Document Imaging Security and Forensics Ecosystem Considerations

Steven Simske, Margaret Sturgill, Guy Adams, Paul Everest

HP Laboratories
HPL-2010-156

Keyword(s):

Security, Forensics, Color Tiles, 3D Bar Codes, Document Fraud, High-resolution Imaging

Abstract:

Much of the focus in document security tends to be on the deterrent - the physical (printed, manufactured) item placed on a document, often used for routing in addition to security purposes. Hybrid (multiple) deterrents are not always reliably read by a single imaging device, and so a single device generally cannot simultaneously provide overall document security. We herein show how a relatively simple deterrent can be used in combination with multiple imaging devices to provide document security. In this paper, we show how these devices can be used to classify the printing technology used, a subject of importance for counterfeiter identification as well as printer quality control. Forensic-level imaging is also useful in preventing repudiation and forging, while mobile and/or simple scanning can be used to prevent tampering - propitiously in addition to providing useful, non-security related, capabilities such as document routing (track and trace) and workflow association.

External Posting Date: October 21, 2010 [Fulltext] Approved for External Publication
Internal Posting Date: October 21, 2010 [Fulltext]
To be published in ACM DocEng 2010, Manchester UK

Document Imaging Security and Forensics Ecosystem Considerations

Steven Simske, Margaret Sturgill
Hewlett-Packard Labs
3404 E. Harmony Rd.
Fort Collins CO 80528 USA
{Steven.Simske,Margaret.Sturgill}@hp.com

Guy Adams
Hewlett-Packard Labs
Filton Rd. Stoke Gifford
Bristol UK BS34 8QZ
Guy.Adams@hp.com

Paul Everest
Hewlett-Packard Co.
1000 NE Circle Blvd.
Corvallis OR 97330 USA
Paul.Everest@hp.com

ABSTRACT
Much of the focus in document security tends to be on the deterrent — the physical (printed, manufactured) item placed on a document, often used for routing in addition to security purposes. Hybrid (multiple) deterrents are not always reliably read by a single imaging device, and so a single device generally cannot simultaneously provide overall document security. We herein show how a relatively simple deterrent can be used in combination with multiple imaging devices to provide document security. In this paper, we show how these devices can be used to classify the printing technology used, a subject of importance for counterfeiter identification as well as printer quality control. Forensic-level imaging is also useful in preventing repudiation and forging, while mobile and/or simple scanning can be used to prevent tampering — propitiously in addition to providing useful, non-security related, capabilities such as document routing (track and trace) and workflow association.

Categories and Subject Descriptors
I.4.1 [**Image Processing and Computer Vision**]: Digitization and Image Capture–*Scanning*. I.4.6 [**Image Processing and Computer Vision**]: Segmentation. K.6.5 [**Management of Computing and Information Systems**]: Security and Protection — *Authentication*.

General Terms
Algorithms, Security

Keywords
Security, Forensics, Color Tiles, 3D Bar Codes, Document Fraud, High-resolution Imaging.

1. INTRODUCTION

1.1 Document Fraud
Document fraud, including altered, forged and other counterfeit documents [1], is a prevalent and growing concern for businesses. Document fraud is a broad suite of problems, including the intentional alteration of the information in the document (tampering), the copying or unapproved creation of security documents (forging), and obstructing the ability to establish a document's authenticity (repudiation).

One means of addressing document fraud is to use secure printing workflows, where user authentication is required during scanning and printing. However, even when this can be enforced system-wide, it does not prevent document tampering or re-use. To address this, unique information must be associated with the document while it exists in physical form (e.g. on paper, label, packaging); not just while in electronic form.

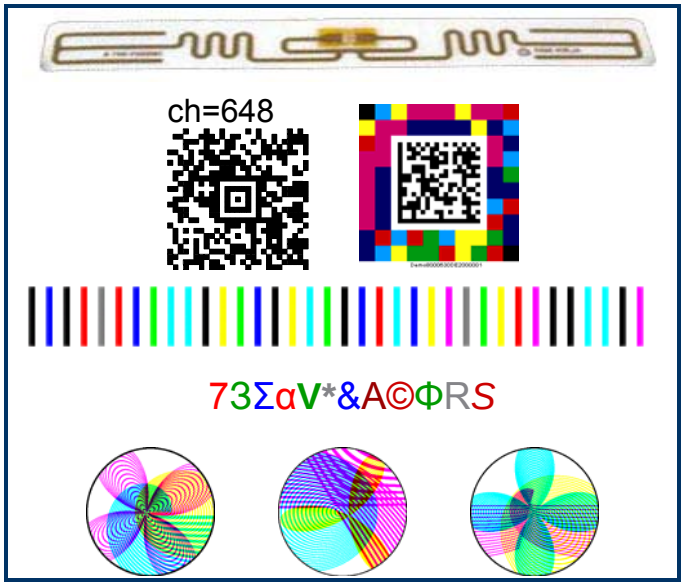

Figure 1. Examples of overt marks, described in reading order (top to bottom, left to right by row). Smart label RFID and antenna (top); Aztec 2D barcode (left, second row, which encodes 648 binary "characters"; ch=648) and 3D color barcode/2D DataMatrix barcode/microtext cluster (right, second row); color bars (third row), variable-colored glyphs (fourth row) and variable guilloches (bottom row).

Appropriately designed, a printed document security approach can simultaneously address tampering, forging and repudiation. Historically, much of the focus in document security has tended to be on the deterrent — the physical (printed, manufactured) item placed on a document, which is often used for routing (e.g. from one point to another in a workflow) in addition to security purposes. Deterrents include the overt (visible, relatively easy to validate), the covert (hidden or non-obvious, often requiring a human or special equipment to validate) and the forensic

DocEng '10, September 21–24, 2010, Manchester, UK.

(providing statistically difficult to reproduce, item unique data). Overt marks include a variety of barcodes, graphics, variable text/character strings, and other variable data printing (VDP) marks, as illustrated in Figure 1. Importantly, variable data printing can be accomplished concomitantly with variable RFID to create an electronic/image hybrid smart label. Overt deterrents should be easy to identify and use. Because of their familiarity as information-carrying marks, there are a wide variety of barcode technologies available. We will review these in the next section.

Covert marks, on the other hand, are frequently used as part of a "check list" — if the covert marks are not there, a "warning" is created and a downstream process — further inspection, notification of supply chain partners, etc. — is activated. Covert marks can also encode unique data (or replicate/hybridize overt data), furthering the payload density (bits per unit area) of the security feature. Covert marks that carry information payload can be used just like overt marks in a security workflow.

