International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:11, No:7, 2017

Fusion of Shape and Texture for Unconstrained Periocular Authentication

D. R. Ambika, K. R. Radhika, D. Seshachalam

***Abstract*—**Unconstrained authentication is an important component for personal automated systems and human-computer interfaces. Existing solutions mostly use face as the primary object of analysis. The performance of face-based systems is largely determined by the extent of deformation caused in the facial region and amount of useful information available in occluded face images. Periocular region is a useful portion of face with discriminative ability coupled with resistance to deformation. A reliable portion of periocular area is available for occluded images. The present work demonstrates that joint representation of periocular texture and periocular structure provides an effective expression and poses invariant representation. The proposed methodology provides an effective and compact description of periocular texture and shape. The method is tested over four benchmark datasets exhibiting varied acquisition conditions.

***Keywords*—**Periocular authentication, Zernike moments, LBPV, shape and texture fusion.

I. Introduction

AUTHENTICATION of an individual under naturalistic states of daily life is a building block for most personalized automated systems. Commercial applications including residential security, voter verification and banking using ATMs require verification of an identity in an affective and pose variant environment. Development of an affect and pose resistant biometric systems is usually formulated by using facial features. Adequate facial representation is central to achieve reliable authentication. Affective expressions for instance happy, anger, surprise, disgust and sadness cause significant occlusions in the lower part of the face, specifically in the mouth and nose region, as illustrated in Fig. 1 (a). The resulting deformations cause a significant impact on the performance of the system. Compensation, normalization and image alignment algorithms need to be executed to incorporate deformation and pose invariance. An alternate solution is to utilize a more robust portion of face for authentication.

Periocular region, a portion of face around the eye, is currently drawing significant attention towards development of authentication systems in challenging environments [1]-[5]. The region is particularly useful when portions of faces are occluded. The region is explored by researchers for challenging iris and face images [6], [7]. The sample images in Fig. 1 (b) show that representable periocular information is available under pose variant condition. Enormous work is carried out to investigate the performance of periocular verification in an unconstrained environment, exhibiting variations in illumination, pose and occlusions [8]-[11]. Barroso et al. studied the impact of facial expressions in the periocular region for biometric applications [12]. The paper reports that multiple sample enrolment is useful to improve the performance of expression variant authentication. Smereka et al. proposed Periocular Probabilistic Deformation Models (PPDM) to approximate periocular distortions using local patch level spatial translations [13]. Bakshi et al. proposed a Phase-Intensive Local Pattern method of feature extraction for periocular images exhibiting variation in pose and illumination [14]. PCA coefficients of LBP features are explored in an unconstrained environment by Uzzair et al. [8]. A comparison of iris and periocular regions for a smartphone environment using Binarized-Statistical-Image-Features (BSIF) and Sparse-Reconstruction-Classifier (SRC) is explored in the literature. The results report that periocular images outperform iris images with a lowest EER of 4.17% [15]. Fusion of face and periocular information is investigated for a smartphone environment using Scale Invariant Feature Transform (SIFT), Speeded Up Robust Features (SURF) and BSIF. A verification rate of 94.91% is attained using the fusion scheme [16]. A robust method of shape representation is proposed for periocular images using FEM based Laplace-Beltrami eigenvalues [17].

In this work, joint representation of texture and shape characteristics of the periocular region is proposed for verification in an unconstrained environment. The periocular region is associated with both structural and textural information. Fig. 2 shows the dominant features encompassing the region. Eye folding, eyelashes wrinkles and fine lines contribute as periocular descriptors. In addition to texture, characterizing shape of eyebrow, eye folding, eye socket and shape of eye near tear duct assists in constructing robust discriminative periocular feature vectors. In order to take the benefit of both the types of characteristics effectively, two feature extraction schemes are implemented in the current work to perform texture and shape characterization independently. The textural components are computed locally using the LBP variance (LBPV) method proposed by Zhenhua et al. [18]. The LBPV algorithm incorporates local contrast information during representation of local texture and achieves rotation invariant feature extraction. Periocular shape representation is formulated using computation of Zernike

