images. In some cases a digitally scanned image can meet the threshold definition requirements of a "legal duplicate" if the document can be properly authenticated. Forensic tools that help establish the origin, authenticity, and the chain of custody of such digital images are essential to a forensic examiner.

The same holds true for printed material, which in many cases is a direct accessory to many criminal and terrorist acts. Examples include forgery or alteration of documents used for purposes of identity, security, or recording transactions. In addition, printed material may be used in the course of conducting illicit or terrorist activities. Examples include instruction manuals, team rosters, meeting notes, and correspondence. In both cases, the ability to identify the specific device or type of device used to print the material in question would provide a valuable aid for law enforcement and intelligence agencies.

There are various levels at which this source identification problem can be addressed. One may want to find the particular device which generated the image or one might be interested in knowing only the make and model of the device.

As summarized in [1], a number of robust methods, involving characterization of sensor noise, have been proposed for source camera identification. In [2, 3], techniques for classification of images based on their sources, scanned, camera generated, and computer generated, are presented.

Forensic characterization of a printer involves finding intrinsic features in the printed document that are characteristic of that particular printer, model, or manufacturer's products. This is referred to as the *intrinsic signature*. The intrinsic signature requires an understanding and modeling of the printer mechanism, and the development of analysis tools for the detection of the signature in a printed page with arbitrary content [4].

In this paper we present a brief overview of some recent advances in source identification of both scanned images and printed documents. Results are presented for our methods based on intrinsic signatures of the devices.

2 Characterization of Scanners

Pioneering work in the development of source camera identification was performed by Lukas et al. using techniques involving the imaging sensor's pattern noise [5, 6]. Our method extends the use of sensor noise for source scanner identification by replacing correlation detector by statistical features and support vector machine (SVM) classifier. Since sensor pattern noise is estimated using the simple averaging method, further improvements in results may be obtained by using the improved method for sensor noise estimation presented in [6]. Extensive experimentation on a large set of scanners and different scanning scenarios show the effectiveness of our proposed scheme.

2.1 Statistical Feature Extraction

The basic structure of the scanner imaging pipeline and various types of noises involved in the scanning process are explained in [7]. Pattern noise refers to any spatial pattern that does not change significantly from image to image. It has been successfully used for source camera identification in [5, 6]. In this study the sensor noise is modeled as the sum of two components, a random component and a fixed component. The random component is that portion of the noise which changes from image to image and varies over a period of time, while the fixed component is that portion of the noise which remains constant from image to image. The fixed component can be considered as a signature of the imaging sensor and can be used for source scanner identification. The challenge is to separate the fixed component from the random component of the noise, and then design suitable classifiers based on appropriate features. Since both CCD based and CIS based scanners suffer from similar noise sources, the proposed scheme is expected to work on both types of scanners. This is demonstrated by extensive experimentation.

Selection of relevant features from the sensor noise is the key to accurate and robust source scanner identification. The features selected should satisfy the following requirements:

- Independent of image content
- Capture the characteristics of a particular scanner and if possible differ amongst scanners of the same make and model
- Independent of scan area, that is, they should be able to characterize the source scanner even if the images are placed at different positions on the scanner's glass plate

An image scanned twice with the same scanner, but at different non-overlapping locations on the scanner bed will contain different PRNU due to variations in the manufacturing process such as doping concentration. The proposed scheme uses PRNU based scanner fingerprints. The fixed component of sensor noise is caused by PRNU as well as noise-like characteristics remaining after the post processing steps. To generate the final image, a number of non-linear operations are performed on the values read by the sensor array. Thus, the statistical properties of the fixed component of noise are expected to remain same irrespective of the image placement on the scanner bed. This is the reason behind using the statistical features of the fixed component of sensor noise for source scanner identification. Our experimental results suggest that this is true.

The scanned images are generated by translating a linear sensor array along the length of a document or image. Each row of the resulting digital image is generated by the same set of sensor pixels. Thus, for scanned images, the average of all the rows of sensor noise will give an estimate of the fixed "row-pattern". Averaging will reduce the random component while at the same time enhancing the fixed component of the noise. Along with using different statistical features of the "row-pattern", similar features are estimated along the column direction as well for comparing them with statistics along the row direction.

