

Counterfeit Detection Using Paper PUF and Mobile Cameras

Chau-Wai Wong and Min Wu
University of Maryland, College Park, USA

Abstract—This work studies the paper authentication problem by exploiting optical features through mobile imaging devices to characterize the unique, physically unclonable properties of paper surfaces. Prior work showing high matching accuracy either used a consumer-level scanner for estimating a projected normal vector field of the surface of the paper as the feature for authentication, or used an industrial camera with controlled lighting to obtain an appearance image of the surface as the feature. In comparison, past explorations based on mobile cameras were very limited and have not had substantial success in obtaining consistent appearance images due to the uncontrolled nature of the ambient light. We show in this work that images captured by mobile cameras can be directly used for authentication by exploiting the camera flash to create a semi-controlled lighting condition. We have proposed new algorithms to demonstrate that the microscopic normal vector field of a paper surface can be estimated by using multiple camera-captured images of different viewpoints. Our findings can relax the restricted imaging setups to enable paper authentication under a more casual, ubiquitous setting of a mobile imaging device, which may facilitate duplicate detection of paper documents and merchandise packaging.

Keywords—Anti-Counterfeit, Paper Physically Unclonable Features (PUFs), Mobile Cameras, Photometric Stereo

I. INTRODUCTION

Merchandise packaging and valuable documents such as tickets and IDs are common targets for counterfeiters. Low-cost surface structures have been exploited for counterfeit detection by using their optical features. The randomness of the surface makes the structures physically unclonable or difficult to clone to deter duplications. Such surface structures can be extrinsic by adding ingredients such as fiber [1], [2], small plastic dots [1], air bubble [1], powders/glitters [3] that are foreign to the surface; and the surface structures can also be intrinsic by exploring the optical effect of the microscopic roughness of the surface, such as the paper surface formed by inter-twisted wood fibers [3]–[7].

In this paper, we focus on the intrinsic property of the paper surface for counterfeit detection and deterrence, and seek to find a more casual, ubiquitous imaging setup using consumer-level mobile cameras under commonly available lighting conditions. The previous work in [3]–[5] shows that the microscopic roughness of the paper surface can be optically captured by consumer-level scanners and industrial cameras, both under controlled lighting conditions in the form of image appearance rendered according to the physical law of light reflection at the paper surface. The appearance images, and the subsequent normal vector field of the surface estimated from the appearance images, can achieve satisfactory authentication results. However, recent work in [3], [7] also shows that if

the ambient lighting is not well controlled, the image appearance alone has not achieved satisfactory authentication results. Instead, features based on the intensity gradient of visually observable dots are less sensitive to the change of lighting and may be used for authentication at the cost of higher algorithm complexity and moderate discrimination capabilities [7].

Two requirements may facilitate paper authentication via mobile cameras. First, the mobile-captured images should be comparable in resolution and contrast to those captured by scanners. Second, lighting should be controlled to render a desirable image appearance of the paper. The first requirement is qualitatively confirmed by comparing the acquired images from scanners and mobile cameras. Images acquired in both ways do have significant intensity fluctuations within small neighborhoods of pixels. The second requirement can be fulfilled by activating the flash next to the camera lens on mobile devices. As the relative position of the flash is fixed with respect to the lens, the appearance of the surface can be reasonably expected for a given position between the camera and the paper.

As we shall show in this paper, exploring camera flash can significantly improve the authentication performance of appearance images, and more importantly, can allow for the estimation of the normal vector field in fine surface details. By knowing the estimated location of the lens, the direction of incident light for every pixel of the paper can be calculated. Then, the normal vector of a particular pixel can be estimated by using the fully diffuse reflection model [4], [8] with a special treatment on the non-uniform intensity in camera images due to different distances from pixel locations to the flash. To the best of our knowledge, this is the first work using mobile cameras to obtain an effective estimate of the normal vector field of the paper surface to enable authentication.

The paper is organized as follows. In Section II, we review light reflection models, the method for paper surface registration, and the method for paper authentication. In Section III, we examine the authentication performances when restricted imaging setups such as scanners and industrial cameras are used. In Sections IV and V, we examine the authentication performances under the more flexible setup using mobile cameras with built-in flash. In Section VI, we conclude the paper and discuss future directions.

II. BACKGROUND AND PRELIMINARY STUDIES

A. Paper in Microscopic View and Light Reflection Models

The uniqueness of the inherent 3-D structure of the paper surface exploited for authentication purposes [3]–[5] is due to overlapped and inter-twisted wood fibers. Fig. 1(a) shows a topographic map of a 1mm-by-1mm patch of a copy paper surface [9] captured by a confocal microscope in which the

Originally published at IEEE WIFS 2015. Updated to correct typographical errors; no substantive changes made.

