18th International Conference on Information Fusion
Washington, DC - July 6-9, 2015

Fusion of Face and Periocular Information for Improved Authentication on Smartphones

Kiran B. Raja R. Raghavendra Martin Stokkenes Christoph Busch
Gjøvik University College, Gjøvik, Norway
{kiran.raja; raghavendra.ramachandra; martin.stokkenes; christoph.busch}@hig.no

***Abstract*—Authentication on smartphones for high valued transactions in e-banking applications require authentication methods that are both more robust and more convenient than simple PIN based authentication. Although a two factor authentication approach can increase the system security, the second factor has to be robust enough not to be compromised by simple attacks. Face and periocular based biometrics have been explored in many works for authentication on smartphone. Multi-modal biometrics always have an edge over the unimodal biometrics. In this work, we explore various multi-modal biometrics employing feature level and score level fusion to improve the biometric performance of the system. An extensive number of experiments are conducted using two different devices - Samsung Galaxy S5 smartphone and Samsung Galaxy Note 10.1 tablet. The database consisting of 46 subjects has provided a performance with best Genuine Match Rate (GMR) of 95.52% for smartphone and 96.65% GMR for tablet. The obtained scores for feature level and score level fusion advocates the use of a fused approaches over a unimodal biometric authentication on the smartphones.**

I. Introduction

Smartphones are employed as a form of authentication device in many secure applications such as e-banking and mobile commerce. The risk of compromising simple knowledge-based authentication factors such as PINs and alpha-numeric keys instigated the use of biometric characteristics as a means of authentication. Major smartphone manufacturers such as Apple and Samsung have started providing integrated sensors in the mainstream smartphones employing fingerprint for authentication. Such an authentication form has been widely used for authenticating the owner of the device. Increased acceptance of biometric characteristics for authentication has gained the trust of the general public, which has further motivated Apple to launch an e-vallet application like ApplePay [1].

It is interesting to note that smartphones without integrated fingerprint sensors can still be used as biometric sensor by employing the embedded cameras available in the smartphone. Smartphone cameras are well exploited for many biometric applications such as fingerphoto recognition [1], [2]. Earlier works have also obtained high quality biometric samples corresponding to face from smartphones [3], [4]. Inspired by success of various biometric characteristics as modes of authentication on smartphone, in our earlier work, we have explored pericoular image based authentication on smartphones [5]. During the acquisition process of periocular images the face image is a side-product, as the field of view of camera is large and periocular images are to be segmented from the entire face image. Thus the facial characteristics can be exploited as supplementary information for improving a mere periocular based authentication system.

In this work, we explore multi-modal biometrics by using both face and periocular information to make the system more robust and accurate. We first evaluate the baseline performance of each of the characteristics independently and evaluate the improvement in the verification performance when both modalities face and periocular are fused. Another set of experiments are conducted in this work to evaluate the performance of various fusion techniques such as feature level fusion and comparison score level fusion.

In the rest of the paper, we present the layout of the authentication system for smartphones in Section II and we discuss the database used in this set of experiments in Section III. In Section II-A, we discuss the employed feature extraction techniques in this work. Section IV presents the experiments performed in this work along with obtained results. Section VIII provides the summary of the work.

II. Framework of Smartphone Based Authentication

The authentication system for smartphones using face and periocular information is depicted in Figure 1. The system uses both the information from the face and periocular region. Any subject can be enrolled in the authentication system by presenting the face image to either the frontal camera or the rear camera of the smartphone. Once the face is captured by the system, the entire face images and the two segmented periocular images provide features that are stored in the enrollment database. In a similar fashion, probe images for authentication are obtained using frontal or rear camera of the smartphone. For a detailed description of the authentication system in smartphones, the reader can refer [5], [6].

A. Feature Extraction and Comparision

The obtained images are used to extract the features using three well known state-of-art feature extraction techniques : Scale Invariant Feature Transform (SIFT) [7], Speeded Up Robust Features (SURF) [8] and Binarized Statistical Image Features (BSIF) [9]. SIFT features have been well explored for face recognition [10] and periocular recognition [11], [12]. In a similar manner, SURF has been studied for its effectiveness in face and periocular recognition[13], [11], [12]. BSIF has been explored earlier for biometric applications by our works [5], [14]. Further, all these features are well explored for the smartphone based biometric applications and features have demonstrated the suitability for such applications [5], [6].