Forensic marks need not be intentionally written. Quantitative forensic metrics can be garnered directly from the printing signature of the printer [2], from the unpredictable interaction of the ink with the substrate [3] due to the preferential wicking of aqueous inks along the longitudinal direction of the cellulose in many substrates [4], and even from the unique structure of the substrate itself [5]. Forensic marks uniquely identify (with a measurable statistical probability) a single item since they represent characteristics not easily replicated or copied. Unlike overt and covert marks, forensic marks are used primarily for identification purposes (and not for information-carrying).

A set of multiple deterrents are not always reliably read by a single imaging device, and so cannot simultaneously provide all of the document security tasks described above. We herein investigate how a relatively simple deterrent can be used in combination with multiple imaging devices to meet all of the document security requirements. Forensic-level imaging can be used to prevent repudiation and forging, while scanning can be used to detect tampering in addition to providing useful, non-security related, capabilities such as document routing (track and trace) and workflow association. In addition, we show how this set (or "ecosystem") of imaging devices can be used to classify the printing technology used, a subject of importance for counterfeiter identification as well as printer quality control.

1.2 Prior Art

Barcodes are one of the most familiar of all printed information-carrying marks (ICMs). In addition to the one-dimensional (1D) universal product code (UPC) barcodes [6] often read at point-of-sale, two-dimensional (2D) barcodes such as the Data Matrix [7] and Aztec [8] barcodes shown in Figure 1 have become familiar for use in couponing, extended packaging and other mobile retailing applications. One of us (Simske) has participated in the GS1 efforts to define extended packaging and mobile retail approaches using barcodes [9][10].

The GS1 recommendations are intended to help standardize an already crowded field. The Open Mobile Alliance [11], for instance, has provided facilitation of "global user adoption of mobile data services by specifying market driven mobile service enablers that ensure service interoperability across devices, geographies, service providers, operators, and networks while allowing businesses to compete through innovation and differentiation" for several years.

With the growing prevalence of 2D barcodes, new higher-density 3D barcodes — that is, 2D barcodes with more than one bit per module — are being invented and productized. Microsoft offers the high-capacity color barcode (HCCB), a triangular-module based color mark, which can be used for mobile applications in addition to high-density data carrying [12][13]. The Colorzip [14] color barcode is applicable to a variety of mobile content-accessing workflows.

Research on color barcodes is active and varied. Bulan et al. [15] are focused on using multiple color planes and halftone dot orientations to provide densities as high as 3000 bytes/in^2. Villán et al. [16] provide several multilevel (grayscale or color) barcode approaches providing up to 2400 bytes/in^2 information density. Limiting the color set to the primary printing colors — cyan, magenta and yellow, or CMY — and using a variety of error-correcting code (ECC) approaches, Mayer et al. provide a large footprint information-containing 3D print code [17].

Previously, we have focused on using print-scan (PS) cycle information to significantly improve security payload density (PD); that is, the amount of security information that can be printed in a given area. Structural pre-compensation (StrPC) is used to optimize the relative size of the ICMs over the background. We have used StrPC to anticipate the "module gain" of 2D barcodes over their background [18] and to reduce the deleterious effect on of the spread of inkjetted ink into the substrate — which otherwise creates color overlap on the boundaries of color modules [17]. The extent of StrPC needed depends on the ink spread into/on the surface of the substrate.

Spectral pre-compensation (SPC) is used to optimize the colors in a 3D color barcode prior to printing. We have shown that SPC is a highly effective means of increasing the payload density (PD), especially on lower image quality substrates such as office paper [19]. In many cases, SPC can double the PD. However, SPC is tied to a single (printer, imager) combination.

1.3 Authentication

Figure 2. 3D barcode "color tile" configuration used for the test sheets. The individual tiles, of which there are 64 in total, sized 5-10 pixels on a side at 600 dpi.

The primary role of a barcode is to embed tacit information (a bit string of a given length) that can be later extracted using imaging hardware — a barcode reader, scanner, digital camera, phone camera, or other inspection device, for example. We used a color tile (3D color barcode) as described in [19] as our ICM (Figure 2). A color tile ICM with 56 payload-bearing elements carries 144 bits of data — equivalent to a 96-bit GS1 SGTIN with an additional 50% payload.

Authentication consists of imaging the barcode and then using an algorithm to decode the embedded information. Authentication accuracy in the case of a color tile is defined as interpreting all 56 colors correctly — that is, with no ECC. Authentication accuracy,

therefore, provides a direct measure of the image quality and is used to determine the deployment settings — size, shape, StrPC and SPC among them — for the color tile.

Previous work has shown that, for the Dr. CID [3] and scanner [19] imaging hardware we use in this paper, printing the color tiles so that the individual modules range from 5-10 pixels (at 600 dots/inch, or dpi, print resolution) on a side is sufficient for testing authentication accuracy.

1.4 Forensics

We have developed two systems for "forensic" image analysis. The first is a hardware/software combination called Dr. CID (an acronym for the Dyson relay CMOS imaging device). Dr. CID is a USB-powered mini-appliance [3] that is capable of resolving spatial features of less than 5 microns with 1:1 magnification. The device incorporates a single Dyson relay lens in series with a mirror and a low cost 3.2 μm/pixel, 3 Mpixel CMOS color image sensor. With a self-contained (white LED, though other LEDs can be readily accommodated) illumination source, this device affords the capture of individual typed characters along with their printing parasitics. The Dr. CID provides a modulation transfer function (MTF) of ≈ 3.5 microns in color. This is achieved in a handheld "contact" use model (with a field of view of approximately 5 x 6 mm) with a uniform diffuse illumination source, with no unwanted reflections. Dr. CID successfully resolves the USAF 1951 resolution chart's group 7, element 2 at 25-30% contrast, equating with an MTF of 288 lines/mm = 3.5 μm (7257 dpi). The inherent 1:1 magnification of DR CID also means that the variability of the scale of the images can be tightly controlled by the tolerances of the lens, further improving image analysis.

The second forensic system is software-based. Named the image-based forensic system, or IBFS [20], it is an imaging system that typically uses a small set of pre-classified ("training") images for initial training, and thereafter adaptively classifies and aggregates images from multiple sources as they join the population to be classified. The system can also work without any training images whatsoever. Multiple classes of images are identified, and can be compared for proximity based on a weighted distance approach. The system currently uses a set of 420 features which are pruned based on correlation procedures to a smaller (typically 60-120) features. This filtered set of features, or feature signature, is particularly important when there are training sets, as it is used for the clustering of non-training images thereafter.

2. EXPERIMENTS PERFORMED

In this section, we describe the experiments performed. First, we describe what we printed and scanned to simulate the manufacturing and imaging aspects of using information-containing marks for security, including authentication and forensics. Next, we describe the tests performed using a "general" image forensics system.

