Enhanced Geometric Reflection Models for Paper Surface Based Authentication

Runze Liu[1], Chau-Wai Wong[1], Min Wu[2]

{rliu10, chauwai.wong}@ncsu.edu; minwu@umd.edu

[1]North Carolina State University, USA, [2]University of Maryland, College Park, USA

Abstract

Paper under the microscopic view has a rough surface formed by intertwisted wood fibers. Such roughness is unique on a specific location of the paper and is almost impossible to duplicate. Previous work has shown that commodity scanners and cameras are capable of capturing such intrinsic roughness in term of surface normal vectors for security and forensics applications. In this paper, we examine several candidate mathematical models for camera captured images of paper surfaces and compare the modeling accuracies with reference to the measurement by the confocal microscopy. Experimental results show that the model with distinct intensity bias for images captured from different viewpoints can provide the closest result to the confocal measurement. We discover that high-frequency subbands of reconstructed 3D surfaces are more powerful than the norm map in describing the uniqueness of a physical surface. We show through a practical paper surface based authentication system that incorporating these findings can improve the discrimination performance.

1. Introduction

Paper surfaces under the microscopic view has fuzzy random appearance caused by inter-twisted wood fibers [1–3, 5, 6, 9, 10, 15, 17–19]. Such randomness is intrinsic and physically unclonable, which makes the paper surface an ideal candidate for anti-counterfeiting applications. Using the microscopic roughness of surfaces, important documents, concert tickets, and packages can be uniquely identified without relying on other extrinsic means.

Two major approaches have been taken to exploit the paper surface for authentication. The first approach [1, 2, 9, 10, 15] directly uses the camera/scanner captured image or derived features of the image to characterize the paper. The second approach [3, 18, 19] characterizes the paper using the normal vector field that quantifies the microscopic roughness estimated using several images of the patch. The physical laws of light reflection is employed to derived the normal vectors at the surface. In this paper, we focus on

the second approach that uses the normal vector field as the matching feature. Previous work used scanners [3, 18] and mobile cameras [19] to capture the reflected light and estimate the normal vector fields for authentication.

It is desirable to have a deeper understanding in the models and algorithms for estimating such physical feature using mobile devices because of their widespread use. The estimated normal vector field from the camera-captured images can achieve satisfactory authentication results [19]. However, the rudimentary model resulted in coarse normal vector estimates when verified against the confocal microscopy results [19]. In addition, the understanding of the normal vector field's key information toward satisfactory authentication performance is lacking.

In this paper, we propose new mathematical models that takes into account both the light reflection and the image acquisition processes to estimate the normal vector field. We reconstruct 3D microscopic paper surfaces from normal vector fields and decompose the surfaces into different spatial-frequency subbands. The surfaces and their frequency subbands estimated from different models are verified against those measured from the confocal microscopy. The proposed models and new features will be tested in a practical authentication system to measure the performance gain.

The paper is organized as follows. In Section 2, we review the fundamentals on light reflection and tools that will be used. In Section 3, we propose new mathematical models. In Section 4, we examine the discrimination performance of the proposed models at different subbands. In Section 5, we evaluate the performance gain in a practical engineering system when our findings are incorporated. In Section 6, we conclude the paper and discuss future work.

2. Background and Preliminaries

2.1. Diffuse Reflection Model and Baseline Method for Norm Vector Estimation

Under the fully diffuse model, the perceived intensity l_r at location p is modeled as follows [3, 16, 19]:

$$l_r(p) = \lambda \cdot l(p) \cdot n(p)^T v(p), \qquad (1)$$

where $n = (n_x, n_y, n_z)$ is the normal direction of the paper surface at the microscopic level, $v(p)$ is the incident light direction arriving at p, $l(p)$ is the strength of the light and λ characterizes the physical capability of reflecting the light.

The authors of [19] decomposed the perceived intensity of a paper patch under a point source into two components: i) a smooth component, named the macroscopic intensity, with a mild spatial change mainly due to the varying distance between each pixel and the point source, and ii) a highly fluctuating component due to the inconsistent orientation of the paper surface at the microscopic level. In order to estimate the microscopic paper surface that leads to the fluctuating component, the authors approached the problem by compensating the effect of the macroscopic intensity, and then reformulating the problem into a linear regression problem.