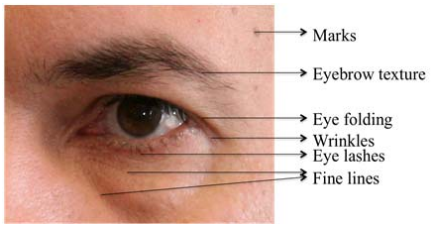
D. R. Ambika is currently working as Assistant Professor in the Department of Electronics and Communication Engineering, BMSCE, India (e-mail: amusatish@gmail.com).
K. R. Radhika is currently working as Professor and Head in the Department of Information Science Engineering, BMSCE, India.
D. Seshachalam is currently working as Professor and Head in the Department of Electronics and Communication Engineering, BMSCE, India.

moments [19]. Zernike moments are useful to describe image discontinuities at various scales and directions. The orthogonal properties of Zernike moments are very useful for shape classification unlike geometric moments [20]. The constructed texture and shape periocular features are fused at the feature level. Support Vector Machines (SVM) classifier is subsequently used to classify the identity of an individual as either genuine or an imposter. Extensive experiments are carried out using the proposed methodology for four uncontrolled acquisition environments, namely, varied facial expression, pose, camera standoff distances and illumination. The performance of the system is evaluated using four benchmark databases: UBIPr dataset, CMU AMP facial expression dataset, Cross-eyed periocular dataset under both visible spectrum and NIR spectrum and UBIRISv2 dataset. The rest of the paper is organized as follows: Section II presents the details of the proposed methodology. Section III discusses the experimental details and results. Section IV concludes the paper.

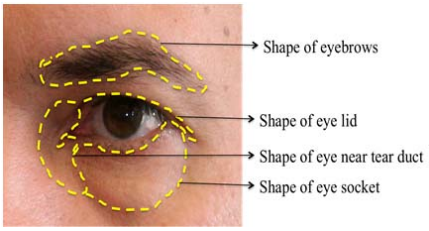
(a)

(b)

Fig. 1 (a) Samples of affective facial images from CMU AMP dataset (b) Samples of pose varied periocular images from the UBIPr dataset



(a)



(b)

Fig. 2 Illustration of (a) texture and (b) shape features present in the periocular region

II. Proposed Methodology

The proposed methodology for unconstrained periocular authentication is performed in four different stages: segmentation of periocular region, periocular texture feature extraction, periocular shape feature extraction and classification using fused features.

A. Segmentation of Periocular Region

Images available from the benchmark datasets include different sizes of the periocular area for each subject. In order to obtain a uniform sized periocular region, an automated segmentation method based on Daugman's integro-differential operator is implemented over periocular datasets [21]. The limbic boundary of iris is determined using the integro-differential operator defined in (1). The operation performs the circular edge detection, controlled by the radius, r, and the center of the detecting circle, (x_0, y_0).

$$\max(r, x_{0,} y_0) \left| G_\sigma(r) * \frac{\partial}{\partial r \oint_{r,x_0,y_0}} \frac{I(x,y)}{2\pi r} ds \right| \quad (1)$$