2.1 Printing and Imaging Specifications

We tested several thermal ink-jet and several dry electrophotography (DEP) — or laserjet — based printers, along with several substrates (office paper, soft gloss and glossy paper), searching for the combination of (IJ+substrate) and (LJ+substrate) that would give us the **poorest** results. Poor results were sought so that we could determine the sensitivity of the authentication and forensic imaging approaches to printing imperfections and gain insight into the role of inspection in both authentication and forensics.

The poorest results for an inkjet printer were obtained using no pre-compensation — that is, no StrPC or SPC — on HP Multi-Purpose Plain Office Paper (hereafter "Office Paper") substrate. An additional reason for not applying SPC is that the use of two different imaging devices precludes it (SPC is tuned to a single type of imaging device). We used the HP Photosmart C6280 inkjet all-in-one as the inkjet printer (hereafter "IJ"). Glossy substrates provided PD > 2400 bytes/in^2 even without pre-compensation. The poorest results for the laserjet printers tested also coincided with the use of Office Paper; however, an additional defect was noted. On one of our HP 3600 color laserjet (hereafter "LJ") printers, there was a substantial color plane misalignment, with the yellow toner cartridge printing roughly 1/300th of an inch "higher" on the page than the magenta toner cartridge. As a consequence, we had a different print-related defect for the IJ and LJ printers chosen:

(1) The IJ printer and Office Paper substrate result in substantial ink bleed between neighboring tiles, which is clearly evident along the overall periphery of the ICM.

(2) The LJ printer, while displaying much less pronounced printing parasitics (small toner satellites around the periphery and overlapping neighboring tiles), had substantial (2 pixels at 600 dpi) color plane mis-registration (most notable in the Yellow color plane), which would be expected to significantly affect authentication.

2.2 Test Sheets

The experiments performed use a color tile deterrent described in [19]. The color tile information-carrying mark (ICM) consists of six data-carrying colors: red (R), green (G), blue (B), cyan (C), magenta (M) and yellow (Y). In addition, two black (K) tiles and one of each of the six color squares are placed in the upper left and lower right corners to aid in registration (these are termed "non-payload indicia", or NPI), as shown in Figure 2. White (W) surrounds the color tile deterrent and is used for segmentation purposes. The white interior of the ICM is typically used to carry a 2D barcode, but is not printed for our test purposes.

All images are printed and scanned in 8-bits/channel, 3-channel RGB-space, and the individual pixels P(i,j) in the images are designated as having R, G and B values in the triplet {r,g,b}. We used a static color array for the purposes of deploying the IBFS as part of our testing [20]. The color tiles are arranged as in Figure 2. The tiles are arranged on a 10x10 grid, so that if necessary, the deterrent can support 100 tiles.

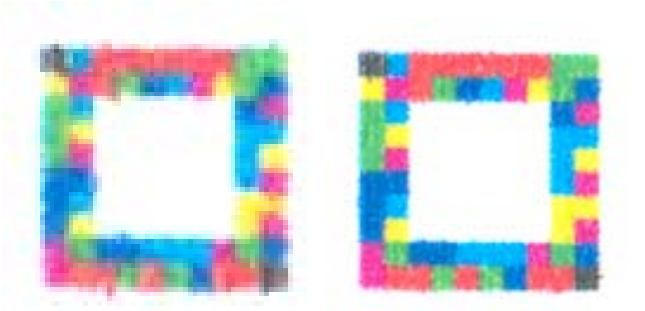

Figure 3. IJ printer + 8200 scanner at 5x5 and 10x10 pixel (at 600 dpi) size (left, right).

Test sheets of multiple color tile deterrents were printed for both printers (IJ and LJ). Individual tiles were sized from 5x5 to 10x10 pixels, and all printing was performed at 600 dots/inch (dpi). The 5x5 pixel tiles correspond to a PD of 4650 bytes/in^2. Due to differences in the overall deterrent size, the test sheets included 165, 150, 140, 140, 130 and 117 full deterrents, respectively, for individual tile sizes of 5x5, 6x6, 7x7, 8x8, 9x9 and 10x10 pixels. Thus, from 6552-9240 individual tiles were tested per sheet. The color tile deterrents range from 1/6 x 1/6 (5x5 pixel tiles) to 1/3 x 1/3 (10x10 pixel tiles) inches on a side, separated by 1/2 inch of white space in each direction. All tests were performed in parallel on the IJ and LJ printers. Once printed, the HP Scanjet 8200 (hereafter "8200") and Dr. CID imager were used to capture the color tile images of the exact same printed ICMs.

The scanner and, separately, Dr. CID imager were used to capture the same (first image on the page in reading order) image ten separate times. These "same" images are, ideally, identical, but in reality provide us with a measure of a systemic variance, the imaging variance: $\sigma^2_{imaging}$.

Figure 3 shows representative color tiles printed using the IJ printer and captured using the 8200. The left image has 5x5 pixel modules (original size at 600 dpi); the right image has 10x10 pixel modules. The images show very high contrast — the Office Paper background has effectively disappeared into pure white — and illustrate the pseudopod-like protrusions from the otherwise square periphery of the ICMs due to ink wicking along the longitudinal axes of the substrate fibers [4].

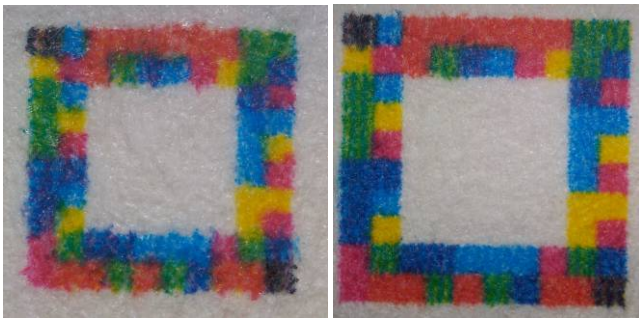

Figure 4. IJ printer + Dr. CID imager at 5x5 and 10x10 pixel (at 600 dpi) size (left, right).

Figure 4 illustrates the images captured with the Dr. CID imager. The contrast is much lower, with the Office Paper showing as light gray and the substrate roughness and variability obvious. This additional variance in the image we term $\sigma^2_{substrate}$. Because of the high-contrast, we can safely define $\sigma^2_{substrate}$ = 0 for scanning with the 8200.