It was shown that paper's macroscopic intensity $\tilde{y}(p)$ is spatially dominated by $l(p)$, namely,

$$\tilde{y}(p) \approx \mathbb{E}[l_r(p)] = \lambda \cdot l(p) \cdot m_z \cdot v_z(p) \qquad (2)$$

where m_z is mean of the z component of the normal vectors, $v_z(p)$ is the z component of the incident light direction vector and its value is usually slightly less than 1. With the equality established in Eq. (2), the term containing the arriving light intensity $\lambda l(p)$ may be estimated by:

$$\widehat{\lambda l(p)} = \tilde{y}(p)/(m_z \cdot v_z(p)). \qquad (3)$$

Substituting Eq. (3) into Eq. (1), the effect of the smooth component can then be compensated and the remaining knowns and estimated values are now in a linear form of the three coordinates of normal vector $n(p)$, namely,

$$\zeta(p) \approx n(p)^T v(p) \qquad (4)$$

where $\zeta(p) \triangleq l_r(p)/\widehat{\lambda l(p)}$. To solve for the three unknowns in $n(p)$ and to account for some bias for a particular location p, at least four equations, which correspond to four camera-captured photos of the paper surface, are needed. Normal vectors for different locations are calculated separately and there is no spatial smoothness assumption imposed. We refer to this baseline model as Model 0.

2.2. 3D Surface Reconstruction From Normals

To reconstruct a 3D surface from a normal vector field, an intuitive approach is to spatially integrate the gradient field, but such approach can be sensitive to noise and the gradient may be non-integrable due to discontinuities of the surface [8]. We therefore use a method named shapelet proposed in [11] to reconstruct 3D microscopic paper surfaces. In this method, a set of filters, namely, the gradients of the shapelets are convolved with gradient field. The filtered versions of the gradient field are summed up to form the reconstructed 3D surface. Such reconstruction process is highly robust to noise.

2.3. Difference-of-Gaussian Representation

The difference-of-Gaussian (DoG) representation [13, 14] of a function on a 2D spatial grid, such as a topographical map or an image, allows us to investigate the behaviors of the function around different spatial-frequency subbands. We decompose the reconstructed topographical map of the paper surface into different spatial-frequency subbands to investigate the surface trend and the spatial changes of the surface. A DoG representation with N levels is constructed by taking the differences of Gaussian-blurred images. We define the nth subband as follows:

$$L_n = G_n - G_{n+1}, \ n = 1, ..., N \qquad (5)$$

where G_1 is defined to be the original image, G_{N+1} is defined to be a constant intensity image with the intensity to be the average intensity of the original image, and G_n, $n = 2, ..., N$, is the result of blurring the original image by a Gaussian filter with the standard deviation $\sigma^{n-1}, \sigma > 1$.

2.4. Measure of Discrimination Performance

We use the hypothesis testing framework [12,19] to evaluate discrimination performance of a system. The null hypothesis H_0 is that the test paper surface does not match the reference surface, and the alternative hypothesis H_1 is that the test surface matches the reference surface. The Pearson's correlation between the test and the reference surfaces quantifying the degree of match is used as a test statistic. We repeatedly collect the correlation values under H_0 and H_1, and estimate the probability density functions by calculating the histograms for H_0 and H_1. Thresholding is applied to calculate the probability of false alarm, P_F, and the probability of miss detection, P_M. The discrimination capability is measured by the receiver operating characteristic (ROC) curve and more compactly, by the equal error rate (EER).

When designing a practical engineering system such as an authentication system, we can decide whether the paper surface captured during a test session matches its record in a reference database using the thresholding rule. We calculate the correlation, and compare it against a predefined threshold τ that controls the tradeoff between P_F and P_M. The surface being examined is only considered authentic when the correlation is larger than τ.

3. Proposed Models for Estimating Normals

In this paper, we explicitly model factors that may lead to the bias in intensity such as the ambient illumination and cameras' internal brightness/contrast adjustment processes. We propose models with intercepts in addition to the diffuse reflection component for estimating the normal vector field.

