On Smartphone Camera based Fingerphoto Authentication

Anush Sankaran*, Aakarsh Malhotra*, Apoorva Mittal*, Mayank Vatsa, Richa Singh
IIIT Delhi, India
{anushs, aakarsh11002, apoorva11029, mayank, rsingh}@iiitd.ac.in

Abstract

Authenticating fingerphoto images captured using a smartphone camera, provide a good alternate solution in place of traditional pin or pattern based approaches. There are multiple challenges associated with fingerphoto authentication such as background variations, environmental illumination, estimating finger position, and camera resolution. In this research, we propose a novel ScatNet feature based fingerphoto matching approach. Effective fingerphoto segmentation and enhancement are performed to aid the matching process and to attenuate the effect of capture variations. Further, we propose and create a publicly available smartphone fingerphoto database having three different subsets addressing the challenges of environmental illumination and background, along with their corresponding live scan fingerprints. Experimental results show improved performance across multiple challenges present in the database.

1. Introduction

Over the past few years, there has been a rapid growth in the applications a smartphone can provide. Users access and store a lot of confidential and personal data on their cell phones. Therefore, the security of such devices is of vital importance. Existing authentication methods such as pins, passwords, and patterns are inconvenient for users and susceptible to attacks as well. Incorporating biometric based authentication in smartphones can be a suitable alternative. Among the existing biometric modalities, fingerprint is a popular biometric requiring minimal co-operation from the user. In this research, we are exploring the use of fingerprint based authentication in smartphones. The approaches considered for smartphone based fingerprint recognition can be broadly classified into two categories: (i) fingerprint based authentication and (ii) fingerphoto based authentication. Fingerprint recognition is performed with the use of specially designed fingerprint sensors. An embedded sensor (external or attached within the display unit) is used to capture the fingerprint and minutia based approaches are used for matching. For example, Samsung S5 and Apple iPhone 5s have fingerprint sensors for authentication [2]. Adding an optical or capacitive sensor to a smart-phone adds a lot to the cost of the device. However, with improvements in technology, almost every smart phone has a good resolution camera. According to TimAhonen's Phone Book, in 2014, about 89% of all the cameras in use are on mobile devices [1]. Fingerphoto based authentication, as shown in Figure 1, utilizes the camera to capture a photo of the finger which can be used for authentication. Using smartphone camera for fingerphoto capture can provide a cost effective and secure method for user authentication.

(a) (b)

Figure 1. (a) Sample image showing the fingerphoto capturing mechanism using a smart-phone camera, and (b) captured fingerphoto image.

Researchers have explored fingerphoto recognition in the literature and Table 1 summarizes these approaches. The existing research has primarily focused on fingerphoto preprocessing techniques such as quality enhancement [14], pitch correction [19], and foreground segmentation [11]. For matching, only minutia based algorithms have been explored [11, 14, 19]. Li et al. [13] showed that these algorithms do not yield good results on fingerphoto images [13]. The major challenges in fingerphoto matching can be grouped into three major categories:

- Preprocessing and segmentation: As shown in Figure 1, fingerphoto images can be captured at any time and

*Equal contributions from the authors.

Table 1. A literature survey of existing fingerphoto recognition algorithms.

Research	Capture Type	Problems Addressed	Database	Comments
Lee et al., 2005 [11]	Mobile camera	Segmentation, Ridge orientation extraction	2 subsets of 400 and 840 images	Database not available
Lee et al., 2008 [12]	Mobile camera	Yaw, roll, pose estimation	Samsung DB-I, II, III, IV	Controlled illumination fingerprints. Database not available.
Stein et al., 2012 [19]	Mobile camera	Segmentation, pitch correction, quality enhancement	41 subjects using two mobiles	Matching performed real time within the mobile. Database not available.
Li et al., 2012 [14]	Mobile camera	Illumination, resolution variation	2100 fingerphoto using three mobiles	Manual segmentation performed. Database not available.
Li et al., 2013 [13]	Mobile camera	Quality estimation	2100 fingerphoto using three mobiles	Manual segmentation performed. Database not available.
Stein et al., 2013 [18]	Video from mobile camera	Enhancement, fingerprint spoofing	990 fingerphoto images, 66 finger videos	Uniform background and illumination. Database not available.
Proposed	Mobile camera, optical sensor	Background, illumination, segmentation, enhancement, feature representation	3 sets, 128 classes with over 5100 images	Database made publicly available, baseline protocol, and results provided.

at any place. This leads to the challenge of variations in background and illumination.

