Periocular Region-based Person Identification in the Visible, Infrared and Hyperspectral Imagery

Muhammad Uzair, Arif Mahmood, Ajmal Mian and Chris McDonald

Computer Science & Software Engineering
The University of Western Australia
35 Stirling Highway, Crawley WA, 6009, Australia
uzair@csse.uwa.edu.au, {arif.mahmood, ajmal.mian,chris.mcdonald}@uwa.edu.au

Abstract

Face recognition performance degrades significantly under occlusions that occur intentionally or unintentionally due to head gear or hair style. In many incidents captured by surveillance videos, the offenders cover their faces leaving only the *periocular region* visible. We present an extensive study on periocular region based person identification in video. While, previous techniques have handpicked a single best frame from videos, we formulate, for the first time, periocular region based person identification in video as an image-set classification problem. For thorough analysis, we perform experiments on periocular regions extracted automatically from RGB videos, NIR videos and hyperspectral image cubes. Each image-set is represented by four heterogeneous feature types and classified with six state-of-the-art image-set classification algorithms. We propose a novel two stage inverse Error Weighted Fusion algorithm for feature and classifier score fusion. The proposed two stage fusion is superior to single stage fusion. Comprehensive experiments were performed on four standard datasets, MBGC NIR and visible spectrum [1], CMU Hyperspectral [2] and UBIPr [3]. We obtained average rank-1 recognition rates of 99.8, 98.5, 97.2, and 99.5% respectively which are significantly higher than the existing state of the art. Our results demonstrate the feasibility of image-set based periocular biometrics for real world applications.

Keywords: Periocular Biometric, Partial Facial Occlusion, Image-set Classification, Hyperspectral Face Recognition, Face and Iris Biometric

Figure 1: Examples where periocular biometrics are more effective than the full face biometrics.

1. Introduction

Face recognition is a well studied area and existing techniques perform well in controlled environments. However, the performance of face recognition algorithms degrades if the face is partially covered due to head gear or hair style. The complete face of offenders is often not visible in surveillance videos. Helmets, that additionally cover faces, are commonly used by motorcyclists and sportsmen. Helmets and skiing masks have been exploited by offenders to cover their faces during crimes. Finally, in many countries, women partially cover their faces due to religious and cultural reasons. In most of these cases, the region around the eyes, the *periocular region* is the only visible biometric (see Fig. 1).

Besides robustness to occlusion, periocular region based person identification offers some advantages over the full face biometrics as it is least affected by expression variations, aging effects [4, 5] and the changes due to growth of male facial hair. Moreover, full face recognition performance degrades significantly in the presence of pose variations whereas the periocular region based identification is likely to perform better in the case of extreme pose changes when only one eye is visible. In these cases, the single visible periocular region can be used or mirrored and matched to the opposite side periocular region in the database.

Periocular biometrics also offer some advantages over the iris biometric [6, 7, 8]. Periocular biometrics can be acquired from a larger distance by surveillance cameras whereas iris scanning requires the subjects to open their eyes fully and look into a scanner from a very close range. The latter may not be socially acceptable to users [9, 10]. In many practical scenarios, the subject may not be cooperative or willing, or even aware that he/she is being imaged. For example,
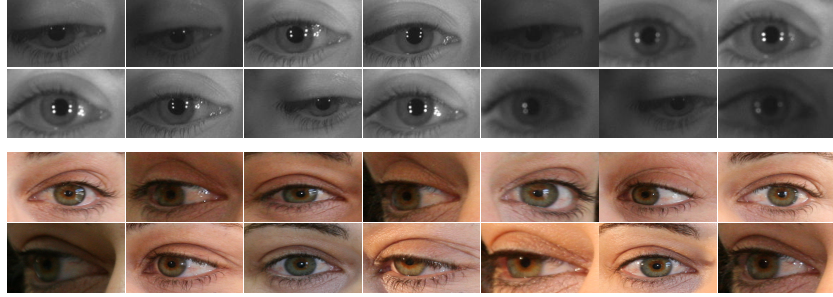
Figure 2: Appearance variations of the periocular region of the same subject in the MBGC NIR video (top row) and UBIPr (bottom row) databases

in banks and shopping malls, valued customers may need to be identified without causing inconvenience to them.

