

Survey of Periocular Recognition Techniques

Joke A. Badejo¹, Adekunle A. Akinrinmade¹ and Emmanuel Adetiba^{1,2}

¹Department of Electrical and Information Engineering, Covenant University, Ota, Nigeria.

²Institute for Systems Science, Durban University of Technology, Durban, South Africa

Received 28 January 2019; Accepted 24 October 2019

Abstract

Periocular recognition has evolved over the years and has been shown to possess discriminative features for personal identification either as a stand-alone trait or when fused with other modalities such as face and iris, especially in unconstrained scenarios. It has a number of advantages, first being the region could be easily cropped from existing face images. Secondly, unlike the iris, its capturing process is less intrusive. This area of the face is easily captured with ease from the subject, for example by surveillance cameras. Thirdly, in crime scenes where masks hide the face or in cases where the subject's face is covered due to religious or cultural beliefs, the periocular region could still be captured providing useful details, that is, it is robust to face occlusion and is least affected by expression change. This survey relates the various techniques employed for periocular recognition at different stages: segmentation approaches, image preprocessing methods, feature extraction and matching algorithms. This survey is meant to facilitate a quick grasp of the development in this area for interested students, researchers as well as enthusiasts in the field of biometrics or any related application area. Information about the various databases used for performance evaluation of these techniques as well as the performance indicators is also provided.

Keywords: convolution neural networks, periocular biometrics recognition, texture descriptors

1. Introduction

The periocular region of the face is the immediate surrounding of the eye. It includes the eyebrows, upper and lower eyelids, sclera, skin texture and color, blood vessels, tear duct, inner corner, outer corner, upper and lower eye folds and eye lashes, as shown in Fig.1 [1]. Periocular recognition, like similar biometric recognition systems, can be adopted as a means of personal authentication to ensure authorized access to sensitive information as required for smartphones [2] [3], for access control to sensitive units like switch rooms in organizations and in financial and payment applications where user identification is necessary for optimal security [4].

Periocular biometrics has proven to be an alternative modality to face and iris, offering discriminative features from a face region, as shown in Fig.1. The periocular biometric modality captures information from the skin texture, shape of eyebrow, density of eyelashes, structure of blood vessels, depth of eye socket, curvature of eye boundary and eye corners to distinguish a person from another [1] [5]. It has a number of advantages, first being the region could be easily cropped from existing face images. Secondly, unlike the iris, its capturing process is less intrusive. This area of the face is easily captured with ease from the subject, for example by surveillance cameras. Thirdly, in crime scenes where masks hide the face or in cases where the subject's face is covered due to religious or cultural beliefs, the periocular region could still be captured providing useful details, that is,

it is robust to face occlusion and is least affected by expression change [6].

The periocular region is sometimes used as a stand-alone trait or in fusion with the face or iris modalities for improved recognition performance [5]. The human face is a physiological trait that contains a number of features such as eyes, nose ears, lips, hairs, cheeks, chin, wrinkles and skin, which can be employed for recognition and as such attracted a lot of attention in research. However, traditional face recognition systems perform optimally for images captured under controlled conditions. Their performance degrades with images captured under challenging scenarios such as varying expression, pose, lightning conditions and occlusion [6].

On the other hand, most commercially deployed iris recognition systems work best with images captured at a close distance with special imaging devices such as the infrared sensors and under constrained scenarios. This makes iris recognition intrusive, requiring the high-level user cooperation, to reduce the difficulty of iris segmentation [2] [3] [7] [8].

A number of research has been carried out in periocular recognition with successful and promising results, with increasing interest [8]. Periocular biometrics has also been successfully employed for soft biometrics analysis such as gender and ethnicity classification with impressive results, even under unconstrained scenarios [9].

Periocular recognition systems consist of three sequential stages, as summarized in Fig.2. The first is the preprocessing stage, where the image captured by the sensor goes through a number of tuning usually by color conversion, filtering, illumination normalization, rotation, cropping, resizing, periocular region segmentation, division of the image into sub-regions etc. The second, which is the feature extraction

*E-mail address: joke.badejo@covenantuniversity.edu.ng

ISSN: 1791-2377 © 2019 School of Science, IHU. All rights reserved.

doi:10.25103/jestr.125.24

stage, is concerned with representing the discriminative features from the preprocessed images in a compact form. These features are usually extracted densely over the entire image or from specific key points in the image. The compact representation is passed into the third and final stage, for classification or matching of images.

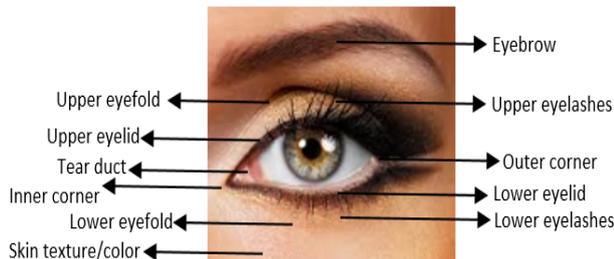


Fig. 1. Periocular image, with labelled important features.

This survey presents an overview of the various techniques adopted by researchers, within this current decade, for periocular recognition. This survey is meant to facilitate a quick grasp of the development in this area for interested students, researchers, enthusiasts in the field of biometrics or any related application area. The remainder of this paper covers a description of various periocular image databases in Section 2, periocular image preprocessing in Section 3, periocular image representation in Section 4, periocular template matching in Section 5. An extensive summary of periocular recognition research methodologies, performance indicators and image data adopted is reported in Tab. 1.

2. Databases for periocular recognition research

Unlike the face and iris where research has been more extensive, the periocular research is relatively newer [8] and so periocular databases are relatively few. Notwithstanding, a number of periocular experiments are carried out by extracting this region from publicly available face and iris databases. Some of the common databases employed in periocular research are discussed in this section. Refer to Tab. 1 for a comprehensive list.

Cross-Sensor Iris and Periocular (CSIP) database [2] contain 2,004 images from 50 subjects acquired under ten unique setups simulating various acquisition scenarios using the front and rear cameras of four different smartphones with and without flash. The mobile devices used were Sony Ericsson Xperia Arc S running Android 2.3.4, Apple iPhone 4 running iOS 7.1, THL W200 running Android 4.2.1 and Huawei Ideos X3 (U8510) running Android 2.3.3. The cameras have different resolutions which ranged from 640x480 to 3264x2448. The images are characterized by varying illuminations, rotation, chromatic distortions, defocus, off-angle and occlusion.

UBI periocular recognition (UBIPr) database [10] contains 10,950 unconstrained images (522 being periocular regions) captured from 261 persons, 104 of which were in two sessions in visible spectrum, varying in scale, head pose, illumination, occlusion, distance, eyeball movement and pigmentation. The camera distance varies from 4m to 8m at intervals of 1m while the resolutions vary from 501x401 to 1001x801 pixels. Images were stored in sRGB format with 54.4% of the subjects being males while 45.6% were females.

Face and Ocular Challenge Series (FOCS) database [11]

contain NIR images obtained from videos of subjects on-the-move. A total of 9,581 images were obtained from 136 subjects. The images are in different degrees of illumination, occlusion, and specular reflections. Majority of the images are of degraded quality due to sensor noise and out-of-focus blur.

Indian Institute of Technology (IITD) Multispectral Periocular (IMP) database [12] contains images in three spectrums. The first, NIR spectrum images, were captured with a close-up iris scanner. The second, visible spectrum images were captured at 1.3m with a digital camera while the night vision images were captured using a handy cam in night mode. It has a total of 620 images from 62 subjects.

BATH database [13] contains grayscale NIR images developed by the University of Bath, UK. It has three versions: Iris DB 400 (containing 16,000 images of 400 subjects), Iris DB 800 (containing 8000 images of 200 subjects) and Iris DB 1600 (containing 32,000 images of 800 subjects). The images were captured using IrisGuard AD-100 camera model with Dual-Eye autofocus under controlled conditions.

CASIA (IRISv3-Lamp) database [14] developed by Academy of Sciences, Beijing, China, contains 16,212 NIR images of 411 subjects at resolutions of 640x480. The images captured using OKI IRISPASS-h camera contain specular highlight noises and vary in illumination. CASIA-IrisV4 extends V3 and contains six subsets having over 54601 grayscale NIR images from 1,800 true subjects and 1,000 artificial subjects.

UBIRIS versions 1 and 2 [15] [16] were developed by the University of Beira, Portugal. It has two versions, version1 contains 1877 visible spectrum images of 241 subjects in RGB color model and resolutions of 800x600. Its version2 is similar but contains 11,102 unconstrained images of 261 subjects captured on-the-move, at varying distances and with non-orthogonal views at resolutions of 400x300 in sRGB format.