- Feature representation: Due to the unconstrained nature of image capture, the representation should be invariant to translation, rotation, and pose.
- Database: Another major challenge faced by the research community is the lack of a publicly available database. Researchers have reported the results on private databases and it is not possible to reproduce the experiments[1].

With the above analysis, this research proposes a three-fold contribution: (i) propose a segmentation and enhancement algorithm for fingerphoto images captured using smartphones, (ii) propose a novel Scattering Network (ScatNet) based feature representation [17] and matching algorithm for fingerphoto images, and (iii) create a public fingerphoto database to study and analyze the two important challenges affecting fingerphoto recognition: (a) illumination variations and (b) complex background information.

2. Proposed Fingerphoto Matching Algorithm

As shown in Figure 2, the proposed fingerphoto matching algorithm consists of three major steps: (i) fingerphoto segmentation, (ii) fingerphoto enhancement, and (iii) ScatNet based feature representation and matching.

2.1. Fingerphoto Segmentation

Fingerphoto images usually contain highly varying background information. Across the illumination and background, we observe that the skin color of the finger is uniform and serves as a distinguishing feature. Following this observation, we propose the adaptive skin color thresholding for fingerphoto segmentation. The RGB image is converted into the corresponding CMYK scale and the Magenta (M) component is thresholded using the Otsu's method [16]. This provides a binary mask representing the skin region of the fingerphoto. The mask contains both false positive and false negative errors. To remove the false positive errors, the largest connected component is found using the standard run-length encoding technique [7]. To remove the false negative errors, image opening operation is performed twice with a square structuring element.

Once the skin color is segmented, we need to determine the exact ROI of the binary mask. Starting from the middle row of the image, the left most true pixel is extracted. This pixel is traced in both upward and downward directions till a true pixel of the current row is more than 10 pixels away from the preceding row. The left top and bottom points of the ROI are obtained using this trace. The right side boundary is also traced in a similar manner. This provides both the boundaries and a rectangular ROI is cropped using the four coordinates. Figures 3 (a), (b), (c) show the output at different stages of the segmentation algorithm.

2.2. Fingerphoto Enhancement

Fingerphoto enhancement is essential to normalize the effect of illumination variation, blurriness, and to improve the contrast between ridge and valley structure in a fingerprint. Segmented image obtained from the previous step is converted to gray scale and median filtering is applied to remove any speckle noise introduced during capture. Histogram equalization is performed to address the illumination variation and the resulting image is sharpened to im-

[1] HKPU low resolution database [10] is publicly available. However, it is not captured using smart-phone cameras and has minimal variations in terms of illumination and background.

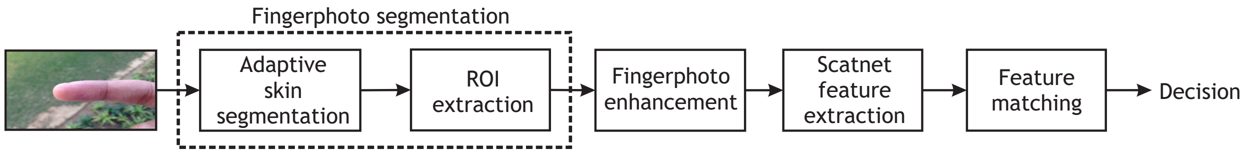

Figure 2. Steps involved in the proposed fingerphoto matching algorithm.

prove the contrast between ridge and valley structures (reducing the blur). Sharpening is performed by subtracting the Gaussian blurred image ($\sigma = 2$) from the original image itself. Thus, the low frequency components are removed and only the high frequency components (such as edges) are preserved, making the ridges more prominent. Figure 3(d) shows a sample of fingerphoto enhancement.

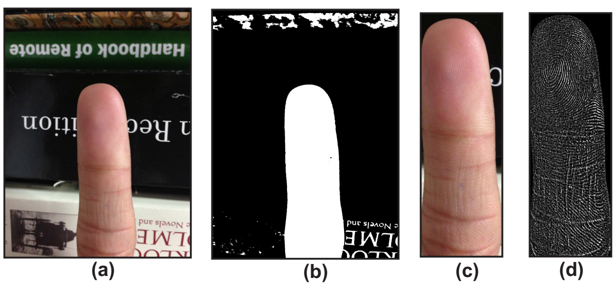

Figure 3. Demonstrating the output at different stages of preprocessing. (a) original fingerphoto, (b) adaptive skin thresholding, (c) ROI extraction, and (d) enhanced image showing the ridge-valley structure.

