Printer-scanner identification via analysis of structured security deterrents

Matthew D. Gaubatz, Steven J. Simske

HP Laboratories
HPL-2009-370

Keyword(s):
device identification, quality assurance, security printing, color tile deterrents

Abstract:

Device identification, the ability to discern the (separate) devices by which a document was produced and/or imaged, can be leveraged in the design of quality assurance (QA) systems as well as the practice of forensic analysis. It is shown that QA metrics associated with printed security markings provide a useful approach for performing multiple device identification, i.e., printer-scanner identification. While some previous methods have focused on properties of sensors to extract signatures from general image data, the proposed approach leverages the highly structured nature of color tile deterrents to predict device (combination) signatures based on a limited amount of information. Constraints introduced by the deterrent structure yield a relatively simple classification strategy with strong performance using a 10-dimensional feature vector. Sixteen printer-scanner combinations (composed from 4 printers and 4 scanners) are tested using this method. Results illustrate device signature prediction performance that is competitive with current state-of-the-art approaches based on physical models of the devices involved.

External Posting Date: December 17, 2009 [Fulltext] Approved for External Publication
Internal Posting Date: December 17, 2009 [Fulltext]

Presented at IEEE WIFS 2009, London, UK, Dec. 7-9, 2009.

PRINTER-SCANNER IDENTIFICATION VIA ANALYSIS OF STRUCTURED SECURITY DETERRENTS

Matthew D. Gaubatz and Steven J. Simske

Hewlett-Packard, Co.
{matthew.gaubatz,steven.simske}@hp.com

ABSTRACT

Device identification, the ability to discern the (separate) devices by which a document was produced and/or imaged, can be leveraged in the design of quality assurance (QA) systems as well as the practice of forensic analysis. It is shown that QA metrics associated with printed security markings provide a useful approach for performing multiple device identification, i.e., *printer-scanner* identification. While some previous methods have focused on properties of sensors to extract signatures from general image data, the proposed approach leverages the highly structured nature of color tile deterrents to predict device (combination) signatures based on a limited amount of information. Constraints introduced by the deterrent structure yield a relatively simple classification strategy with strong performance using a 10-dimensional feature vector. Sixteen printer-scanner combinations (composed from 4 printers and 4 scanners) are tested using this method. Results illustrate device signature prediction performance that is competitive with current state-of-the-art approaches based on physical models of the devices involved.

Index Terms— device identification, quality assurance, security printing, color tile deterrents

1. INTRODUCTION

Device identification is the process by which the type of hardware used to create/capture an image is detected via image analysis. While this functionality provides advantages in forensic studies of images data, it can also be a useful tool in quality assurance (QA) tasks, such as those that perform device-specific inspection, whether or not it is in a security context. For any system that analyzes barcodes [1], color tile deterrents [2], or other printed security markings, it is helpful to know on what device the markings were produced in order to calibrate the system appropriately. Furthermore, if the inspection system is to be deployed in any real-world ecosystems, it is reasonable to assume that different imaging devices will be used in different situations. *Simultaneous device identification* strategies, i.e., methods that can determine on what device a security marking was printed *and* on what device the printed marking was imaged, yield a number of benefits including (1) ease of deployment, (2) the ability to perform on-the-fly device-specific calibration and (3) potential to generate a level of trust between the core of an inspection system and an attached capture device.

Much of the work related to the problem of capture device identification has focused on cameras. Some particularly effective approaches have been based on analyzing artifacts introduced by sensor noise [3, 4]. More recent approaches have applied sensor-noise modeling to *scanner* identification [5, 6]. One of these is of particular interest because it demonstrates performance that is robust under various enhancement operations as well as JPEG compression artifacts [6]. Printer characterization has been achieved by a variety of techniques, including analysis of periodic banding artifacts resulting from half-tone patterns [7] and computation of a distance transform [8]. Another method that has demonstrated strong performance has been based on analyzing print defects surrounding glyphs [9]. This approach is of particular interest not only because of its classification performance, but also due to its inherent efficiency; very little content is required in order to make a very accurate classification decision. Reviews of identification techniques for capture devices and printers can be found in [6, 10].