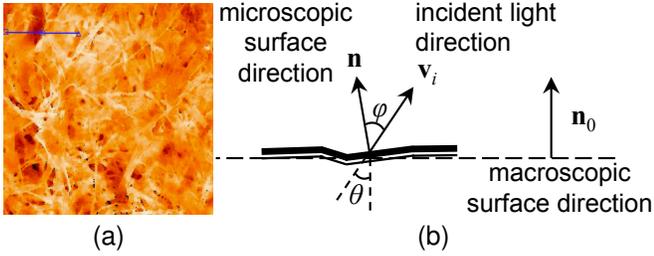


Fig. 1: (a) A topographic map of a 1mm-by-1mm patch of a paper surface captured by a confocal microscope, reproduced from [9]. The pseudo-color represents the elevation of the fibers in z -direction. (b) Microscopic view of a particular spot on a paper surface. Note that φ and θ are not co-planar in most cases. All vectors are unit vectors.

imperfect surface comprising of fibers is clearly shown. The appearance of the surface under a light source is a rendering according to the physical law of light reflection.

There are two fundamental types of reflection models, namely, specular versus diffuse [8]. The perceived intensity due to the mirror-like specular reflection is mainly dependent on the angle between the directions of the reflected light and the eye/sensor, whereas the perceived intensity due to diffuse reflection is mainly dependent on the angle between the directions of incident light and the normal of the microscopic surface. Most surfaces are combinations of both surface types.

Research in paper science [10] found that a piece of paper at different regions can have different dominant reflection types, but previous work [3], [4] treating paper as a fully diffuse surface led to satisfactory results for the authentication purpose. So, in this work, we assume the fully diffuse reflection model with the understanding that the majority of locations follow this model, and the remaining locations are outliers under this model.

Fig. 1(b) shows the surface normal direction of a particular spot in a microscopic view and an incident light direction. The perceived intensity of the fully diffuse reflection model [4], [8]

$$l_r = \lambda \cdot \underbrace{l}_{\propto \cos^\kappa \theta} \cdot \underbrace{\mathbf{n}^T \mathbf{v}_i}_{=\cos \varphi} \quad (1)$$

depends on the angle φ between normal direction of the surface at microscopic level, $\mathbf{n} = (n_x, n_y, n_z)$, and the direction where the incident light coming from, $\mathbf{v}_i = (v_{i,x}, v_{i,y}, v_{i,z})$; the strength of the light at the current spot, l ; and the albedo, λ , characterizing the physical capability of reflecting the light [8]. In our work, we assume λ to be constant over the whole paper patch. In the literature on reflection models [8], l can be modeled in proportion to $\cos^\kappa \theta$, where κ is a positive number accounting for the effect of energy fall-off according to the inverse-square law, the effect of foreshortening, *etc.*, and θ is the angle of incidence. In the case of a scanner, θ is a factory-specified design parameter relating to the position of the linear light source and therefore fixed for every pixel location; in the case of a camera, θ 's are generally different for neighboring pixel locations.

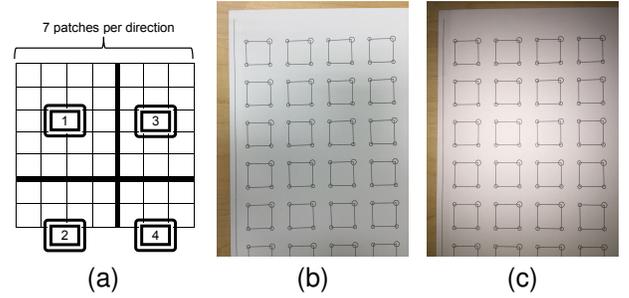


Fig. 2: (a) Camera locations for capturing a piece of cotton paper consisting of 49 square patches. Image captured at top-left (position#1) under fluorescent light (b) without flash (database 505), and (c) with flash (database 501). Capturing device: iPhone 6.

B. Patch Registration

We use the design of a registration container from our recent work [3], as shown in Fig. 2(b), to facilitate precise registration in experiments. Considering 600 pixels per inch printing resolution, our container is a square box of 400-by-400 in pixels, the line width is 5 pixels, and there are four circles at the corners. A preliminary alignment based on four boundaries can be achieved using a Hough transform, and subpixel resolution refinement with perspective transform compensation is then carried out based on the circle markers. Lens location relative to the captured surface in the world coordinate system can be readily calculated from the estimated perspective transform matrix, and then the direction of incident light at every pixel location is known.