FERETv4 database was developed by National Institute of Standards and Technology (NIST) [17]. It contains 14,126 noisy unconstrained visible spectrum images from 1,191 subjects with different facial expressions and orientations at resolutions of 768x512, 384x256 and 192x128 in RGB color model. It was one of the four databases used in [1].

Mobile Iris CHallenge Evaluation (MICHE I and II) database [37] [43] contain visible spectrum eye images captured using different mobile devices. MICHE-I is the core of version II and contains a total of 3,148 fully annotated samples from 75 subjects captured using three mobile devices in unconstrained scenarios. The smart phones used were Galaxy Samsung IV (GS4) using Google Android Operating System at a resolution of 1080x1920, Phone5 (IP5) using Apple iOS Operating System at a resolution of 960x1280 and Galaxy Tablet II (GT2) at resolution of 640x480. Images are characterized by light reflections, occlusion, out-of-focus blur, off-axis gaze, variance in illumination and device-specific artifacts. MICHE-II has samples from 86 individuals [4].

CASIA 3D FV1 database [35] contains 4,624 scans from 123 subjects in unconstrained scenario (varied expression, pose and illumination) captured with non-contact 3D digitizer, Minolta Vivid 910. It contains both the 2D color image as well as the 3D facial model. The image resolution is 256x256.

Face Recognition Grand Challenge (FRGC) database [19] contains 44,278 high resolution (1200x1400) visible spectrum still face images from 568 subjects captured at

different recording sessions with varying expressions and illumination. It also contains images captured in controlled scenarios (controlled lighting condition, fixed distance from camera and neutral expression). The dataset includes gender and ethnicity tags for 466 of the individuals. The database was used in [9] [18].

Multiple Biometrics Grand Challenge (MBGC) v2 is a face database of both visible spectrum images and NIR video datasets. The visible spectrum dataset [20] consists of 91 subjects having 1 - 6 image sets per subject and between 23 - 352 images per set. Its NIR video dataset [9] was obtained from 114 individuals captured on-the-move at the rate of 15 frames per second with a resolution of 2048x2048 having 1 - 12 image sets per subject and between 6 - 48 images per set. The stored images extracted from the video are characterized by varying NIR illumination, occlusion, motion-blur, off-axis eye angles and sensor noise. Database was used in [18].

CMU Hyperspectral (CMU-H) database [33] consists of 764 videos of 54 subjects captured in 1 - 5 sessions. The images are of resolution 640x480 in 65 spectral bands varying from 450 - 1100nm at intervals of 10 nm using three halogen lamps. There was no occlusion and no variation in expression and pose. It was used in [6].

ND 2004-2005 iris image dataset, used in the work of [26]

has a total of 64,980 images from 356 different persons. Both left and right eyes giving 712 unique irises. There are 158 females and 198 males whose ages ranged between 18 and 75 years. Some of the images are characterized by occlusion, blur, off-axis view, artifacts introduced by use of contact and cosmetic lenses while some part of the eye were not captured in some of the image [32]. BioSec database [27] contains infrared iris images from both eyes of 200 subjects, each person was captured four times in two sessions leading to a total of 3200 database images. The 480x640 resolution images were captured using an LG Iris Access EOU3000 camera. Database was used in the experiments of [26]. BioSec database [27] contains infrared iris images from both eyes of 200 subjects, each person was captured four times in two sessions leading to a total of 3200 database images. The 480x640 resolution images were captured using an LG Iris Access EOU3000 camera. Database was used in the experiments of [26].

MobBIO [28] database contains 1,680 visible spectrum iris images from 105 individuals captured using Asus Eee Pad Transformer TE300T Tablet at a fixed distance and under two light conditions but varying conditions of occlusion and eye orientation. Each volunteer contributed 16 images (8 left and 8 right). Database was used in the experiments of [26].

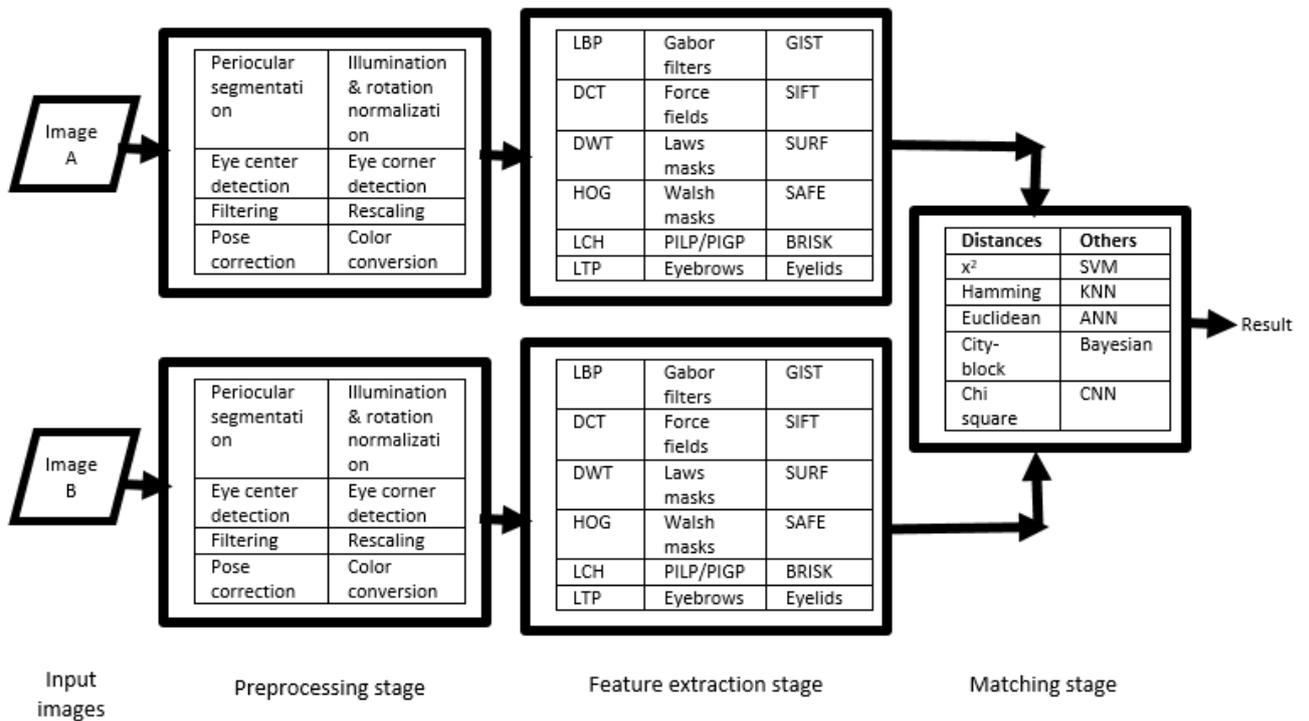


Fig. 2. Periocular recognition stages and relevant techniques.

Images of Groups (IoG) database [34] contains 28,231 facial images taken from 5,080 different subjects at different life stages grouped into seven labelled age groups in uncontrolled scenarios: people sitting, standing, lying down, varying occlusion and facial expressions. Many of the images are of low resolutions.

Cross-Spectral Iris/Periocular Dataset (Cross-Eyed DB) [38] was developed to facilitate periocular and iris recognition research across multiple spectrums, it contains 5,600 eye images captured from 175 subjects. Half of these were NIR images while the other half were VW (visible wavelength) images, each category contains half the left and half right side of the eyes. Each eye capture simultaneously took both NIR and VW spectrum of the image in one shot. Database was divided into 3 subsets: ocular, iris and periocular.

Visible Light Mobile Ocular Biometric (VISOB) dataset [41] consists of images from 550 subjects in the first visit and 290 subjects in the second visit captured under three lightning conditions: day light, dim and office light. Images were captured using three mobile phones: Iphone 5s, Oppo model N1 and Samsung Galaxy Note 4. Images were taken under unconstrained conditions of varying occlusion, blur, makeups, off gaze angles and illumination. It contains a total of 80,194 enrollment images and 77,942 validation images.

3. Periocular image segmentation and preprocessing methods

Most research experiments on periocular biometrics make use of face databases, especially in uncontrolled environment, the periocular region first needs to be cropped from the face region. There are two major ways to achieve this: manual segmentation and automatic segmentation. Some researchers manually segment the periocular region by taking advantage of the ground truth eye centers provided with the database as reference points [8] [9].