Figure 5 illustrates representative samples printed with the laserjet (LJ) printer and scanned with the 8200 scanner. Again, the Office Paper background is effectively eliminated by the high contrast. The color plane mis-registration is illustrated by the yellow band along the top of the images — more pronouncedly for the image on the left, since it is magnified by a factor of 2 in comparison to the image on the right. Additionally, the halftoning used to create the red, green and blue tiles is quite evident. Note that halftoning is also used for the IJ printing, but the ink spread into the substrate greatly reduces its visual impact after printing.

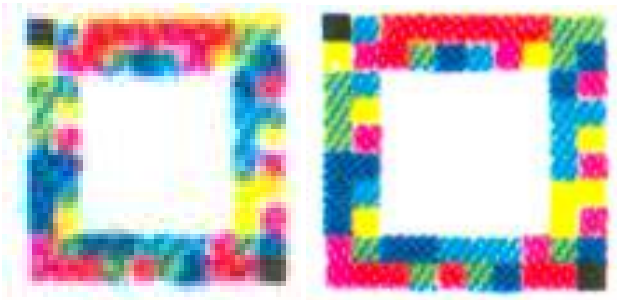

Figure 5. LJ printer + 8200 scanner at 5x5 and 10x10 pixel (at 600 dpi) size (left, right).

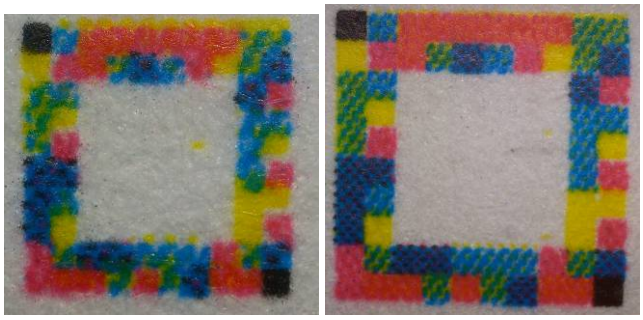

Figure 6. LJ printer + Dr. CID imager at 5x5 and 10x10 pixel (at 600 dpi) size (left, right).

Figure 7. Up-close 5x5 pixel size (at 600 dpi original image printing specifications) captured using the Dr. CID imager showing the ink wicking (IJ printing, left) and the color plane mis-registration (LJ printing, right).

Figure 6 illustrates representative examples of the LJ prints after being captured with the Dr. CID imager. As with Figure 4, there is lower contrast than the scanned images. The Office Paper and the appreciable $\sigma^2_{substrate}$ are also evident. Figure 7 illustrates the different printing-related "defects" on the IJ and LJ printers.

2.3 Imaging and Authentication

All test pages were printed using both the IJ and LJ printers, and then scanned 600 dpi, 24 bit color, default settings, using the HP ScanJet 8200 and, separately, the Dr. CID imager. The images were then segmented using custom software [19]. Non-white (non-W) pixels were defined as those having at least one {r,g,b} value < 128 (after contrast adjustment in the case of the Dr. CID imager) and then run-length smeared using 1/150 of an inch (original size) run-length gap. Connected components were formed, and then prepared for authentication.

Each connected component (presumably a single 10x10 tile ICM as in Figure 2) was divided into 10x10 sections, which correspond to the individual tile zones. The corner tiles were checked for black pixels to ensure orientation was correct, and the individual tile images (100/deterrent) were eroded by 15% along each edge, resulting in, for example, 7x7 pixel, 24-bit color images for each tile in a 10x10 pixel deterrent. Note that this level of erosion is insufficient to compensate for color plane mis-registration below 10x10 pixel size for the LJ prints. The colors of these tiles were determined as described next and compared to the printed sequence of tiles.

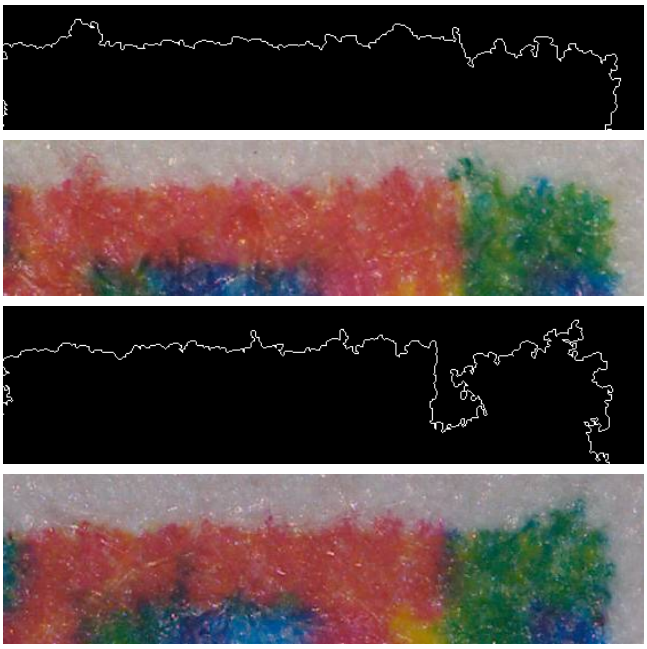

Figure 8. High-resolution perimeters (above) of portions (below) of two Dr. CID captured images for 5x5 pixel size (at 600 dpi original image IJ printing specifications) color tile ICMs. As described in [3], the perimeters account for ink wicking in addition to dark zones within the marks, etc. The most effective matching parameters are 5-15 standard deviations apart for imaging the same mark twice versus imaging two different marks.

The authentication approach assigned the color of each tile based on the minimum of the angular distance of the mean hue of the tile sub-segment and the hues of the six color NPI. That is, the minimum absolute hue difference between the tile sub-segment's mean {r,g,b} value and the hue of the NPIs' {**r,g,b**} values assigned that NPI's color to the tile. Hue angle of {R,Y,G,C,B, and M} is {0,60,120,180,240, and 300}, although the actual colors after scanning are somewhat different (red and magenta tend to move closer together after printing and scanning, for example). Authentication accuracy was determined based on the percentage of 56-tile sequences that were correctly read in their entirety (that is, no sequence errors). Even one sequence error was considered an "authentication" failure. Thus, the reported accuracy is "full deterrent accuracy", and is appropriate since no error code checking (ECC) is used.

2.4 Forensic Evaluation

Even the smallest (tile sizes of 5x5 pixels at 600 dpi) ICMs are huge in comparison to individual typed characters (such as the letter "a" used in [3]). Therefore, forensic analysis of the perimeters of the color tiles was readily accomplished, as shown in Figure 8.

We have previously shown the forensic (less than 10^{-9} probability of false match) capability for the Dr. CID on individually printed single characters [3], so it is unremarkable that this approach leads to forensic-level identification of the perimeters of individually printed tiles. Our focus in this paper, instead, was to determine if we could use the IBFS and Dr. CID to perform forensic analysis of the entire ICM, excepting the perimeter. To that end, we used the IBFS and different sets of the images in a series of 10 tests described next.

