# High Resolution Imaging for Forensics and Security

*Guy Adams [1], Stephen Pollard [1], Steve Simske [2]; Hewlett Packard Labs; Bristol, UK[1] & Fort Collins, Colorado, USA[2].*

## Abstract

*Printing provides innate forensic capabilities useful for product security as a consequence of the microscopic stochastic nature of the printing process itself and ink/substrate interaction during printing. This is especially true for substrates with a high degree of surface roughness/porosity, such as office paper, recycled paper, cardstock and packaging. Further imperfections are incurred during high speed printing, which taxes the limitations of the printing processes. These imperfections, consistent with reduced print quality, can be used serendipitously to provide a unique identifier for any printed symbol. This paper describes the hardware design for an imaging device that can analyze, with 7600 lines/inch resolving capability over a relatively large field of view of 6.6 x 4.9mm, any printed mark—from character to glyph to outline of an image—with high mark specificity. Combined with image analysis software written to describe the interface, or boundary, between ink-covered and ink-free substrate, this device, dubbed the Dr. CID (Dyson Relay CMOS Imaging Device [1]), can provide simultaneous image authentication and forensics.*

Keywords: high resolution; forensics; anti-counterfeiting; inspection; print parasitics

## Introduction

The volume of counterfeit goods worldwide continues to grow. OECD figures from 2007 indicate the worldwide costs to be around $250B and rising year on year. The presence of fake items undermines confidence and is damaging to individual brands as well as to entire sectors. Certain affected industries like the pharmaceutical and electrical items etc, suffer from the added problem that there is a safety risk with counterfeit goods and a direct impact on human health. Furthermore, organized crime is attracted to counterfeiting because of the significant returns for little investment and lack of effective policing.

With globalization, there is increasing pressure to address the issue as the items that are being counterfeited continue to expand into areas that affect safety. However the internet and current offering of cloud based services means that, the hitherto issue of added cost of the system means that barrier to adoption is lowering. Once item level track and trace systems (or mass serialization) are in place [2] then the added cost of item specific data that provides authentication is low. Our proposal to use printed elements such as 2D or 3D barcodes further enhances the route to mass adoption with a minimum of disruption to the product flow.

Looking at figure 1, the images of small 6pt characters printed on a thermal inkjet printer (TIJ) clearly show small aberrations, the majority of which can be described at a sub 10µm level. It is important to note that this microscopic range is smaller than the smallest addressable single droplet firing of a TIJ printer thus, even if it was scanned with a specialist scanner at sufficient resolution (most mainstream scanners have a maximum optical resolution of 1200dpi, or 21µm/dot) it would not be possible to address and deposit drops that would re-create the irregularity especially as the substrate will interact with the ink in a similar random way to the original as well as having to attempt to pre-compensate the drops for the random errors in the nozzles and time of flight. Comparing a like for like region of the two letters in figure 1, which were printed successively on the same line of a single page by a printer, it is clearly visible that the random elements create a 'fingerprint' [3], [4].

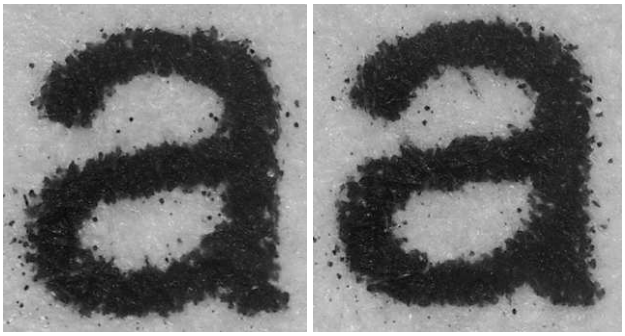

***Figure 1.*** *Thermal inkjet signatures. 6pt letters printed successively on an HP K5400 inkjet printer (40x magnification).*

***Figure 2.*** Images from different stock with an HP D7260 TIJ printer – top left is uncoated, top right is office, bottom left is photo and bottom right is craft paper.

