

FLATBED SCANNER IDENTIFICATION BASED ON DUST AND SCRATCHES OVER SCANNER PLATEN

Ahmet Emir Dirik

Polytechnic Institute of NYU
Electrical & Computer Engineering Dept.
Brooklyn, NY, 11201, USA

*Husrev Taha Sencar, Nasir Memon**

TOBB Economics and Technology University
Computer Engineering Dept.
Sogutozu, Ankara, 06560, Turkey.

ABSTRACT

In this paper, a novel individual source scanner identification scheme is proposed. The scheme uses traces of dust, dirt, and scratches over scanner platen on scanned images to characterize a source scanner. The efficacy of the proposed scheme is substantiated with experimental analysis. The robustness of the scheme to the JPEG compression is also investigated. Experimental results show that proposed scheme could be used to match a scanned image to its source.

Index Terms— Image analysis, Object detection.

1. INTRODUCTION

One of the key problems in digital image forensics is the analysis of a media object with the purpose of identifying its source acquisition device. Therefore, in recent years, several methods have been proposed to identify source digital camera and scanner. This is realized by identifying unique device characteristics that leave measurable traces on images. Essentially, these characteristics are due to manufacturing imperfections component technologies, component failures, and defects introduced during use of devices. To date, various source identification techniques have been proposed [1, 2, 3, 4, 5].

In this paper, a new approach to source scanner identification is proposed. The basis of our method is the appearance of dust and scratches in scanned images due to dusty or defected flatbed scanner platen. It is a known fact that, after a new flatbed scanner has been used for a while, the platen will be contaminated with dust particles and paper debris¹. In many cases, these particles introduce scratches over the surface of platen due to crude use. The accumulation of dust and scratches result with localized defects over the scanned image that are imperceptible but nevertheless measurable. Figures 1 and 2 display scanned images wherein black and white spots show the effects of such artifacts. Most typically, dust particles reveal themselves as dark spots on the image, whereas,

glass scratches cause bright and white spots due to light reflection.

Although the positions of dust particles and paper debris on scanner platen may change over time, some of them will stay still. On the other hand, the positions of the scratches on the scanner pane will not change under any condition. The scratches and dusts which are strongly adhered to the scanner platen will essentially create a unique pattern associated with the scanner. Since those scratch and debris positions will be relatively fixed, their positions can be used as a fingerprint of the scanner, and the random nature of the dust position pattern makes the fingerprint unique to the device. In Fig. 2, the dust and scratch positions on two scanned images are shown. It can be seen from the figure that relative blemish positions remain same for two different scans. This observation is the basis of the proposed identification scheme.

Various solutions have been proposed to reduce the impact of dust and other debris on the scanned image [6]. These include built in mechanisms to locate dust/dirt positions over the platen. Once the dust position pattern is created, the scanned image is post-processed to correct defected regions. However, probably, the most effective way for removal of dust and dirt particles is manual cleaning of glass platen. However, users can scratch the platen if they do not use the right cleaning chemical solution. Moreover, if the platen already comprises some physical defects, the manual cleaning would not solve the problem. Therefore, the impurities on scanner pane could be used in favor of forensic analysis to identify the scanning device from the traces of scratch and dusts in a scanned image.

The rest of the paper is organized as follows: In Section 2 and 3 a model based dust spot detection method and its use in scanner identification is explained in detail. The efficacy of the proposed method is substantiated by experimental results in Section 4. The robustness of the proposed scheme to compression is explained in Section 5. Finally, a discussion is presented in Section 6.

*N. Memon with Polytechnic Institute of NYU, Computer & Information Science Dept., Brooklyn, NY, 11201, USA

¹<http://www.vad1.com/photo/dirty-scanner/>

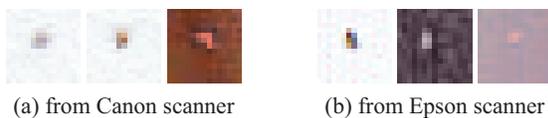


Fig. 1. Blemishes over scanned images due to dirty/scratched platen. Scan resolution: 300dpi, figure dimensions: 15x15 pixels.