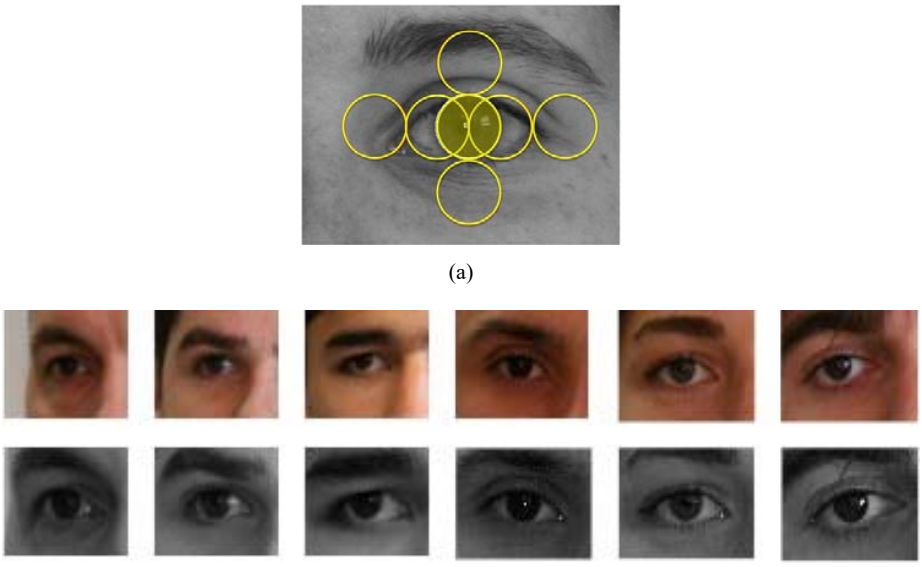
$\boldsymbol{G_\sigma(r)} = \frac{1}{\sqrt{2\pi\sigma}} \boldsymbol{exp}^{\frac{-(r-r_0)^2}{2\sigma^2}}$ represents radial Gaussian with center $\boldsymbol{r_0}$ and standard deviation σ. Convolving the periocular image, I(x, y), with Gaussian operator smoothens the image and highlights edge information. Circular integral is normalized with respect to its perimeter by dividing with a factor, 2πr. On determining the iris diameter, the width and height of periocular region is defined using (2). The corresponding definition of periocular region bounded by eyebrows above and cheekbone below is represented in Figs. 3 (a) and (b) shows samples of automatically segmented periocular region for UBIRISv2 and UBIPr dataset.

$$\begin{aligned} periocular_{width} &= 4 \times iris_{diameter} \\ periocular_{height} &= 3 \times iris_{diameter} \end{aligned} \quad (2)$$

B. Periocular Texture and Shape Feature Extraction

Feature extraction is performed over segmented periocular regions to extract texture and shape characteristics useful for identity verification. Texture features are extracted using LBPV method. LBPV combines local variance component

with LBP patterns, as shown in Fig. 4, and is coupled with the ability of rotation invariance. Initially, the input image is normalized to remove the global intensity using the function given in (3) for an image size m × n with constants a and b.



(a)

(b)

Fig. 3 (a) Localization of iris region (b) Subsequent segmentation of periocular region

Uniform LBPV patterns with discontinuities ≤ 2, are extracted over local neighborhoods of P sampling points on a circle of radius R. Variance values are computed for P sampling points from each local region and accumulated into the LBP bin to obtain LBPV features. (4) is used to compute the joint histogram of local binary pattern and local variance values [18]. The variance value computed from (5) is used as an adaptive weight to determine the joint histogram using (6):

$$\mathrm{LBPV}_{P,R}(k) = \sum_{i=1}^{M}\sum_{j=1}^{N} w(\mathrm{LBP}_{P,R}(i,j),k) \quad (4)$$

$$\mathrm{VAR}_{P,R} = \frac{1}{P}\sum_{p=0}^{P-1}(i_p - u)^2, u = \frac{1}{P}\sum_{p=0}^{P-1}(i_p) \quad (5)$$

$$w(\mathrm{LBP}_{P,R}(i,j),k) = \begin{cases} \mathrm{VAR}_{P,R} \text{ if } \mathrm{LBP}_{P,R}(i,j) = k \\ 0 \text{ otherwise} \end{cases} \quad (6)$$