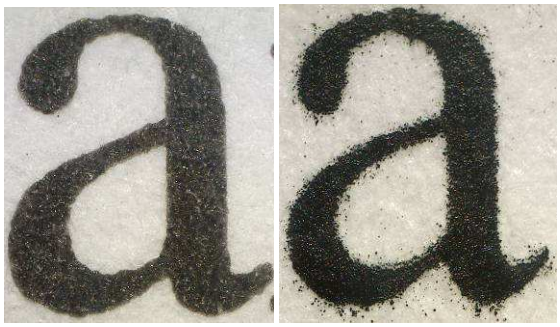

***Figure 3.*** *Images from different stock – HP Indigo 5000 (LEP) plain paper left and HP LJ P4515 right.*

The images in figures 2 and 3 show how the microscopic variation is linked to substrate and print quality. It is also worth noting the types of aberration – from random imperfections in the boundary from the wicking of the ink along exposed paper fibers [5] due to the ink and paper interaction, to droplet tails and some boundary imperfections arising from the random variation in ink ejection and flight. The image in figure 3, printed on an Indigo liquid electrophotography printer, reveals a more controlled process with fewer variations and the office laser image has stray toner particles near the boundary. The relationship between print quality and structures suitable for providing forensic matching is inversely proportional.


***Figure 4.*** *Dr CID Prototype*

## Approach

USB-powered and approximately the size of a marker pen (figure 4), the Dr. CID (Dyson Relay CMOS Imaging Device [1]) is easy to calibrate, focus and use. The use of a Dyson Relay configured lens provides a high resolution and large field of view with only one refractive surface and one reflective element (figure 5).

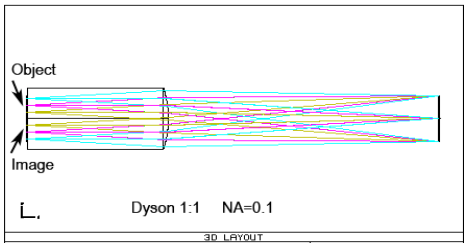


***Figure 5.*** *Basic Dyson Relay Lens Configuration*

The suitability of this design is revealed when we look at the required resolution. To resolve sub 10µm features the spatial sampling should be in the low µm range. Currently most mainstream CMOS image sensors have pixels in the 2-5µm range thus we can explore a class of 1:1 optics. The design illustrated below uses a mainstream 3MP 3.2µ pixel sensor from Aptina.

To control the depth of field the lens is designed to be used in contact mode with the output optical path separated from the input and turned through 90 degrees so as to move the image sensor away from the input plane (figure 6). In order to provide illumination, which is essential with this use model, an LED is configured to provide uniform diffuse illumination internally with no unwanted internal reflections.

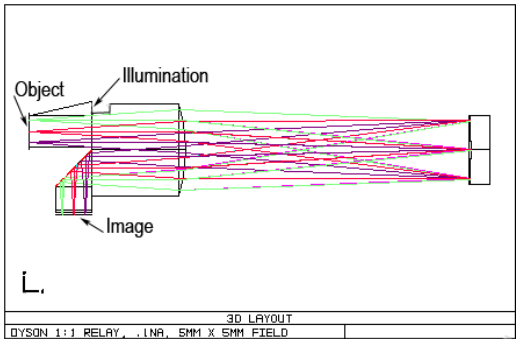


***Figure 6.*** *Modified Dyson Relay Lens Configuration*

## Design

The design was constructed and simulated using a software package from Zemax. The following approximate equations for numerical aperture and diffraction limit were used to define the requirements of the lens.

$$A_A = \frac{1}{2A_N} \tag{1}$$

$$d = 1.22\lambda A_A \tag{2}$$

Using a numerical aperture ($A_N$) of 0.105 in (1) the angular aperture ($A_A$) is 5. Using (2) and a wavelength ($\lambda$) of 550nm, the diffraction limit ($d$) is 3.2µm.

Thus the resolution of the lens was matched to the pixel size of the sensor. Close to diffraction-limited performance was measured by

simulation over the relatively large field of view along with low chromatic aberrations and distortion which are also characteristics of this type of lens design. In figure 7, the plots of modulation transfer function (MTF) for various points across the filed of view are closely distributed near the diffraction limited curve.