A key difference between previous approaches and the proposed algorithm is that it is not designed to work with *arbitrary* image data. Rather, it leverages the fact that printed security markings represent highly structured information. The structured nature of this data imposes constraints on the possible outputs captured by a scanner; the proposed algorithm takes advantage of this fact to perform simultaneous device identification using limited amounts of data. This approach is similar in spirit to the glyph-based printer classification procedure [9], though a larger number of effective device classes are addressed and the total number of pixels involved in the input data is considerably smaller. Instead of explicitly focusing on the physical modeling of the production and capture systems, the proposed approach generates features based on QA metrics designed to analyze various aspects of performance of the security markings themselves.

This paper is organized as follows. Section 2 describes inspection systems and reviews color tile deterrents in the context of the device identification problem. Features based on

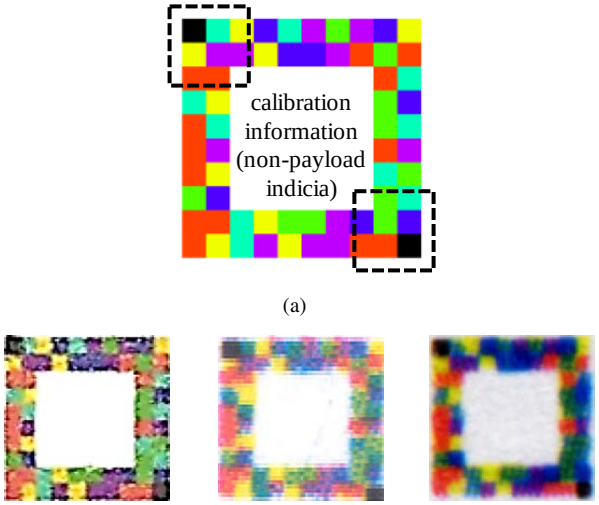

Fig. 1. Example of a color-tile deterrent (top) where the non-payload indicia (NPI), used for calibration during authentication, are indicated with dotted lines. The bottom images represent the same color deterrent produced with three different printer-scanner combinations (HP AiO C3180 + HP AiO C3180, Xerox Phaser 7760 + HP ScanJet 5550c, HP LaserJet 4550 + Epson 1660 Photo). Each deterrent is composed of 8x8 pixel tiles, and was scanned at 600 dots-per-inch (dpi).

these data, and approaches used to classify them are presented in Section 3. Experimental results are discussed in Section 4, and Section 5 concludes the paper.

2. FORENSICS AND QA VIA COLOR TILES

In the context of this work, an *inspection system* could be considered any system that analyzes printed and imaged document data. Often composed of a scanner attached to a computer, a key functionality it provides is forensic analysis of image data. Nevertheless, such systems are versatile enough to provide a range of other QA features, such as the ability to monitor different aspects of print quality or the effectiveness of printed security deterrents (designed for track and trace, authentication, or data embedding). It is thus important to note that many analytical techniques proposed in support of forensic services are also useful in this setting as well.

The device identification strategy herein relies on the presence of a *color tile* security deterrent. Color tiles have initially been proposed as an authentication strategy [2]. Since they demonstrate a high data density [11], these deterrents enable security services, including authentication, for packaging and labels. This work focuses on utilizing the presence of such a deterrent to implement printer-scanner identification. An example of a color tile is illustrated in Figure 1, which includes the original digital representation of the deterrent as well as several variants produced with different printer-scanner combinations. A unique identifier is encoded by setting the colors of each of the 56 payload sub-regions to one of six different choices; authentication can be performed by any device capable of successfully decoding this unique identifier. The corner tiles are used strictly for calibration purposes, and do not carry any information, i.e., they are *non-payload indicia* (NPI).

It has been shown that a variety of no-reference distortion metrics can be used to predict the outcome of this authentication procedure [12], and are designed to analyze different aspects of the decoding process. Along with other QA metrics and qualification parameters described in the next section, these measurements form feature vectors that are used for device identification. The NPI play a key role in computing these measurements.