2.3. ScatNet based Feature Representation

For feature extraction, researchers have explored the representation techniques (minutiae) that are well established for fingerprint matching. It is to be noted that the resolution and clarity of fingerprint images is significantly high compared to fingerphoto images. Therefore, it is challenging to accurately extract these features from fingerphoto images. Li et al. [13] also showed that minutiae based matching does not perform well for fingerphoto recognition.

As discussed earlier, the challenges in fingerphoto matching include illumination, noise, translation, and rotation. With the effect of illumination and background noise being addressed in the preprocessing stage, it is important to find a feature representation that is translation and rotation invariant. Many of the feature extraction techniques discard the high-frequency components from an image as high-frequency components are unstable under deformation [15]. However, as can be observed in Figure 3, the high frequency band of the enhanced image contains the ridge pattern. To retain and effectively encode these properties, we propose feature representation for fingerphoto images using ScatNet features [4]. It has been shown that the ScatNet features are good for extracting texture patterns in images [17]. The enhanced fingerprint has a good ridge-valley texture and it is our assertion that ScatNet features can effectively represent these local patterns.

Scattering Networks (ScatNet)[2] is a filter bank of wavelets that produces a representation which is stable to local affine transformation. Let x be any signal in $\mathbb{R}^2$ (an image, in this case) and $\phi_J(u) = 2^{-2J}\phi(2^{-J}u)$ be a low pass averaging filter. A locally affine invariant representation is obtained by the following convolution:

$$S_0 x(u) = x \star \phi_J(u) \tag{1}$$

This representation is translation invariant upto 2^J pixels and also loses all the high frequency information. To make additional use of the high information as well, a high frequency wavelet bank ψ is constructed by varying the rotation parameter θ and the scale 2^j. The high-frequency, quadrature phase, complex wavelet filterbank is given as $\psi_{\theta,j}(u) = 2^{-2j}\psi(2^j r_{-\theta} u)$. Thus, the wavelet-modulus transform for high frequency components are obtained by:

$$|W_1|x = (x \star \phi(u), |x \star \psi_{\lambda_1}(u)|) \tag{2}$$

These high frequency filters provide additional information to the features obtained in Equation 1. Convolving the wavelet transform coefficients with a low pass filter produces an affine invariant representation of the high frequency components, as follows

$$S_1 x(u, \lambda_1) = |x \star \psi_{\lambda_1}(u)| \star \phi_J(u) \tag{3}$$

These coefficients are called the first-order scattering network coefficients and represent the concatenation of all the filter responses in the wavelet bank $\psi_{\lambda_1}(u)$. If four different frequency bands are chosen in λ_1, the overall response S_1 is the concatenation of wavelet responses of all four filters. Higher-order scattering network coefficients can be obtained by recursively constructing deeper wavelet filter banks as follows:

$$|W_2||x \star \psi_{\lambda_1}(u)| = (|x \star \psi_{\lambda_1}(u)| \star \phi, ||x \star \psi_{\lambda_1}(u)| \star \psi_{\lambda_1}|) \tag{4}$$

[2]MATLAB toolbox publicly available at http://www.di.ens.fr/data/software/scatnet/