We first propose a model with distinct intercept for each image. For image k, $k = 1, \ldots, M$, the intensity of ac-

quired image at pixel location p is modeled as:

$$y^{(k)}(p) = \lambda l^{(k)}(p)n(p)^T v(p) + \beta_0^{(k)}(p). \quad (6)$$

We name this model with distinct intercept over k and p Model 1. To estimate the normal vector $n(p)$, we follow a similar procedure as in Section 2.1: We first compensate the smooth component in terms of $\lambda l^{(k)}(p)m_z$ and $\beta_0^{(k)}(p)$ and then reformulate the problem into a linear regression problem. By taking an expectation over the randomness of the normal vector, we can obtain a macroscopic intensity field that contain the global factors, i.e., the strength of the arriving light, the overall intensity bias, and the effect of the camera direction:

$$\tilde{y}^{(k)}(p) \approx \mathbb{E}\big[y^{(k)}(p)\big] = \lambda l^{(k)}(p)m_z v_z(p) + \beta_0^{(k)}(p), \quad (7)$$

where all symbols are similarly defined as in Section 2.1. In order to estimate $\lambda l^{(k)}(p)m_z$ and $\beta_0^{(k)}(p)$, we make an explicit assumption that these two terms are roughly constant over a small neighborhood of the spatial grid, e.g., a 4-connected set. When the context is clear and the order of the four neighboring points does not matter, we denote the four points as $p_1, \ldots, p_4$ and denote the center point as p_0. Therefore, by explicitly applying the spatial smoothness constraint, we set up the overdetermined system as follows:

$$\begin{bmatrix} \tilde{y}^{(k)}(p_0) \\ \vdots \\ \tilde{y}^{(k)}(p_4) \end{bmatrix} \approx \begin{bmatrix} v_z^{(k)}(p_0) & 1 \\ \vdots & \vdots \\ v_z^{(k)}(p_4) & 1 \end{bmatrix} \begin{bmatrix} \lambda l^{(k)}(p_0)m_z \\ \beta_0^{(k)}(p_0) \end{bmatrix}. \quad (8)$$

Solving this linear system using least-squares gives us the estimates for $\lambda l^{(k)}(p)m_z$ and $\beta_0^{(k)}(p)$. Substituting the estimates into Eq. (6), we may setup another system of linear equations using M images:

$$\begin{bmatrix} \tilde{y}^{(1)}(p) \\ \vdots \\ \tilde{y}^{(M)}(p) \end{bmatrix} - \begin{bmatrix} \widehat{\beta_0^{(1)}(p)} \\ \vdots \\ \widehat{\beta_0^{(M)}(p)} \end{bmatrix} \approx \begin{bmatrix} \widehat{\lambda l^{(1)}(p)}v^{(1)T}(p) \\ \vdots \\ \widehat{\lambda l^{(M)}(p)}v^{(M)T}(p) \end{bmatrix} \begin{bmatrix} n_x(p) \\ n_y(p) \\ n_z(p) \end{bmatrix} \quad (9)$$

in which the normal vector at location p can be obtained by least-squares.

We propose a more stringent Model 2 by assuming that the bias in intensity such as the ambient illumination and cameras' internal brightness/contrast adjustment processes has the same effect for all images, namely, $\beta_0^{(1)}(p) = \ldots = \beta_0^{(M)}(p)$, so that the model becomes:

$$y^{(k)}(p) = \lambda l^{(k)}(p)n(p)^T v(p) + \beta_0(p). \quad (10)$$

In this scenario, $\beta_0(p)$ can be jointly estimated using all M

images with the linear system shown as follows:

$$\begin{bmatrix} \tilde{y}^{(1)}(p_0) \\ \vdots \\ \tilde{y}^{(1)}(p_4) \\ \vdots \\ \tilde{y}^{(M)}(p_0) \\ \vdots \\ \tilde{y}^{(M)}(p_4) \end{bmatrix} \approx \begin{bmatrix} v_z^{(1)}(p_0) & 0 & \ldots & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ v_z^{(1)}(p_4) & 0 & \ldots & 0 & 1 \\ \vdots & & & & \vdots \\ 0 & \ldots & 0 & v_z^{(M)}(p_0) & 1 \\ \vdots & & & \vdots & \vdots \\ 0 & \ldots & 0 & v_z^{(M)}(p_4) & 1 \end{bmatrix} \begin{bmatrix} \lambda l^{(1)}(p_0)m_z \\ \vdots \\ \lambda l^{(M)}(p_0)m_z \\ \beta_0(p_0) \end{bmatrix}.$$
$$(11)$$

Finally, the normal vector at each location p could be estimated following the similar procedure as described for Model 1.