The procedure to extract features from one of the color channels is described here and is applied to all three channels separately to get the complete feature vector. 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 a denoising filter on I. Then as in [5], $I_{noise} = I - I_{denoised}$. 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}), respectively. 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}). Similarly, $\rho_{col}(j)$ denotes the value of correlation between the average of all the columns ($\widetilde{I}^c_{noise}$) and j^{th} column of the noise (I_{noise}) (Equation 1 for rows and similar for columns).

$$\widetilde{I}^r_{noise}(1,j) = \frac{1}{M} \sum_{i=1}^{M} I_{noise}(i,j); \qquad \rho_{row}(i) = \mathbf{C}(\widetilde{I}^r_{noise}, I_{noise}(i,.)) \qquad (1)$$

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 statistical properties of ρ_{row}, ρ_{col}, $\widetilde{I}^r_{noise}$ and $\widetilde{I}^c_{noise}$ capture the essential properties of the image which are useful for discriminating between different scanners. The mean, standard deviation, skewness and kurtosis of ρ_{row} and ρ_{col}

are the first eight features, extracted from each color channel of the 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 channel of the input image is given by (2) which is representative of relative difference in periodicity along row and column directions in the sensor noise. Values of ρ_{col} corresponds to the case when each element is generated independently and so f_{15} will have high positive value. Some images from low-quality scanners or images which have undergone post-processing operations such as high down-sampling or JPEG compression, may be an exception to this.

$$\mathbf{f}_{15} = \left(1 - \frac{\frac{1}{N}\sum_{j=1}^{N}\rho_{col}(j)}{\frac{1}{M}\sum_{i=1}^{M}\rho_{row}(i)}\right) * 100 \tag{2}$$

By extracting these 15 features from each of the three color channels, a 45 dimensional feature vector is obtained for each scanned image. To capture the three color channels, some scanners use three different linear sensors while others use a single imaging sensor in coordination with a tri-color light source. To capture this difference among scanners of different make and models, six additional features are used. These features are obtained by taking mutual correlations of $\widetilde{I}^r_{noise}$ for different color channels (same for $\widetilde{I}^c_{noise}$). Hence, in total each scanned image has a 51 dimensional feature vector associated with it.

In our previous work [7] on source scanner identification from images scanned at native resolution of the scanners, a recently developed anisotropic local polynomial estimator for image restoration based on directional multiscale optimizations [8] was used for denoising. In this study, a denoising filter bank comprising of four different denoising algorithms: LPA-ICI denoising scheme [8], median filtering (size 3×3) and Wiener adaptive image denoising for neighborhood sizes 3×3 and 5×5, is used. Using a set of denoising algorithms helps to better capture different types of sensor noises. These denoising algorithms are chosen based on the performance of the complete filter bank in scanner identification. Each denoising algorithm is independently applied to each color band of an image. The features extracted from individual blocks of the filter bank are concatenated together to give the final feature vector for each scanned image. Hence, each scanned image has a 204 dimensional feature vector associated with it. To reduce the dimensionality of the feature vectors, linear discriminant analysis (LDA) is used and a ten dimensional feature vector is obtained for each image. Each component of the ten dimensional feature vector is then a linear combination of the original 204 features. Finally a SVM classifier is trained and tested on these ten dimensional feature vectors.

2.2 Experimental Results

The following set of scanners are used in our experiments (label, make/model, DPI): { $\{S_1$, Epson Perfection 4490 Photo, 4800$\}$, $\{S_2$, HP 6300c-1, 1200$\}$, $\{S_3$, HP 6300c-2,1200$\}$, $\{S_4$, HP 8250, 4800$\}$, $\{S_5$, Mustek 1200 III EP, 1200$\}$, $\{S_6$, Visioneer OneTouch 7300, 1200$\}$, $\{S_7$, Canon LiDe 25, 1200$\}$, $\{S_8$, Canon Lide
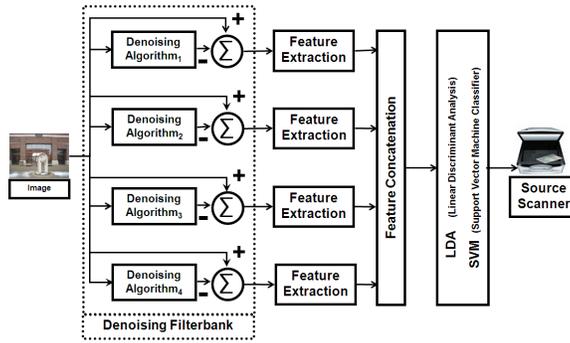
Fig. 1. Feature-based scanner identification.