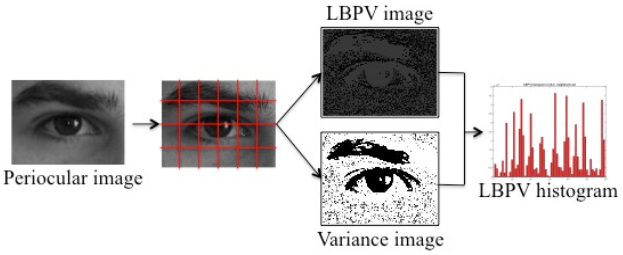


Fig. 4 LBPV block diagram

Shape characterization of the periocular region is performed using Zernike moments. Zernike moments are rotational invariant yet, dependent on translation and scaling. The variation is compensated by translating the centroid of periocular image to center of region of interest. A square window is placed such that its size is maximized without distortions and is used as the periocular template for the experiment. Radius is co-scaled by average radius value, determined using normalized radial length vector [22]. Zernike moments, $z_{nm}(\rho)$, are computed over the preprocessed image using (7). $z_{nm}(\rho)$ represents the projection of periocular image function, $f(x,y)$ onto the orthogonal basis functions. The orthogonal polynomials on a unit disk are defined using (8), where the radial polynomial is given by (9). A twelfth order Zernike basis function is used to compute 49 Zernike moments as periocular shape descriptors. Fig. 5 shows reconstruction of a periocular image using Zernike moments up to an order 12. The figure shows that a twelfth order Zernike basis function is able to describe the important shape characteristics of the eye, eyebrows and the socket near the tear duct most appropriately. The magnitude components of the Zernike moments are constructed as feature vectors. A feature-level fusion of 59-bin LBPV texture histogram with 49-element shape vector is performed to construct joint representation of periocular texture and shape.

$$z_{nm}(\rho) = \frac{n+1}{\pi}\sum_x\sum_y f(x,y)V^*_{nm}(\rho,\theta)$$
$$x^2 + y^2 \leq 1, V^*_{nm}(\rho,\theta) = V^*_{n,-m}(\rho,\theta) \quad (7)$$

$$V^*_{nm}(\rho,\theta) = R_{nm}(\rho)e^{jm\theta} \quad (8)$$

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{n-|m|}{2}} \frac{(-1)^s[(n-s)!]\rho^{n-2s}}{s!(\frac{n+|m|}{2}-s)!(\frac{n-|m|}{2}-s)!} \quad (9)$$

Fig. 5 Reconstructed periocular images using Zernike moments of order up to 12

C. Classification

Classifying the identity of an individual as either genuine or an imposter is attained using the Support Vector Machine (SVM) classifier. SVM is a powerful, supervised learning algorithm. The method is extensively used for binary classification problems. The algorithm determines a decision surface known as the hyper-plane that separates the data into two classes [23]. Input to SVM algorithm is a set of labeled training data, $\{x_i, y_i\}$. Training data is represented as a vector x_i with n number of observations, where i represents the index of the subject. $y_i = -1$ or $+1$ is the label of each data x_i. The separating hyperplane is determined using (10). w is normal to hyperplane and $|b|/w$ is the perpendicular distance from hyperplane to origin.

$$x_i \cdot w + b \geq +1 \text{ for } y_i = +1 \qquad (10)$$
$$x_i \cdot w + b \leq -1 \text{ for } y_i = -1$$

The classifier is trained using shape and texture feature vectors generated over the training dataset. During testing, the feature vectors of probe images are given to the classifier and the output class labels associated with the identity are obtained.

III. Experiments and Results

A. Databases

Four benchmark datasets exhibiting varied pose, expression, illumination and presence of occlusions are used to validate the proposed method of feature extraction.

The UBIPr dataset includes periocular images acquired under an uncontrolled imaging environment. The dataset embeds variability factors in subject-camera distances, levels of illumination, pose and occlusions [24]. The dataset is acquired using Canon EOS 5D camera over 261 subjects in two different sessions. A total of 10,950 periocular images are available with varied image resolutions in '.bmp' format.

The CMU AMP face expression dataset consists of facial images of 13 subjects, each with 75 images. The images are collected in the same lighting condition and allowed only changes in human expressions. The expressions included naturalistic affective behavior: neutral, anger, happy, surprise and disgust images. The images are of 64×64 resolution [25]. Periocular regions corresponding to both left and right eyes are used as the region of interest. Fig. 6 shows a sample periocular images used for the experiment.

Fig. 6 Samples of periocular images enclosing regions around both eyes of CMU AMP face images

Fig. 7 Samples of NIR and visible images of cross-eyed dataset

Cross-eyed periocular dataset consists of eye images from 120 subjects, each with eight images. Both visible and NIR spectrum images with resolution 900×800 are included in the dataset. The images are captured with a custom made dual spectrum imaging sensor, which acquires near infrared and color images synchronously. Any possible influence of the pupil is eliminated by placing an elliptical mask of neutral color. The images are acquired from varied distances between 1.5 meters and 2 meters. A realistic indoor environment with realistic an illumination condition is considered during the acquisition [26]. The dataset exhibits varied eye colors. The samples of cross-eyed periocular images captured in the both visible and NIR spectrum are shown in Fig. 7.