4. Experimental Results

4.1. Dataset and Experimental Setup

For the analysis conducted in this section, we use a publicly available dataset on which the confocal microscopy related work in Section VII of [19] was conducted. The dataset contains measurements by a confocal microscope, and images of a paper surface acquired by two commodity scanners and a mobile camera. Measurements from the confocal microscope leads to a topographic map with high spatial resolution. We follow the procedure mentioned in Section VII.C of [19] that calculates the surface direction over the squared coverage of each working pixel to generate a 200-by-200 normal vector field. We use it as the physical reference since the confocal measurement is precise. For the scanner acquired images, we estimate norm maps using the improved method mentioned in Section III.A of [19].

The images of the paper surface by mobile camera were acquired when a flashlight is activated under a normal indoor ambient light environment in 20 different camera locations. In order to evaluate the statistical behaviors of a hypothetical, representative large dataset (i.e., the statistical population) using the limited data from the dataset, we apply the idea of bootstrapping [4, 7] to estimate a set of resampled normal vector fields. Specifically, we use five images randomly chosen from the 20 images to estimate one resampled normal vector field, and repeat such random sampling with replacement 100 times to prepare for the data for the subsequent evaluation for models. Five is the second smallest number needed to estimate a normal vector field. Although the choice of using five images will reveal pessimistic results on the discrimination performance, it reasonably mimics application scenarios that allow capturing fewer images and are with limited computational resources.

Visualizing norm maps or normal vector fields can reveal the surface orientation at each pixel location. For example, when examining a neighborhood of pixels, one can measure the spatial consistency of surface orientations. In
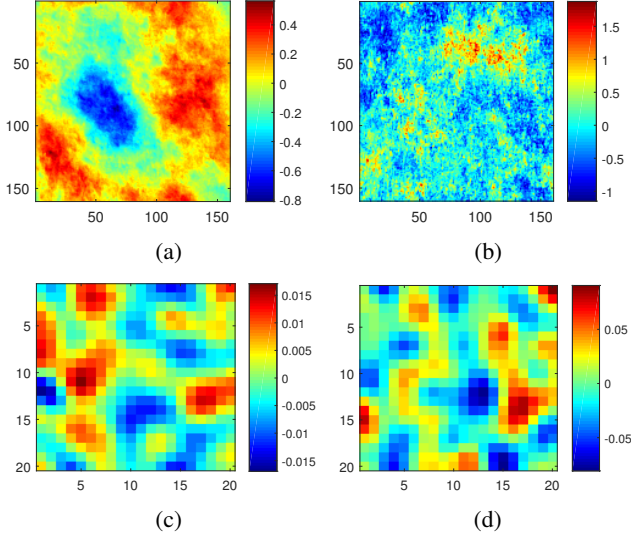
Figure 1: Top view of a reconstructed surface by (a) mobile camera with normal vector field estimator derived from Model 1 and (b) confocal microscope. (c) and (d) are SUBBAND #2 of a zoomed-in region of (a) and (b), respectively, exhibiting higher similarity than that between (a) and (b). The colorbar illustrates the relative height of the surface with $84.5\ \mu m$ as the unit.

comparison, 3D surfaces reconstructed from normal vector fields are more appealing to human eyes and quantitative understandings can be better achieved using off-the-shelf image/surface analysis tools. In this paper, we employ the DoG representation that allows us to separately analyze the discrimination performance at different frequency subbands.

4.2. Match Using Full and Subbands of 3D Surface

We use shapelet reviewed in Section 2.2 to reconstruct 3D microscopic paper surfaces from normal vector fields. Fig. 1(a) and Fig. 1(b) show a reconstructed surface by Model 1 and the reference by the confocal microscope, respectively. The two reconstructed 3D surface appears differently at the patch scale: the surface by Model 1 has a large valley in the middle and high peaks around it, whereas the confocal surface is relatively flat at the patch scale and is spatially "busy" with many peaks. The discrepancy in the surface trend may have masked other features that are consistent between the two surfaces.

Next, we the decompose the surface into different spatial-frequency subbands using a DoG representation of 10 levels. High spatial-frequency fluctuations of the microscopic surface are captured in subbands with small index numbers. Every subband of the surfaces reconstructed from the camera estimated and from confocal estimated norm maps is examined.