There are three main techniques usually employed in literature for automatic detection and/or segmentation of the periocular region [44]. The first method is to initially detect the face with a face detector such as the Viola-Jones (VJ) algorithm [45]. Here the overall performance of periocular

recognition depends on the accuracy of the face detector. This method was employed in [23] [46].

The second method is to directly detect the eye in the image either using the iris, pupil or eye corners. Pupil centers could also be detected by application of Haar features via weak classifiers followed by binarization and contour processing [6], eye corners were detected by use of Canny edge detectors, morphological operations and Harris corner detector [35]. In the iris segmentation approach, the periocular region is determined based on the position of the iris. Correlation filters are employed to detect the eye center [11] [47]. Eye detection could also be performed using Gabor filters [48] or from convolutions with 1D filters tuned to detect circular symmetries [26] [31].

Lastly, the periocular region could be detected using sub-face parts such as eyebrow, sclera or VJ sub-part detectors. The eyebrow could be detected directly from a face image using Local Eyebrow Active Shape Model (LE-ASM) followed by graph-cut based segmentation [24] [46]. The sclera region was detected after iris segmentation using the HSV/YCbCr color spaces [49] [50].

A common preprocessing step is division of the periocular region into sub-blocks in rows and columns for global feature analysis, 6x8 [10], 7x5 [2], multiple sub-block configurations were used in [4] [8]. Images are rescaled to suitable dimensions: 100x160 [18], 128x168 [3], 49x19 [8], 20x30 [6], 32x49 [40], image cropping [42] and image padding [39]. Conversion from one color space to another or to grayscale and color correction [2] [3] [9] [18]. Histogram normalization or equalization [9] [18]. Automatic exclusion of images in the database of poor quality using global or average intensity thresholding [9] [18]. Rotation correction by horizontal alignment using the eye centers or corners [6] [8] [23] [29] [30]. Correction of illumination variance or photometric normalization [2] [29] [30] [35]. Gaussian blurring applied to smoothen the variation across local pixel intensities [23]. Removal of short line edges in image, for example eyelashes by filtering using morphological operations or filtering with 1D rank filters [31] [35].

Table 1. Extensive Summary of Research Works on Periocular Biometric Recognition Techniques

Publication Year	Authors	Databases	Samples	Subjects	Image Preprocessing	Feature Extraction	Matching or Classification	Accuracy (min – max)
2010	Woodard et al [16]	FRGC [17] MBGC NIR [18]	44,278 NA	568 114	Scaling, grayscale/color space conversion, gray value thresholding	LBP, Red-green color histogram	City-block function & Bhattacharya coeff.	91% 87%
	Bharadwaj et al [19]	UBIRIS v2 [12]	11.102	261	Contrast normalization & image division into sub-blocks	CLBP and GIST	X ² distance	73.65%
2011	Dong et al [20]	FRGC [17] MBGC NIR [18]	44,278 NA	568 114	Manual eyebrow segmentation	Eyebrow shape	MD, LDA, SVM	75-97% 90-96%
	Park et al [21]	FRGC 2.0 [17]	44,278	568	Partitioning into sub-blocks and Gaussian blurring	HOG, LBP and SIFT	Euclidean distance	87.32%
2012	Padole et al [9]	UBIPr [9]	10,950	261	Cropping, sub-block division	HOG, LBP, SIFT	Least square, ANN	NA
	Lyle et al [8]	FRGC v2 [17] MBGC NIR [18]	44,278 NA	568 114	Color space/grayscale conversion, histogram normalization & global intensity thresholding	LBP, HOG, DCT & LCH	SVM & ANN	89-97.3%
2014	Le et al [22]	AR [10] MBGC [10, 18]	> 4,000 3,482	126 437	Grayscale conversion	Eyebrow shape	Chi-square & Cosine distances	54.4-76.0% 85.0-71.3%

	Alonso-Fernandez et al [23]	BioSec [24] MobBIO [25]	3,200 1,680	200 105	Symmetry filters	Gabor filters	X ² distance	NA
	Xu, et al [26]	FRGC [17]	44,278	568	eye alignment, eye coordinate normalization, illumination processing & cropping	SIFT, SURF, LBP, DCT, DWT, LoG, FFT, Gabor filters, Walsh & Laws' masks.	Cosine, Manhattan & Euclidean distances	53.2-61.2%
	Nie et al [27]	UBIPr [9]	10,950	261	Rotation/illumination normalization & retinex image enhancement	DSIFT & CRBM features	CNN,ITML/SVM	50.1% EER=0.064
2015	Alonso-Fernandez et al [28]	BioSec [24] CASIAv3 [6] IIT Delhi v1 [10] MobBIO [25] UBIRIS v2 [12] ND-0405 [29]	3,200 16,212 2,240 1,680 11,102 64,980	200 411 224 105 261 356	Symmetry filters	Gabor filters	X ² distance	NA
	Bakshi et al [3]	BATH [3] CASIA v3 [6] UBIRIS v2 [12] FERET v4 [13]	32,000 16,212 11,102 14,126	800 411 261 1,191	NA	PILP	Euclidean distance	92.47- 99.87%
	Santos et al [4]	CSIP [4]	2,004	50	Noise correction, device-specific color correction, grayscale conversion & sub-blocks division	HOG, LBP, ULBP, SIFT and GIST	ANN, distance-ratio based score & X ² distance,	DEC=2.331 AUC=0.934 EER=0.145
	Uzair et al [2]	MBGC v2 [18] VW [18] MBGC [18] NIR v2 [18] CMU-H [10] UBIPr [9]	3,482 NA NA 10,950	437 114 54 261	Adaptive thresholding, binarization, contour processing, eye center normalization & scaling	Raw pixel, LBP, PCA of pixels, PCA of LBP	DCC, MMD, MDA, AHISD, CHISD & SANP	97.2-99.8%
2016	Castrillón-Santana et al [7]	IoG [30]	28,231	5,080	Eye coordinate normalization, cropping, rescaling, 48 different sub-block configurations	WLD, LBP, ULBP, LOSIB, LTP	SVM	76.24- 92.46%
	Ambika et al [31]	Basel dataset Bmsce Artec 3D CASIA 3D FV1 [31]	NA NA 4,624	200 NA 123	Canny edge & Harris corner detectors, morphological operations & isotropic smoothing for photometric normalization	Laplace–Beltrami shape descriptors	Euclidean distance	80-97.5%
2017	Ahmed et al [5]	MICHE – II [15]	NA	86	Color space/Grayscale conversion, iris radius normalization, cropping & rescaling	MB-TLBP	Hamming & Chi-Square distances	EER=1.27% FRR=2.56% FAR=0.001
	Aginako et al [6]	MICHE II [15]	NA	86	Scaling and use of two different sub-block partition configurations	LSP, WLD, LPQ, HOG, NILBP, LBP, ULBP, LGP, LTP & LOSIB	K-NN, Bagging, Random Forest, Naive Bayes and C4.5	91.07%
	Sequeira et al [32]	Cross-eyed [32]	5,600	175	NA	LBP, Gabor features, LPQ, HOG,	correlation scores, region-bounded matchers, Bayesian	EER: 0.82-12.25%

	Zhao et al [33]	UBIRIS.v2 [12] FRGC [17] FOCS [10] CASIA.v4-dist. [6]	11.102 44,278 9,581 >54601	261 568 136 1000+1800	Resizing and padding	SIFT, SAFE, SURF Global and semantic features (gender/eye position (left/right))	CNN, Bayesian technique	Joint	82.43-98.90%
	Ahuja et al [34]	VISOB [35] [36] MICHE-II [15]	NA NA	550+290 86	Iris/periocular segmentation and resizing	OpenFace lib. 128-D features, Visobnet 1024-D features, Root SIFT, 512-D CNN-based features	CNN, KNN squared L2 distance, Cosine similarity		EER=0.053
2018	Zhao et al [36]	UBIPr [9] FRGC [17] FOCS [10] UBIRIS.v2 [12] VISOB [35] [36] CASIA.v4-dist. [6]	10,950 44,278 9,581 11.102 NA >54601	261 568 136 261 550+290 1000+1800	Resizing, scaling and cropping	128 D CNN features	CNN, Euclidean distance, DSC		EER=0.0147

4. Periocular image representation methods

The features used for periocular recognition could be broadly divided into two: global features and local features. The global approach extracts properties describing the whole region of interest (ROI) and performs global image analysis while the local approach first selects crucial locations, tagged key points, in the image. The properties in the neighborhood pixels of such key points are extracted based on some local analysis on them [44]. Other feature extraction methods which do not fall within these two broad groups are discussed in a separate category in this section.