The UBIRIS.v2 database facilitates evaluation of biometric algorithms for non-ideal imaging conditions [27]. The dataset includes 11,102 images of over 261 subjects imaged on-the-move, at-a-distance in visible wavelength spectrum using Canon EOS 5D camera. The images exhibit challenges in

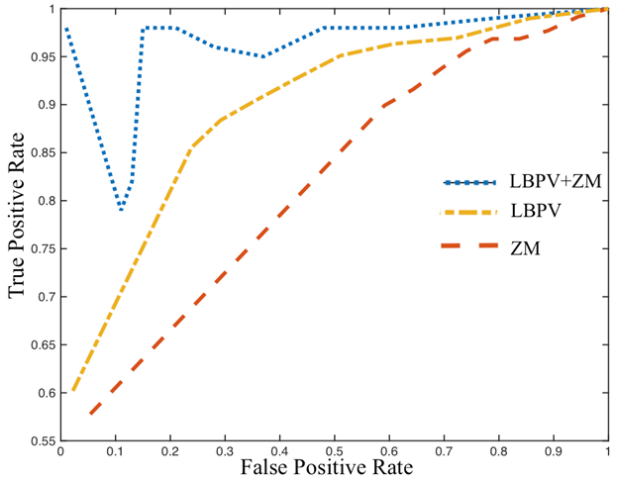
scale, occlusion and illumination. Occlusions include flapped eyelids, eyelashes, hair and spectacles. Imaging at variable lighting conditions emulates an unconstrained environment. Five images enclosing sufficient periocular region is used as the experimental dataset for each subject.

Periocular images of each dataset are separated as training and testing images. The UBIPr dataset provides 15 images with varied pose and distances for each subject. Five images exhibiting different poses and distances along with a frontal view which is used as the training set for each subject. The remaining 10 images are grouped as testing samples. The CMU AMP dataset provides more than 10 samples for each expression of every subject. Four samples of each expression: neutral, happy, anger and surprise are used as the training images and the remaining 51 images are used as the probe. The cross-eyed dataset provides eight NIR and eight visible wavelength images for each subject. Three samples from each spectrum are grouped as the training set for each subject. Testing data includes the other five images for each subject. Five images with useful periocular region in UBIRISv2 dataset are divided as two training images and three testing images. Each training image is compared against every probe image. As a result, a subject with five training and 10 probe sets, results in 50 genuine scores. Imposter scores for each subject will correspond to $(\text{total number of subjects} - 1) \times 50$. The results are reported in terms of ROC curves, plot of false positive rate versus true positive rate (Genuine Match Rate (GMR)).

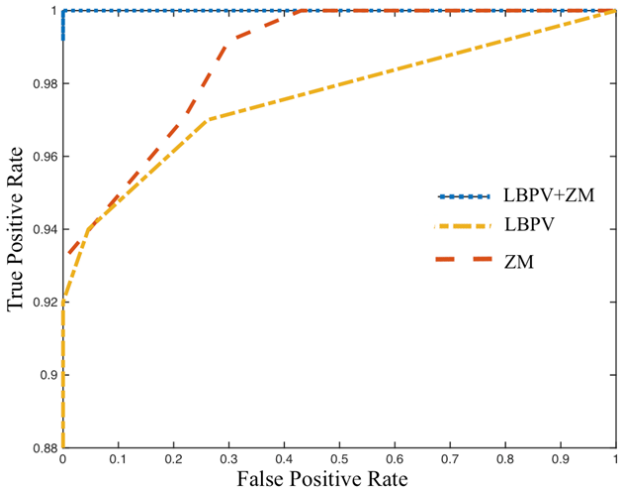
TABLE I
PERIOCULAR VERIFICATION PERFORMANCE IN TERMS OF GENUINE MATCH RATE FOR DIFFERENT ACQUISITION CONDITIONS