[1]https://www.apple.com/apple-pay

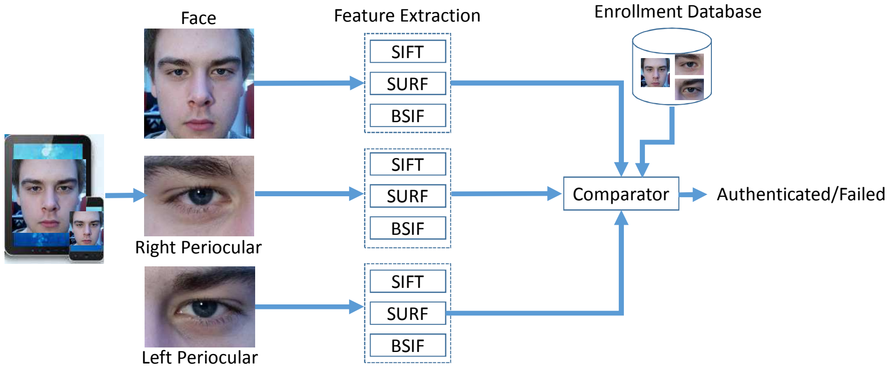


Fig. 1: Multi-modal authentication system for smartphones

TABLE I: Specifications of different hardware in this work

Device	Operating System	Screen Size	Back Camera	Front Camera
Samsung Galaxy S5	Android v4.4.2	1080 x 1920 pixels 5.1 inches	16 MP, 5312 x 2988 pixels	2 MP
Samsung Galaxy Note 10.1	Android v4.4.2	800 x 1280 pixels, 10.1 inches	5 MP, 2592 x 1944 pixels	1.9 MP

The extracted features from reference image and probe image are compared using Bhattacharya distance between histograms [15] for BSIF and Fast Approximate Nearest Neighbor Search [16] for SIFT and SURF features. The obtained comparison score is used as genuine score and impostor score.

III. Database

The authentication system described in Section II, was used to acquire a new database of 46 subjects. The database was constructed using two devices of which the first is a smartphone - Samsung Galaxy S5 and the other is a tablet - Samsung Galaxy Note 10.1. The detailed specifications of the devices and the embedded cameras are given in Table I. The data was collected from both cameras (frontal and rear) for both devices. Additionally, data was collected by a trained operator using the rear camera of both the devices.

A. Reference Image Acquisition

A set of 5 reference images were obtained for each subject in a single session, which were captured 2-3 minutes apart under relatively constrained scenario fulfilling sufficient amount of illumination of the face. All the images acquired are presented to the user to choose the images to be enrolled into the database. This visual inspection by the user works as a secondary quality check after the primary quality factors of the face region are satisfied in accordance with OpenCV face detector [17].

B. Probe Image Acquisition

In order to simulate the real life authentication scenario, where the user is expected to authenticate himself under various unconstrained conditions, we have collected the probe data over 10 different sessions. The sessions are spread in a wide span of time period and the probe data is acquired in 10 different non-constrained conditions. The quality of the acquired image is evaluated by the OpenCV based face detector and additionally the visual input from the user. Thus for each subject the database contains 10 probe images.

C. Experimental Protocol

TABLE II: Details of the database in this work

Camera	Total Subjects	Reference Image	Probe Image	Total Images	Geniune Comparisons	Impostor Comparisons
Smartphone Samsung S5						
Front	46	5	10	690	2300	105800
Back	46	5	10	690	2300	105800
Back Assisted	46	5	10	690	2300	105800
Smartphone Samsung S5						
Front	46	5	10	690	2300	105800
Back	46	5	10	690	2300	105800
Back Assisted	46	5	10	690	2300	105800

The database consists of images collected from 46 subjects using two different devices. The total number of images in the database is provided in the Table II. In order to obtain the baseline performance of the system, each reference image is compared against each probe image. Thus for each subject with 5 reference images and 10 probe images, we obtain a set of 50 genuine scores. As there are 46 subjects in the database, we obtain a total of $45 \times 50 = 2250$ impostor scores for each subject. The total number of genuine and impostor scores for each set of data corresponding to each different device is presented in Table II.