4.1 Global periocular features

4.1.1 Gabor Filters

Gabor filters are employed in texture analysis to determine the presence of specific frequencies in specific directions in an image. They are applied using a set of different frequencies and orientation. In the works of [26] [31] [51], the image was divided into grids and a set of 5 frequency channels and equally spaced 6 orientations were used to sample the local power spectrum at each grid positioned in the eye center. The output from the grids were concatenated to form the feature vector. Bharadwaj et al [21], applied Gabor filter with 4 scales and 8 orientations to their normalized image to obtain its spatial envelope. It was also used in [29]. Gabor filters are good tools for texture representation and discrimination.

4.1.2 Walsh Masks

Walsh masks are convolution filters whose elements are only +1 and -1 used to capture local characteristics of an image such as contrast. A set of Walsh masks consist of N^2 $N \times N$ masks ($N=3,5,7,9,\dots$), where N corresponds to the N -sample long discrete version of the Walsh function combined in all possible pairs. Walsh masks filters were used to approximate the Walsh-Hadamard transform in the work of [46] for age invariant recognition based on periocular biometrics. They were also employed in Local Walsh-Transform Binary Pattern approach for periocular identification in [29].

4.1.3 Histogram of Gradients

Histogram of Oriented Gradients (HOG) is an appearance based feature extraction for object detection by representation of the gradient orientations in images relying on the concept that objects and shapes in images could be described by the spread of its intensity gradients [52]. The method works by counting the occurrence of gradient orientations cells of images. The given image is divided into smaller rectangular grid cells, the gradient orientation and magnitude is computed for each pixel within each cell which represents the histogram for that cell. The histogram for the whole image is built by the concatenation of all the cell histograms. HOG was applied as one of the descriptors in the works of [8] yielding the best and fastest result for a grid of 7x6 cells. It was also one of the features used by [2] [4] [9] [10] [23] and was one of the methods utilized by Halmstad University, Sweden participants in the Cross-Eyed2017 competition [38].

4.1.4 GIST descriptors

GIST descriptors are used to represent an image in low dimensions containing enough details to identify the scene in the image using five components which are naturalness (edge distribution in vertical and horizontal orientation), roughness (size of largest prominent object in image), ruggedness (extent to which contours in image deviate from horizontal), openness (extent of the presence of reference points), and expansion (gradient depth of space within image) representing the dominant spatial structure of the image scene. To obtain the GIST descriptor of a given image, it is convolved with N Gabor filters at different scales and orientations to produce N feature maps of same dimensions as image which are combined to produce the feature vector [53]. GIST descriptors reduced computational costs and yielded the most benefits of the works carried out by [2]. Bharadwaj et al also employed score level fusion of global GIST of length 1536 with local circular LBP in [21].

4.1.5 Phase Intensive Global Pattern

Phase Intensive Global Pattern (PIGP) basically involves convolving an image with a set of kernels at 4 phase-tilt angles to obtain the variation in intensity in the neighborhoods of pixels. The technique was applied to periocular recognition

by Bakshi et al [54].

4.1.6 Laws' Masks

Laws' masks also known as energy filters are a set of 2D kernels constructed from different combinations of five vectors each containing five elements. These kernels can be convolved with images in order to extract its texture properties such as spot, level, ripple, edge and wave. It was used in [29].

4.1.7 Local Phase Quantization

Local phase quantization (LPQ) is a texture analysis method that exploits the quantized phase spectrum property of the discrete Fourier transform (DFT) to achieve texture classification which is insensitive to centrally symmetric blur such as out-of-focus blur, motion-blur and atmospheric turbulence blur. To obtain the LPQ code, the 2-D DFT of the image is first computed using an $N \times N$ window ($N=3,5,\dots$) kernel, the four low-frequency coefficients for the real and imaginary part of the transform were quantized and combined into an 8-bit binary code. Codes from all pixels are concatenated resulting in a histogram of codes or LPQ code. [55]. LPQ was employed in the work of [4]. It was also used by Indian Institute of Technology (IIT) Indore, India in the Cross-Eyed2017 competition [38].

4.1.8 Force Field Transform

Force Field Transform (FFT) works on an image by regarding its pixels as an array of particles each generating a spherically symmetric force field on other pixels in the array similar to the inverse square law of gravitational force. This force is considered to be a vector depending on pixel intensity and position so that the total force acting on a pixel of unit intensity at a certain location is the vector sum of all the forces due to the other pixels contained in the image relative to their positions. The transform is a powerful averaging tool for reduction of noise effects in an image. [56]. This technique was used in the experiments of [29].

4.1.9 Discrete Wavelet Transform

Discrete Wavelet Transform (DWT) is a feature extraction technique obtained by use of non-continuous single level 2D DWT with respect to the Haar wavelet, leading to decomposition of approximation coefficients at a given level in four components which are the approximation at the next level, and details in the vertical, horizontal and diagonal orientation which could be treated as features [29] [57].

4.1.10 Discrete Cosine Transform

Discrete Cosine Transform (DCT) are 2D masks of sizes $N \times N$ (N must be an odd number) applied to an image to obtain N^2 coefficients representing a combination of the vertical and horizontal frequencies. These coefficients are linearly combined with the source mask used to obtain the features representing the image [29]. DCT was applied on LBP features in the work of [9] for soft biometric classification using the periocular region.

4.1.11 Local Binary Patterns

Local Binary Patterns (LBP) are good tools for texture classification because they can identify edges, line ends, corners, spots and other patterns. It is computed for every pixel in a given image by considering a 3×3 neighborhood around it. The value of each pixel's 8 neighbors is assigned a '0' or '1' according to whether the intensity of that neighbor is lower or greater than the reference pixel. These binary

values are then concatenated into an 8-bit string whose decimal value is assigned to the reference pixel to indicate the texture of the image at that point. Each pixel can thus have one of 256 possible values and all values for the pixels in a region of interest are quantized into an 8-bin histogram. LBP has a number of variants, some of them are: Circular LBP (CLBP) which is similar to the traditional LBP but the neighbors are considered within a distant, R from the reference pixel [1] [21]. Uniform LBP (ULBP), so labeled if the 8-bit string formed has a maximum of at most two transitions from 1 to 0 or vice-versa. All non-uniform patterns are assigned a single label while a separate label is assigned to each uniform pattern thus reducing the number of possible values from 256 to 59. ULBP is rotation insensitive [2]. The Local Ternary Pattern (LTP) reduces sensitivity to noise, it makes use of 3 values for coding (-1, 0, +1) and a threshold value. All values less than the difference of the reference pixel and the threshold are assigned '-1', those greater than the sum are assigned '+1' while those in-between are assigned '0'. The 8-bit string thus formed is used to construct an upper and lower pattern used to compute the local grayscale difference [58]. In the Median Binary Pattern, rather than threshold the neighboring pixel against the value of the center pixel, the localized 8-bit string is obtained by thresholding against the median values of all neighboring pixels to make the descriptor resistant to noise. In some literature, the thresholding is done against the local mean [59]. LBP was used in the works of [10] with a bin of 32, [2] [8] [18] with a bin of 59, [6] [9] [23] [29]. Its variant, CLBP was used in [21]. MB-TLBP, using bi-cubic interpolation was obtained for two pairs of parameters: a radius of 3 with 12 pixel points and also a radius of 6 with 24 pixel points. This resulted in a combined feature vector of size 2,070 [3] and LTP was used in [4] [8].

4.1.12 Weber Local Descriptors

Weber local descriptor (WLD) was inspired by the theory of human perception which observes that perception depends on the size or intensity of the original stimulus and the change in that stimulus. The Weber constant is thus the ratio of the change in intensity to the original intensity. Thus, WLD is a measure of change in intensity with reference to a center pixel and the gradient orientation. WLD was used in the work of [4] [8].

4.1.13 Local Colour Histograms

Local Color Histogram (LCH) is a descriptor that shows how the colors in an image are distributed. To obtain it, each color dimension is quantized into discrete ranges, the number of ranges represent the bin of the color histogram and a count of the pixels that fall in the ranges are computed [60]. Color histograms could be 1D, 2D or 3D depending on the number of combination of space used for the computation. Global color histogram (GCH) is the application of this technique to an image as a whole while LCH refers to its application to subsets of an image. RGB images were converted to the HSI color space in the work of Ahmed et al [3], local color histogram was used in [9] while Woodard et al experimented with RGB and HSV color spaces and their sub-spaces and various bin configurations, the red-green (RG) color space was observed to give the optimal result with a 4×4 bin configuration. This was used for the computation of the 2D color histogram [18].