IBFS Test 1) For each tile size, perform IBFS aggregation analysis on all 20 IJ and all 20 LJ images (10 of the same and the 10 different ICMs) captured with the Dr. CID imager. No training set is used.

IBFS Test 2) For each tile size, perform IBFS aggregation analysis on all 127-175 (depending on size of tile) IJ images and all 127-175 LJ images (10 of the same and the 117-165 different ICMs) captured with the 8200 scanner. No training set is used.

IBFS Test 3) For each tile size, perform IBFS aggregation analysis on all 20 IJ images captures with the Dr. CID imager. 5 of the "same" images are the training set.

IBFS Test 4) For each tile size, perform IBFS aggregation analysis on all 20 LJ images captures with the Dr. CID imager. 5 of the "same" images are the training set.

IBFS Test 5) For each tile size, perform IBFS aggregation analysis on all 20 IJ images captures with the Dr. CID imager. 5 of the "different" images are the training set.

IBFS Test 6) For each tile size, perform IBFS aggregation analysis on all 20 LJ images captures with the Dr. CID imager. 5 of the "different" images are the training set.

IBFS Test 7) For each tile size, perform IBFS aggregation analysis on all 20 IJ images captures with the 8200 scanner. 5 of the "same" images are the training set.

IBFS Test 8) For each tile size, perform IBFS aggregation analysis on all 20 LJ images captures with the 8200 scanner. 5 of the "same" images are the training set.

IBFS Test 9) For each tile size, perform IBFS aggregation analysis on all 20 IJ images captures with the 8200 scanner. 5 of the "different" images are the training set.

IBFS Test 10) For each tile size, perform IBFS aggregation analysis on all 20 LJ images captures with the 8200 scanner. 5 of the "different" images are the training set.

For tests 1-2, correctly classified (that is, aggregated) samples will be assigned to groups containing no images from the pool for the other printing technology; that is, groups that are all IJ or all LJ only. For tests 3-10, correctly classified/aggregated samples will be assigned to the training set if they are of the same type, and assigned to a different set (again, non-mixed) if they are of a different type.

3. RESULTS

Here, we define the sources of image distortion, then evaluate the imaging for utility in authentication and forensic evaluation via batch inspection using the IBFS. Imaging is performed using the USB-powered microscope (Dr. CID) and a desktop scanner.

3.1 Image Distortion Due to Printing and Imaging (Scanning or Dr. CID Capture)

As described above, we identified two types of printing defects during our selection of the final (printer, substrate) combinations for full testing. We name the ink-spread variance $\sigma^2_{ink-spread}$ and the color-plane mis-registration variance $\sigma^2_{CPmis-R}$. Taking into account the variances described above — $\sigma^2_{imaging}$ and $\sigma^2_{substrate}$, we can generalize to the following:

1) The LJ printer has high $\sigma^2_{CPmis-R}$ and some $\sigma^2_{ink-spread}$.

2) The IJ printer has high $\sigma^2_{ink-spread}$ and some $\sigma^2_{CPmis-R}$.

3) The Dr. CID provides high $\sigma^2_{substrate}$.

4) The scanner provides low or no $\sigma^2_{substrate}$.

5) Comparing the 10 "same" samples defines $\sigma^2_{imaging}$.

3.2 Authentication

Table 1. Authentication results (correct/total) for LJ printing. The Dr. CID imager consistently achieves an equivalent authentication accuracy to the scanner at one pixel smaller size (e.g. 95% overall accuracy at size 9x9, whereas the scanner has 90.6% overall accuracy at size 10 x 10).

	Dr. CID		Scanner	
Size of Tile	**Same**	**Different**	**Same**	**Different**
5 x 5	0/10	0/10	0/10	3/165
6 x 6	0/10	3/10	0/10	6/150
7 x 7	0/10	6/10	0/10	31/140
8 x 8	4/10	5/10	1/10	44/140
9 x 9	10/10	9/10	0/10	63/130
10 x 10	10/10	10/10	10/10	105/117

The authentication data are given in Tables 1 and 2. For both types of printers, the Dr. CID imager provided 100% authentication for smaller tile sizes (and thus provided higher payload density). However, the effect was much less pronounced for the LJ (Table 1) than for the IJ (Table 2), presumably as a consequence of the $\sigma^2_{CPmis-R}$ associated with the LJ.

Table 2. Authentication results (correct/total) for IJ printing. The Dr. CID imager achieves 100% accuracy at tile size 5x5, whereas the scanner does not achieve 100% accuracy until the tile size is 10x10 (meaning ¼ the payload density).

	Dr. CID		Scanner	
Size of Tile	**Same**	**Different**	**Same**	**Different**
5 x 5	10/10	10/10	0/10	9/165
6 x 6	10/10	9/10	2/10	20/150
7 x 7	10/10	10/10	7/10	53/140
8 x 8	10/10	9/10	7/10	84/140
9 x 9	10/10	10/10	10/10	111/130
10 x 10	10/10	10/10	10/10	117/117

3.3 Forensic Evaluation

Forensic evaluation consists of the IBFS Tests 1-10, as described above. Table 3 shows the results for Test 1 (Dr. CID imager, all IJ and LJ samples together, no training data). At 5x5 tile size, only 11 of 40 (27.5%) of the samples are appropriately aggregated. At 6x6 tile size, 45% of the samples are appropriately aggregated. At size 7x7 or above, the accuracy is 0% since all samples belong to combined groupings of IJ and LJ images.

Table 3. (IBFS Test 1) Dr. CID imaged, combined laserjet (LJ) and inkjet (IJ) printer samples (n=20 each) results for image classification when all 40 samples are analyzed simultaneously (no training samples), for each of the 6 sizes investigated. Number of aggregates in each grouping is indicated in parentheses.

Size of Tile	IJ → IJ	LJ → LJ	(IJ+LJ)→Combined
5 x 5	11 (1)	0	29 (3)
6 x 6	0	18 (2)	22 (2)
7 x 7	0	16 (3)	24 (3)
8 x 8	0	5 (1)	34 (5)
9 x 9	0	0	40 (4)
10 x 10	0	0	37 (4)

Table 4 presents the results for IBFS Test 2 (8200 Scanner, all IJ and LJ samples together, no training data). At all tile sizes, there is 100% accuracy of aggregation (no classification errors). All samples are assigned to clusters consisting solely of samples printed using the same printer. Thus, the scanned images provide more accurate classification in spite of containing only 1/144 as much data (they are scanned at 600 dpi, while the Dr. CID images are captured at 7200 dpi).