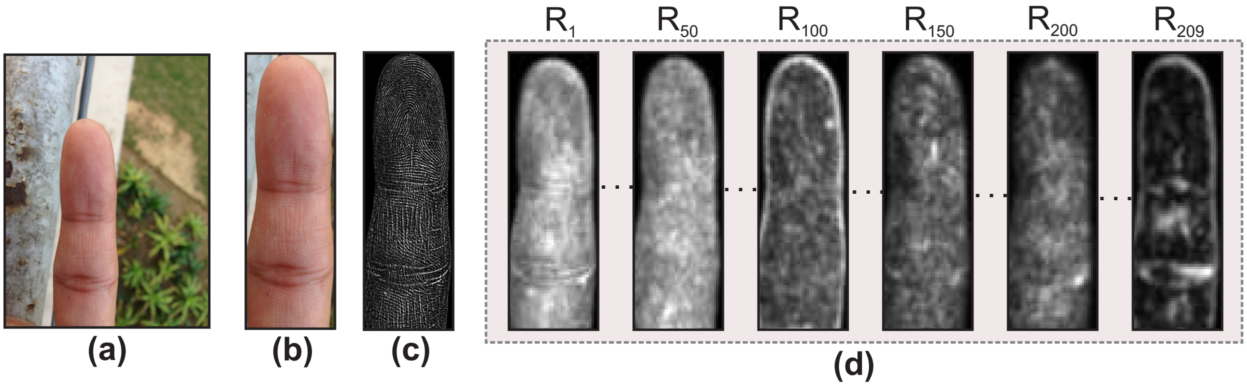

Figure 4. (a) Original image in natural outdoor environment, (b) segmented image using algorithm in Section 2.1, (c) enhanced image using algorithm in Section 2.2, and (d) ScatNet feature representation of the enhanced image. R_1 to R_{209} is the second order ScatNet representation (S_2), which is the concatenation of responses obtained from a wavelet filter bank of 209 filters.

These higher order ScatNet coefficients provide a more stable translation and rotation invariant representation for the fingerphoto images, as shown in Figure 4. The effective representation for a fingerphoto is the concatenation of all n-order responses such as $\{S_0, S_1, \ldots, S_n\}$, where n is chosen as two in this research. Also, as these filters are pre-designed, calculating a ScatNet representation is convolving these filters over the image. Thus, it is easy to extract ScatNet features using only the computation power of smartphones.

2.4. Feature Matching

Let P and G represent the $1 \times N$ length vectorized ScatNet representations of the probe and the gallery fingerphoto images, respectively. To match these features, we demonstrate two different matching scenarios.

- *L1 Distance* between the two ScatNet features is given as follows:

$$d(P, G) = \sum_{i=1}^{N} (P_i - G_i) \tag{5}$$

- *Learning based method:* A supervised binary classifier g:$X \rightarrow Y$ can be learnt to classify a pair of ScatNet feature representations (P, G) as a match or non-match pair. The feature set X is the difference of the two feature representations $(P - G)$ and the classification labels Y are $\{$*Match, Non-match*$\}$. From the difference of representations, the supervised classifier learns whether an image pair is a match or a non-match pair.

3. IIITD Smartphone Fingerphoto Database

Since there is no publicly available fingephoto database, we prepared a new database. The database consists of 128 classes with over 5100 images and is called as the IIITD Smartphone Fingerphoto Database[3]. All the images are captured using Apple iPhone5 at $8MP$ resolution. Flash LEDs are turned off while the auto-focus is kept ON. To analyze the effect of individual challenges such as illumination and background, the database is divided into three subsets. A summary of the dataset is provided in Table 2 and Figure 5 shows sample images corresponding to the three subsets.

- **Set I - White Background:** Fingerphoto images are captured in both indoor (controlled illumination) and outdoor (with uncontrolled lighting) environment with white background, as shown in Figure 5(a). The two subsets (*WI* and *WO*) show the effect of varying illumination with a constant uniform white background. Each subset has 8 images of right index and right middle fingers corresponding to 64 subjects, resulting in *64 subjects* $\times$ *2 fingers* $\times$ *2 lighting variations* $\times$ *8 instances = 2048* images for Set I.

- **Set II - Natural Background:** Fingerphoto images are captured in both indoor and outdoor environment, allowing any natural background to be present, as

[3]The database is made publicly available to the research community to encourage further research: http://iab-rubric.org/resources/spfd.html

Table 2. A summary of the multiple subsets and their variations in the IIITD smartphone fingerphoto database.

Background	Illumination		Classes	Images
Set I (White)	Indoor	(*WI*)	128	1024
	Outdoor	(*WO*)	128	1024
Set II (Natural)	Indoor	(*NI*)	128	1024
	Outdoor	(*NO*)	128	1024
Set III (Live scan sensor)			128	1024