4.1.14 Local Oriented Statistics Information Booster

LOSIB relies on the concept of LBP but improves upon it by computing the mean of all the absolute differences between

the reference pixel and its 8 neighbors using a 3x3 window in order to obtain the local oriented statistical content for the given image [61]. It was used in [8].

4.1.15 Local Salient Patterns

Local Salient Patterns (LSP) similar to LBP but rather makes use of the largest absolute difference between the reference pixel and its eight neighbors in a 3x3 window to obtain the features with a view to reducing the influence of noise [62]. It was employed in [4].

4.1.16 Local Gradient Patterns

Local Gradient Patterns (LGP) uses the same concept as LBP but computes its features using gradient information rather than pixel values. [63]. It was employed in [4].

4.2 Local periocular features

The algorithms for extracting local features are characterised by detection and/or selection of highlighted image locations, tagged key points, in the image. The properties in the neighborhood pixels of the key points are extracted based on some local analysis on them [44]. These algorithms are reviewed here.

4.2.1 Scale Invariant Feature Transformation

Scale Invariant Feature Transformation (SIFT) works on an image by first locating key points in the image using a difference of Gaussian (DoG) function to filter the image with two different scales for a few number of octaves and observing where the extrema is found. It is insensitive to scaling, translation and rotation [10]. Every point extracted is represented using its coordinates, scale and orientation. The array of information is then normalized through an affine transformation and processed with the image to obtain a feature descriptor for it [1] [29] [64]. SIFT was iteratively applied using adaptive thresholding until the number of key points was a maximum of 200 [10], it was also used to obtain descriptors for key points in [2] [23] [29]. Participants from both Halmstad University and Norwegian Biometrics Laboratory employed SIFT on the Iris/Periocular dataset in the Crossed-eye competition, 2017 [38].

4.2.2 Speeded Up Robust Features

Speeded Up Robust Features (SURF) locates key points in an input image based on the Hessian matrix, builds a box over its axis, perform affine transformation and using Haar wavelet responses performs feature extraction over a 4x4 sub-region [29]. A comparative analysis was performed using SURF and PILP in [1] [64]. It was one of the features used by both Halmstad University and Norwegian Biometrics Laboratory participants on the Iris/Periocular dataset in the Crossed-eye competition, 2017 [38].

4.2.3 Symmetry Assessment By Feature Expansion

Symmetry Assessment by Feature Expansion (SAFE) works by first detecting key points in an image, then using harmonic functions, estimates the presence of a variety of symmetric curves around the neighborhood of the key points. These highly symmetric curve functions are used to describe the neighborhoods about the key points in bands of concentric circles with varying radii. [44]. Also used for periocular recognition in [31] and by the Halmstad University participants in the Cross-Eyed2017 competition [38].

4.2.4 Binary Robust Invariant Scalable Key Points

Binary Robust Invariant Scalable Key points (BRISK) is a

descriptor obtained by the concatenation of results of brightness comparison tests carried out by applying a sampling pattern of 60 points equidistant on concentric circles about the key points. The sampling pattern is then rotated about its origin with respect to the gradient angle around the key point thus achieving rotation invariance. Comparison of intensities is then performed between all the possible short-distant pixel pairs of the sampling pattern and a value of 0 or 1 is assigned according to whether the pixel value is less than or greater than the reference giving rise to a feature vector of 512 bits for each key point [44]. This technique was applied by [64] for periocular recognition.

4.2.5 Oriented FAST and Rotated BRIEF

Oriented FAST and Rotated BRIEF (ORB) is a feature extraction that improves the BRIEF (Binary Robust Independent Elementary Features) descriptor by combining it with the FAST corner detector since BRIEF is susceptible to rotation. In this combined effort, the dominant rotation of the key point is initially obtained by first order moments and then steers the BRIEF descriptor accordingly. [44]. This technique was applied by [64].

4.2.6 Phase Intensive Local Pattern

Phase Intensive Local Pattern (PILP) first obtains key points in the input image by convolving it with filters of sizes 3x3, 5x5, 7x7 and 9x9, each at 4 phase-tilt angles yielding four filtered images per kernel which are then subjected to extrema detection. The gradient orientation histogram is then computed for each key point and concatenated to obtain the feature vector for the image. This approach was used in the work of [1].

4.3 Other features for periocular recognition

Asides global and local periocular features, other features such as facial marks like freckles, scars, moles [65], wrinkles, skin pores [18], the shape of the eyelids and eyebrow have also been utilized for periocular recognition. For example, Dong et al [22], employed nineteen eyebrow shape-based features which were extracted and grouped into three categories, these were the global shape features (GSF), critical point features (CPF) and local area features (LAF). The GSF included features such as rectangularity (similarity with a rectangle), eccentricity and isoperimetric quotient (similarity with a circle). The LAF included the area percentage, that is, the percentage of the area of eyebrow contained in sub-blocks of the eyebrow minimum bounding box. Eight such features were obtained by dividing the bounding box into 4 equal blocks horizontally then 4 equal blocks vertically. The CPF includes points in the eyebrow such as the eyebrow centroid, rightmost end, leftmost end, highest point and so on. In the work of Le et al [24], the eyebrow shape was obtained using graph-cut based segmentation, the vicinity around the eye was divided into 3 regions using Local Eyebrow Active Shape Model (LE-ASM) by employing 64 landmarks to determine the boundaries: foreground (representing the assumed eyebrow shape), foreground (the assumed skin area) and unclassified region (in-between the background and foreground region). Segmentation basically involved correctly classifying pixels in the unclassified region into foreground or background by use of pairs of neighboring pixels around them considering the intensity difference (edge weight) and shortest path amongst the pixel pairs. The Laplace-Beltrami shape descriptors, which was based on geometric attributes of the periocular region was used in the experiments of Ambika et

al, it was obtained using the Laplace–Beltrami operator having Riemannian metric that describes the global shape and allows for measurement of angles and distances on the surface for topology-preserved mapping. The triangular mesh obtained in their preprocessing stage was analyzed using adjacency and stiffness matrices to obtain the periocular eigen spectrum. The eigen values thus obtained were processed through a sequence of scaling, logarithmic transform and averaging to produce the shape descriptors [35]. Researchers in [49] extracted eyelid shape descriptors in fusion with LBP as features for periocular biometrics, employing the use of eyelid statistics such as accumulated curvature, shape context, Elliptical Fourier Descriptors (EFD) for periocular recognition.

5. Matching methods for periocular template matching

A number of distance metrics (which tends to match feature vectors by computing how close or far apart they are from each other) are used for classification and matching in periocular biometrics such as Euclidean distance, X^2 distance, Chi square distance and so on, besides these some common classifiers include support vector machines (SVM), K-Nearest Neighbors (K-NN), Bagging, Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Bayesian Networks and so on. In this paper, they are broadly classified into three: Distance metrics, CNN and others.

5.1 Distance Metrics

In the matching stage, Santos et al used the X^2 distance for the concatenated HOG, LBP, ULBP and GIST feature vectors [2]. This was also used to match GIST and CLBP features by Bharadwaj et al [21] before fusing both results by applying the weighted sum rule. X^2 distance was also used for comparison of the normalized magnitude of complex values of the Gabor response vectors and image pairs having the lowest distance were matched [26]. Euclidean distance was one of the classifiers used by Xu et al [29], it was also used to compute the matching score for HOG and LBP features in the works of Park et al [23]. Verification of periocular images was by computation of Euclidean distance between the shape descriptor of the probe image with those of the images in the gallery. Authentication was considered positive if the Euclidean distance between both image vectors falls within a specified threshold [35]. In the work of Bakshi et al, for every key point in the probe image, its key point descriptor was compared with the key point descriptor of every key point in the gallery image using Euclidean distance. Key points are considered a match if the Euclidean distance between them falls within a certain threshold [1]. Many distance metrics were tested by Woodard et al and the city block distance and Bhattacharya coefficient were found to perform the best for the LBP histogram and color histogram respectively. Both features were combined at match score level by summing with min-max normalization [18]. In the works of Ahmed et al, the Hamming distance and Chi-Square distance were used for iris and periocular matching respectively. Both results were then fused using a weighted sum of scores determined experimentally from the test samples. The best weights for iris and periocular were found to be 0.55 and 0.45 respectively [3]. Chi-square distance was used to compute the cost of matching points on one eyebrow with another [24]. Normalized Cosine distance and Manhattan distance were used for computation of similarity scores in [29], Cosine distance was also used for classification of the computed

Procrustes feature vectors in [24], Minimum Distance (MD) and Linear Discriminant Analysis (LDA) were used to train and carry out recognition and gender classification using the eyebrow features [22]. At the matching stage, Halmstad University, Sweden employed Bayesian Networks, Indian Institute of Technology (IIT) Indore, India used Cosine similarity and correlation scores while neighborhood-bounded matching was utilized by Norwegian Biometrics Laboratory, NTNU, Norway [38].