The results are reported in terms of Equal Error Rate (EER) and Genuine Match Rate (GMR) at various False Match Rate (FMR) [18]. The GMR is defined using False Non Match Rate (FNMR) (%) as:

$$GMR = 1 - FNMR$$

TABLE III: Biometric performance in terms of Genuine Match Rate and Equal Error Rate for unimodal approaches.

Camera	Feature Extraction	Face		Right Periocular		Left Periocular	
		GMR @ FMR=0.01%	EER	GMR @ FMR=0.01%	EER	GMR @ FMR=0.01%	EER
Smartphone - Samsung S5							
Back	SIFT	76.36	5.18	57.34	6.63	45.35	7.07
	SURF	45.03	10.21	70.56	**6.55**	59.28	6.20
	BSIF	**87.55**	**4.65**	**75.86**	7.01	**76.00**	**5.80**
Back Assisted	SIFT	88.43	1.88	65.43	5.00	70.83	5.03
	SURF	52.91	5.13	**84.04**	**4.00**	**80.22**	**4.86**
	BSIF	**94.39**	**1.61**	79.00	5.56	72.09	5.73
Front	SIFT	70.73	5.54	17.43	10.92	39.91	9.60
	SURF	28.78	10.93	52.00	9.17	47.32	**8.81**
	BSIF	**84.82**	**2.70**	**58.21**	**9.08**	**56.86**	10.17
Tablet - Samsung Note 10.1							
Back	SIFT	92.83	2.62	31.43	9.40	61.91	8.40
	SURF	81.83	3.34	77.26	6.01	74.30	7.54
	BSIF	**94.61**	**2.43**	**77.39**	**5.91**	**77.30**	**6.75**
Back Assisted	SIFT	95.57	**1.81**	30.83	8.79	48.52	6.87
	SURF	79.91	1.96	71.30	5.31	51.09	5.47
	BSIF	**96.65**	2.03	**82.91**	**5.00**	**85.78**	**4.78**
Front	SIFT	93.26	1.77	51.43	7.73	66.21	7.32
	SURF	75.26	3.5	**66.6**	**6.39**	**76.69**	**5.72**
	BSIF	**94.30**	**0.91**	53.91	8.78	49.43	9.74

For the simplicity of the work, we present the GMR at a $FMR = 0.01\%$. GMR for other values can be obtained from the presented graphs.

IV. Results

This section presents the results obtained in this work. In the first part, we present the baseline results for a single biometric characteristic based verification and in the second part, we present the results of the multi-biometric characteristics based verification of fused features. Further, in the last part we evaluate the verification performance when the comparison scores obtained from different features are fused.

V. Unimodal Verification Performance

Unimodal verification performance in this work is studied using face, left and right periocular images. As discussed previously, we employ SIFT, SURF and BSIF features for each of these biometric characteristics. The obtained results for various features and biometric characteristics are provided the Table III. It can be noted from Table III that the face based verification performs better than either left or right periocular based features. BSIF features has performed consistently well for face and left periocular features. The overall observation is that the GMR obtained for face based verification with both devices and three different acquisition modi is much higher than the verification scores obtained by employing a single region such as left or right periocular image alone. Nevertheless, the EER for the left and right periocular image based verification indicates that the system can be used in low security two factor authentication applications.

Another important factor to note is that the images obtained from the frontal camera perform as good as the images obtained from rear camera of the devices. General patterns in the result also suggest that the performance obtained from the Samsung Galaxy Note tablet is higher as compared to results obtained with a smartphone. This can be attributed to the fact that the camera is centrally placed on side corresponding to longest side of tablet. The placement allows the users to align the face with reference to the camera introducing lesser variations in pose and angle.