Fig. 1(c) and Fig. 1(d) show SUBBAND #2 of a zoomed-in region of the surface by Model 1 and the reference by the confocal microscope, respectively. The two zoomed-in regions are similar that both have small valleys in the middle with peaks around, and there are peaks in the up-right corners.

To illustrate more clearly how well the subbands from camera estimation match those from the confocal measurement, we show in Figs. 2(b)–(d) representative slices in the x direction from SUBBANDS #1 to #3. Fig. 2 reveals that subbands of high spatial-frequencies have many more overlapped peaks and valleys than the "full spectrum" curves from the original surfaces. This is consistent with what we observed earlier in this subsection.

We then quantitatively examine at each subband the correlation between camera estimation and confocal reference. The bootstrap distributions of the correlation values in matched cases (H_1) and unmatched cases (H_0) are shown in Fig. 3. In matched case, the correlation values are relatively large, i.e., around 0.3 at high frequency subbands, and for unmatched cases the correlation values are around 0. The distributions of correlation values for H_0 and H_1 are narrower and farther away from each other at higher frequency subbands, suggesting a better discrimination performance, i.e., the capability of describing the uniqueness of physical surfaces. Using the thresholding rule, we plot the ROC curves for all subbands to reveal the discrimination capability under all decision threshold values. EER is annotated on each ROC curve for easy comparison.

Fig. 4 shows the EER as a function of the subband index when the correlation values are believed to follow Gaussian and Laplacian distributions, respectively. Both Models 1 and 2 outperform the baseline model at high frequency subbands. As the index of the subband decreases, the performance using Models 1 and 2 increases. Model 1 with EER at 10^{-5} (Gaussian) and 10^{-3} (Laplacian) at the highest frequency subband has the best discrimination capability.

Such observation would naturally trigger a question: Can the best discrimination performance be improved by combining neighboring frequency band(s) with the highest frequency band? We therefore calculate the cumulative subbands of reconstructed surfaces and evaluate the performance. Fig. 5 shows the EER values as a function of accumulated subband index when the correlation values are Gaussian and Laplacian, respectively. The plots reveal that the smallest accumulated subband has the best performance, indicating combining other subbands to the highest frequency subband cannot improve the performance.