3. FEATURE VECTORS AND CLASSIFICATION STRATEGY

This section introduces notation and formulae used to describe feature vector components. It also summarizes how the resulting feature vectors are used for classification.

3.1. Notation

Each deterrent is made up of a number of different tiles. Let i index each tile, j index each pixel, and $k \in \{R, G, B\}$ index each color channel of each pixel. Pixels in the distorted deterrents are given by $\hat{\mathbf{p}}_{i,j} = [\hat{p}_{i,j,R}, \hat{p}_{i,j,G}, \hat{p}_{i,j,B}]$. The index c represents the pixel colors in the deterrent, given by $\mathbf{p}_c \in \{\mathbf{R}, \mathbf{G}, \mathbf{B}, \mathbf{C}, \mathbf{M}, \mathbf{Y}\}$, where boldface indicates the vector nature of these colors. The operators $\text{avg}_n\{s_n\}$ and $\text{std}_n\{s_n\}$ denote the first and (square-root) second moments, respectively, empirically computed over any set $\{s_n\}$.

3.2. QA Metrics as Features

A number of different quality grading algorithms are described, all of which are used as feature vector entries. Acronyms are given with each definition for convenience. The NPI play an important role in defining the features since every deterrent has 6 NPI tiles of predefined colors. The mean component pixel values for all pixels in each NPI tile of color c serve as a first set of features:

$$\text{MC (mean color) } \hat{\mathbf{p}}_c = (\text{avg}_j\{\hat{p}_{i_c,j,R}\}, \text{avg}_j\{\hat{p}_{i_c,j,G}\}, \text{avg}_j\{\hat{p}_{i_c,j,B}\}). \quad (1)$$

With 6 NPI color and 3 color channels, these first moment component values account for 18 features. The second moment component values account for another 18.

Once the colors of the NPI tiles are determined, the color of all other tiles in the deterrent can determined by calculating the difference between the mean tile color and each one of the

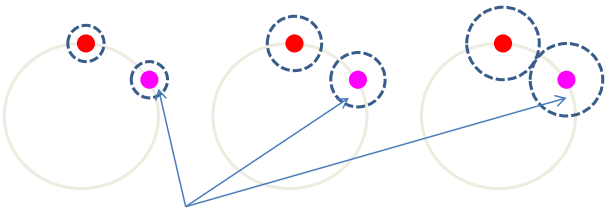

Fig. 2. Illustration of two target colors in hue space, red and magenta, associated with color tile deterrents. The large circle represents the continuum of colors that can be represented in hue space. The dotted lines indicate differently-sized neighborhoods which include hues of other tiles with the same target color. The printer-scanner combination used to process a deterrent has a significant effect on these neighborhoods. The mean colors $\text{avg}_j\{p_{i,j,k}\}$ in each color class fall within them. Authentication errors essentially occur when neighborhoods "overlap", and the mean color variation (MCV) metric measures this effect.

mean NPI colors; this distance can be measured in a vector space (such as RGB) or in hue space. The color of a candidate tile is given as the color of the NPI tile measured to be closest to the candidate. The next two features consist of two metrics that measure different aspects of variation in tiles grouped by assigned color. One measures the mean variation over each collection of tiles associated with a given color:

$$\text{MCV (mean color variation)} = \underset{k}{\text{avg}}\{(\sum_c n_c \cdot \underset{\text{tiles } i \text{ of color } c}{\text{std}}\{\underset{j}{\text{avg}}\{\hat{p}_{i,j,k}\}\})(\sum_c n_c)^{-1}\}, \quad (2)$$

where n_c is the number of tiles i of color c, that is, $n_c = \#\{i; c = \text{argmin}_d ||\hat{\mathbf{p}}_d - \text{avg}_j\{\hat{\mathbf{p}}_{i,j}\}||\}$, and $\hat{\mathbf{p}}_c$ is given by (1). The MCV metric is particularly useful for predicting authentication performance, because when the variation of colors gets large enough, tiles are misclassified and authentication can fail. Figure 2 illustrates this behavior. If only one tile i of color c is present in the deterrent, the standard deviation of $\text{avg}_j\{\hat{p}_{i,j,k}\}$ is 0. A metric measuring the variation of color means (VCM), which can be thought of as a dual of MCV, is used as well.