Testing the prototype lens with a resolution chart showed that the actual resolution was marginally worse than the theoretical 7600dpi diffraction limited simulation. In particular, light scattering and leakage as well as the use of a color sensor will contribute to this reduction.

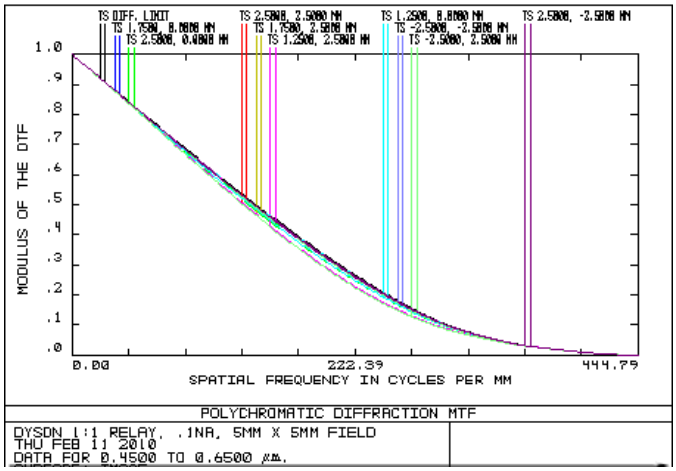


***Figure 7. Simulated resolution plot***

Further modeling has shown that the design is scalable to both higher resolution (1.8µm) and larger fields of view (50mm)

## Image Analysis

The software accompanying the Dr. CID hardware has been broadened to enable a number of security workflows. These include image matching when connected to a database, robustness to image distortion, and structured crafting of the software to take advantage of imaging simplification engendered by the structure of security printing deterrents.

### *A. Approach 1*

One approach we have developed [6] uses a series of metrics based on perimeter characteristics to compare glyphs. Here, the images captured by the DR CID are analyzed using the following steps: (1) a contrast-insensitive thresholding algorithm to binarize the image; (2) segmentation into connected components, or "regions"; (3) perimeter determination; and (4) a wide array of perimeter descriptor calculations.

The thresholding algorithm is simple, and is implemented in such a way as to provide consistent behavior despite differences in contrast, exposure, etc., between different imaging devices. The threshold therefore consists of finding the 5% and 95% points in the image intensity histogram, $H\{Int_I\}$ and setting the threshold, $T_I$, as:

$$T_I = H\{Int_I\}|_{5\%} + 0.5*[H\{Int_I\}|_{95\%} - H\{Int_I\}|_{5\%}] \quad (3)$$

After thresholding the image, the connected components, or regions, are identified and the appropriate region is selected as the glyph of interest (based on size, shape, location, etc.). The perimeters are then created, as shown in Figure 8. The shape descriptors for the perimeter are next determined. In the original implementation, the centroid of the region of interest is computed, and the perimeter is divided into sections by angle (e.g. 0.5º increments from 0º to 360º around the perimeter results in 720 "pie pieces"). For each angular section, the minimum radius, maximum radius, complexity (number of changes in direction of the perimeter in radial direction with the glyph centroid as the origin), shared elements (number of perimeter points in the section), uncertainty (number of perimeter line segments in the section), and neighborhood uncertainty (moving average of the uncertainty to account for minor—i.e. less than 0.25º—differences in alignment of the two images with the angular sectioning) are computed.

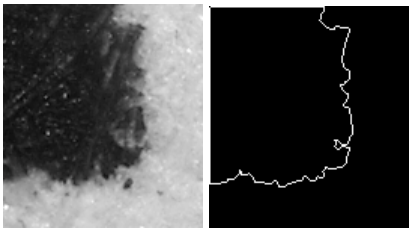

***Figure 8. Perimeter Map***

When two images are to be directly compared, the second image is scaled to the first image to match connected component size. This "normalization" corrects for any difference in focal length between two DR CID devices; difference in height of the DR CID devices over the glyph during image capture; and difference in size of the glyph, e.g. due to font, ink gain, etc. differences. Additional normalization procedures have recently been developed [8], and the set of metrics used to compare two images augmented by grayscale and edge-directionality metrics.