70, 1200}, $\{S_9$, OpticSlim 2420, 1200}, $\{S_{10}$, Visioneer OneTouch 7100, 1200}, $\{S_{11}$, Mustek ScanExpress A3, 600}}. Experiments are performed on images scanned at the native resolution of the scanners as well as on images scanned at lower resolution, such as 200 DPI. This paper presents only results for 200 DPI images since images are generally scanned at lower resolution to meet constraints on storage space, scanning time, and transmission bandwidth. Experiments on native resolution images show similar accuracies.

The experimental procedure for source scanner identification using statistical features of sensor noise is shown in Fig. 1. The features are mapped to the range $[-1, 1]$ before SVM classification to achieve higher classification accuracy. The mapping is decided by the values of features in the training set and same mapping is applied on the features in the testing set. A radial basis function is chosen as the kernel function and a grid search is performed to select the best parameters for the kernel. To generate the final confusion matrices, SVM training and testing steps are repeated multiple times using a random selection of training and testing sets.

In the next few experiments, the effectiveness of the proposed scheme is shown for heavily sub-sampled (200 DPI) images. The scheme proposed here has good performance on 200 DPI images (which corresponds to scaling by 17% to 4% for native resolutions of 1200 DPI to 4800 DPI). 108 images are scanned at 200 DPI using each of the eleven scanners. All these scanned images are saved as uncompressed TIFF images. Unless stated otherwise, for each experiment 80 randomly selected images from each class are used for training and the remaining images are used for testing.

Degradation in the characteristics of sensor noise due to heavy down-sampling prevents successful separation of images scanned from the two scanners of exact same make and model as demonstrated by our initial experiments. For all further experiments performed on images scanned at 200 DPI, scanners S_2 and S_3 are treated as a single class. The bar graph shown in Fig. 2 shows the classification accuracies of the multi-class classifier for 200 DPI TIFF images. The proposed algorithm has an average classification accuracy of 99.9% among ten scanner models.
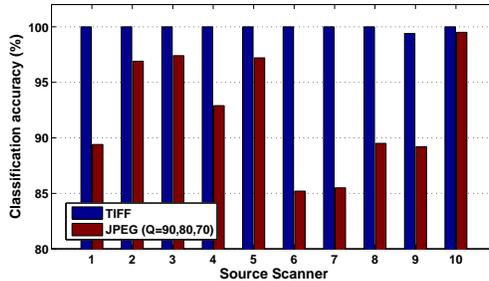
Fig. 2. Classification accuracies of multi-class classifiers. (Dedicated classifier for TIFF and general classifier for JPEG).

To further check the robustness of the proposed scheme for scanner model identification, a SVM classifier is trained without images from scanner S_3 and tested on images from the scanner S_3 only. This has classification accuracy of 95% for the testing images from scanner S_3. A similar classifier designed by training without images from scanner S_2 and testing on images from S_2 only, gives a classification accuracy of 97%. These results imply that even in the absence of the training data from a particular scanner, the proposed scheme can identify the scanner model as long as training data for another scanner of the same make and model is available.

Another aspect of robustness is independence from scanning location. That is, even when the image is placed at a random unknown location on the scanner bed, source scanner identification should still be possible. The images used in all the earlier experiments were scanned from the "default" scanning location (generally marked at the top left corner) of the scanner. For this experiment another 108 images are scanned from scanner S_{11}, with their location on the scanner's bed slightly translated horizontally and vertically between each scan. A SVM classifier is trained using the images scanned from "default" location and tested using images from the "random" locations only. This has a classification accuracy of 100% for "randomly" placed images. This suggests that the proposed scheme for scanner model identification is independent of the scanning location.