4.3. Match Using Norm Map

For comparison purpose, we also examine the discrimination performance when the norm map is used as the matching feature. When assuming correlation values are Gaussian distributed, EER values are around 10^{-1}, 10^{-4}, and 10^{-3} for Models 0, 1, and 2, respectively. When assuming correlation values are Laplacian distributed, EER values
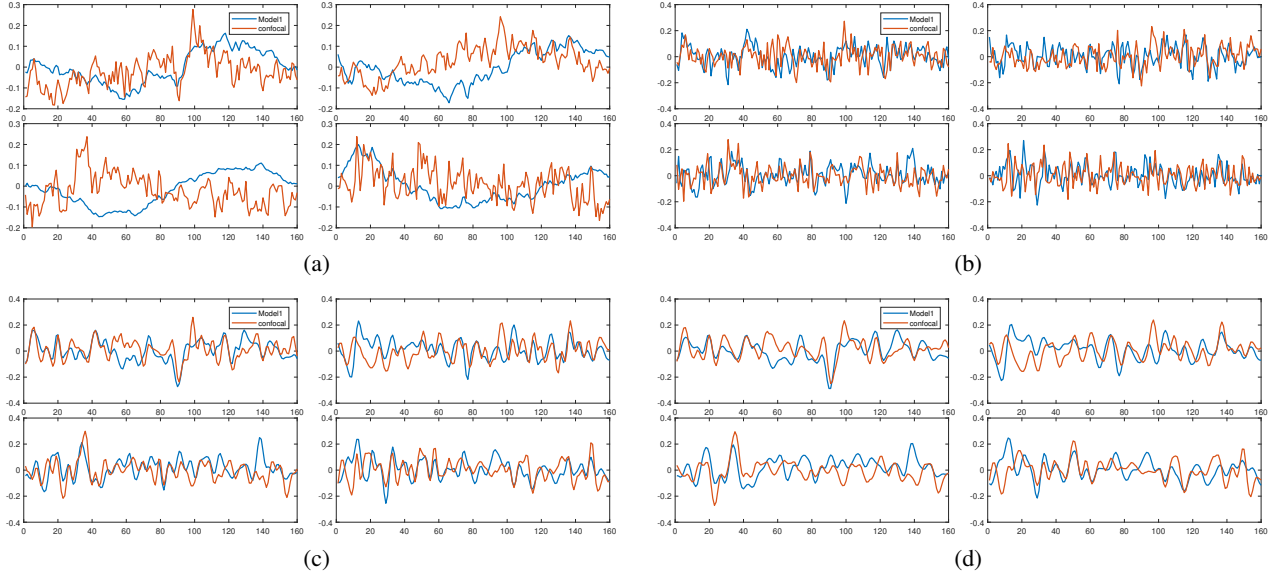
Figure 2: Representative slices in x direction from (a) reconstructed surfaces, (b) SUBBAND #1, i.e., the highest frequency subband, (c) SUBBAND #2 and (d) SUBBAND #3. High spatial-frequency subbands have many more overlapped peaks and valleys than the "full spectrum" curves from the original surfaces.
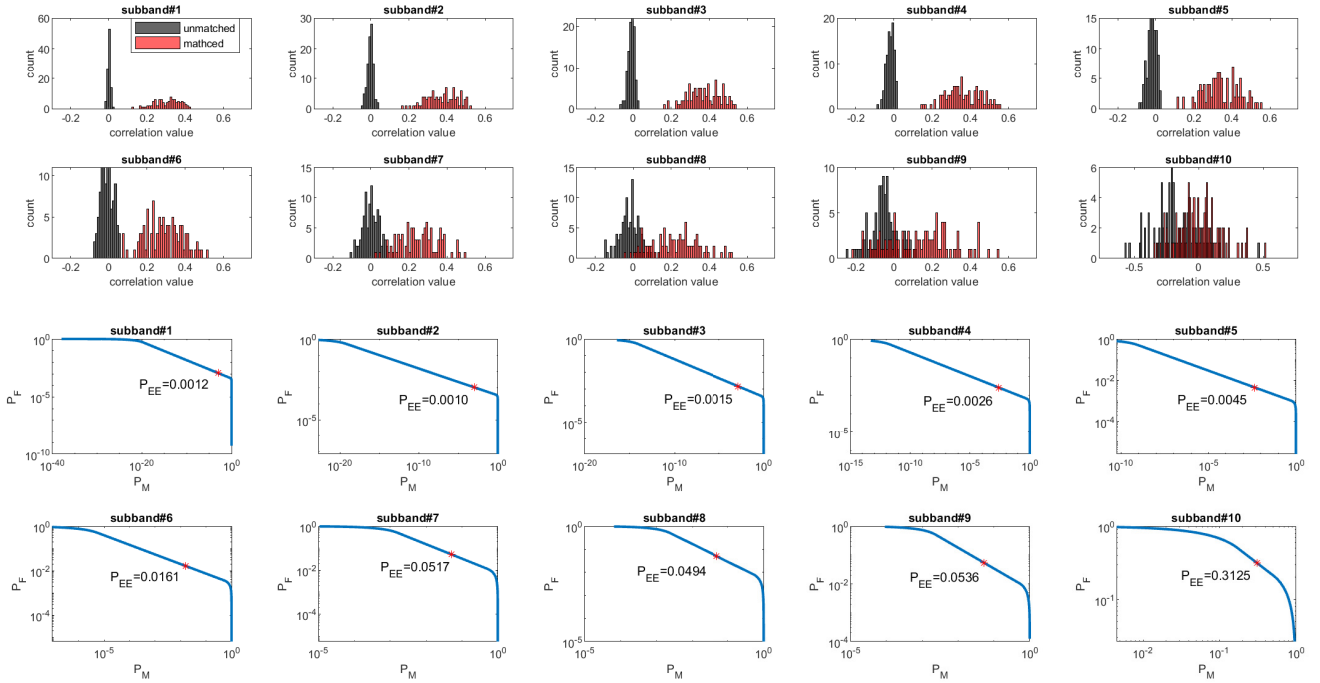

Figure 3: Distributions of correlation values for matched cases (H_1) and unmatched cases (H_0) at different subbands for Model 1, and corresponding ROC curves when assuming correlation is Laplacian distributed. Higher spatial-frequency subbands (those with smaller indices) generally are more powerful in describing the uniqueness of physical surfaces.

are around 10^{-1}, 10^{-2}, and 10^{-2} for Models 0, 1, and 2, respectively. We observe that using the highest frequency subband as the matching feature performs better than using the norm map by one order of magnitude. This implies that the highest frequency subband is physically more discriminative than the norm map.