C. Authentication Test

We focus on optical characteristics of physically unclonable features (PUFs) and approach the PUF verification problem as an image authentication problem commonly formulated as hypothesis tests. The null hypothesis H_0 corresponds to incorrectly matched pairs of test and reference patches whereas the alternative hypothesis H_1 corresponds to correctly matched pairs. The optimal decision rule maximizing the statistical power is the likelihood-ratio test: rejects H_0 if $\frac{f_1(\mathbf{x})}{f_0(\mathbf{x})} \geq \tau$ holds, where \mathbf{x} represents the test patch, f_0 and f_1 are the probability density functions under null and alternative hypotheses respectively, and τ is a threshold. As a proof of concept, we consider a simple hypothesis testing model differentiating a known reference image \mathbf{w} against all other images as follows:

$$\begin{cases} H_0 : \mathbf{x} = \mathbf{e}_0, & \mathbf{e}_0 \sim N(m\mathbf{1}, \Sigma_0), \\ H_1 : \mathbf{x} = \mathbf{w} + \mathbf{e}_1, & \mathbf{e}_1 \sim N(\mathbf{0}, \sigma_1^2 \mathbf{I}). \end{cases} \quad (2)$$

Here, normally distributed \mathbf{e}_0 stochastically represents any acquired image with a non-degenerate covariance matrix Σ_0 for image content and acquisition noise, $\mathbf{1}$ is an all 1 vector with the same dimension as \mathbf{x} , m corresponds to a value at the center of the linear range of the digital representation of luminance ($m = 128$ for intensity in the range $[0, 255]$), \mathbf{w} deterministically represents the reference image, and \mathbf{e}_1 is the image acquisition noise (white Gaussian, with constant variance σ_1^2). When the patch \mathbf{x} is represented by the x - or y -component of the normal vector field, the above under the hypothesis test setup is still valid with $m = 0$. Sample

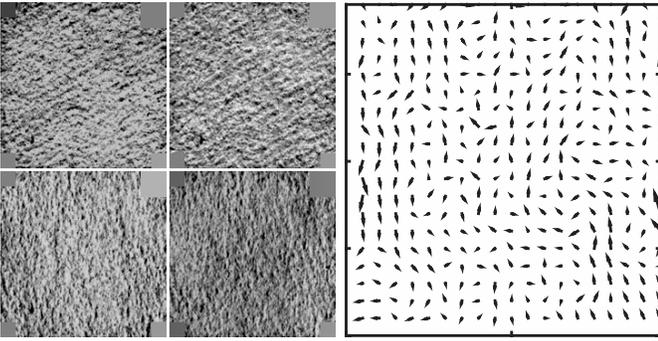


Fig. 3: A surface of cotton paper scanned from 4 perpendicular orientations and the resulting estimated norm map.

correlation coefficient $\hat{\rho}(\mathbf{w}, \mathbf{x})$ against a threshold is used as the decision rule in this paper.

III. PAPER AUTHENTICATION USING SCANNERS AND INDUSTRIAL CAMERAS

A. Norm Maps by Scanners

Clarkson *et al.* [4] assumes the fully diffuse model as described in Eq. (1) to estimate the projected normal vectors at all integer-pixel locations of the surface (*aka* the norm map, which does not contain the information in z -direction) by using images scanned from 4 different orientations of the paper: 0° , 90° , 180° , and 270° . Without knowing the exact direction of incident light, an estimate of one component can be obtained as the difference between two scans in exactly opposite directions, canceling the effect of the unknown incident direction of the scanner light.

In order to provide a solid best-effort comparison with alternatives to be discussed in later sections, we first carry out improvement over those in [3], [4] by removing the global bias for x - and y -components of the estimated norm map. The removal of the bias improves the accuracy of the length estimate of the vector. Using the improved technique, we estimate norm maps for 49 independent PUFs. The acquisition procedure is repeated using two Epson scanners: Perfection 2450 and GT-2500. Sample patches for scanner 2450 and the resulting norm map estimate are shown in Fig. 3.

Correlation analyses are carried out between norm maps estimated from the *same* scanner as well as from *different* scanners. Normal vector's length, x and y -components were used for correlation analysis. Fig. 4 reveals that the distributions of sample correlation coefficient $\hat{\rho}$ for correctly matched patches (H_1) and incorrectly matched patches (H_0) are generally far apart, suggesting a very good authentication performance of the norm map under the well-controlled acquisition condition. This is especially true when patches are obtained from the same scanner. When reference patches and test patches are from different scanners, features have slightly lower performances.