To investigate the robustness of the proposed scheme under JPEG compression, TIFF images from all the scanners are compressed at three different quality factors Q = 90, 80 and 70. For using a dedicated classifier on distorted images, we need to know the particular post-processing applied on the image of unknown origin. In some cases, this a priori information is available or can be obtained by using other forensic methods. For example, it may be possible to know the JPEG quality factor by an analysis of quantization table embedded in the JPEG image. But in general, an image of unknown origin is provided for forensic examination without reliable knowledge of post-processing operations applied on it. And in some cases, revealing the post-processing operations applied on the unknown image is one of the goals of forensic study. Thus, there is need for a general classifier which does not need to know the quality factors of training and testing images.
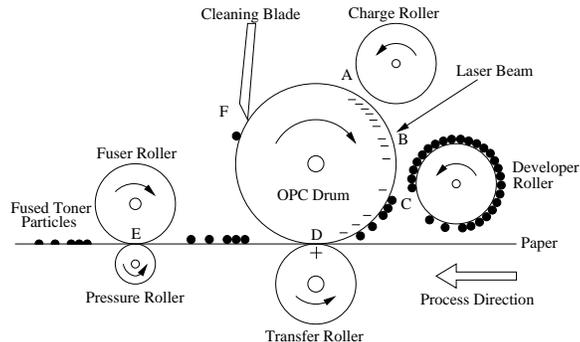
Fig. 3. Diagram of the EP process.

To design a robust general classifier, the TIFF images compressed at three quality factors are grouped together and 80 randomly chosen images from each scanner class are used for training the classifier. The remaining $(11 * 108 * 3 - 10 \times 80 =)$ 2764 images are used for testing the classifier. The bar graph shown in Fig. 2 shows the classification accuracies of general multi-class classifier. The proposed algorithm has an average classification accuracy of 92.3% among ten scanner models. The proposed scheme gives high classification accuracy even without knowing the JPEG quality factors of training or testing images.

3   Characterization of Printers

Techniques that use the print quality defect known as *banding* in electrophotographic (EP) printers as an intrinsic signature to identify the model and manufacturer of the printer have been previously reported in [9, 10]. We showed that different printers have different sets of *banding frequencies* that are dependent upon brand and model. This feature is relatively easy to estimate from documents with large midtone regions. However, it is difficult to estimate the banding frequencies from text. The reason for this is that the banding feature is present in only the process direction and in printed areas. The text acts as a high energy noise source upon which the low energy banding signal is added.

One solution for identifying intrinsic signatures in text, previously reported in [4], is to find a feature or set of features which can be measured over smaller regions of the document such as individual text characters. If the print quality defects are modeled as a texture in the printed areas of the document then texture features can be used to classify the document. These types of features can be more easily estimated over small areas such as inside a text character.

An understanding of the EP (laser) printing process is necessary in order to gain insight into the types of features that can be used to describe these printers. The first thing to note is that in the printed output from any printer there exist defects caused by electromechanical fluctuations or imperfections in the print mechanism. Because these "print quality defects" are directly related to the printer mechanism, they can also be viewed as an intrinsic signature of

the printer. The major components of these intrinsic signature are stable over time and independent of the consumables in the printer.

Figure 3 shows a side view of a typical EP printer. The print process has six steps. The first step is to uniformly charge an optical photoconductor (OPC) drum. Next a laser scans the drum and discharges specific locations on the drum. The discharged locations on the drum attract toner particles which are then attracted to the paper which has an opposite charge. Next the paper with the toner particles on it passes through a fuser and pressure roller which melt and permanently affix the toner to the paper. Finally a blade or brush cleans any excess toner from the OPC drum.

In EP printing, some causes of the artifacts in the printed output are fluctuations in the angular velocity of the OPC drum, gear eccentricity, gear backlash, and polygon mirror wobble. These imperfections in the printer are directly tied to the electromechanical properties of the printer and create corresponding fluctuations in the developed toner on the printed page. The fluctuations in the developed toner can be modeled as a texture. Since the mechanical properties which contribute the most to these fluctuations, such as gear ratios, do not change over time, they can be used reliably to intrinsically identify a printer. In the following sections, two techniques for identifying the printer which created a document are described.