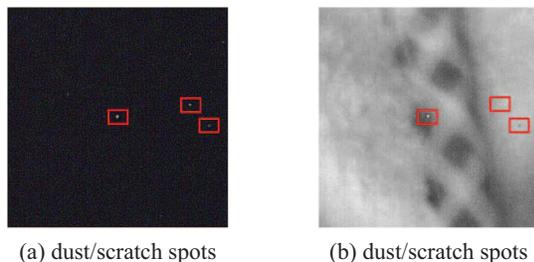


Fig. 2. Dusts and scratches over platen can be used as a unique fingerprint

1.1. Related work

Although, a variety of source identification schemes for digital cameras have been proposed so far, there are only a couple of published works related to scanner device identification [3, 4, 7]. The work on source digital camera identification essentially utilize a static noise component of the imaging sensor during matching an image to its source [2, 8, 9]. In [4] and [7], two scanner identification methods are proposed based on a similar methodology which utilizes imaging sensor noise characteristics. Crucially, imaging sensors suffer from two different noise components: dark signal non-uniformity (DSNU) and photo response non-uniformity (PRNU). DSNU causes pixel to pixel variations among pixels when the sensor is not illuminated. Whereas, PRNU noise becomes significant when the sensor is illuminated. PRNU and DSNU, together, add a structured noise pattern over the acquired image. This pattern is unique and can be used for individual source device identification. Nevertheless, PRNU based identification does not work well for scanners. Different from digital cameras, digital scanners can utilize significant noise reduction methods to compensate the effects of DSNU and PRNU noise by recording the pixel offsets while the light is off and individual pixel gain values when they are illuminated [10]. During scanner calibration, based on recorded offsets and individual gains, DSNU and PRNU can be compensated so as to get a flat image output when there is nothing over the scanner platen. Therefore, after noise calibration, it is very unlikely to detect and use the PRNU noise of the scanner sensors for forensic applications with high accuracy.

Because of these shortcomings, the authors in [4, 7] used some statistical characteristics of sensor noise and deployed machine learning methods to identify scanning device. However, general sensor noise statistics does not provide precise

device identification as PRNU noise based methods for digital cameras. Similar to [4, 7], PRNU based scanner identification is also investigated in [3]. The authors in [3] also note that PRNU based scanner identification is possible only under certain conditions but it provides less accuracy in comparison to digital camera identification.

2. DETECTION OF DUST AND STRATCH POSITIONS

As it can be seen in Fig. 2, scratch and dust positions (white and black dust spots) create a fixed pattern. To capture this pattern for source identification, a model based detection method is utilized similar to the blemish detection method in [5]. To simplify the problem, we ignore large image defects caused by lint, hair, etc, and focus on relatively small defects that are imperceptible and more persistent. It is seen from observations that these defects could be modelled with a dark spot surrounded with a white background or vice versa (see Fig. 2, and 1). Another observation about dust and scratch defects is that they cause high gradient along vertical and horizontal axes at defected locations.

To detect the positions of dust and scratches, scanned images are first filtered with a high pass filter. High pass filtering eliminates redundant image details and highlights the regions with high gradient values which comprise edges, textures, and dust and scratch regions. To be able to detect both scratch and dust spots with the same model we obtain the absolute values filtered image coefficients. Then, to separate dusts and scratches from other details, such as edges and high frequency components, a model as shown in Fig. 3 is searched all over the high-pass filtered image through normalized cross correlation (NCC) [11]. Finally, the NCC output is applied to an empirically determined threshold to select local maxima regions. The regions that yield high cross-correlations are deemed to be dust and scratch locations.



Fig. 3. Dust/scratch model for high pass filtered scanned image. White region dimensions: 5x5 pixels.

3. SCANNER IDENTIFICATION

To realize scanner identification, a dust and scratch template of a scanner has to be first generated using the detection algorithm described above. For this, scanner platen could be scanned under completely black and/or white background. While white background scan helps detecting positions of dust and dirt particles over the platen, scanning under black background helps detecting both dusts and some type of

scratches that completely reflect the incoming light and shine as small light spots. In the rest of the paper, we will only consider scanning black background for the sake of ease. However, the results of the white background scan can also be incorporated with it.

To generate a scanner template, solely two different scans of completely black background is sufficient. This could be realized by opening up the scanner lid and scanning the platen surface. The scanned image comprises just dusts and scratches since only these impurities over the platen reflect the incoming light back to the scanner sensor. Once, black scans are obtained, likely dust and scratch positions are detected. However, it should be noted that due to the vertical and horizontal scanner head position shifts, dust and scratch positions detected in two black images may not align properly. To compensate for the shifts between the two scanned images, the dust and scratch positions of the two images are matched with respect to each other through cross correlation. The scanner dust and scratch template is finally generated by taking Hadamard product of the scanned images that are correctly aligned. Thus, noisy components in scanned images are suppressed and matched dusts and scratches are highlighted.