Dataset	Variability factors	Total images	Test images	Genuine Match Rate 0.01% FAR
UBIPr	Pose, occlusions, camera standoff distances, illumination	800	500	97.8%
CMU AMP	Facial expressions	975	663	100.0%
Cross-eyed visible spectrum	Camera standoff distances, illumination, eye color	960	600	99.1%
Cross-eyed NIR spectrum	Camera standoff distances, illumination	960	600	87.8%
UBIRISv2	Camera standoff distances, illumination, occlusions	1,305	783	85.0%

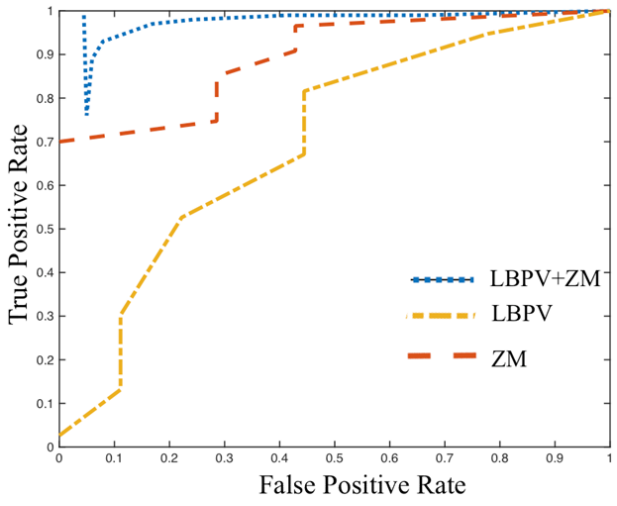
Table I lists the performance of the proposed methodology for different datasets. The table indicates that feature-level fusion in the proposed method presents impressive results for periocular images under an unconstrained environment. Figs. 8 and 9 show the ROC curves obtained for different datasets. A set of ROC curves are plotted for each dataset indicating three scenarios, namely - single shape feature representation using Zernike moments, single texture feature representation using the LBPV method, and the joint representation of both the periocular shape and texture using the presented approach.



(a)

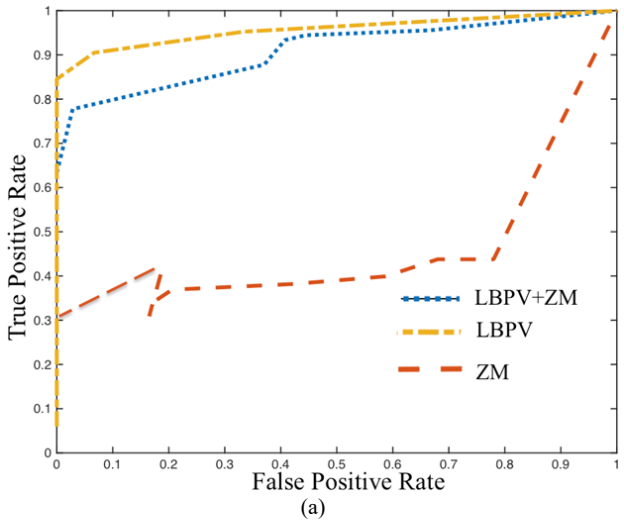


(b)

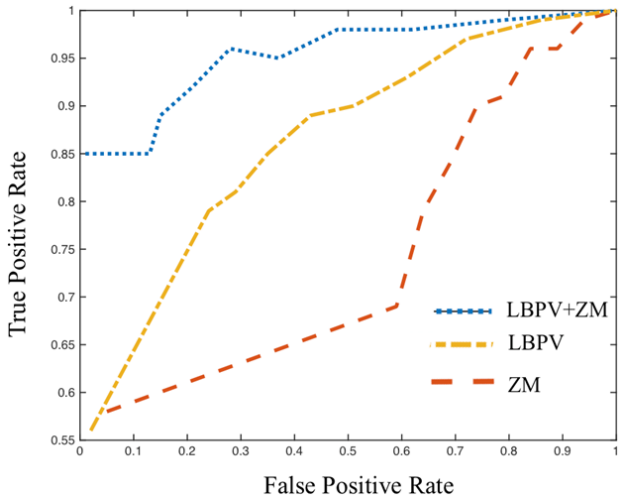


(c)