3.1 Halftone Images

In EP printing, the major artifact in the printed output is *banding* which is defined as those artifacts that are due to quasiperiodic fluctuations in process direction parameters in the printer. These are primarily due to fluctuations in the angular velocity of the OPC drum and result in non-uniform scan line spacing. This causes a corresponding fluctuation in developed toner on the printed page. The appearance of the banding artifact is alternating light and dark bands perpendicular to the process direction (the direction the paper passes through the printer). The main cause of banding is electromechanical fluctuations in the printer mechanism, mostly from gear backlash. Because these fluctuations are related to the gearing, the *banding frequencies* present in the printed page directly reflect mechanical properties of the printer.

To estimate the banding frequencies of an EP printer, test pages with midtone graylevel patches created using a line fill pattern were printed and analyzed. These patterns were printed on a set of EP printers and then each patttern was scanned at 2400dpi. Each scanned image, $img(i, j)$, was then projected horizontally to produce $proj(i) = \sum_j img(i, j)$ shown in Fig. 4a. Fourier analysis of the projections was then obtained such as in Fig. 4 which shows spikes at 132 cycles/inch and 150 cycles/inch. Table 1 shows a list of printers and their principle banding frequencies as found by this method.

Detection and measurement of the banding signal in documents with large midtone regions, such as those with graphic art, can easily be done using methods similar to that used to produce Tab. 1.
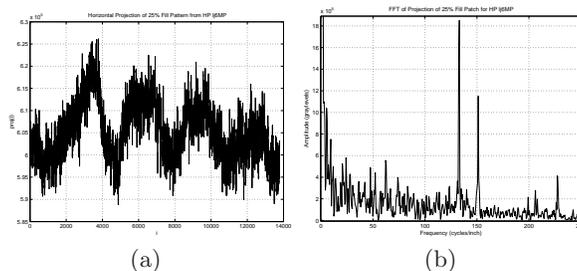
Fig. 4. (a) Projection of 25% fill pattern from HP LaserJet. (b) FFT of the projection showing peaks at 132 and 150 cycles/inch.

Printer Model	Principal Banding Frequencies (cycles/inch)
HP LaserJet 5MP	37, 74
HP LaserJet 6MP	132, 150
HP LaserJet 1000	27, 69
HP LaserJet 1200	69
HP LaserJet 4050	51, 100
Samsung ML-1450	16, 32, 100, 106

Table 1. Banding frequencies.

3.2 Forensic Characterization of Printed Text

Detection of the banding signal in text is difficult because the power of the banding signal is small with respect to the text, and because only a limited number of cycles of the banding signal can be captured within the height of one text character. Instead, all the print quality defects, including the banding, are lumped together and considered a texture in the printed regions of the document. Features are then extracted from this texture to be used as an intrinsic signature of the printer.

Features are extracted from individual printed characters, in particular the letter "e"s in a document. The reason for this is that "e" is the most frequently occurring character in the English language. Each character is very small, about 180x160 pixels and is non-convex, so it is difficult to perform any meaningful filtering operations in either the pixel or transform domain if the area of interest is only the printed region of each character. To model the texture in the printed regions, graylevel co-occurrence texture features, as well as two pixel based features are used [4].

Graylevel co-occurrence texture features assume that the texture information in an image is contained in the overall spatial relationships among the pixels in the image. This is done by first generating the Graylevel Co-occurrence Matrix (GLCM). This is an estimate of the second order probability density function of the pixels in the image. The features are statistics obtained from the GLCM as described in [4].