$$\text{VCM (variation of color means)} = \underset{k}{\text{avg}}\{\underset{c}{\text{std}}\{\underset{\text{tiles } i \text{ of color } c}{\text{avg}}\{\underset{j}{\text{avg}}\{\hat{p}_{i,j,k}\}\}\}. \quad (3)$$

The last two features are metrics computed from measurements averaged over all pixels in the deterrent, mean tile variation (MTV) and mean tile entropy (MTE), which represent the per-pixel standard deviation and entropy, respectively, averaged over each color component. (The formulae for metrics are straightforward and can be found in [12].) Thus, the most general form of the proposed approach yields a length-40 feature vector: the (3) mean color channel values from each 6 NPI tiles, the second moments of the same sets of color channel values, and the four metrics mentioned above (MCV, VCM, MTV and MTE).

3.3. Classification Engine

Many state-of-the-art device identification strategies use either support vector machines (SVMs) or k-nearest-neighbor (k-NN) classifiers, and both methods are evaluated in the following section. The proposed approach is in fact simple enough that a 1-NN classifier can be used to represent the data. The LibSVM package [13] was used to implement the tested SVM classifiers. Unless otherwise specified, radial basis functions were used for kernels.

4. RESULTS AND DISCUSSION

Test data to evaluate the proposed algorithm was created by placing 117 randomly generated color tiles (with 8×8 pixel sub-regions) on a series of test sheets. Sheets were printed and scanned at 600 dots-per-inch (dpi) by one of 16 different printer-scanner combinations. Printers tested included the HP Photosmart C3180 All-In-One (AiO), HP LaserJet 4550, Xerox Phaser 7760 and Xerox Phaser 6250; scanners used were the HP Photosmart C3180 AiO, HP ScanJet 5550c, Epson 1660 Photo and Epson Perfection 1640SU. Color tiles were segmented out of each printed and scanned test sheet, and classification was performed based only on the segmented information. Half of the collected samples were (randomly) used for training and the other half was used for testing. Results reported represent the mean statistics achieved over ten iterations of the training/testing process. Table 1 includes key symbols used to represent printer-scanner combinations in subsequent results.

4.1. Device Identification Accuracy

Device identification accuracy achieved with the proposed features was evaluated by using a classifier to separate the test data samples into 16 respective classes; classification was performed with an SVM for this purpose. The confusion matrix associated with this approach is given in Table 2. Clearly, the method achieves accurate device identification using features computed from only the color tile region. In particular, the prediction accuracy of this method is 0.97. This result is in rough agreement with results previously achieved in a printer-only classification framework [9]. It should be noted that if the classification is performed in a two-step process, i.e., using two SVM classifiers to separately determine the printer and scanner choices, the overall accuracy achieved is almost identical. While there are practical reasons to entertain this type of implementation, this paper is more focused

scanner model	printer model			
	HP Photosmart C3180 AiO	HP LaserJet 4550	Xerox Phaser 7760	Xerox Phaser 6250
HP Photosmart C3180 AiO	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{1,4}$
HP ScanJet 5550c	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{2,4}$
Epson 1660 Photo	$C_{3,1}$	$C_{3,2}$	$C_{3,3}$	$C_{3,4}$
Epson Perfection 1640SU	$C_{4,1}$	$C_{4,2}$	$C_{4,3}$	$C_{4,4}$

Table 1. Printer-scanner combination key.