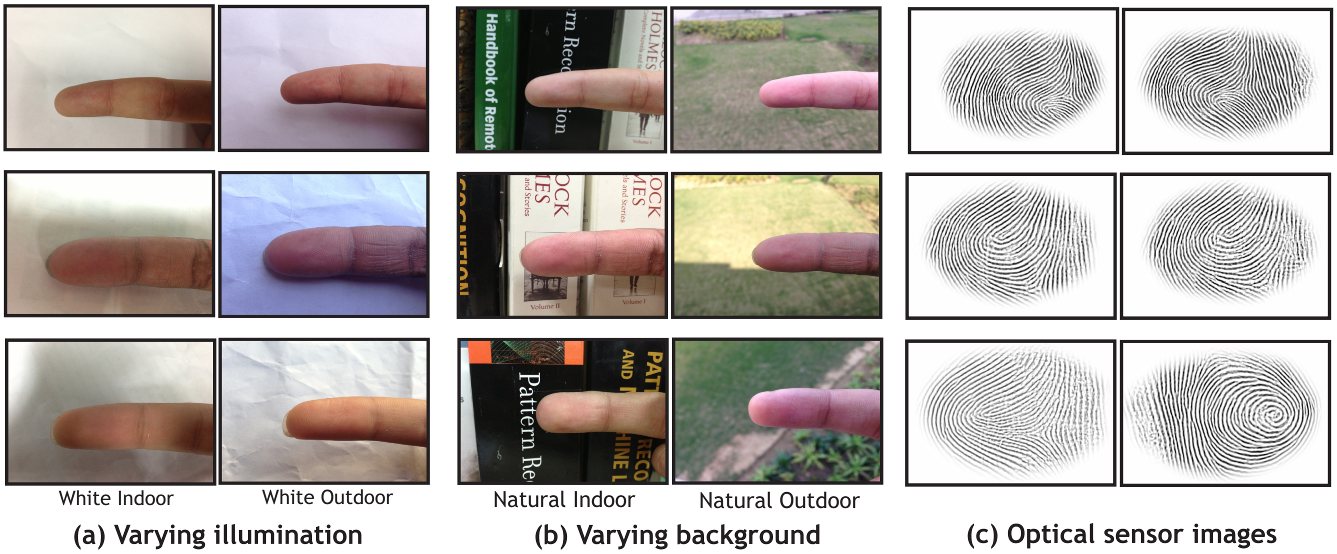

Figure 5. Sample images showing the challenges of illumination and background variations addressed in IIITD smartphone fingerphoto database. Multiple samples also show the amount of intra-class variations and blur noise present in the database. (Figure best viewed in color and under zoom).

shown in Figure 5(b). The subset *NI*, shows the effect of background variation under controlled illumination while the subset *NO*, shows the effect of varying both background and illumination together. Set II also has *2048 images*.

- **Set III - Live Scan:** In online banking applications, there can be a scenario where a fingerphoto has to be matched with a background gallery of live-scan fingerprints. To evaluate the performance in such scenarios, a database of live-scan fingerprints is captured using Lumidigm Venus IP65 Shell fingerprint sensor. Similar to Set I and Set II, 8 images of right index and right middle finger are captured for 64 subjects, thereby having a total of *64 subjects* $\times$ *2 fingers* $\times$ *8 instances* = *1024 images*.

4. Experimental Results

Since the primary purpose of fingerphoto is verification, two main experiments are performed:

- *Fingerphoto-to-fingerphoto comparison (E-I):* White indoor images in *WI* subset are fixed as gallery while the other three subsets $\{WO, NI, NO\}$ are used as probe images, independently. Assuming *WI* as the most stable capture of fingerphoto images, *WI-WO* matching shows the impact of illumination, *WI-NI* shows the impact of background variation, and *WI-NO* shows the impact of illumination and background together on the matching performance. To set a baseline matching, *WI* set is split randomly into 50% gallery and 50% probe, and the results are shown.
- *Fingerphoto-to-fingerprint comparison (E-II):* In this experiment, all four subsets of fingerphoto $\{WI, WO, NI, NO\}$ are independently matched against the gallery of live scan fingerprints (Set III). An additional experiment is performed where the testing database of WI is matched against the live scan gallery to compare with the baseline of *E-I*.