VI. Feature level fusion based verification performance

Face region comprises of periocular information which can be used independently or jointly with the face information. Non-standard and unconstrained face acquisitions always suffer from non-standard pose, illumination and angle. Even under such non-standard conditions, the use of the periocular region has proven to perform substantially or equivalently well as compared to face based verification [11]. In the similar terms, our earlier work has confirmed the performance of periocular information for authentication on smartphones [5]. However, there are no works evaluating the performance of the fusion of features from three different biometric characteristics on a smartphone. Thus, in this work we explore feature level fusion for our multi-biometric system.

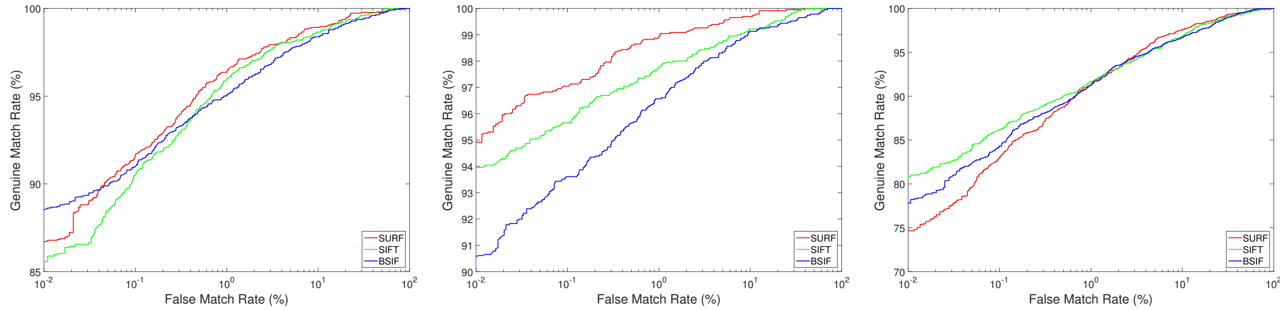
If BSIF features from the face region, left periocular region and right periocular region are denoted by f_b, lp_b and rp_b, then the final feature vector is obtained by concatenation of features obtained from three characteristics given by fv_b:

$$fv_b = f_b \sqcup lp_b \sqcup rp_b$$

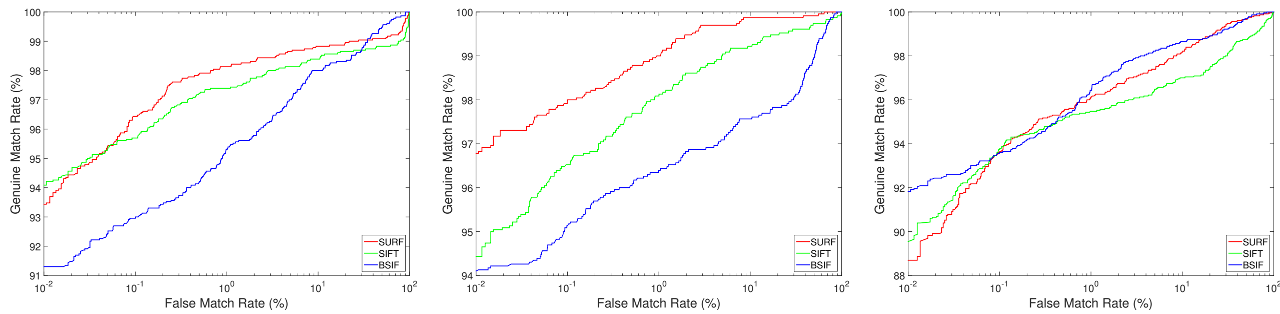
Along the same lines, if the features from SIFT are denoted by subscript s and SURF by u, the final feature vectors can be given as below:

$$fv_s = f_s \sqcup lp_s \sqcup rp_s$$
$$fv_u = f_u \sqcup lp_u \sqcup rp_u$$

The fused feature vector from different characteristics such as face and pericocular region for BISF, SIFT and SURF



(a) Samsung S5 - Back Camera (b) Samsung S5 - Back Camera (assisted) (c) Samsung S5 - Front Camera



(d) Samsung Note - Back Camera (e) Samsung Note - Back Camera (assisted) (f) Samsung Note - Front Camera

Fig. 2: ROC curves for various multi-biometric recognition approaches fusing face and periocular features with the feature level fusion method on Samsung S5 and Samsung Note

TABLE IV: Biometric performance in terms of Genuine Match Rate and Equal Error Rate for multi-biometrics using feature level fusion.