5. Practical Authentication System

In this section, we examine a practical paper surface based package/label/document authentication system that uses mobile cameras to capture test images and uses scanners to capture the reference. We did not use confocal mi-
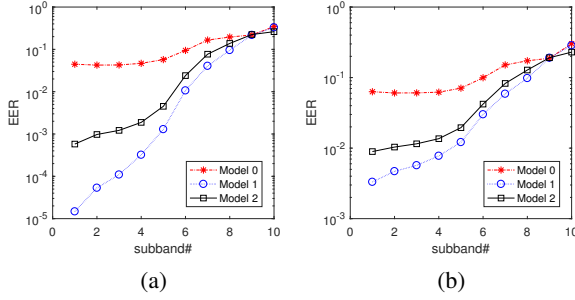
Figure 4: Discrimination capability in terms of EER as a function of subband index for (a) Gaussian and (b) Laplacian distributed correlation values. Two proposed models perform significantly better than the baseline model at high-frequency subbands.
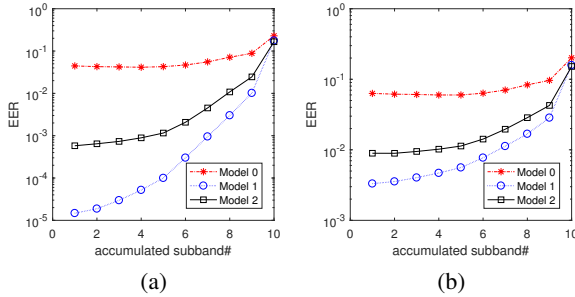

Figure 5: Discrimination capability in terms of EER as a function of accumulated subband index for (a) Gaussian and (b) Laplacian distributed correlation values. When viewed with Fig. 4, it reveals that combining other subbands to the highest frequency subband cannot improve the best performance.

croscope for capturing the reference because it can be difficult to automate and expensive for commercial applications. In comparison, scanners are easy to automate and inexpensive, and has been shown in [3, 18, 19] to have satisfactory performance when used to capture the reference. Cameras may also be used to capture the reference but may lead to lower performance.

In order to use reconstructed surfaces for authentication, we must have the normal vector field that contains the z-component ready. For precise reference produced by scanners, only x- and y-components are available and these two quantities are scaled. To estimate the z-component, one should properly rescale the x- and y-components and use the relationship $n_z = \left(1 - n_x^2 - n_y^2\right)^{1/2}$ to obtain the estimate. An intuitive approach is to match their probability distributions with those of their counterparts from the normal vector field. In our experiment, we rescale the x-component by the scanner such that its standard deviation becomes the same as that by the confocal microscope. The y-components is scaled similarly.

Section 4 reveals that using high frequency subbands as the matching feature with improved models could lead to a better match with the physical measurement. We examine in this section if they can lead to a better performed engineer-
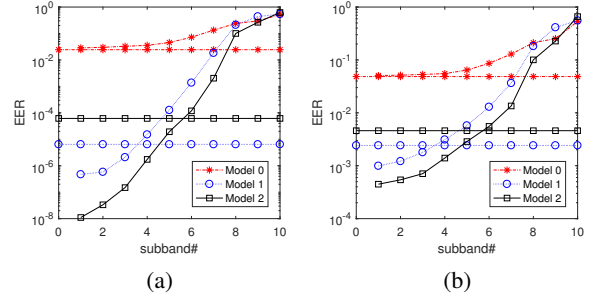

Figure 6: Authentication performance in terms of EER as a function of subband index for (a) Gaussian and (b) Laplacian distributed correlation values. Horizontal lines correspond to the performance when the normal vector instead of the reconstructed surface is used as the matching feature.

ing system. That is, whether they can outperform systems using the norm map directly as the discriminative feature.