actual devices	predicted devices															
	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{1,4}$	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{2,4}$	$C_{3,1}$	$C_{3,2}$	$C_{3,3}$	$C_{3,4}$	$C_{4,1}$	$C_{4,2}$	$C_{4,3}$	$C_{4,4}$
$C_{1,1}$	**1.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$C_{1,2}$	0.00	**1.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$C_{1,3}$	0.00	0.00	**0.99**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$C_{1,4}$	0.00	0.00	0.00	**0.99**	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
$C_{2,1}$	0.00	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$C_{2,2}$	0.00	0.00	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$C_{2,3}$	0.00	0.00	0.00	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$C_{2,4}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.99**	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
$C_{3,1}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.98**	0.01	0.00	0.01	0.00	0.00	0.00	0.00
$C_{3,2}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.93**	0.04	0.01	0.00	0.02	0.00	0.00
$C_{3,3}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	**0.95**	0.00	0.00	0.00	0.04	0.01
$C_{3,4}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	**0.90**	0.00	0.00	0.00	0.08
$C_{4,1}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00
$C_{4,2}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	**0.96**	0.00	0.00
$C_{4,3}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	**0.95**	0.00
$C_{4,4}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	**0.95**

Table 2. Probability of actual class as a function of predicted class, when using an SVM-based version of the proposed approach with radial basis functions. The accuracy associated with this approach (the mean diagonal probability) is 0.97.

actual devices	predicted devices															
	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{1,4}$	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{2,4}$	$C_{3,1}$	$C_{3,2}$	$C_{3,3}$	$C_{3,4}$	$C_{4,1}$	$C_{4,2}$	$C_{4,3}$	$C_{4,4}$
$C_{1,1}$	**0.85**	0.00	0.02	0.12	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$C_{1,2}$	0.01	**0.95**	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$C_{1,3}$	0.02	0.01	**0.93**	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$C_{1,4}$	0.09	0.01	0.01	**0.89**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$C_{2,1}$	0.00	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$C_{2,2}$	0.00	0.00	0.00	0.00	0.00	**0.94**	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$C_{2,3}$	0.00	0.00	0.00	0.00	0.00	0.03	**0.96**	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$C_{2,4}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.99**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$C_{3,1}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.61**	0.01	0.01	0.00	0.37	0.00	0.00	0.01
$C_{3,2}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	**0.44**	0.10	0.11	0.01	0.20	0.06	0.07
$C_{3,3}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.08	**0.55**	0.04	0.00	0.03	0.26	0.03
$C_{3,4}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.02	0.00	**0.46**	0.04	0.03	0.00	0.41
$C_{4,1}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	**0.95**	0.00	0.00	0.00
$C_{4,2}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.24	0.01	0.02	0.00	**0.65**	0.03	0.05
$C_{4,3}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.36	0.02	0.00	0.02	**0.57**	0.01
$C_{4,4}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.04	0.24	0.01	0.12	0.08	**0.40**

Table 3. Probability of actual class as a function of predicted class, achieved with a two-stage approach using implementations of the sensor-noise-based scanner identification procedure in [6] and a version of the glyph-based printer identification algorithm in [9]. The accuracy associated with this two-stage method is 0.76.

on demonstrating the separability of deterrents using the proposed features.

Table 3 illustrates the results achieved with a two-stage approach that uses two state-of-the-art device identification algorithms (with different classifiers and different features) to determine the scanner and printer associated with each deter-

rent. The scanner identification algorithm is the sensor-noise-based approach in [6], which analyzes statistics that take into account properties such as noise correlation, and uses an SVM classifier. The printer identification algorithm measures aspects of joint histograms associated with pixels located near one another [9] and performs classification with a 5-NN classifier; in the original paper this procedure is applied on regions with glyphs, but herein, color tile regions are considered instead. The scanner identification and printer identification accuracies are 0.84 and 0.89, respectively, and the overall accuracy of this approach is 0.76.

Given that the pre-existing methods involved [6, 9] were designed for other applications (general photographic data, or images with glyphs, respectively) *and* for use with larger blocks of data, the accuracy achieved is reasonable. Because the new method proposed herein is designed to address a narrower class of images, however, it leverages the associated constraints to create a simple, effective solution. The resulting advantage is considerable, since the accuracy of the proposed approach is about twenty percent higher and the resulting error is considerably smaller. The caveat that accompanies this performance improvement is that it can be achieved only in the presence of color tile security deterrents, and not with any general image data.