In this work we present a comprehensive study into the feasibility of the periocular region as a biometric for person identification. We consider realistic scenarios such as automatic detection of periocular regions, the availability of videos rather than a single handpicked frame and propose a two stage inverse Error Weighted Fusion scheme to achieve the state of the art results.

Previous works on periocular biometrics are mostly based on single image matching [11, 12, 13, 14, 15, 16] where a single best frame per subject is handpicked to form the gallery. Each image in the query sequence is matched with the gallery to find the best match. Such techniques inherently suffer from the lack of information since only one image cannot contain all possible variations in the periocular region of an individual (Fig. 2). For example, if the gallery image is taken from a frontal view with centered eyeballs and the query images have pose variations with eyeballs at the extreme ends, recognition performance will significantly degrade.

In contrast, we formulate periocular region based identification as an image-set classification problem. Each set contains multiple periocular images of the same person and represents a wide range of variations including different eyelid positions, illuminations, eyeball movements and poses. We observe that the left and the right periocular regions of the same subject possess more similarities to each other than to other subjects. Therefore, the reflection of one region is combined with the other to form a single large set per identity. We believe that this strategy better suits linear modeling techniques. We construct a gallery from image-sets of different identities. The query set also contains multiple periocular images of the same person and is assigned the label of the nearest gallery set. Compared to single image matching, set-to-set matching offers significantly more information

3

because multiple images in the same set model the appearance variations [17].

We extract four types of features from each periocular region and combine six state-of-the-art image-set classification techniques including the Affine Hull based image-set Distance (AHISD) [18], Convex Hull based image-set Distance (CHISD) [18], Discriminative Canonical Correlation (DCC) [19], Manifold-Manifold Distance (MMD) [20], Manifold Discriminant Analysis (MDA) [21], and Sparse Approximated Nearest Point (SANP) distance [22]. A brief overview of these techniques is given in Section 2. Note that our objective is to show the feasibility of image-set based classification for periocular biometrics as opposed to an unbiased comparison of the image-set classification algorithms.

Our preliminary results were published in [23]. Here, we extend our work in a number of directions. We present a detailed literature survey of the periocular recognition techniques. We test different imaging modalities beyond the visible spectrum, such as infra-red and hyperspectral image cubes. We develop a fully automatic video based periocular region detection system for all three modalities and investigate the use of different classifier fusion techniques. Rigorous experiments are performed on more datasets and higher recognition rates are observed compared to our previous work [23] and existing state-of-the-art methods. Brief results on the four databases are shown in Table 1.

2. Survey of Periocular Biometrics Research

Periocular region includes the iris [10], eyes, eyelids, eye lashes, and part of the eyebrows [13]. Recognition using periocular biometrics is an emerging research area. Initial feasibility studies were done by Park et al. [9, 10] who detected the iris in the visible spectrum images and extracted features at fixed grid positions around the center of the limbus circle. Experiments on a subset of the FRGC v2.0 [1] dataset showed that their algorithm's performance depends on the accurate iris detection and is sensitive to the eyeball movements.

Table 1: Average and standard deviation of the identification and verification rates at 0.001 FAR (%) of the proposed approach on four public databases. Detailed results are in Section 7.1.

Database	Subjects	Identification	Verification
UBIPr [3]	342	99.5±0.5	99.9±0.1
MGBC v2 VS video [1]	91	98.5±1.0	97.8±0.5
MGBC v2 NIR video [1]	114	99.9±0.1	99.8±0.2
CMU HS [2]	48	97.2±1.0	96.8±1.0

4

The use of Local Binary Patterns (LBP) [24] to represent the texture of the periocular region has been investigated by several researchers. Miller et al. [25] used city block distance to classify LBP features. Their work was extended by Adams et al. [15] using genetic algorithms to select the optimal subset of LBP features. They experimented with manually generated periocular data from a subset of high resolution frontal images of FRGC [1] with neutral expression, less appearance variations and controlled illumination. Woodard et al. [13] used the LBP features and color histograms to represent the local appearance of the periocular region. They achieved better performance by using city block distance for matching LBP features and Bhattacharya coefficient for matching color histograms. They performed experiments on periocular data generated from two databases, FRGC v2 face [1] and MBGC v1 NIR video [26]. In the case of MBGC NIR dataset, the complete videos were not used. Manually extracted periocular regions from the first two frames of the video were used as gallery and periocular region from a random frame was chosen as a probe. The periocular regions had the same scale and no appearance variations.