Fig. 8 ROC curves for (a) UBIPR dataset (b) CMU AMP facial dataset (c) cross-eyed spectrum dataset



(a)



(b)

Fig. 9 ROC curves for different datasets: (a) Cross-eyed NIR dataset (b) UBIRISv2 dataset

From the results, it is observed that description of periocular region using texture and structural attributes generates more discriminative features than using only one of the two attributes. A significant enhancement is observed in the authentication performance for pose variant images of UBIPr dataset from Fig. 8 (a). An authentication accuracy of 97.8% is obtained using the joint representation, while single feature representation achieves <60% accuracy. For expression of variant images, an improvement by 6% at ≈< 0.1%. FAR is attained as shown in the ROC curves of Fig. 8 (b). Although an elliptical mask is employed to mask the sclera and iris information of the cross-eyed periocular images, sufficient shape characteristics around the eye is encapsulated by Zernike moments. Fig. 8 (c) shows that Zernike moment results outperform LBPV results. Coupling the shape features with texture largely increases the performance for the visible spectrum periocular images more than using single feature extraction. However, results for the NIR images of the cross-eyed dataset, shown in Fig. 9 (a), indicates that the LBPV texture extraction method works better than incorporating shape description. Further, a significant upsurge of authentication accuracy is witnessed for UBIRISv2 dataset. Fig. 9 (b) shows that the proposed methodology achieves an increase in accuracy by ≈30%.

IV. Conclusions

The present work proposes to fuse texture and shape characteristics of periocular region to construct a robust feature vector resistant to variations in acquisition environment. The experimental results show that the fusion mechanism has outperformed the performance of single-feature-based authentication using either only texture or shape. Combining LBPV features with Zernike moments achieves significant enhancement in authenticating an individual under varied facial expressions, pose and camera standoff distances. Experiments carried out reveals that an accuracy of 100% GMR is attained for expression variant images, emphasizing the joint contribution of shape and texture for appropriate periocular region representation.

Acknowledgement

The authors would like to thank Hugo Proenca, Chandrashekar Padole and Ana Sequeira for facilitating the provision of datasets used in the current work.

References

[1] R. Raghavendra, K. B. Raja, and C. Busch, “On comparison score fusion of deep autoencoders and relaxed collaborative representation for smartphone based accurate periocular verification,” 19th International conference on Information Fusion (FUSION), pp. 2221 – 2228, 2016.

[2] J. M. Smereka, B. V. K. V. Kumar, and A. Rodriguez, “Selecting discriminative regions for periocular verification,” in IEEE International Conference on Identity, Security and Behavior Analysis (ISBA), 2016.

[3] G. Santos, E. Grancho, M. V. Bernardo, and P. T. Fiadeiro, “Fusing iris and periocular information for cross-sensor recognition,” Pattern Recognition Letters, vol. 57, pp. 52 – 59, 2014.

[4] F. Alonso-Fernandez, A. Mikaelyan, and J. Bigun, “Comparison and fusion of multiple iris and periocular matchers using near-infrared and visible images,” IWBF 2015 – 3rd International Workshop on Biometrics and Forensics, March 2015.

[5] S. Bakshi, P. K. Sa, and B. Majhi, “Phase intensive global pattern for periocular recognition,” Annual IEEE India Conference (INDICON), pp. 1 – 5, December 2014.

[6] M. V. Samarth Bharadwaj, Himanshu. S. Bhatt and R. Singh, “Periocular biometrics: When iris recognition fails,” IEEE Transactions, 2010.

[7] R. Jillela and A. Ross, “Matching face against iris images using periocular information,” Proc Intl Conf Image Processing, ICIP 2014, 2014.

[8] M.Uzair, A. Mahmood, A. Mian, and C. McDonald, “Periocularregion-based person identification in the visible, infrared and hyperspectral imagery,” Neurocomputing, Elsevier publication, vol. 149, pp. 854 – 867, August 2014.

[9] J.-X. F. and M. Savvides, “Unconstrained periocular biometric acquisition and recognition using cots ptz camera for uncooperative and non-cooperative subjects,” 2012.