4.2. Reduced Complexity Classification

To illustrate the flexibility of the proposed approach, as well as the appropriateness of the feature set, a subset of the presented features and a very simple classifier is constructed. In order to train this classifier, the mean feature vector for each class is computed by linearly averaging all samples in the class. The classifier then assigns the class of a candidate feature vector as that associated with the mean feature vector closest to the candidate feature vector. (This classifier is equivalent to a 1-NN classifier, trained with a single feature vector per class.) The reduced feature set consists only of the first order red and green NPI statistics coupled with the four QA metrics in [12]; the statistics associated with this pair of colors yields the best performance of any pair in conjunction with the QA metrics. The accuracy associated with this reduced complexity approach is 0.87. This performance indicates that the candidate samples are highly separable in the feature space, since this simple classifier outperforms the more sophisticated two-stage approach tested earlier.

5. CONCLUSION

A device classification scheme was presented that leverages existing printed security deterrents coupled with existing quality assurance metrics designed to predict the outcome of an authentication procedure. Preliminary results indicate that this method yields a classification accuracy competitive with other state-of-the-art approaches, and suggest that highly accurate printer-scanner identification is indeed possible by devoting a small portion of a printed document to a structured signal. In fact, a classifier based only on the distances to mean feature vectors (from each class) can improve upon the separability of device combinations achieved using state-of-the-art methods designed for more general data. It is important to consider that in a security ecosystem, the benefit associated with the presented ideas is not necessarily just the ability to perform simple device identification on images created by ever-growing combinations of devices, but also the ability to identify images that were produced outside the ecosystem. Future work will involve investigation of meta-algorithmic classification patterns [14] for improved accuracy.

6. REFERENCES

[1] Microsoft Research, *High Capacity Color Barcode Technology*, 2009, available at `http://research.microsoft.com/en-us/projects/hccb/about.aspx`.

[2] S. J. Simske and J. S. Aronoff, "Spectral pre-compensation and security deterrent authentication," *JIST*, vol. 51, no. 1, pp. 86–95, 2007.

[3] J. Lukas, J. Fridrich, and M. Goljan, "Determining digital image origin using sensor imperfections," in *Proc. SPIE International Conference on Image and Video Communications and Processing*, 2005, vol. 5685, pp. 249–260.

[4] J. Lukas, J. 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.

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

[6] N. Khanna, A. K. Mikkilineni, and E. J. Delp, "Scanner identification using feature-based processing and analysis," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 1, pp. 123–139, 2009.

[7] G. N. Ali, P-J Chiang, A. K. Mikkilineni, J. P. Allebach, G.T.-C. Chiu, and E. J. Delp, "Intrinsic and extrinsic signatures for information hiding and secure printing with electrophotographic devices," in *Proc. IS&Ts NIP19*, September 2003, vol. 19, pp. 511–515.

[8] W. Deng, Q. Chen, F. Yuan, and Y. Yan, "Printer identification based on distance transform," in *Proc. First International Workshop on Intelligent Networks and Intelligent Systems*, November 2008, pp. 565–568.

[9] A. K. Mikkilineni, P-J Chiang, G. N. Ali, G. T.-C. Chiu, J. P. Allebach, and E. J. Delp, "Printer identification based on textural features," in *Proc. IS&Ts NIP20*, October/November 2004, vol. 20, pp. 306–311.

[10] 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 Investigations*, vol. 3, pp. 17–28, 2006.

[11] S. J. Simske, M. Sturgill, and J. S. Aronoff, "Effect of copying and restoration on color barcode payload density," in *Proc. ACM DocEng 2009*, September 2009, pp. 127–130.

[12] M. D. Gaubatz, S. J. Simske, and S. Gibson, "Distortion metrics for predicting authentication functionality of printed security deterrents," in *Proc. ICIP (to appear)*, 2009.

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

[14] S. J. Simske, D. W. Wright, and M. Sturgill, "Meta-algorithmic systems for document classification," in *Proc. ACM DocEng*, 2006, pp. 98–106.