Bharadwaj et al. [27] performed score level fusion of global GIST features and local circular LBP features of the periocular regions. They used the visible spectrum eye images of UBIIRIS v2 [28] dataset which contained slight appearance variations. Xu et al. [16] proposed Walsh Transform based local binary patterns (WLBP). Periocular region containing both eyes were cropped using the detected iris centers. They performed experiments on the FRGC v2 [1] database and achieved a verification rate of 17.0% at 0.1% FAR. Woodard et al. [14] simultaneously utilized the iris and periocular biometrics by performing score-level fusion. LBP features were extracted from the periocular region while the iris texture was encoded using the Gabor filters. Their experiments showed that the periocular recognition performed better than iris recognition for the MBGC v1 NIR portal videos.

Hollingsworth et al. [12] investigated human performance on the periocular recognition task. By presenting unlabeled pairs of NIR periocular images to different humans, they found that humans can recognize the periocular region with 92% accuracy. They also calculated the performance of three computer algorithms on the periocular recognition task and found that the performance of humans and computer was similar. Boddeti et al. [29] used probabilistic matching based on Optimal Trade-off Synthetic Discriminant Function correlation filter to classify periocular regions. They performed experiments on the Face and FOCS database [30] which contains periocular region appearance variations in the form of illumination, blur and off-angle iris. They reported better recognition rates of the ocular

5

Table 2: Recognition Rates (RR) of the existing state of the art periocular biometric recognition techniques.

Author	Year	Database	#Subj	RR
Park et al.[10]	2009	Proprietary	30	80.80
Miller et al.[25]	2010	FRGC, FERET	464	89.70
Adams et al.[15]	2010	FRGC, FERET	464	92.16
Woodard et al.[14]	2010	MBGC v1 NIR	85	96.50
Woodard et al.[13]	2010	MBGCv1NIR, FRGCv2	85, 410	87.00, 91.20
Xu et al. [16]	2010	FRCG	466	53.20
Bhar. et al. [27]	2010	UBIRIS v2 [28]	261	73.60
Park et al.[9]	2011	FRGCv2	568	87.32
Boddeti et al.[29]	2011	FOCS[30]	136	94.20
Pauca et al.[32]	2011	COIR[32]	285	95.00
Padole and Proenca[3]	2012	UBIPr[3]	342	38.00†
Xu et al. [31]	2012	Compass[31]	342	60.00†
Fernandez and Bigün [33]	2012	CASIAv3[34], BioSec[35]	249, 200	5.66, 13.08‡
Oh et al. [11]	2014	UBIRIS v1	131	6.89‡

† Verification rate at 0.01 False Accept Rate   ‡ Equal Error Rate

regions than the iris on the FOCS database.

Xu and Savvides [31] employed different schemes to correct the illumination and pose variations of the periocular regions. Class dependent Feature Analysis was used to classify the WLBP features [16]. Pauca et al. [32] used SIFT features for classification of the periocular regions and also introduced the COIR database. Padole et al. [3] introduced the UBIPr database which contains scale and pose variations as well as non-uniform illumination variations. They observed that the eye corners offer a better reference for grid construction compared to the iris center. Fernandez and Bigün [33] used retinoscopic sampling to generate grids of various geometry centering the pupil. Gabor features calculated from the grids were used for periocular region classification. Oh et al. [11] combined periocular and sclera biometrics using score level fusion. They used structured random projections for feature extraction from periocular regions and a modified LBP operator for feature extraction from sclera region. The obtained improved equal error rates on UBIRIS v1 dataset.

Table 2 summarizes the major works in the area of periocular biometrics. Existing research has investigated the use of texture or point features calculated from

a single handpicked periocular image. Therefore, these methods do not model the periocular region variations including partial occlusion, eyeball and eyelid movements, pose and illumination conditions. In order to mitigate some of these challenges, we propose an image-set based approach for periocular biometric recognition. For each subject, the gallery may contain one or more image-sets per subject. An image-set will contain multiple periocular region images of the same subject thereby modeling a wide range of variations.