The perimeter-based approach resolves differences between camera-to-camera variance and glyph-to-glyph (or character-to-character) variance. Ongoing research has further improved the metrics used for the perimeter-based approach, improving the odds of a false match from less than 1 in $10^3$ to less than 1 in $10^9$. This provides a forensic level security for a single glyph [8] rather than batch-level forensic security (e.g. requiring four glyphs instead of 1 using the method of [4]).

### *B. Approach 2*

Using a 3D tile matrix (2D matrix with 6 colors in figure 9), we have shown [7] that it is possible to provide both authentication from the payload of the color tiles, and additionally through analyzing the irregularities in the perimeter a second, higher confidence, level of validation.

The high resolution image capture of Dr CID revealed that the payload density is print-limited. This means that the payload density cannot be increased by reducing the tile size below 4x4 pixels at 600dpi (5000 bytes/in$^2$) because of limitations, such as bleed, in the print process. The comparison with a flat bed scanner showed a peak payload density of 9x9 pixels which equates to 1440 bytes/in$^2$. A model-based approach to finding the location

and extent of individual color tiles in the image was used. The last step in determining the payload is to use the averaged hue angle for each tile to determine the RGBCMY value.

The secondary level of validation is provided by the spatial variation in the edge of the perimeter that can be seen in figure 9. A profile of the perimeter was extracted from the image and then divided into 40 blocks (the number of tile edges). The sum squared error (SSE) was computed for each of the 40 blocks and then each block assigned an integer value based on the deviation from overall mean of the SSE. These shape warp descriptors (SWD) then form a compact sequence that can be compared to other sequences by a modified form of Hamming distance called the shape distortion encoding difference (SDED) in order to determine the similarity.

***Figure 9.*** ***3D barcode matrix - 2.04mm on a side***

Test sheets with between 117 and 165 3D barcodes with sizes from 5x5 to 10x10 pixels (at 600ppi) were analyzed and showed that the smallest size had a confidence accuracy of 99.99%. For the largest tile sizes, the SDED becomes more uniform, in turn reducing the discrimination power. Even these larger sizes still had less than a 1.3 in $10^6$ chance of 4 or more of being a false positive.

### *C. Approach 3*

The discussion so far in this paper has been centered on Dr CID which is based around a small area image sensor intended for mobile use. However, the use of an area sensor when applied to a strip/continuous flow of packaging or labels is not ideal as the strip would have to be mechanically paused for image capture at these resolutions. Using a high power strobe to freeze the motion is not practical as the strobe burst would have to be sufficient to freeze less than 5µm of motion as well as synchronized to the strip flow. At a strip speed of 1M/S the exposure would have to be less than 5µs. The requirements for capturing high resolution data from a strip based process means that a line scan camera is best suited. However, because the strip is moving, unless the line capture rate of the sensor is perfectly synchronized to the speed, image distortion will be generated. There will be additional distortion due to skew of the axis of the image sensor to the flow. This potential skew makes the use of time delay integration (TDI) sensors unworkable as the skew means that the image will not pass across a single pixel column, thus introducing blur.

We have developed a novel dynamic time warping approach that matches the perimeters of printed character even under quite severe distortion [8]. Figure 10 shows the test setup used where a high speed line scan camera with 5µm pixels (Aviiva UM8) and 1:1 Schneider precision optics is mounted above the output of an HP K5400 TIJ printer. Combined with optical triggering of the frame grabber using black trigger bars at the edge of the page, this system enables the image capture of a 60mm wide strip of the printed output as it exits the printer.