All the segmented images are resized to a standard size of 400×840. Second-order ScatNet representations are used in this research, resulting in a feature representation of length $1,097,250$ per fingerphoto image. Due to the high dimensionality of the representation, PCA is applied, preserving 99% Eigen energy, to get a more succinct representation of fingerphotos. Two supervised classifiers are used, Neural Network and Random Decision Forest (RDF) [8]. RDF is known to yield good results for high-dimensional data [3]. The performance of ScatNet based matching results is compared with (i) minutiae based matching - minutiae are extracted using VeriFinger SDK and matched using MCC descriptor [5] and (ii) CompCode feature based matching [9]. The results of both fingerphoto-to-fingerphoto comparison (*E-I*) and fingerphoto-to-fingerprint comparison (*E-II*) are shown in Table 3[4]. The major conclusions that can be drawn from the results are:

- By comparing the performance of multiple algorithms on both *E-I* and *E-II*, it is observed that ScatNet matching performs better than MCC descriptor based minu-

[4] For more graphical display of results, kindly visit: http://iab-rubric.org/resources/spfd.html

Table 3. Equal Error Rate (%) of the proposed algorithm and comparison with existing algorithms.

	Gallery	Probe	MCC descriptor	CompCode	ScatNet + L1	ScatNet + NN	ScatNet + RDF
E-I	White Indoor (WI)	WO	22.12	**6.90**	18.83	**7.51**	**5.07**
		NI	21.33	**5.02**	19.75	27.32	**7.45**
		NO	21.52	**5.31**	18.98	13.12	**3.65**
	WI/2	WI/2	37.25	**6.61**	28.42	32.89	**6.00**
E-II	Live-scan (LS)	WI/2	31.01	21.07	49.51	20.54	**5.53**
		WI	29.92	14.58	19.38	**15.60**	**7.07**
		WO	**12.92**	14.74	18.95	23.34	**7.12**
		NI	18.05	**10.60**	18.59	17.02	**10.43**
		NO	**12.76**	**11.38**	19.18	17.42	**10.38**

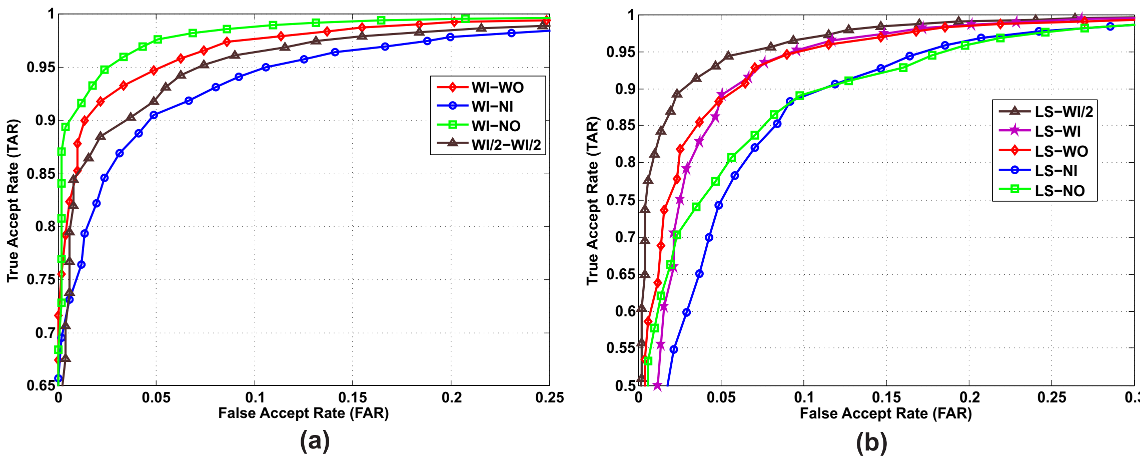

Figure 6. ROC curves using ScatNet features and RDF based matching method for (a) fingerphoto-to-fingerphoto comparison (E-I) and (b) fingerphoto-to-fingerprint comparison (E-II).

tia matching and Gabor filter based CompCode matching. While using ScatNet features, learning based matching provides Equal Error Rate (EER) in the range of $(3-10)\%$, which is much better compared to distance based matching. The ROC curve for the best matching scenario of ScatNet + RDF is shown in Figure 6.

- For the first set of experiment *E-I*, with white indoor fingerphoto as gallery, natural outdoor images provide the best matching results, across all the algorithms. That is, capturing probe fingerphoto images in natural outdoor setting is observed to be most suitable and preferred method. Similarly, for experiment *E-II*, with live scan fingerprint images as gallery, white outdoor images are found to be the best matching probe images.