To determine the match of a given scanned image to a scanner, likely dust and scratch positions are detected using the proposed model. Obtained image dust and scratch positions are then correlated with the scanner template (generated as described above). If the dust and scratch pattern extracted from the given image matches with the scanner template, i.e., yields a correlation above a predetermined threshold, it is assumed that the given image is created with the scanner in question. To exemplify this, two cross-correlation outputs are given in Fig. 4. Fig. 4.a displays the case where the image is acquired by the scanner in question, and the correlation is high. Whereas, in Fig. 4.b, image and scanner are unrelated and the resulting correlation is not significant.

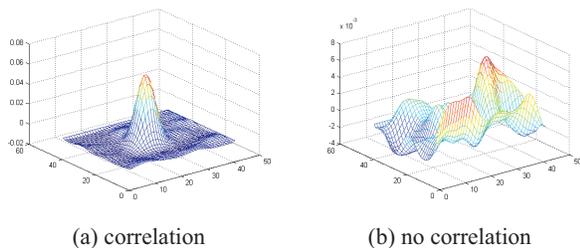


Fig. 4. Cross-correlation results for scanner identification. (a) the image dust/scratch positions are matched with the scanner dust/scratch template. (b) there is no matching.

4. EXPERIMENTAL RESULTS

To test the efficacy of the proposed method, we conducted an experiment with two different scanners (Epson Perfection

1250 and Canon Canoscan LiDE90). The Epson platen had traces of some scratches and dusts due to several years of use. The Canon scanner was brand new. During one month use, the Canon scanner’s platen was contaminated by dust/dirt and it was never cleaned. To test the identification scheme, first, the dust and scratch templates of the scanners were created as explained in Sec. 3. Then, with each scanner, several images were scanned and compared with scanner templates. As an example, in Fig. 5, the cross correlation values of two images scanned from Epson and Canon scanners are given. In the figure, each image was scanned with both scanners and compared with two dust / scratch templates. While NCC values for Canon-Canon and Epson-Epson tests are 0.065 and 0.082, respectively; cross correlations for Canon-Epson and Epson-Canon tests are lower than 0.010. Hence, there is an order of magnitude difference between the matching and non-matching cases.



<p>(a) Scanned with Canon Canoscan LiDE90 corr. with Canon : 0.065 corr. with Epson : 0.005</p>	<p>(b) Scanned with Epson Perfection 1250 corr. with Epson : 0.082 corr. with Canon : 0.008</p>
--------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------

Fig. 5. Cross correlation values of images scanned with Epson and Canon scanner

To determine a decision threshold on cross correlation value, we tested scanner templates on two different image sets. The first set includes images scanned with the two available scanners. The second set is composed of images acquired from a different scanners and digital cameras. For each image, dust and scratch positions are estimated and correlated with the scanner templates. The scanning resolution in this experiment was fixed to 300 dpi and scanner templates were also created at this resolution. The cross-correlation values for Epson and Canon scanner dust and scratch templates on two different image sets are given in Fig. 6. It can be seen from the figure that the proposed scheme could be used for scanner identification.

5. ROBUSTNESS TO COMPRESSION

In this section, we investigate the proposed method’s robustness to JPEG image compression. For this, scanned image sets (Canon and Epson) are compressed with JPEG at quality factor 50. Then, dust and scratch positions are also estimated from the scanned images. Cross-correlation results obtained between dust and scratch positions from original images and

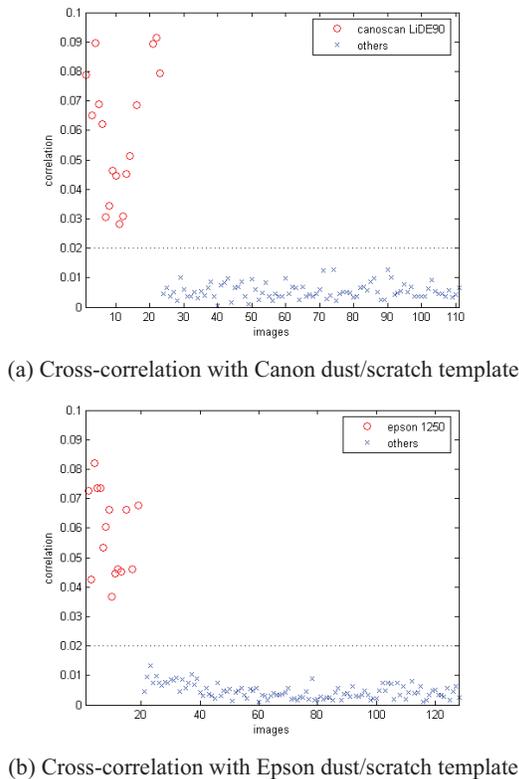


Fig. 6. Cross-correlation values between scanner templates and estimated dust/scratch spots in input images