Table 4. (Test 2) Scanner (HP 8200) image, combined laserjet (LJ) and inkjet (IJ) printed samples (n=20 each) results for image classification when all 40 samples are analyzed simultaneously (no training samples), for each of the 6 sizes investigated. Number of aggregates in each grouping is indicated in parentheses.

Size of Tile	IJ → IJ	LJ → LJ	(IJ+LJ)→Combined
5 x 5	175 (9)	175 (10)	0
6 x 6	160 (7)	160 (10)	0
7 x 7	150 (7)	150 (10)	0
8 x 8	150 (7)	150 (10)	0
9 x 9	140 (7)	140 (10)	0
10 x 10	127 (1)	127 (6)	0

Table 5. IBFS Test 3. See text for details of the test (IJ printer, Dr. CID imager, training set of 5 "same" images). Data presented as (number correct)/(number of images).

Size of Tile	"Same"	"Different"	Accuracy
5 x 5	5/5	0/10	0.333
6 x 6	5/5	1/10	0.4
7 x 7	5/5	0/10	0.333
8 x 8	5/5	0/10	0.333
9 x 9	5/5	0/10	0.333
10 x 10	2/5	0/10	0.133

Table 6. IBFS Test 4. See text for details of the test (LJ printer, Dr. CID imager, training set of 5 "same" images). Data presented as (number correct)/(number of images).

Size of Tile	"Same"	"Different"	Accuracy (%)
5 x 5	5/5	0/10	0.333
6 x 6	5/5	5/10	0.667
7 x 7	5/5	1/10	0.4
8 x 8	5/5	0/10	0.333
9 x 9	5/5	1/10	0.4
10 x 10	5/5	3/10	0.533

The results for IBFS Tests 3-4 are shown in Tables 5 and 6. Since the $\sigma^2_{imaging}$ is expected to be much lower for the "same" images than for the "different" images, it is not surprising that the "different" images aggregate with the "same" images when they are used for training. IBFS Tests 5-6 (Tables 7 and 8) demonstrate that the converse is true — when half of the more variable "different" images are used as the training set, then the "same" images generally aggregate with the "different" images. These results are clearly different from the perimeter-related approach shown in Figure 8 and reported in [3], and relate to an "averaging" of the image characteristics rather than identifying unique aspects.

Table 7. IBFS Test 5. See text for details of the test (IJ printer, Dr. CID imager, training set of 5 "different" images). Data presented as (number correct)/(number of images).

Size of Tile	"Same"	"Different"	Accuracy (%)
5 x 5	0/10	5/5	0.333
6 x 6	0/10	5/5	0.333
7 x 7	5/10	5/5	0.667
8 x 8	0/10	5/5	0.333
9 x 9	0/10	4/5	0.267
10 x 10	0/7	3/5	0.25

Table 8. IBFS Test 6. See text for details of the test (LJ printer, Dr. CID imager, training set of 5 "different" images). Data presented as (number correct)/(number of images).

Size of Tile	"Same"	"Different"	Accuracy (%)
5 x 5	0/10	4/5	0.267
6 x 6	0/10	5/5	0.333
7 x 7	1/10	5/5	0.4
8 x 8	0/10	5/5	0.333
9 x 9	0/10	5/5	0.333
10 x 10	4/10	2/5	0.4

Table 9. IBFS Test 7. See text for details of the test (IJ printer, 8200 scanner, training set of 5 "same" images). Data presented as (number correct)/(number of images).

Size of Tile	"Same"	"Different"	Accuracy (%)
5 x 5	5/5	165/165	1.000
6 x 6	5/5	149/150	0.994
7 x 7	0/5	125/140	0.862
8 x 8	0/5	136/140	0.938
9 x 9	5/5	130/130	1.000
10 x 10	0/5	109/117	0.893

The tests performed in Tables 5-7 were repeated using the 8200 scanner in place of the Dr. CID imager in Tables 9-11 ("IBFS Tests 7-9" — note that the results for repeating Table 8 using the 8200 are not presented as they are identical to those of Table 11). The mean accuracy for all the aggregations using the 8200 scanner and training on a set of 5 "same" images is 0.948 (nearly 95% accuracy) for the six different tile sizes presented in Table 9. In Table 10, the results for the LJ experiments (still using the 8200 scanner and training on 5 of the "same" images) are presented. While the overall mean accuracy was substantially lower (88.8%) than for the IJ samples in Table 9, the results are still substantially better than any of the IBFS results using the Dr. CID imager (Tables 5-8).

Table 11 provides the results when using the 8200 scanner for the imaging and using half of the "different" images as the training set. Since the training sets are very large, it is not remarkable that all of the other samples aggregated with the training samples. The

mean accuracy was 87.4% overall — 0% for the "same" set and 100% for the "different" set.

Table 10. IBFS Test 8. See text for details of the test (LJ printer, 8200 scanner, training set of 5 "same" images). Data presented as (number correct)/(number of images).

Size of Tile	"Same"	"Different"	Accuracy (%)
5 x 5	0/5	161/165	0.947
6 x 6	0/5	134/160	0.812
7 x 7	0/5	107/140	0.738
8 x 8	5/5	140/140	1.000
9 x 9	0/5	125/130	0.926
10 x 10	0/5	110/117	0.902

Table 11. IBFS Test 9. See text for details of the test (IJ printer, 8200 scanner, training set of 5 "different" images). Data presented as (number correct)/(number of images).

Size of Tile	"Same"	"Different"	Accuracy (%)
5 x 5	0/10	83/83	0.892
6 x 6	0/10	75/75	0.882
7 x 7	0/10	70/70	0.875
8 x 8	0/10	70/70	0.875
9 x 9	0/10	65/65	0.867
10 x 10	0/10	59/59	0.855

4. DISCUSSION AND CONCLUSIONS

4.1 Authentication

The authentication results (Tables 1-2) demonstrate that the Dr. CID imager provides higher payload density. We have earlier [19] defined the algorithms for obtaining the SPD-fA-PL, which is the security payload density, at full authentication, using a piecewise linear model for the authentication data. This "single metric" for comparison is valuable for comparing authentication accuracy in general between two or more imaging devices or workflows. For the data in Table 1, the Dr. CID provided the higher density for both printing technologies: 4650 bytes/in^2 for the IJ printing, and 1440 bytes/in^2 for the LJ printing. These are considerably higher than the SPD-fA-PL values estimated for the scanner — 1440 bytes/in^2 for the IJ printing, and 1160 bytes/in^2 for the LJ printing. The Dr. CID imager supports a higher payload density due to the fact that it provides 144 times as many individual pixels as the scanner — having 12X the resolution in each direction. Moreover, the Dr. CID has true 7200 dpi resolution, so the extra pixels are not due to oversampling. The authentication algorithm is based on comparing the mean hue of the sub-segmented tile regions to the mean hues of the six non-payload indicia (NPI); therefore, convergence to the actual hue is governed by the Central Limit Theorem. Convergence in hue, then, is $\propto 1/\sqrt{n}$ where n = 144. Since the z-value, or $\sigma/\sqrt{n}$, is indicative of the variability in the mean estimate, this implies that unless the ratio $(\sigma_{Dr.CID}/\sigma_{Scanner})$, is ≥ 12, the Dr. CID imager will in general provide better authentication accuracy (for color tiles of a given size) than the 8200 scanner. The numerator of this ratio is a function of $\sigma^2_{substrate}$ and the variability of the printing, while the denominator is generally a function of the printing variability only.