Fig. 6 shows the EER as a function of the subband index when the correlation values are believed to follow Gaussian and Laplacian distributions. Horizontal lines correspond to the performance when the norm map instead of the reconstructed surface is used as the feature. The plots reveal that the discrimination capability has significantly improved when the practical authentication system uses the improved models and high-frequency subbands of the reconstructed surface as the feature. When assuming correlations are Gaussian distributed, EER values are improved for about one and four orders of magnitude for Models 1 and 2 at SUBBAND #1, respectively. In contrast, no improvement is observed for the baseline model even the highest frequency subband of the reconstructed surface is used as the matching feature. When assuming correlations are Laplacian distributed, EER values are also improved at SUBBAND #1 for the proposed models but not for the baseline model.

6. Conclusion and Future Work

In this paper, we have shown that the improved models taking into account the effect of ambient lights and cameras' brightness/contrast adjustment processes can provide better modeling accuracies with reference to the measurement by the confocal microscopy. We have discovered that the high-frequency subbands of the reconstructed surface is a better discriminative feature than the norm map. When such discovery is incorporated into the design of a practical engineering system, it can improve the authentication performance. For future work, we plan to build a large dataset containing confocal measurements, scanner images, and camera images with a variety of paper types, camera models, and acquisition conditions. We also plan to explore the effect of the specular reflection at the paper surface. Such studies could lead to more accurate models for camera captured images and may improve the performance of a practical authentication system.

References

[1] F. Beekhof, S. Voloshynovskiy, O. Koval, R. Villán, and T. Pun. Secure surface identification codes. In *Security, Forensics, Steganography, and Watermarking of Multimedia Contents X*, volume 6819, page 68190D. International Society for Optics and Photonics, 2008. 1

[2] J. D. Buchanan, R. P. Cowburn, A.-V. Jausovec, D. Petit, P. Seem, G. Xiong, D. Atkinson, K. Fenton, D. A. Allwood, and M. T. Bryan. Forgery: "Fingerprinting" documents and packaging. *Nature*, 436(7050):475, 2005. 1

[3] 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*, pages 301–314, Berkeley, CA, May 2009. 1, 6

[4] J. L. Devore. *Probability and Statistics for Engineering and the Sciences*. Cengage Learning, 2015. 3

[5] 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*, pages 119–123, Trieste, Italy, Sept. 2013. 1

[6] 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*, page 90280T, San Francisco, CA, Feb. 2014. 1

[7] B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. Monographs on Statistics & Applied Probability. Chapman & Hall/CRC, 1994. 3

[8] R. Framkot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (4):439–451, 1988. 2

[9] Y. D. Kariakin. Authentication of articles, Oct. 7 1997. Patent WO9724699A1. 1

[10] C. Kauba, L. Debiasi, R. Schraml, and A. Uhl. Towards drug counterfeit detection using package paperboard classification. In *Pacific Rim Conference on Multimedia*, pages 136–146. Springer, 2016. 1

[11] P. Kovesi. Shapelets correlated with surface normals produce surfaces. In *IEEE International Conference on Computer Vision*, pages 994–1001, Beijing, China, Oct. 2005. 2

[12] B. C. Levy. *Principles of Signal Detection and Parameter Estimation*. Springer, 2008. 2

[13] T. Lindeberg. Scale-space theory: A basic tool for analyzing structures at different scales. *Journal of applied statistics*, 21(1-2):225–270, 1994. 2

[14] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004. 2

[15] R. Schraml, L. Debiasi, C. Kauba, and A. Uhl. On the feasibility of classification-based product package authentication. In *IEEE Workshop on Information Forensics and Security*, pages 1–6. IEEE, 2017. 1

[16] R. Szeliski. *Computer Vision: Algorithms and Applications*. Springer, 2010. 1

[17] 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*, pages 43–48, Tenerife, Spain, Dec. 2012. 1

[18] C.-W. Wong and M. Wu. A study on PUF characteristics for counterfeiting detection. In *Proc. IEEE International Conference on Image Processing*, pages 1643–1647, Quebec City, Canada, Sept. 2015. 1, 6

[19] C.-W. Wong and M. Wu. Counterfeit detection based on unclonable feature of paper using mobile camera. *IEEE Transactions on Information Forensics and Security*, 12(8):1885–1899, Apr. 2017. 1, 2, 3, 6