Figure 5 shows the block diagram of the printer identification scheme for text documents proposed in [4]. Given a document with an unknown source, referred
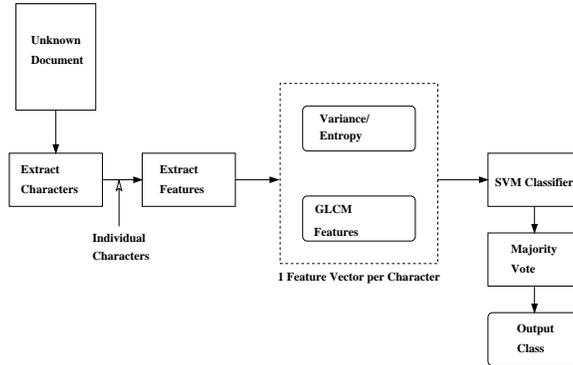
Fig. 5. Diagram of printer identification.

train\test	lj5m	lj6mp	lj1000	lj1200	E320	ml1430	ml1450	hl1440	1250w	14e	Output class
lj5m	296	2	0	1	0	1	0	0	0	0	lj5m
lj6mp	1	256	6	0	17	0	0	15	5	0	lj6mp
lj1000	2	2	284	12	0	0	0	0	0	0	lj1000
lj1200	7	2	2	289	0	0	0	0	0	0	lj1200
E320	0	0	0	0	300	0	0	0	0	0	E320
ml1430	1	0	0	0	0	299	0	0	0	0	ml1430
ml1450	0	0	0	0	0	0	300	0	0	0	ml1450
hl1440	0	28	0	0	0	5	2	259	6	0	hl1440
1250w	0	0	0	0	0	0	0	3	292	5	1250w
14e	0	0	0	0	0	0	0	17	67	216	14e

Fig. 6. Classification results using 300 "e"s from 12 point Times text documents.

to as the *unknown document*, this process can be used to identify the printer that created it. For testing purposes, the Forensic Monkey Text Generator (FMTG) described in [9] is used to create random documents with known statistics to be classified.

The first step is to scan the document at 2400 dpi with 8 bits/pixel (grayscale). Next, all the letter "e"s in the document are extracted. A set of features are extracted from each character forming a feature vector for each letter "e" in the document. Each feature vector is then classified individually using a support vector machine (SVM) classifier.

The SVM classifier is trained with 5000 known feature vectors. The training set is made up of 500 feature vectors from each of 10 printers. Each of these feature vectors generated are independent of one another.

Let Ψ be the set of all printers $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$. 10 printers are used in our work: { {Brother HL-1440, 1200dpi}, {HP lj4050, 600dpi}, {HP lj1000, 600dpi}, {HP lj1200, 600dpi}, {HP lj5M, 600dpi}, {HP lj6MP, 600dpi}, {Minolta 1250W, 1200dpi}, {Okidata 14e, 600dpi}, {Samsung ML-1430, 600dpi} }. For any $\phi\epsilon\Psi$, let $c(\phi)$ be the number of "e"s classified as being printed by printer ϕ. The final classification is decided by choosing ϕ such that $c(\phi)$ is maximum. In other words, a majority vote is performed on the resulting classifications from the SVM classifier.
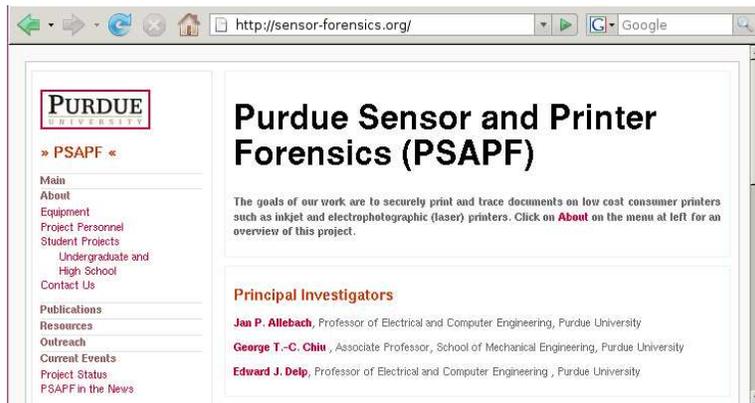
Fig. 7. Sensor forensics web site at Purdue University.

Using this process with the GLCM feature set, a high classification rate can be achieved between the 10 printers. A classification matrix showing these results using 300 letter "e"s from 12 point Times documents is shown in Fig. 6. Each entry of the matrix is the number of "e"s out of the 300 in the test document that were classified as the printer listed at the heading of its column.