5.2 Other classifiers

The SVM works on the principle of finding an optimal hyperplane that maximally classify samples into their correct classes. For example. It was one of the three methods used to train and carry out recognition and gender classification using the eyebrow features by Dong et al [22], it was used for periocular-based gender classification in the works of Castrillón et al [8] and also in the work of Lyle et al for gender and ethnicity classification [9]. Joint Bayesian was used in the work of [39]. K-nearest neighbors (K-NN), Bagging, Random Forest, Naive Bayes and C4.5 classifiers were used in the works of [4]. Six classifiers were used for classification: DCC (Discriminative Canonical Correlation), MMD (Manifold-Manifold Distance), MDA (Manifold Discriminant Analysis), AHISD (affine hulls image-set dissimilarity), CHISD (convex hulls image-set dissimilarity) and SANP (sparse approximated nearest points). A number of classifiers are used for each image in the set each generating its own similarity vector. These are then fused into one vector by experimenting with different fusion rules such as the median, sum, product, max, min and majority voting. The proposed label from each specific rule is further fused using an error weighted fusion (EWF) technique [6].

Some research work used ANN of varying configurations for periocular classification tasks, the second classification method used by Padole et al was an ANN having a single hidden layer of three neuron using the sigmoid transfer function while the output layer had only one neuron [10]. This configuration was used by Jamie et al but they experimented with 15, 30, 60 and 120 neurons in the hidden layer [9]. The ANN configuration in the work of Santos et al had two hidden layers of nodes 11 and 6 for the first and second hidden layers respectively and one output neuron trained using back-propagation was used for the final recognition score [2].

6. Convolutional neural networks, CNN-based periocular recognition

More recently, researchers adopt the end-to-end CNN framework, with varying architecture and configurations, for periocular recognition to achieve improved performance. In [39], Zhao et al made use of CNN for classification with semantic assistance (SCNN) comprising of a main structure and a branch structure. Layers 1 to 6 of the main structure was made up of 3 sets of convolution layer and pooling layers which made use of the Rectified Linear Unit (ReLU) transfer function, connected to layer 7, the fully connected (FC) layer. The branch structure was a similar architecture but had only two sets of convolution and pooling layers, with layer 5 being its FC layer. The main CNN was trained with the identities of the periocular images while the semantic or branch structure augments the main by complementing classification via training with the gender and the right/left eye positions. In their more recent work [42], they paid more attention to the eyebrow and eye region for periocular recognition using

attention-based CNN which adopted a multiple-glance structure; a mechanism which identifies the region of interest by utilizing a fully convolutional network (FCN-peri) [66] and an attention mechanism CNN (AttNet) consisting of a series of four convolution layers and a FC layer 5, the first convolution layer extracts low level features like height, depth, width while the other convolution layers were split into two branches. The top branch processing global periocular features while the bottom branch incorporates the region of interest (eyebrow and eye) in the second and fourth layer so that higher emphasis is placed on those regions. The overall feature used was the global feature in conjunction with the features from the eyebrow and eye region. The work of Ahuja et al examined verification using a fusion of the periocular and iris trait by applying a hybrid of two CNN models: model 1 and model 2. Model 1 consist of series convolution and max-pooling layers ending in a FC layer which was trained using face and periocular images from external source in a transfer learning approach. It had three parts, the first used the OpenFace library to generate feature vectors of length 128 and computes their similarity score using the squared L2 distance. The second part used Visobnet to obtain feature vectors and of length 1024 and compares them using Cosine similarity. The third extract key points of identical dimension and orientation in the iris images using Root SIFT feature descriptors and computes similarity scores using KNN. Outputs from the three parts were normalized and fused by taking the average of the three scores. Model 2 had a similar structure to model 1 but was trained using data from an internal source, MICHE II. It generated a feature vector with 512-dimension which were compared using Cosine similarity. Finally, the scores from both models were fused by computing their average [40].

7. Performance indicators of periocular recognition research

Impressive results have been obtained in researches on periocular-based recognition, when fused with other modalities, it has been found to further increase accuracy. The values found in literature should be regarded as approximates since researchers use subset of databases or exclude images of poor qualities in the database. The approach used for eyebrow segmentation achieved an F-Measure accuracy of 99.4% on the MBGC database for 50 subjects having a total of 200 images. On a small subset of the of the AR database, an accuracy of 76.0% was achieved and an accuracy of 54.4% for a larger subset of the database. On a small subset of the of the MBGC database, an accuracy of 85.0% was achieved and an accuracy of 71.3% for a larger subset of the database [24]. Woodard et al, Lyle et al and Dong et al both made use of the FRGC and MBGC databases. A recognition accuracy of 91% was achieved on 410 individuals on the FRGC database while a recognition accuracy of 87% was obtained for 85 subjects on the MBGC NIR database [18]. Recognition rate accuracies of 90% and 75% were achieved on MBGC and FRGC databases respectively while gender classification accuracies of 96% and 97% were achieved on MBGC and FRGC databases respectively [22]. Jamie et al worked on the FRGC dataset of 4232 images from 404 subjects, baseline classification accuracies of 97.3% and 94% were achieved for gender and ethnicity (Asian and non-Asian) respectively while classification accuracies of 90% and 89% respectively were achieved for the MGBC dataset of 350 images of 60 subjects [9]. Xu et al, also worked on the FRGC database and

obtained rank-1 recognition accuracy of 53.2% for a fusion of DWT and LBP while a verification rate accuracy of 61.2% was achieved when their best extraction technique was fused with the Kernel Correlation Feature Analysis (KCFA). This was a commendable achievement relative to the difficulty and size of the database used (FRGC: Experiment 4 of NIST's FRGC, ~128 million) [29]. Park et al used a smaller subset of the same FRGC database, their work on 1136 probe and 1136 gallery periocular images from the FRGC 2.0 database achieved 87.32% rank-1 recognition accuracy using the periocular modality by a fusion of 3 different matchers (HOG, LBP and SIFT) [23]. Ahmed et al and Aginako et al both worked on the MICHE II database, in the experiment of the former, an equal error rate (EER) of 2.74%, a false rejection rate (FRR) of 9.13% when the false acceptance rate (FAR) was set to 0.001 was achieved. When the periocular result was fused with the iris matching result, the overall result was significantly improved giving an EER value of 1.22% and a FRR of 2.56% at an FAR of 0.001 [3]. A fusion of the iris and periocular modalities gave an overall recognition rate accuracy of 91.07% with the LPQ descriptor and K-NN classifier [4].

The work of Bharadwaj et al on the UBIRIS v2 database using GIST descriptor as a stand-alone resulted in rank-1 identification accuracies of 63.34% and 61.64% were obtained for the right and left periocular regions respectively while using of the CLBP descriptor as a stand-alone, rank-1 identification accuracies of 54.30% and 52.82% were obtained for the right and left periocular regions respectively. The fusion combining both left and right regions yielded an accuracy of 70.82% for GIST and an accuracy of 63.77% for CLBP. The proposed algorithm fusing both descriptors with left and right periocular region gave a rank-1 identification accuracy of 73.65% [21]. Bakshi et al using PILP features for periocular verification achieved an accuracy of 99.87% on the BATH (University of Bath) database, 99.62% on CASIAv3 (Chinese Academy of Sciences) database, 95.51% on UBIRISv2 (University of Beira) database and 92.47% on cropped periocular region of the FERETv4 (National Institute of Standards and Technology) database. [1]. Various experiments were carried out by Santos et al on the CSIP image dataset with different parameters. Their best result was obtained with the following performance metrics: decidability (DEC) = 2.331, area under curve (AUC) = 0.934, the equal error rate (EER) = 0.145 [2]. The experiment by Castrillón et al was on use of periocular region for gender classification using the GROUPS database. The performances of the 6 descriptors were evaluated at trade-off between margin and error, $C = 1$ and $\gamma = 0.07$. The best accuracies obtained for the descriptors were as follows: HOG = 83.02%, ULBP = 80.31%, LBP = 76.24%, LTP = 80.08%, WLD = 82.20% and LOSIB = 76.45%. The best fusion result gave an accuracy of 82.91% for the combination of HOG, ULBP, LTP and WLD. Further fusion with the shoulder, face and head region gave an improved accuracy of 92.46% [8]. The minimum authentication accuracy in the work of Ambika et al was with the Basel dataset [36] under pose variation where they achieved 80% while their best accuracy for neutral conditions was on the Bmsce Artec 3D database [35] where they achieved 97.5%. Rank-1 periocular region recognition rates of 99.8%, 98.5%, 97.2%, and 99.5% were achieved for MBGC NIR, MBGC VW, CMU Hyperspectral and UBIPr datasets respectively [6]. The best performance in cross-eyed competition for periocular method was obtained by Halmstad University with an EER value of 0.082 on the cross-eyed database [38].