[10] M. Castrillo´n-Santana, J. Lorenzo-Navarro, and E. Ramo´n-Balmaseda, “On using periocular biometric for gender classification in the wild,” Pattern Recognition Letters, Elsevier, pp. 1 – 9, 2015.

[11] J. C. Monteiro and J. S. Cardoso, “Periocular recognition under unconstrained settings with universal background models,” Proceedings of the International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS), 2015.

[12] E. Barroso, G. Santos, L. Cardoso, C. Padole, and H. Proenc ̧a, "Periocular recognition: how much facial expressions affect performance?" Journal of Pattern Analysis and Applications, vol. 19, no. 2, pp. 517 – 530, 2016.
[13] J. M. Smereka, V. N. Boddeti, and B. V. Kumar, "Probabilistic deformation models for challenging periocular image verification," IEEE Transactions on Information Forensics and Security, vol. 10, no. 9, pp. 1 – 17, 2015.
[14] S. Bakshi, P. K. Sa, and B. Majhi, "Anovelphase-intensive local pattern for periocular recognition under visible spectrum," Biocybernetics and Biomedical Engineering, vol. 35, no. 1, pp. 30–44, 2015.
[15] K. B. Raja, R. Raghavendra, and C. Busch, "Empirical evaluation of visible spectrum iris versus periocular recognition in unconstrained scenario on smartphones," Annual Summit and Conference (APSIPA) Asia-Pacific Signal and Information Processing Association, 2014.
[16] K. B. Raja, R. Raghavendra, M. Stokkenes, and C. Busch, "Fusion of face and periocular information for improved authentication on smartphones," 18th IEEE International Conference on Information Fusion (FUSION), 2015.
[17] D R Ambika, Radhika K R and D. Seshachalam, "Periocular authentication based on fem using laplace–beltrami eigenvalues," Pattern Recognition journal, vol. 50, no. C, pp. 178 – 194, February 2016.
[18] D. Z. Zhenhua Guo, Lei Zhang, "Rotation invariant texture classification using lbp variance (lbpv) with global matching," Pattern Recognition journal, vol. 43, pp. 706 – 719, 2010.
[19] K. A and Y. H. Hong, "Invariant image recognition by zernike moments," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, no. 5, pp. 489 – 497, 1990.
[20] E. Sariyanidi, V. Dagli, S. C. Tek, and M. Go ̈kmen, "Local zernike moments: A new representation for face recognition," 19th IEEE International conference on Image Processing, 2012.
[21] J. Daugman, "How iris recognition works," IEEE transactions on circuits and systems for Video technology, vol. 14, no. 1, pp. 21 – 30, 2004.
[22] A. Wiliem, V. K. Madasu, W. W. Boles, and P. K. Yarlagadda, "A feature based face recognition technique using zernike moments," in RNSA Security Technology Conference 2007, P. Mendis, J. Lai, E. Dawson, and H. Abbas, Eds. Australian Homeland Research Centre, 2007, pp. 341–355.
[23] B. E. Boser, I. M. Guyon, and V. Vapnik, "A training algorithm for optimal margin classifiers," pp. 144 – 152, 1992.
[24] C. N.Padole and H. Proenc ̧a, "Periocular recognition: Analysis of performance degradation factors." in ICB, 2012, pp. 439–445.
[25] R. Goh, L. Liu, X. Liu, and T. Chen, "The cmu face in action (fia) database," Proceeding AMFG 2005 Proceedings of the Second international conference on Analysis and Modelling of Faces and Gestures, pp. 255 – 263, 2005.
[26] A. F. Sequeira, L. Chen, J. Ferryman, F. Alonso-Fernandez, J. Bigun, K. B. Raja, R. Raghavendra, C. Busch, and P. Wild, "Cross-eyed - cross-spectral iris/periocular recognition database and competition," 15th International conference of the Biometrics Special Interest Group (BIOSIG 2016), 2016.
[27] H. Proenca, S. Filipe, R. R. Santos, J. Oliveira, and L. A. Alexandre, "The ubiris.v2: A database of visible wavelength iris images captured on- thepmove and at-a-distance," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 8, pp. 1529–1535, August 2010.