In [11], the performance of this printer identification technique was tested for other font sizes, font types, paper types, and age difference between training and testing data sets. The classification results in these cases remains near 90% except for the case where the training data is older than the testing data, in which case the classification rate is near 70%.

4   Conclusion

Forensic characterization of devices is important in many situations today and will continue to be important for many more devices in the future. We have presented an overview of current characterization techniques for scanners and printers. For images scanned at low resolutions such as 200 DPI, it is possible to successfully identify the scanner model, and groups scanners of the same make and model as a single class (Figure 2). An average classification accuracy of 99.9% is obtained among eleven scanners of 10 different models.. The proposed scheme performs well even with images that have undergone JPEG compression with low quality factors. The design of suitable noise features and use of a denoising filter bank which can capture different kinds of scanning noises results in consistently high classification accuracy of the proposed scheme. Likewise, accurate printer identification from printed documents has been shown to be possible using simple image processing and classification techniques. These techniques are shown to be robust against a variety of variables such as document or printer age and paper type, which might be encountered in an actual forensic examination. All of our work can be found at our web site (Fig. 7), `http://sensor-forensics.org/`.

References

1. Khanna, N., Mikkilineni, A.K., Martone, A.F., Ali, G.N., Chiu, G.T.C., Allebach, J.P., Delp, E.J.: A survey of forensic characterization methods for physical devices. Digital Investigation **3** (2006) 17–28
2. Khanna, N., Mikkilineni, A.K., Chiu, G.T.C., Allebach, J.P., Delp, E.J.: Forensic classification of imaging sensor types. In: Proceedings of the SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents IX. Volume 6505., SPIE (2007) 65050U
3. Khanna, N., Chiu, G.T., Allebach, J.P., Delp, E.J.: Forensic techniques for classifying scanner, computer generated and digital camera images. In: Proceedings of the 2008 IEEE International Conference on Acoustics, Speech, and Signal Processing, Las Vegas, NV (March 2008)
4. Mikkilineni, A.K., Chiang, P.J., Ali, G.N., Chiu, G.T.C., Allebach, J.P., Delp, E.J.: Printer identification based on graylevel co-occurrence features for security and forensic applications. In: Proceedings of the SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents VII. Volume 5681., San Jose, CA (March 2005) 430–440
5. Lukas, J., Fridrich, J.J., Goljan, M.: Digital camera identification from sensor pattern noise. IEEE Transactions on Information Forensics and Security **1**(2) (June 2006) 205– 214
6. Chen, M., Fridrich, J., Goljan, M., Lukas, J.: Determining image origin and integrity using sensor noise. Information Forensics and Security, IEEE Transactions on **3**(1) (March 2008) 74–90
7. Khanna, N., Mikkilineni, A.K., Chiu, G.T.C., Allebach, J.P., Delp, E.J.: Scanner identification using sensor pattern noise. In: Proceedings of the SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents IX. Volume 6505., SPIE (2007) 65051K
8. Foi, A., Katkovnik, V., Egiazarian, K., Astola, J.: A novel local polynomial estimator based on directional multiscale optimizations. In: Proceedings of the 6th IMA Int. Conf. Math. in Signal Processing. Volume 5685. (2004) 79–82
9. Mikkilineni, A.K., Ali, G.N., Chiang, P.J., Chiu, G.T., Allebach, J.P., Delp, E.J.: Signature-embedding in printed documents for security and forensic applications. In: Proceedings of the SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents VI. Volume 5306., San Jose, CA (January 2004) 455–466
10. Ali, G.N., Chiang, P.J., Mikkilineni, A.K., Chiu, G.T.C., Delp, E.J., Allebach, J.P.: Application of principal components analysis and gaussian mixture models to printer identification. In: Proceedings of the IS&T's NIP20: International Conference on Digital Printing Technologies. Volume 20., Salt Lake City, UT (October/November 2004) 301–305
11. Mikkilineni, A.K., Arslan, O., Chiang, P.J., Kumontoy, R.M., Allebach, J.P., Chiu, G.T.C., Delp, E.J.: Printer forensics using svm techniques. In: Proceedings of the IS&T's NIP21: International Conference on Digital Printing Technologies. Volume 21., Baltimore, MD (October 2005) 223–226