Very good results were realized using convolutional neural networks (CNN), Zhao et al using SCNN achieved rank 1 verification accuracies of 82.43% on the UBIRIS v2 database, 91.13% on the FRGC database, 96.93% on the FOCS database and 98.90% on the CASIA v4 database [39]. In their more recent experiment using attention-based CNN model achieved an EER of 0.0226, 0.0859, 0.0768, 0.490, 0.1005 and 0.0147 (closed-world) on the UBIPr, FRGC, FOCS, CASIA v4 distance, UBIRIS v2 and VISOB databases respectively [42]. Ahuja et al's hybrid CNN model performed best when the training and testing samples were taken from same mobile devices, giving an AUROC value of 0.986 and an EER value of 0.053 [40]. On the UBIPr database, Nie et al obtained their best verification rate for a fusion of CRBM, LBP, HOG and DSIFT using SVM classifiers was for an EER of 0.064 while they achieved a periocular recognition accuracy of 50.1% using the weighted-product fusion of CRBM, LBP, HOG and DSIFT, a good result considering the relatively larger size of database used [30].

8. Suggestions for further research

Although the periocular region has many discriminative features, a number of such features are yet to be fully explored. Future research works consider features such as the properties of the upper and lower eye folds (shape, thickness and average distance apart), the eyelashes (length, intensity, density, direction and distribution), the sclera (texture and color), the tear duct (texture, shape and size) and eye shape [67]. Research works can also extend the periocular-based age estimation or classification solutions. More accurate periocular segmentation techniques are anticipated with more research in the use of eye corners which are more stable unlike the moving iris on the moving eyeball within moving eyelids. Discovery of more robust descriptors, classifiers and fusion techniques with reduced computation cost favoring recognition of video images of subjects captured on-the-move and mobile-based recognition systems. More importantly, the availability of newer image databases simulating extended real-world application scenarios and

captured with a variety of sensors will facilitate further research developments.

9. Conclusions

A number of algorithms used for periocular recognition were reviewed. It is clear that the periocular region possesses good discriminative attributes found useful as a stand-alone trait for periocular-based recognition with good results and have been found to augment other traits like the face and iris especially under difficult conditions. The periocular trait has also proved useful for ethnicity and gender classification. Research in this area is relatively new compared with the face and iris trait and is expected to continue to draw more research interests. The various methods for segmentation of the periocular region from face databases using the eye centers, eye corners and eyebrow were analyzed. The various preprocessing methods to deal with unconstrained scenarios for improved recognition accuracies were reviewed as well as the principles behind the various techniques used as descriptors together with the different classification methods and the commendable performance accuracies achieved. These have been found to be usually as a result of the fusion of two or more feature descriptors. The various nature of public databases used in literature have also been reviewed. Definitely there is room for improvements and this paper serves to equip researchers with insight needed in this field and to inspire more experiments with improvements through fusion and other novel methods.

Acknowledgement(s)

This research is fully sponsored by Covenant University Centre for Research, Innovation and Development (CUCRID), Covenant University, Ota, Nigeria.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License



References

- [1] S. Bakshi, P. K. Sa, and B. Majhi, "A novel phase-intensive local pattern for periocular recognition under visible spectrum," *Biocybernetics and Biomedical Engineering*, vol. 35, pp. 30-44, 2015.
- [2] 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, 2015.
- [3] N. U. Ahmed, S. Cvetkovic, E. H. Siddiqi, A. Nikiforov, and I. Nikiforov, "Combining iris and periocular biometric for matching visible spectrum eye images," *Pattern Recognition Letters*, vol. 91, pp. 11-16, 2017.
- [4] N. Aginako, M. Castrillón-Santana, J. Lorenzo-Navarro, J. M. Martínez-Otzeta, and B. Sierra, "Periocular and iris local descriptors for identity verification in mobile applications," *Pattern Recognition Letters*, vol. 91, pp. 52-59, 2017.
- [5] G. Santos and H. Proença, "Periocular biometrics: An emerging technology for unconstrained scenarios," in *Computational Intelligence in Biometrics and Identity Management (CIBIM), 2013 IEEE Workshop on*, 2013, pp. 14-21.
- [6] M. Uzair, A. Mahmood, A. Mian, and C. McDonald, "Periocular region-based person identification in the visible, infrared and hyperspectral imagery," *Neurocomputing*, vol. 149, pp. 854-867, 2015.
- [7] J. A. Badejo, A. A. Atayero, and T. S. Ibiyemi, "A robust preprocessing algorithm for iris segmentation from low contrast eye images," in *Future Technologies Conference (FTC)*, 2016, pp. 567-576.
- [8] M. Castrillón-Santana, J. Lorenzo-Navarro, and E. Ramón-Balmaseda, "On using periocular biometric for gender classification in the wild," *Pattern Recognition Letters*, vol. 82, pp. 181-189, 2016.
- [9] J. R. Lyle, P. E. Miller, S. J. Pundlik, and D. L. Woodard, "Soft biometric classification using local appearance periocular region features," *Pattern Recognition*, vol. 45, pp. 3877-3885, 2012.
- [10] C. N. Padole and H. Proença, "Periocular recognition: Analysis of performance degradation factors," in *Biometrics (ICB), 2012 5th IAPR International Conference on*, 2012, pp. 439-445.
- [11] R. Jillela, A. A. Ross, V. N. Boddeti, B. V. Kumar, X. Hu, R. Plemmons, et al., "Iris segmentation for challenging periocular images," in *Handbook of Iris Recognition*, ed: Springer, 2013, pp. 281-308.
- [12] A. Kumar and A. Passi, "Comparison and combination of iris matchers for reliable personal authentication," *Pattern recognition*, vol. 43, pp. 1016-1026, 2010.
- [13] University of BATH. (2005, November 15). *BATH University Database*. Available: <http://www.bath.ac.uk/news/2005/11/15/iris151105.html>
- [14] Center for Biometrics and Security Research. (2010, December 14). *CASIA Database*. Available: <http://www.cbsr.ia.ac.cn/english/IrisDatabase.asp>
- [15] H. Proença and L. A. Alexandre, "UBIRIS: A noisy iris image