- To study the effect of illumination, it can be observed that outdoor uncontrolled illuminated images provide better matching performance than indoor images. This is due to the reason that outdoor environment provides a complete surrounding illumination, whereas, indoor environment creates a shadow of the smartphone over the finger (due to the position of the camera and hand). As can be observed in Figure 5(a), this shadow-like formation reduces the clarity of ridge patterns in those regions.

- All the results using ScatNet show that natural indoor probe images provide the lowest matching performance. This is because the background information in indoor environment are close to fingerprint skin color, thereby creating challenges in segmentation and enhancement. Thus, in this dataset, the overall indoor illumination and background variations pose a stronger problem to fingerphoto recognition when compared to outdoor illumination and background variations.

5. Conclusion and Future Work

This research work summarizes the major challenges associated with matching fingerphoto images captured using a smartphone camera. A novel ScatNet based affine invariant fingerphoto representation is proposed. Matching is performed using RDF classification based approach and compared with minutiae based and CompCode based methods. A fingerphoto segmentation and enhancement algorithm is proposed to aid the matching process. To be able to study and address different challenges, publicly available IIITD smartphone database is created. The database consists of three sets having 128 classes with more than 5100 images and is made publicly available for research purpose. The results show a considerable performance improvement over the existing algorithms in different experimental settings. Future work could be (i) to create a real-time mobile application for providing user authentication, like the knuckle-print based user authentication proposed by Cheng and Kumar [6], and (ii) to develop cross resolution fingerphoto matching algorithm.

6. Acknowledgement

The authors would like to thank all the volunteers for cooperating in the data collection process. Anush Sankaran is partially supported by the TCS PhD research fellowship.

References

[1] T. Ahonen. Phone book 2012: Statistics and facts on the mobile phone industry, 2012. 1

[2] J. T. Apostolos, W. Mouyos, J. Feng, and D. T. Jeffrey. Fingerprint based smartphone user verification, 2014. US Patent 20,140,310,804. 1

[3] L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001. 5

[4] J. Bruna and S. Mallat. Invariant scattering convolution networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1872–1886, 2013. 3

[5] R. Cappelli, M. Ferrara, and D. Maltoni. Minutia cylinder-code: A new representation and matching technique for fingerprint recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(12):2128–2141, 2010. 5

[6] K. Cheng and A. Kumar. Contactless finger knuckle identification using smartphones. In *BIOSIG*, pages 1–6, 2012. 7

[7] R. M. Haralock and L. G. Shapiro. *Computer and robot vision*. Addison-Wesley Longman Publishing Co., Inc., 1991. 2

[8] T. K. Ho. The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8):832–844, 1998. 5

[9] A.-K. Kong and D. Zhang. Competitive coding scheme for palmprint verification. In *ICPR*, volume 1, pages 520–523, 2004. 5

[10] A. Kumar and Y. Zhou. Contactless fingerprint identification using level zero features. In *CVPRW*, pages 114–119, 2011. 2

[11] C. Lee, S. Lee, J. Kim, and S.-J. Kim. Preprocessing of a fingerprint image captured with a mobile camera. In *Advances in Biometrics, LNCS*, pages 348–355. 2005. 1, 2

[12] D. Lee, K. Choi, H. Choi, and J. Kim. Recognizable-image selection for fingerprint recognition with a mobile-device camera. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 38(1):233–243, 2008. 2

[13] G. Li, B. Yang, M. Olsen, and C. Busch. Quality assessment for fingerprints collected by smartphone cameras. In *CVPRW*, pages 146–153, 2013. 1, 2, 3

[14] G. Li, B. Yang, R. Raghavendra, and C. Busch. Testing mobile phone camera based fingerprint recognition under real-life scenarios. In *NISK*, 2012. 1, 2

[15] S. Mallat. Group invariant scattering. *Communications on Pure and Applied Mathematics*, 65(10):1331–1398, 2012. 3

[16] N. Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975. 2

[17] L. Sifre and S. Mallat. Rotation, scaling and deformation invariant scattering for texture discrimination. In *CVPR*, pages 1233–1240, 2013. 2, 3

[18] C. Stein, V. Bouatou, and C. Busch. Video-based fingerphoto recognition with anti-spoofing techniques with smartphone cameras. In *BIOSIG*, pages 1–12, 2013. 2

[19] C. Stein, C. Nickel, and C. Busch. Fingerphoto recognition with smartphone cameras. In *BIOSIG*, pages 1–12, 2012. 1, 2