Figure 9. Representative swathes of red color tiles (images shown extend across three red color tile modules atop left images of Figures 3 and 4), for Dr. CID (upper) and 8200 scanner (lower). While the two images represent the same physical area, the Dr. CID image was captured at 12X the resolution in both the x- and y-direction. The Dr. CID image shows evidence of (much) higher substrate variability and less saturation (comcomitant with lower contrast overall); however, it does not show drastically higher hue variability.

If we were to target reducing $\sigma_{Dr.CID}$, among our options would be to increase the contrast and/or increase the intensity of the light source. However, this may deleteriously affect the forensic capability of Dr. CID (as shown in Figure 8). Moreover, it is clear from Figure 9 that the hue variability in the Dr. CID images is not too high. In fact, we consistently find $(\sigma_{Dr.CID}/\sigma_{Scanner}) \leq 3$, so that Dr. CID will effectively provide a 4X greater information density for authentication. Serendipitously, this is the ratio of SPD-fA-PL for the IJ authentication data. Such an improvement is not seen for the LJ data, most likely due to the more "systematic" print defect — namely color plane mis-registration and the associated large value for $\sigma^2_{CPmis-R}$ –associated with our LJ prints.

4.2 Forensics

The forensic capability of the Dr. CID imager was already established [3]; the current investigation, unremarkably, shows that color tiles are amenable to perimeter-based forensics. Similar capability was observed for LJ prints (Figure 10).

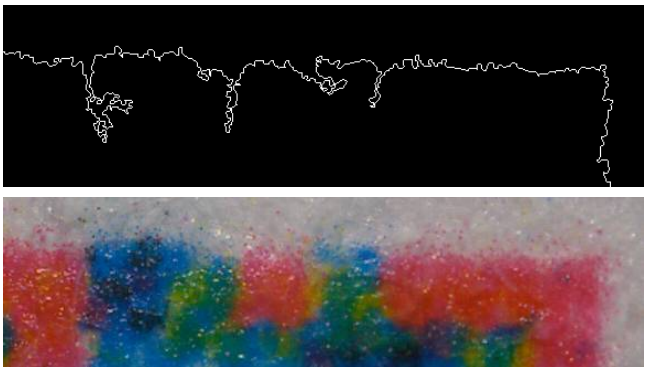

Figure 10. High-resolution perimeter (above) of a portion (below) of a Dr. CID captured image for a 5x5 pixel size (at 600 dpi original image LJ printing specifications) color tile ICM.

Interestingly, however, the image-based forensic service (IBFS), initially described in [20], actually performs better when using the device with lower resolving capability — the 8200 scanner (Tables 3-11). The data in Tables 9 and 10 are also of interest — the lower authentication accuracy of the LJ prints is consistent with the lower accuracy of classification for the LJ prints when using the IBFS.

Based on our analysis of the sources of variance in our tests, it appears that the IBFS provides higher accuracy when using the scanner due to the lack of $\sigma^2_{substrate}$, since all other sources of variance should be equivalent. As a consequence, the IBFS offers a different type of forensic capability. Rather than providing individual image forensics along the lines of Figures 8 and 10 and reference [3], the IBFS provides "batch inspection" results that can lead to forensic level. In order to address this, we take a look at what "inspection" actually means in the context of security imaging.

4.3 What is Inspection?

Having addressed authentication and forensics, the easiest definition for "inspection" may be that inspection is simply all of the imaging information that lies between authentication (intentional, usually mass serialized) and forensics (unintentional, tied to a single object) in the imaging continuum. This is an unsatisfactory definition, however, based on the definition of what inspection is "not". To correct this, we provide two new methods of defining inspection.

1. Statistical. The first definition for inspection is based on statistics. We herein define "forensic" imaging as an item-specific imaging analysis that provides us with less than 1 in 10^9 chance of a false positive, or with $(1\text{-}10^{-9})*100\%$ confidence we have the authentic item associated with the image. On the other side, authentication data is typically reproducible. Color tiles, for example, can be read by hand and re-printed. Thus, we define authentication statistically as having a low probability of "guessing" a correct identifier, but 100% probability of being able to reproduce a correct identifier. From this perspective, then, inspection provides anywhere from 1 in 1 to 1 in 10^9 probability of replication. Batch inspection, then, of multiple items, can provide less than 1 in 10^9 probability of reproducing the **set** of imaging analytics. As this occurs — e.g. in Table 4 wherein the probability of randomly assigning 254-350 images to the correct class is essentially zero — the use of multiple "imaging inspection" steps provides forensic-level analytics.

2. Structural. The second definition of inspection is based on its structural role in the imaging-driven ecosystem. From a structural standpoint, imaging forensics reliably "read" information that is *unintentionally* part of the printing process and is unique to the item — for example, nuances of the substrate [3][5]. Structurally, authentication relates to the intentional placement of information into a printed and/or manufactured mark — this would include the examples in Figure 1 and the various ICMs described in references [6]-[19] and elsewhere. From this perspective, then, inspection — being in the middle again — is defined as information that is *unintentionally* part of the printing process and is *not unique* to the item. This includes printer identification [2][21][22], since these types of analyses can be performed across more than one unique print. As such, the so-called "IBFS" [20] is also an imaging inspection approach. Only when these approaches are applied to a large set of images simultaneously can they be called "forensic", inasmuch as batch inspection of large set of images *structurally* considers the image set (and not a unique image) as its forensic "atomic unit".

4.4 Engineering Documents for a Security Ecosystem

The results of this research are of considerable interest to document security ecosystem architects and investigators. Because, generally, not all aspects of the security ecosystem are known beforehand, it is readily argued that a hybrid approach will be the one most robust to changes in the ecosystem. For this reason, providing printed information that can fulfill a forensic, inspection and authentication role is recommended. Note that we do not discuss track and trace — its capability is herein assumed since the authentication data is usually mass serialized per the relevant track and trace requirements.