B. Appearance Images by Industrial Cameras

Instead of using scanners to capture images with directional linear light and closely placed imaging sensors, Voloshynovskiy *et al.* [5], [6] examined the imaging setup

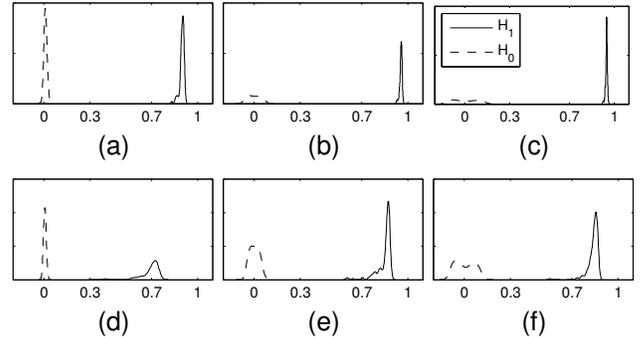


Fig. 4: Estimated PDFs of sample correlation coefficient $\hat{\rho}$ for correctly and incorrectly matched cases. First row: Datasets #2–3 (test) vs. #1 (ref.) of scanner 2450, and second row: Datasets #1–3 (test) of scanner GT vs. #1 (ref.) of scanner 2450. Test statistics: (a)(d) length, (b)(e) x -component of norm vector, (c)(f) y -component of norm vector.

of using two industrial cameras at a high elevation with a semi-controlled lighting condition—a fixed circular ring-shaped light source. The resulting images have a consistent appearance and therefore good authentication performance.

Mobile cameras with a lower resolution were then tested under uncontrolled ambient light [7], whereby the performance was considerably worse than the setup with semi-controlled light. Even using newer mobile cameras such as the iPhone 6 with improved acquisition quality, the authentication performances under uncontrolled ambient light are still limited, as revealed by Fig. 3 of Ref. [3] and Fig. 5(e) of the present paper. One way to improve the authentication performance is to use the intensity gradient-based features of visually observable dots that are less sensitive to the change of lighting, at the cost of increasing the design complexity of the authentication system [7].

IV. PROPOSED PAPER AUTHENTICATION USING IMAGE APPEARANCE UNDER THE CAMERA FLASH

Realizing that the uncontrolled light source may be a major reason for low authentication performance using the appearance images as the feature, we explore a semi-controlled lighting condition with the help of the built-in camera flash of mobile cameras. The relative positions among the light source, lens, and paper patch are known, or at least can be estimated.

The simplest case presented in this section would be using the appearance of patches captured at locations relatively fixed to the lens so that the effect of lighting is the same. A more sophisticated case presented in Section V is to understand the physics of lighting and with multiple appearance images to estimate the normal vector field of the surface for authentication.

The conditions for capturing appearance images are described as follows. Patches are acquired by the built-in cameras of three mobile devices, iPhone 6, 5s, and 5, with and without a flash. The capturing process can happen in a large room with 12 overhead fluorescent light arrays and in a small room with 2 overhead fluorescent light arrays. The device is held by hand roughly in parallel with the surface of a piece of cotton paper and at a height of ~ 15.5 cm. Combinations of capturing conditions and the assigned database ID that will be used in the remaining part of the paper are shown in Fig. 5(a).

Database ID	502	503	506	508	509	511	501	510	505
Lighting	flash only	flash + fluorescent light				flash + blind-filtered daylight		fluorescent light only	
Device	iPhone 6		iPhone 5s		Canon SX230HS		iPhone 6	iPhone 5	iPhone 6
Room Size	small	large	small	large	small				large

(a)

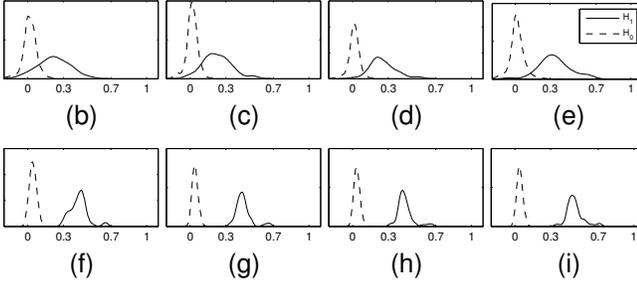


Fig. 5: (a) Database ID and combinations of capturing conditions. Authentication performances of test databases (b) 502, (c) 503, (d) 506, and (e) 505 vs. reference database 505 (fluorescent light only), and test databases (f) 502, (g) 503, (h) 506, and (i) 501 vs. reference database 501 (flash + blind-filtered daylight). Capturing device: iPhone 6.