Camera	Feature	Samsung S5		Samsung Note 10.1	
		GMR@FMR=0.01%	EER	GMR@FMR=0.01%	EER
Back	SIFT	86.74	**2.39**	**93.43**	**1.70**
	SURF	85.57	2.57	85.57	2.57
	BSIF	**88.58**	3.10	91.30	3.47
Back Assisted	SIFT	**94.91**	**1.03**	**96.78**	**0.99**
	SURF	93.95	1.99	93.95	1.99
	BSIF	90.60	2.42	94.13	3.13
Front	SIFT	74.65	**4.26**	88.69	2.95
	SURF	**80.82**	4.69	89.65	3.83
	BSIF	77.82	4.65	**91.82**	**2.30**

characteristics are used to obtain the verification scores as given in Table IV. It can be observed from the Table IV that the fusion of the feature vectors from different modalities improved the verification performance, in terms of both EER and GMR. The gains obtained in GMR supports the fusion of the feature vector for improving the verification performance on smartphones and clear advantage can be viewed with respect to unimodal verification performance. Figure 2 presents the Receiver Operating Characteristics (ROC) for the feature level fusion with features from face, left periocular and right periocular features for both devices.

VII. COMPARISON SCORE FUSION BASED VERIFICATION PERFORMANCE

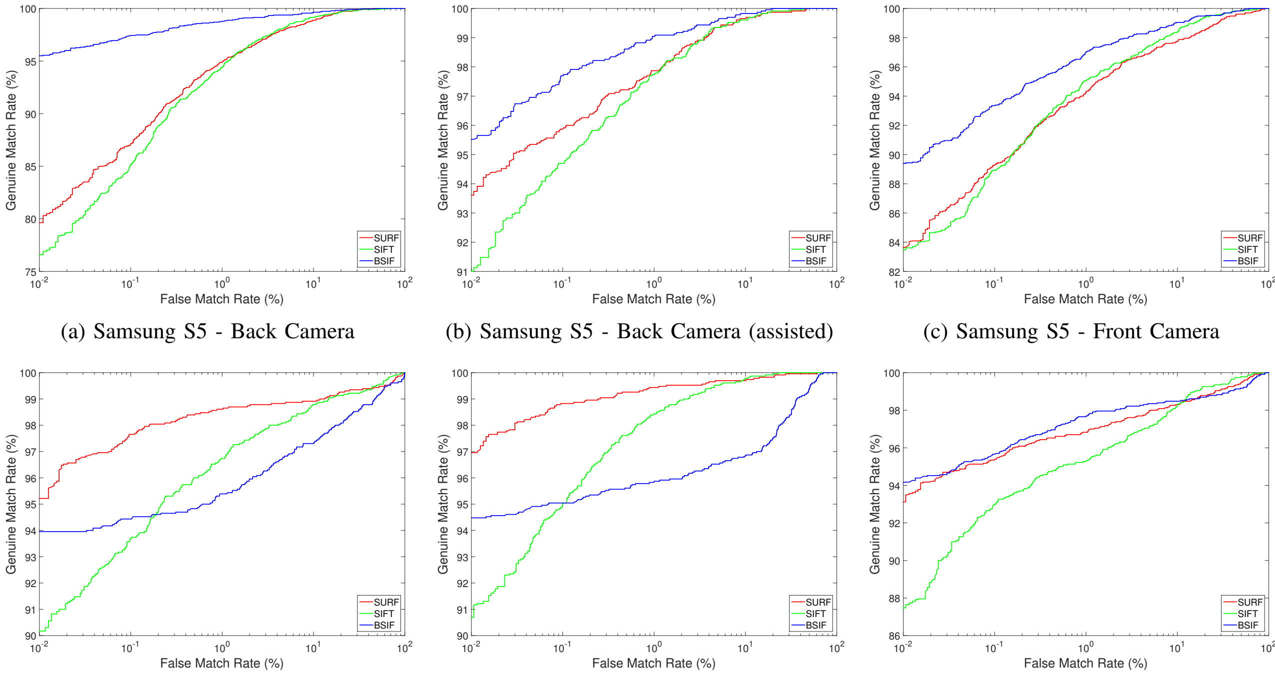
Earlier works have advocated the boost in verification performance when the comparison scores are fused [5], [6]. In order to evaluate the gain in verification performance as compared to the feature level fusion, we carry out score level fusion in this work. Given the score from BSIF comparison for face, left periocular and right periocular region denoted by f_{bc}, lp_{bc} and rp_{bc}, we obtain fused score c_b using the SUM rule as given by:

$$c_b = f_{bc} + lp_{bc} + rp_{bc}$$

In similar terms, the fused scores are obtained by SUM rule for the scores obtained using SIFT features denoted by subscript s and SURF features denoted by u. The fused scores for SIFT and SURF features can can be given as below:

$$c_s = f_{sc} + lp_{sc} + rp_{sc}$$
$$c_u = f_{uc} + lp_{uc} + rp_{uc}$$

Table V provides the verification performance obtained using SUM rule fusion with multi-biometric characteristics. As compared to the single biometric based performance or multi-biometric performance employing feature level fusion, score level fusion further improves the performance. The best GMR 95.5% is obtained for images obtained from back camera, 95.52% is obtained for images obtained from back camera in assisted mode and 89.39% is obtained for images obtained from front camera. In similar terms lower EER can be



(a) Samsung S5 - Back Camera (b) Samsung S5 - Back Camera (assisted) (c) Samsung S5 - Front Camera

(d) Samsung Note - Back Camera (e) Samsung Note - Back Camera (assisted) (f) Samsung Note - Front Camera

Fig. 3: ROC curves for various multi-biometric recognition approaches fusing face and periocular features with the score level fusion method on Samsung S5 and Samsung Note

TABLE V: Biometric performance in terms of Genuine Match Rate and Equal Error Rate for multi-biometrics using score level fusion.

Camera	Feature	Samsung S5		Samsung Note 10.1	
		GMR@FMR=0.01%	EER	GMR@FMR=0.01%	EER
Back	SIFT	76.58	2.75	90.17	2.34
	SURF	79.64	2.91	**95.21**	**1.30**
	BSIF	**95.50**	**1.13**	93.56	3.52
Back Assisted	SIFT	91	1.67	90.69	1.31
	SURF	93.60	1.67	**96.95**	**0.69**
	BSIF	**95.52**	**0.95**	94.47	3.64
Front	SIFT	83.65	3.38	93.13	2.55
	SURF	83.47	3.25	87.47	3.26
	BSIF	**89.39**	**2.26**	**94.17**	**2.04**

observed for verification based on score level fusion. Further, Figure 3 presents the ROC for the SUM score level fusion using face, left periocular and right periocular features for both devices.

VIII. Conclusion

Smartphone authentication for various secure transactions are gaining importance these days. Employing the smartphone cameras, one can use face or periocular based biometrics for authentication. In this work we have explored multi-biometric verification using both feature and score level fusion. An extensive set of experiments conducted on a database consisting of 46 subjects has shown the robust performance in terms of verification accuracy. The best unimodal verification rate was obtained for face based authentication with GMR of 94.39% for Samsung S5 smartphone and GMR of 96.65% for Samsung Note tablet. Availability of three modalities such as face, right periocular and left periocular has motivated us to perform feature level and score level fusion. The best verification rate of 94.91% GMR is obtained for smartphone and GMR of 96.58% is obtained for tablet. The nominal improvement in verification can be observed due to feature level fusion. Further, in this work, we have also explored comparison score level fusion using the scores obtained from face, left and right periocular features. The best GMR of 95.52% if obtained for smartphone and 96.65% GMR is obtained for tablet. The score level fusion has boosted the performance by more than 1% in GMR. It has to be noted that the EER has significantly decreased $(2-3\%)$ in the case of score level fusion as compared to EER obtained when unimodal biometrics is used for authentication.