From this perspective, inspection is the "glue" holding together this security ecosystem. Inspection provides the statistical and structural "bridge" between forensics and authentication, ensuring that the right information can be read/analyzed with the right (statistical) confidence by any set of devices, at any time and any place. Of course, forensic-level confidence will require more items to be imaged, but the confidence levels are shown herein to be predictable.

We have also shown how an inspection-related system, such as the IBFS, can be used to provide forensic-level "batch" inspection. This is in spite of the fact that the IBFS actually works best when the substrate-unique information, described herein as $\sigma^2_{substrate}$, is removed through high-contrast, lower-resolution, desktop scanning in place of high-resolution imaging with the Dr. CID imager. This is an exciting result. It means that a device placed into the security ecosystem for the purpose of "traditional" inspection (image quality assurance, validation of the printing of important regions of interest, etc.) can also be used for (statistical and structural) forensics.

4.5 Conclusions

In this paper, we tried to determine if a single security mark can be used to simultaneously complete multiple security imaging tasks. The first is the correct decoding of the sequence of marks used to embed a unique ID, or "authentication". The second is forensics, which we have previously shown can be performed on

individual printed marks (or "glyphs") using a high-resolution imager [3]. In this paper, we have extended the definition of forensics to batch inspection on statistical grounds (e.g. that less than 1 in 10^9 probability of false match constitutes "forensics"). These results illustrate that a desired level of confidence in a batch of printed items can be attained by using different imaging devices—the difference being in the number of images that must be inspected.

We focused on color tiles primarily because of the wealth of previous published work, although the results are extensible to other marks. The results are also extendible to other printers and imagers. We have recently shown, for example, that this approach works even on aging printing and scanning equipment [23]. This approach is resilient to attack by high-quality printers (none of which have 3 μm addressable resolution, anyway), since the interaction of ink and substrate is not determined by the resolution or quality of the printer, but by the material properties of the ink/substrate interaction. Future work will focus on taking advantage of the repertoire of imaging devices to provide overall "ecosystem" security, where statistical level of confidence can be optimized for a given ecosystem cost.

5. ACKNOWLEDGMENTS

The authors gratefully acknowledge our colleagues who have worked with us on color barcodes over the years, including Jason Aronoff, Matthew Gaubatz, Shawn Gibson, Stephen Pollard, and Juan Carlos Villa. Thanks also to Gary Dispoto for his support of this research.

6. REFERENCES

[1] D.E. Bicknell and G.M. Laporte, "Forged and Counterfeit Documents", in Wiley Encyclopedia of Forensic Science, 3104 pp., June 2009

[2] B. Zhu, J. Wu, and M.S. Kankanhalli, "Print signatures for document authentication," Proc. ACM CCS'03, pp. 145-154, 2003.

[3] S. J. Simske and G. Adams, "High-resolution glyph-inspection based security system", Proc. IEEE ICASSP, pp. 1794-1797, 2010.

[4] C. Skaar, *Wood Water Relations*, Springer-Verlag, NewYork, 283 pp., 1988.

[5] W. Clarkson, T. Wyrich, A. Finkelstein, N. Heninger, J.A. Halderman, and E.W. Felten, "Fingerprinting Blank Paper Using Commodity Scanners", 30th IEEE Symp. Security Privacy, pp. 301-314, 2009.

[6] "Universal Product Code," http://en.wikipedia.org/wiki/Universal_Product_Code, last accessed on 14 April 2010.

[7] "Data Matrix (Computer)," http://en.wikipedia.org/wiki/Data_matrix_(computer), last accessed on 14 April 2010.

[8] "Aztec Code," http://en.wikipedia.org/wiki/Aztec_Code, last accessed on 14 April 2010.

[9] "GS1 MobileCom: Extended packaging pilot handbook," http://www.gs1.org/docs/mobile/GS1_Extended_Packaging_Pilot_Handbook.pdf, 69 pp., last accessed 14 April 2010.

[10] "Mobile in Retail: Getting your retail environment ready for mobile," http://www.gs1.org/docs/mobile/Mobile_in_Retail.pdf, 34 pp., last accessed on 14 April 2010.

[11] "OMA: Open Mobile Alliance," http://www.openmobilealliance.org/, last accessed 14 April 2010.

[12] "High capacity color barcodes (HCCB)", http://research.microsoft.com/en-us/projects/hccb/, last accessed 14 April 2010.

[13] D. Parikh and G. Jancke, "Localization and segmentation of a 2D high capacity color barcode," Proc. IEEE Workshop Appl. Computer Vision (WACV 2008), 6 pp., 2008.

[14] "Colorzip", http://www.colorzip.co.jp/en/, last accessed on 14 April 2010.

[15] O. Bulan, V. Monga, and G. Sharma, "High capacity color barcodes using dot orientation and color separability," Proc. SPIE: Media Forensics and Security XI, vol. 7254, pp.725417-1-7, 2009.

[16] R. Villán, S. Voloshynovskiy, O. Koval, and T. Pun, "Multilevel 2D bar codes: Towards high capacity storage modules for multimedia security and management," IEEE Trans. Info. Forensics Security, vol. 1, no. 4, pp. 405-420, 2006.

[17] J. Mayer, J.C.M. Bermudez, A.P. Legg, B.F. Uchôa-Filho, D. Mukherjee, A. Said, R. Samadani, S. Simske, "Design of high capacity 3D print codes with visual cues aiming for robustness to the PS channel and external distortions," Proc. IEEE Intl. Conf. Image Proc., pp. 105-108, 2009.

[18] M. Vans, S.J. Simske and J.S. Aronoff, "Barcode structural pre-compensation optimization," Proc. IS&T NIP25/DigiFab 2009, pp. 167-169, 2009.

[19] S.J. Simske, M. Sturgill, and J.S. Aronoff, "Effect of Copying and Restoration on Color Barcode Payload Density," Proc. ACM DocEng 2009, pp. 127-130, 2009.

[20] S. Simske, M. Sturgill, P. Everest, and G. Guillory, "A system for forensic analysis of large image sets," Proc. IEEE WIFS 2009, pp. 16-20, 2009.

[21] 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.

[22] M. Gaubatz and S. Simske, "Printer-scanner identification via analysis of structured security deterrents", Proc. IEEE WIFS, pp. 151-155, 2009.

[23] G. Adams, "Hand held Dyson relay lens for anti-counterfeiting", Proc. IEEE IST, pp. 273-278, 2010.