There are a total of 49 distinct square paper patches on a piece of cotton paper used for the experiment. To build a database for a particular capturing condition, each patch is acquired 3 times (hence, 3 datasets) with possibly slight camera rotation and panning. Patches are acquired together with neighboring patches to speed up the capturing process. A total of 4 shots are needed to capture the whole paper, and the camera locations are shown in Fig. 2(a). Boundaries for patches among different shots are shown using thick lines. Fig. 2(b) and (c) are iPhone 6 acquired images for the top-left 20 patches. Notice the luminance non-uniformity due to shadow and the flash. To test the authentication performance, datasets #2 and #3 of a test database are correlated against the dataset #1 of a reference database.

Experimental Results Fig. 5(b)–(i) reveals that the authentication performance is significantly improved with the help of a flash. Fig. 5(e) with limited authentication performance (at 10^{-2} false-alarm rate, the miss rate is greater than 0.1) is the result when only fluorescent light is available, whereas Fig. 5(i) with high authentication performance is the result when the camera flash is superimposed on the ambient light. Other plots in the second row of Fig. 5 reveal that when the patches with a flash are matched against patches without the flash, the authentication performances are limited. Other plots in the third row of Fig. 5 reveal that when both test and reference databases are captured under the flash, the authentication performances are good and the ambient lighting conditions do not have an important role in affecting the performances.

Fig. 6 contains comprehensive results of various combinations of test and reference datasets. The first table reveals that the flash is the dominating factor in the authentication performance whereas the ambient light is not an important factor. The second table reveals that good authentication performance can be achieved across devices of similar imaging modules. Also, it is reasonable to expect limited performance for all combinations of iPhone camera and Cannon camera,

Ref Test	With Camera's Built-in Flash				No flash
	Small Room			Large Room	
	501	502	503	506	505
501	0.48	0.44	0.43	0.40	0.21
502	0.45	0.45	0.47	0.40	0.22
503	0.43	0.47	0.48	0.41	0.19
506	0.42	0.45	0.42	0.43	0.21
507	0.40	0.41	0.40	0.40	0.23
505	0.21	0.24	0.24	0.24	0.32

(a)

Ref Test	iPhone 6		iPhone 5s		iPhone 5	Cannon SX230
	503	506	508	509	510	511
503	0.48	0.41	0.47	0.44	0.36	0.31
506	0.42	0.43	0.44	0.42	0.36	0.29
508	0.46	0.42	0.52	0.47	0.38	0.33
509	0.44	0.41	0.47	0.50	0.37	0.32
510	0.37	0.35	0.38	0.37	0.35	0.24
511	0.28	0.26	0.32	0.30	0.23	0.26

(b)

Fig. 6: Mode of PDF of correlation values for H_1 for various combinations of test databases and reference databases. Results for (a) flash vs. no flash, and (b) between different cameras. (Italic numbers correspond to the cases that H_0 and H_1 PDFs can be perfectly separated.)

because the relative position of the flash module to the lens and the pattern of the flash are different.

V. PROPOSED SURFACE NORM ESTIMATION FOR PAPER AUTHENTICATION USING MOBILE CAMERAS

Since the image appearance is highly dependent on a camera's design parameter, in practice it is not a good feature to be directly used for authentication, as the acquisition device at the user side cannot be limited to a particular model. Realizing that modern mobile cameras have improved in resolution in capturing fine details, we ask this question: Can we estimate the normal vector field by using multiple appearance images, if the issues of camera geometry and lighting can be addressed? Photometric stereo has long been used to reconstruct surfaces using appearance images captured from different perspectives. However, the challenge here is that the scale of an interested surface is much smaller. We therefore have to carefully choose the physical model of light reflection and control the light in experiments to exploit the possibility of obtaining meaningful estimates of the normal vector field.

Examining the images captured under the flash in Fig. 2(c), we observe that there exists a gentle spatial intensity change at large scale (*aka* the macroscopic intensity) with circular-shaped level curves, and this macroscopic intensity should be compensated to reveal the intensity change due to the change of orientation of the microscopic surface.

It can be shown that the macroscopic intensity is proportional to the light strength at the surface, l , and cosine of the incident angle, θ . We approximate the macroscopic intensity by the *averaged perceived intensity* of background pixels over

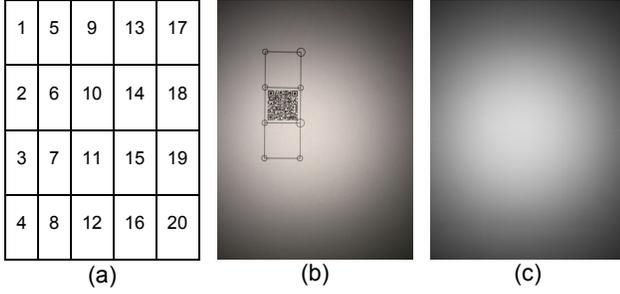


Fig. 7: (a) A total of $M = 20$ indexed camera locations for capturing paper patches, (b) the captured image at location #6 of for Session 5, paper #920, and (c) an estimated macroscopic intensity image obtained from median filtering over a set of images.