Acknowledgements

The authors wish to express thanks to Morpho (Safran Group) for supporting this work, and in particular to Morpho Research & Technology team for the fruitful technical and scientific exchanges related to this particular work.

References

[1] C. Stein, C. Nickel, and C. Busch, "Fingerphoto recognition with smartphone cameras," in *Biometrics Special Interest Group (BIOSIG), 2012 BIOSIG-Proceedings of the International Conference of the*. IEEE, 2012, pp. 1–12.

[2] R. Raghavendra, C. Busch, and B. Yang, "Scaling-robust fingerprint verification with smartphone camera in real-life scenarios," in *Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on*. IEEE, 2013, pp. 1–8.

[3] M. De Marsico, C. Galdi, M. Nappi, and D. Riccio, "Firme: Face and iris recognition for mobile engagement," *Image and Vision Computing*, 2014.

[4] Q. Tao and R. Veldhuis, "Biometric authentication for a mobile personal device," in *Mobile and Ubiquitous Systems-Workshops, 2006. 3rd Annual International Conference on*. IEEE, 2006, pp. 1–3.

[5] Kiran B. Raja, R. Raghavendra, M. Stokkenes, and C. Busch, "Smartphone authentication system using periocular biometrics," in *2014 International Conference on Biometrics Special Interest Group*. IEEE, 2014, pp. 27–38.

[6] Kiran B. Raja, R. Raghavendra, M. Stokkenes, and C. Busch, "Multimodal authentication system for smartphones using face, iris and periocular," in *IEEE International Conf. Biometrics, Phuket, Thailand*, 2015.

[7] D. G. Lowe, "Object recognition from local scale-invariant features," in *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, vol. 2. Ieee, 1999, pp. 1150–1157.

[8] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *Computer Vision–ECCV 2006*. Springer, 2006, pp. 404–417.

[9] J. Kannala and E. Rahtu, "Bsif: Binarized statistical image features," in *Pattern Recognition (ICPR), 2012 21st International Conference on*. IEEE, 2012, pp. 1363–1366.

[10] M. Bicego, A. Lagorio, E. Grosso, and M. Tistarelli, "On the use of sift features for face authentication," in *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW'06. Conference on*. IEEE, 2006, pp. 35–35.

[11] U. Park, R. Jillela, A. Ross, and A. K. Jain, "Periocular biometrics in the visible spectrum," *Information Forensics and Security, IEEE Transactions on*, vol. 6, no. 1, pp. 96–106, 2011.

[12] J. Xu, M. Cha, J. L. Heyman, S. Venugopalan, R. Abiantun, and M. Savvides, "Robust local binary pattern feature sets for periocular biometric identification," in *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*. IEEE, 2010, pp. 1–8.

[13] P. Dreuw, P. Steingrube, H. Hanselmann, H. Ney, and G. Aachen, "Surf-face: Face recognition under viewpoint consistency constraints." in *BMVC*, 2009, pp. 1–11.

[14] Kiran B. Raja, R. Raghavendra, and C. Busch, "Binarized statistical features for improved iris and periocular recognition in visible spectrum," in *2nd International Workshop on Biometrics and Forensics, Malta*. IEEE, 2014.

[15] S.-H. Cha and S. N. Srihari, "On measuring the distance between histograms," *Pattern Recognition*, vol. 35, no. 6, pp. 1355–1370, 2002.

[16] M. Muja and D. G. Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration." in *VISAPP (1)*, 2009, pp. 331–340.

[17] D. G. R. Bradski and A. Kaehler, *Learning Opencv, 1st Edition*, 1st ed. O'Reilly Media, Inc., 2008.

[18] ISO/IEC 19795-1:2006, *ISO/IEC TC JTC1 SC37 Biometrics, Information Technology – Biometric Performance Testing and Reporting – Part 1: Principles and Framework*, International Organization for Standardization and International Electrotechnical Committee, Mar.