their compressed versions are given in Fig. 7. It is seen from the figure that JPEG compression does not effect the identification performance significantly.

6. DISCUSSION

In this paper, a new approach to source scanner identification is introduced. The proposed method uses dust and scratch traces in scanned images to identify their source, and a decision is made on the basis of the degree of match between the positions of scanner platen impurities and detected positions in an image. The efficacy of the proposed method and robustness to the JPEG compression is investigated with experimental tests. In experiments, template and image resolutions are assumed to be the same. In the following work, improved dust and scratch models will be considered, and effects of image resizing and different scanning resolutions to the identification performance will be investigated.

7. REFERENCES

[1] A. Swaminathan, M. Wu, and K. J. Ray Liu, "Non intrusive forensic analysis of visual sensors using output images," *IEEE Transactions of Information Forensics and Security*, vol. 2, no. 1, pp. 91–106, March 2007.

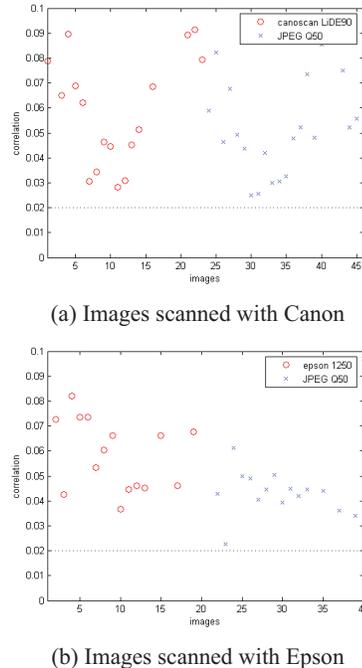


Fig. 7. Identification robustness to JPEG compression. Cross correlations of images compressed with JPEG Q50 are plotted with blue "x" symbol.

[2] J. Lukáš, J. Fridrich, and M. Goljan, "Digital camera identification from sensor noise," *IEEE Transactions on Information Security and Forensics*, vol. 1, pp. 205–214, June 2006.

[3] T. Gloe, E. Franz, and A. Winkler, "Forensics for flatbed scanners," in *Proceedings of the SPIE, Volume 6505*, pp. 650511 (2007).

[4] H. Gou, A. Swaminathan, and M. Wu, "Robust scanner identification based on noise features," in *Proceedings of the SPIE, Volume 6505*, pp. 65050S (2007).

[5] A. E. Dirik, H. T. Sencar, and N. Memon, "Digital single lens reflex camera identification from traces of sensor dust," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 3, pp. 539–552, Sept. 2008.

[6] Donald J. Stavely, Daniel M. Bloom, and et al., "Film scanner with dust and scratch correction by use of dark-field illumination," US Patent, 5969372, 1999.

[7] N. Khanna, A. K. Mikkilineni, G. T. C. Chiu, J. P. Allebach, and E. J. Delp, "Scanner identification using sensor pattern noise," in *Proceedings of the SPIE, Volume 6505*, pp. 65051K (2007).

[8] M. Chen, J. Fridrich, and M. Goljan, "Digital imaging sensor identification (further study)," in *Proceedings of the SPIE, Volume 6505*, pp. 65050P (2007).

[9] M. Chen, J. Fridrich, M. Goljan, and J. Lukáš, "Source digital camcorder identification using sensor photo response non-uniformity," in *Proceedings of the SPIE, Volume 6505*, pp. 65051G (2007).

[10] J. R. Bailey, C. P. Breswick, D. A. Crutchfield, and J. K. Yackzan, "System and method for high-performance scanner calibration," US Patent Application, 2006/0001921 A1, 2006.

[11] J. Lewis, "Fast normalized cross-correlation," in *Proc. of Vision Interface*, 1995, pp. 120–123.