a small neighborhood \mathcal{N} around a pixel location \mathbf{p} :

$$\bar{l}_r(\mathbf{p}) = \frac{1}{|\mathcal{N}(\mathbf{p})|} \sum_{\mathbf{k} \in \mathcal{N}(\mathbf{p})} \lambda \cdot l(\mathbf{k}) \cdot \mathbf{n}(\mathbf{k})^T \mathbf{v}_i(\mathbf{k}) \quad (3a)$$

$$\stackrel{(a)}{\approx} \lambda \cdot l(\mathbf{p}) \cdot \left[\frac{1}{|\mathcal{N}(\mathbf{p})|} \sum \mathbf{n}(\mathbf{k}) \right]^T \mathbf{v}_i(\mathbf{p}) \quad (3b)$$

$$\stackrel{(b)}{\approx} \lambda \cdot l(\mathbf{p}) \cdot \mathbb{E}[\mathbf{n}(\mathbf{p})]^T \mathbf{v}_i(\mathbf{p}) \quad (3c)$$

$$\stackrel{(c)}{=} \lambda \cdot l(\mathbf{p}) \cdot [0, 0, \mu_{n_z}]^T \mathbf{v}_i(\mathbf{p}) \quad (3d)$$

$$= \lambda \cdot l(\mathbf{p}) \cdot \underbrace{\mu_{n_z} \cdot v_{i,z}(\mathbf{p})}_{\cos \theta \text{ at } \mathbf{p}} \quad (3e)$$

where $|\mathcal{N}(\mathbf{p})|$ is the number of pixels in the small neighborhood of \mathbf{p} , (a) follows from the fact that $l(\mathbf{k})$ and $\mathbf{v}_i(\mathbf{k})$ are approximately constant over the small neighborhood, (b) follows from ergodicity, and (c) follows from the assumption that normal vectors are on average pointing straight up, *i.e.*, $\mathbb{E}[n_x] = \mathbb{E}[n_y] = 0$ and $\mathbb{E}[n_z] = \mu_{n_z}$, where μ_{n_z} is a modeling constant between 0 and 1.

For simplicity, we employ median filtering over different shots, and satisfactory estimation results for macroscopic intensity \bar{l}_r were obtained and shown in Fig. 7(c).

With an estimated macroscopic intensity image \bar{l}_r , we define the normalized intensity, $\zeta(\mathbf{p})$, of an image at a particular location \mathbf{p} by compensating the macroscopic intensity, as below:

$$\zeta(\mathbf{p}) \stackrel{\text{def}}{=} \frac{l_r(\mathbf{p})}{\bar{l}_r(\mathbf{p})} \cdot \mu_{n_z} \cdot v_{i,z}(\mathbf{p}) = \mathbf{n}(\mathbf{p})^T \mathbf{v}_i(\mathbf{p}) \quad (4)$$

where \mathbf{n} is the unknown normal vector to be estimated, μ_{n_z} is the unknown modeling constant, l_r is the image acquired under flash, and the terms \mathbf{v}_i , \bar{l}_r , and $v_{i,z}$ are already estimated. Normalized images as shown in Fig. 8 are obtained by dividing the original image captured under flash by the macroscopic intensity image.

In order to quickly examine the correctness of modeling, we carry out parameter estimation using handy off-the-shelf estimators such as least-squares. To obtain meaningful estimates with least-squares, we capture a paper patch at 20 (far greater than 3, the number of unknowns) different camera locations with respect to the paper. Indexed camera positions are shown in Fig. 7(a), and an image captured at location #6 is shown in Fig. 7(b).

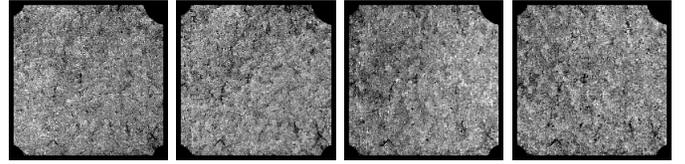


Fig. 8: Normalized images (contrast enhanced) of Session 5 for paper patch #920 captured at camera locations #1-#4.