- database," in *International Conference on Image Analysis and Processing*, 2005, pp. 970-977.
- [16] H. Proença, S. Filipe, R. Santos, J. Oliveira, and L. A. Alexandre, "The ubiris. v2: A database of visible wavelength iris images captured on-the-move and at-a-distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 1529-1535, 2010.
- [17] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, pp. 1090-1104, 2000.
- [18] D. L. Woodard, S. J. Pundlik, J. R. Lyle, and P. E. Miller, "Periocular region appearance cues for biometric identification," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, 2010, pp. 162-169.
- [19] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, et al., "Overview of the face recognition grand challenge," in *Computer vision and pattern recognition, 2005. CVPR 2005. IEEE computer society conference on*, 2005, pp. 947-954.
- [20] P. J. Phillips, P. J. Flynn, J. R. Beveridge, W. T. Scruggs, A. J. O'toole, D. Bolme, et al., "Overview of the multiple biometrics grand challenge," in *International Conference on Biometrics*, 2009, pp. 705-714.
- [21] S. Bharadwaj, H. S. Bhatt, M. Vatsa, and R. Singh, "Periocular biometrics: When iris recognition fails," in *BTAS*, 2010, pp. 1-6.
- [22] Y. Dong and D. L. Woodard, "Eyebrow shape-based features for biometric recognition and gender classification: A feasibility study," 2011.
- [23] U. Park, R. R. Jillela, A. Ross, and A. K. Jain, "Periocular biometrics in the visible spectrum," *IEEE Transactions on Information Forensics and Security*, vol. 6, pp. 96-106, 2011.
- [24] T. H. N. Le, U. Prabhu, and M. Savvides, "A novel eyebrow segmentation and eyebrow shape-based identification," in *Biometrics (IJCB), 2014 IEEE International Joint Conference on*, 2014, pp. 1-8.
- [25] A. Martinez and R. Benavente, "The AR face database, CVC," *Copyright of Informatica (03505596)*, 1998.
- [26] F. Alonso-Fernandez and J. Bigun, "Eye detection by complex filtering for periocular recognition," in *IWBF 2014-2nd International Workshop on Biometrics and Forensics 2014, Valletta, Malta, 27-28th March, 2014*, 2014.
- [27] J. Fierrez, J. Ortega-Garcia, D. T. Toledano, and J. Gonzalez-Rodriguez, "BioSec baseline corpus: A multimodal biometric database," *Pattern Recognition*, vol. 40, pp. 1389-1392, 2007.
- [28] A. F. Sequeira, J. C. Monteiro, A. Rebelo, and H. P. Oliveira, "Mobbio: a multimodal database captured with a portable handheld device," in *Computer Vision Theory and Applications (VISAPP), 2014 International Conference on*, 2014, pp. 133-139.
- [29] 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*, 2010, pp. 1-8.
- [30] L. Nie, A. Kumar, and S. Zhan, "Periocular recognition using unsupervised convolutional RBM feature learning," in *2014 22nd International Conference on Pattern Recognition (ICPR)*, 2014, pp. 399-404.
- [31] F. Alonso-Fernandez and J. Bigun, "Near-infrared and visible-light periocular recognition with gabor features using frequency-adaptive automatic eye detection," *IET Biometrics*, vol. 4, pp. 74-89, 2015.
- [32] K. W. Bowyer and P. J. Flynn, "The ND-IRIS-0405 iris image dataset," *arXiv preprint arXiv:1606.04853*, 2016.
- [33] L. J. Denes, P. Metes, and Y. Liu, *Hyperspectral face database*: Carnegie Mellon University, The Robotics Institute, 2002.
- [34] A. C. Gallagher and T. Chen, "Understanding images of groups of people," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, 2009, pp. 256-263.
- [35] D. Ambika, K. Radhika, and D. Seshachalam, "Periocular authentication based on FEM using Laplace-Beltrami eigenvalues," *Pattern Recognition*, vol. 50, pp. 178-194, 2016.
- [36] P. Paysan, R. Knothe, B. Amberg, S. Romdhani, and T. Vetter, "A 3D face model for pose and illumination invariant face recognition," in *Advanced video and signal based surveillance, 2009. AVSS'09. Sixth IEEE International Conference on*, 2009, pp. 296-301.
- [37] M. Castrillón-Santana, M. De Marsico, M. Nappi, F. Narducci, and H. Proença, "Mobile Iris CHallenge Evaluation II: results from the ICPR competition," in *Pattern Recognition (ICPR), 2016 23rd International Conference on*, 2016, pp. 149-154.
- [38] A. F. Sequeira, L. Chen, J. Ferryman, P. Wild, F. Alonso-Fernandez, J. Bigun, et al., "Cross-eyed 2017: Cross-spectral iris/periocular recognition competition," in *Biometrics (IJCB), 2017 IEEE International Joint Conference on*, 2017, pp. 725-732.
- [39] Z. Zhao and A. Kumar, "Accurate periocular recognition under less constrained environment using semantics-assisted convolutional neural network," *IEEE Transactions on Information Forensics and Security*, vol. 12, pp. 1017-1030, 2017.
- [40] K. Ahuja, R. Islam, F. A. Barbhuiya, and K. Dey, "Convolutional neural networks for ocular smartphone-based biometrics," *Pattern Recognition Letters*, vol. 91, pp. 17-26, 2017.
- [41] A. Rattani, R. Derakhshani, S. K. Saripalle, and V. Gottemukkula, "ICIP 2016 competition on mobile ocular biometric recognition," in *Image Processing (ICIP), 2016 IEEE International Conference on*, 2016, pp. 320-324.
- [42] Z. Zhao and A. Kumar, "Improving Periocular Recognition by Explicit Attention to Critical Regions in Deep Neural Network," *IEEE Transactions on Information Forensics and Security*, vol. 13, pp. 2937-2952, 2018.
- [43] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler, "Mobile iris challenge evaluation (MICHE)-I, biometric iris dataset and protocols," *Pattern Recognition Letters*, vol. 57, pp. 17-23, 2015.
- [44] F. Alonso-Fernandez and J. Bigun, "A survey on periocular biometrics research," *Pattern Recognition Letters*, vol. 82, pp. 92-105, 2016.
- [45] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, pp. 137-154, 2004.
- [46] F. Juefei-Xu and M. Savvides, "Unconstrained periocular biometric acquisition and recognition using COTS PTZ camera for uncooperative and non-cooperative subjects," in *Applications of Computer Vision (WACV), 2012 IEEE Workshop on*, 2012, pp. 201-208.
- [47] G. Mahalingam, K. Ricanek, and A. M. Albert, "Investigating the periocular-based face recognition across gender transformation," *IEEE Transactions on Information Forensics and Security*, vol. 9, pp. 2180-2192, 2014.
- [48] F. Smeraldi and J. Bigun, "Retinal vision applied to facial features detection and face authentication," *Pattern recognition letters*, vol. 23, pp. 463-475, 2002.
- [49] H. Proença, "Ocular biometrics by score-level fusion of disparate experts," *IEEE Transactions on Image Processing*, vol. 23, pp. 5082-5093, 2014.
- [50] K. Oh, B.-S. Oh, K.-A. Toh, W.-Y. Yau, and H.-L. Eng, "Combining sclera and periocular features for multi-modal identity verification," *Neurocomputing*, vol. 128, pp. 185-198, 2014.
- [51] F. Alonso-Fernandez and J. Bigun, "Periocular recognition using retinotopic sampling and gabor decomposition," in *European Conference on Computer Vision*, 2012, pp. 309-318.
- [52] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, pp. 886-893.
- [53] A. Oliva and A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope," *International journal of computer vision*, vol. 42, pp. 145-175, 2001.
- [54] S. Bakshi, P. K. Sa, and B. Majhi, "Phase intensive global pattern for periocular recognition," in *India Conference (INDICON), 2014 Annual IEEE*, 2014, pp. 1-5.
- [55] V. Ojansivu and J. Heikkilä, "Blur insensitive texture classification using local phase quantization," in *International conference on image and signal processing*, 2008, pp. 236-243.
- [56] D. J. Hurley, M. S. Nixon, and J. N. Carter, "Force field feature extraction for ear biometrics," *Computer Vision and Image Understanding*, vol. 98, pp. 491-512, 2005.
- [57] A. Joshi, A. Gangwar, R. Sharma, A. Singh, and Z. Saquib, "Periocular recognition based on Gabor and Parzen PNN," in *Image Processing (ICIP), 2014 IEEE International Conference on*, 2014, pp. 4977-4981.
- [58] J. Ren, X. Jiang, and J. Yuan, "Relaxed local ternary pattern for face recognition," in *ICIP*, 2013, pp. 3680-3684.
- [59] J. Ruiz-del-Solar, R. Verschae, and M. Correa, "Recognition of faces in unconstrained environments: A comparative study," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, p. 184617, 2009.
- [60] C. Gode and A. Ganar, "Image retrieval by using colour, texture and shape features," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 3, 2014.
- [61] O. García-Olalla, E. Alegre, L. Fernández-Robles, and V. González-Castro, "Local oriented statistics information booster (LOSIB) for

- texture classification," in *Pattern Recognition (ICPR), 2014 22nd International Conference on*, 2014, pp. 1114-1119.
- [62] Z. Chai, Z. Sun, T. Tan, and H. Mendez-Vazquez, "Local salient patterns—A novel local descriptor for face recognition," in *Biometrics (ICB), 2013 International Conference on*, 2013, pp. 1-6.
- [63] B. Jun and D. Kim, "Robust face detection using local gradient patterns and evidence accumulation," *Pattern Recognition*, vol. 45, pp. 3304-3316, 2012.
- [64] Ş. Karahan, A. Karaöz, Ö. F. Özdemir, A. G. Gü, and U. Uludag, "On identification from periocular region utilizing sift and surf," in *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European*, 2014, pp. 1392-1396.
- [65] A. K. Jain and U. Park, "Facial marks: Soft biometric for face recognition," in *Image Processing (ICIP), 2009 16th IEEE International Conference on*, 2009, pp. 37-40.
- [66] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431-3440.
- [67] K. Hollingsworth, K. W. Bowyer, and P. J. Flynn, "Useful features for human verification in near-infrared periocular images," *Image and Vision Computing*, vol. 29, pp. 707-715, 2011.