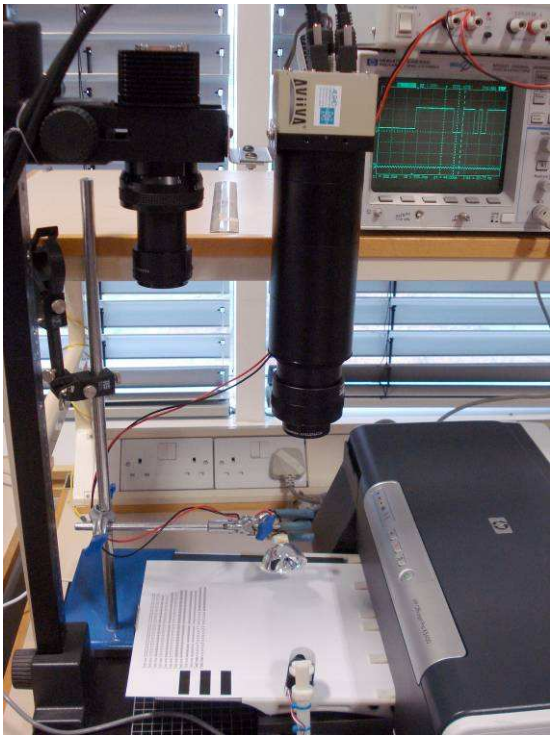


***Figure 10.*** ***In-line capture test bed***

Also, the approach is robust to the fact that the resolution of the line scan camera is slightly lower at 5µm than Dr CID at 3.2µm. This is due to the state of the art of different image sensors for different applications not being the same.

The results show that even under severe distortion it is possible to match the perimeters with high precision which leads to a very low probability of a false positive or negative.

## Discussion

The novel application of Dyson Relay optics has enabled the use of printed aberrations as a robust anti-counterfeiting mechanism. The prototype shows that excellent image capture quality is possible with a system that uses a single refractive surface, mirror and mainstream CMOS image sensor. In combination with appropriate image capture and analysis software we propose that simple printed marks are a suitable platform for a range of product, label and document protection. The approach is

also robust to the different imaging devices that a workflow will typically use and does not require complex or costly calibration.

The combination of the print process and substrate can also provide specific analytic data that describes the magnitude, type and distribution of the errors. In addition, the use of specialist papers with either visible non uniform fibers or exposed fibers that interact with the ink can provide further levels of robustness. This is because the counterfeiter would have to replicate all these elements that are part of random microscopic processes.

## Conclusion

We propose that forensic level authentication is achievable for even a single printed glyph. We define "forensic" as having less than 1 in $10^9$ probability of false matching – i.e. "false positive" identification of a like glyph, or false rejection of same character imaged twice – i.e. a "false negative". We have also shown that this is possible in a realistic workflow where the in-line production image capture devices introduce different image distortions compared to the in-field mobile devices.

In addition, the same hardware that is used for forensic capture during production can be used for print quality analysis.

Combined with mass serialization—the variable data printing of a unique identifier on each printed item—this hardware + software imaging system affords the possibility of a complete, low cost, image-based approach to supply chain, document and label security.

## References


[1] CID, http://en.wikipedia.org/wiki/Criminal_Investigation_Department.

[2] http://www.gs1ca.org/page.asp?intPageID=1395 ?

[3] B. Zhu, J. WU, M. S. Kankanhalli, "Print Signatures for Documant authentication", CCS ACM 2003

[4] L. Hindus, "Image-Based "Fingerprinting"", Advanced Imaging 1998

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

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

[7] S. J. Simske, S. B. Pollard, G. B. Adams, "An Imaging System for Simultaneous Inspection, Authentication and Forensics", IEEE IST 2010, accepted, 2010.

[8] S.B. Pollard, S. J. Simske, G. B. Adams, "Resolving Distortion Between Linear and Area Sensors for Forensic Print Inspection", ICIP 2010, accepted, 2010.


## Author Biography

*Guy is the hardware lead for security printing and imaging project within HP Labs. Guy joined HP Labs in 1996 and has worked on several projects that became successful products such as class leading CMOS image sensors, cameras for PDA's. Guy is always keen to deliver innovative technical solutions and he has more than 10 granted patents and more than 20 pending. Guy is a member of the IET and a Chartered Engineer (UK equivalent of Professional Engineer).*