We estimate the normal vectors at every pixel location for a total of 200×200 pixels. For each pixel location \mathbf{p} , we set up a system of linear equations for solving the normal vector with known or estimated quantities:

$$\underbrace{\begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_M \end{bmatrix}}_{\boldsymbol{\zeta}} = \underbrace{\begin{bmatrix} \mathbf{v}_1 & 1 \\ \mathbf{v}_2 & 1 \\ \vdots & \vdots \\ \mathbf{v}_M & 1 \end{bmatrix}}_{\mathbf{X}} \underbrace{\begin{bmatrix} n_x \\ n_y \\ n_z \\ b \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_M \end{bmatrix}}_{\mathbf{e}}. \quad (5)$$

The unknown parameter $\boldsymbol{\beta}$ contains the normal vector and an intercept capturing any intensity bias, say, due to ambient light, at location \mathbf{p} . The observation vector $\boldsymbol{\zeta}$ consists of normalized intensity values at the collocated position \mathbf{p} from images #1 to # M . The data matrix \mathbf{X} is composed of vectors of incident directions and noise from measurement and/or modeling is modeled by the zero-mean error vector \mathbf{e} .

Experimental Results Two sessions, Session 4 and Session 5, were independently captured in a completely dark environment with an iPhone 6 and its built-in flash for the same paper patch. Two sets of normal vectors were estimated accordingly. The estimates were correlated against 6 slightly different ground-truth norm maps that were obtained from scanners for authentication. (Recall Section III-A shows that the norm maps estimated using the scanner setup are highly consistent) Authentication performances for both sessions are similar and hence only the results for Session 4 are shown in Fig. 9(a). It is observed that the correct matches (H_1) for using x - and y -components as the test statistics are clearly separated from incorrect matches (H_0), which means that the x - and y -components can lead to good authentication performances.

We further investigate a more realistic scenario by allowing ambient light. Recall that we modeled the effect of ambient light using the intercept b in Eq. (5). Session 1 was captured in a room with blind-filtered daylight. Session 6 was captured in a room with low-strength diffuse ambient light from nearly all directions. Detailed authentication results for Sessions 1 and 6 are shown in Fig. 9(b) and (c), which are as good as the results when images are captured without ambient light. It is evident that the estimated normal vector field can give satisfactory authentication performance.

Factor Analysis In this subsection, we examine the factors that may affect the authentication performance.

Number of Images for Normal Vector Field Estimation: In previous experiments, we used 20 images to estimate 4 parameters in order to obtain good estimates with high confidence. Our additional experiments show that even with merely 5 images, the authentication performances are still satisfactory in the sense that the sample correlation values are significantly greater than 0 for correct matches.

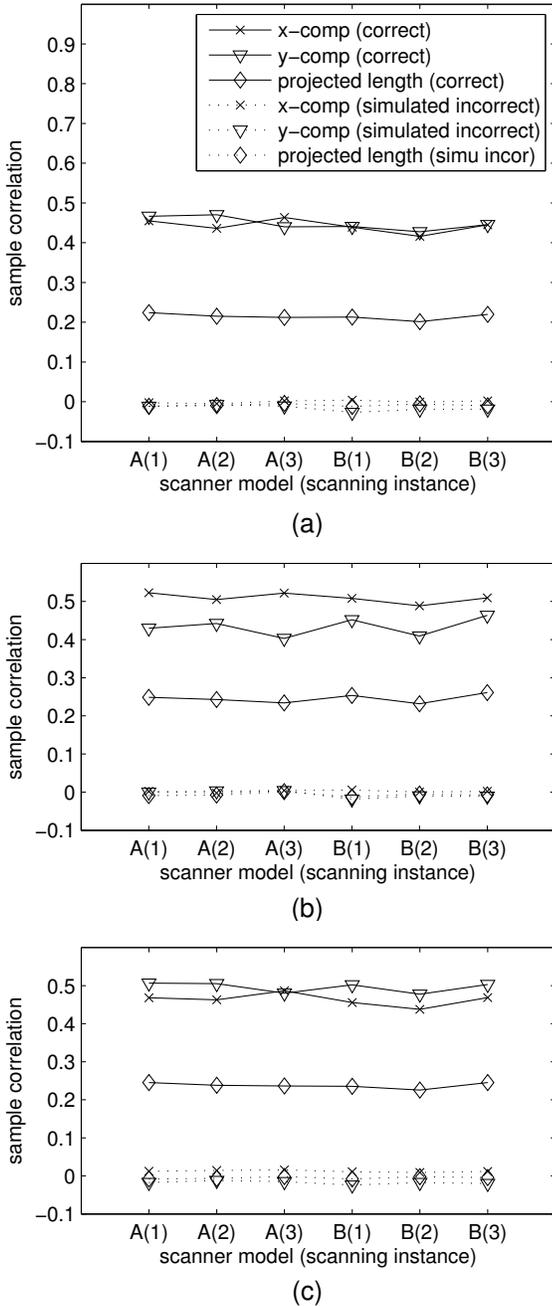


Fig. 9: Authentication performances for (a) Session 4 (totally dark), (b) Session 1 (blind-filtered daylight), and (c) Session 6 (low-strength diffuse ambient light). “A” is scanner 6500, “B” is scanner 2450.

Precision of Estimated Lens Location: The incident light direction \mathbf{v}_i is an essential quantity in obtaining estimates for the normal direction field. In our work, \mathbf{v}_i is itself an estimate from the perspective transform matrix that may be inaccurately estimated, and we would like to check how authentication performance is affected by the imprecise estimate of \mathbf{v}_i .

We carry out a perturbation analysis on the images of Session 4. A 2cm-bias (which corresponds to 2 times the distance between the lens and the flash, about 10° -bias in angle) in horizontal, vertical, and diagonal directions is added to the flash’s estimated location. The correlation is around 0.45 for correct matches when no perturbation is added.

Experimental results show that when the perturbation is in x - (or y -) direction, the x - (or y -) component of the estimated normal vector field has a reduction of about 0.15 in correlation, and the other component has no change. When the perturbation is in diagonal directions, both x - and y -components have a reduction of about 0.1 in correlation. In spite of the reduction in correlation, the correct matches can still be perfectly separated from incorrect matches. Hence, we do not see a strong need to improve the lens location estimator, and the authentication performance is not significantly affected by the 10° -bias of the estimated lens location.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have shown that it is possible to use cameras and built-in flash of mobile devices to estimate the normal vector field that is an intrinsic microscopic feature of the paper surface for authentication purposes. This finding can relax restricted imaging setup to enable paper authentication under a more casual, ubiquitous setting of a mobile imaging device, which may facilitate duplicate detection of paper documents and merchandise packaging.

Further investigation should be carried out for scenarios where the camera is not in parallel with the paper surface. This poses a challenge due to the out-of-focus blur effect. Another direction is to investigate quantitatively the sources of correlation reduction from the > 0.8 range in the scanner setup for correct matches to about 0.4–0.5 in the mobile camera setup. Such a study will elucidate a deeper understanding and potential improvement of the paper PUF authentication.

REFERENCES

- [1] Product Overview on BubbleTag™, Ramdot™, FiberTag™, ProofTag SAS, Retrieved Jan. 2015. <http://www.prooftag.net/>
- [2] Kinde Anti-Counterfeiting Labels, Guangdong Zhengdi (Kinde) Network Technology Co., Ltd., Retrieved Jan. 2015. <http://www.kd315.com/>
- [3] C.-W. Wong and M. Wu, “A study on PUF characteristics for counterfeiting detection,” in *Proc. IEEE International Conference on Image Processing (ICIP)*, Quebec City, Canada, to appear, Sep. 2015.
- [4] W. Clarkson, T. Weyrich, A. Finkelstein, N. Heninger, J. Halderman, and E. Felten, “Fingerprinting blank paper using commodity scanners,” in *Proc. IEEE Symposium on Security and Privacy*, Berkeley, CA, May 2009, pp. 301–314.
- [5] S. Voloshynovskiy, M. Diephuis, F. Beekhof, O. Koval, and B. Keel, “Towards reproducible results in authentication based on physical non-cloneable functions: The forensic authentication microstructure optical set (FAMOS),” in *Proc. IEEE International Workshop on Information Forensics and Security (WIFS)*, Tenerife, Spain, Dec. 2012, pp. 43–48.
- [6] M. Diephuis and S. Voloshynovskiy, “Physical object identification based on FAMOS microstructure fingerprinting: Comparison of templates versus invariant features,” in *Proc. International Symposium on Image and Signal Processing and Analysis (ISPA)*, Trieste, Italy, Sep. 2013, pp. 119–123.
- [7] M. Diephuis, S. Voloshynovskiy, T. Holotyak, N. Stendardo, and B. Keel, “A framework for fast and secure packaging identification on mobile phones,” in *Proc. SPIE, Media Watermarking, Security, and Forensics*, San Francisco, CA, Feb. 2014, p. 90280T.
- [8] S. A. Shafer, “Using color to separate reflection components,” *Color Research & Application*, vol. 19, no. 4, pp. 210–218, 1986.
- [9] “High resolution surface topography FRT MicroProf chromatic aberration sensor,” in *a product sheet by Inventia AB*, Aug. 2012.
- [10] M.-C. Beland and J. M. Bennett, “Effect of local microroughness on the gloss uniformity of printed paper surfaces,” *Applied Optics*, vol. 39, no. 16, pp. 2719–2726, Jun. 2000.