3. Review of Image-set Classification Algorithms

Although image-set classification has been well studied for face recognition, to the best of our knowledge, periocular region recognition has not been formulated before as an image-set classification problem. We argue that a set-based approach is more suitable for periocular biometrics because it can model a number of variations such as pose, position of eyeball and partially open eyelids that are unavoidable in real scenarios. We use six state-of-the-art image-set classification algorithms broadly divided into sample and structure based.

3.1. Sample Based Image-set Classification

Sample based techniques measure the distance between nearest neighbor samples of two image-sets. Let $X = \{x_i\}_{i=1}^n \in \mathcal{R}^{m \times n}$ be an image-set, where $x_i \in \mathcal{R}^m$ is a feature vector and n be the number of feature vectors in a set which may vary across the image-sets. The feature vectors can simply be the image pixel values or some features calculated from the pixels such as the PCA coefficients or LBP features [24]. Each image-set may be considered as a point cloud in $\mathcal{R}^m$. All points in a probe image-set X_p are compared with all points in each gallery set X_g to find the nearest pair of points (x_i, x_j) such that $x_i \in X_p$ and $x_j \in X_g$. If x_i and x_j have zero mean and unit magnitude, the nearest neighbor pair (x_i, x_j) is the one that maximizes the cosine of the angular distance:

$$\max_g \left(\max_{x_i, x_j} X_g^t X_p \right). \tag{1}$$

The probe image-set label is predicted as the label of x_j.

Cevikalp and Triggs [18] considered each image-set as a convex geometric region in $\mathcal{R}^m$. Set dissimilarity was measured by the distance of closest approach between the regions represented by the affine (AHISD) or convex hulls (CHISD). The region contained all the affine combinations $x_p = \sum_{i=1}^n X_{pi} \alpha_{pi}$ and $x_g = \sum_{j=1}^n X_{gj} \alpha_{gj}$, where $\alpha_p, \alpha_g \in \mathcal{R}^m$. For the case of affine hull $\sum_{i=1}^n \alpha_{pi} =$

7

$\sum_{j=1}^{n} \alpha_{gj} = 1$ and for the convex hull $0 \leq (\alpha_{pi}, \alpha_{gj}) \leq 1$. The minimum distance was computed as

$$\min_{g}\Big(\min_{\alpha_p,\alpha_g} ||X_g\alpha_g - X_p\alpha_p||^2\Big). \qquad (2)$$

For the case of affine hull, the minimum distance was computed using least squares while for the case of convex hull, an SVM was used.

Instead of searching the nearest points with dense combinations of samples in the corresponding image-set, Hu et al. [22] proposed that each of the two points should be approximated by a sparse combination from the samples of the respective set. They argued that the sparse approximated nearest points (SANP) will lie close to some facet of the affine hull, and hence, implicitly incorporate structural information of the sets as well. By restricting the SANPs to be close to the facets, this approach can reject outliers resulting in higher accuracy.

3.2. Structure Based Image-set Classification

Structure based techniques represent the underlying structure of an image-set with one or more linear subspaces. Structural similarity of the sets is usually measured using subspace to subspace distance. Kim et al. [19] proposed Discriminative Canonical Correlation (DCC) which performs discriminative learning on canonical correlations between the structures of image-sets. More specifically, a discriminant function is learned that maximized the within-class and minimized the between-class canonical correlations. Image-sets are compared after transforming them optimally by the learned discriminant function.

Wang et al. [20] proposed Manifold-Manifold Distance (MMD) which clustered each image-set into multiple linear local models and represented each model by a linear subspace. The similarity between two sets was defined as the canonical correlation between the nearest local models. In addition, the nearest point distance was also combined with the structural similarity to calculate the final similarity between two sets.

Wang and Chen [21] proposed Manifold Discriminant Analysis (MDA) that uses a hierarchical divisive clustering approach to represent each image-set by multiple local linear models. The local models are transformed by a linear discriminant function where different classes are better separable. The similarity between two sets is calculated as the pair-wise local model distances in the learned embedding space.
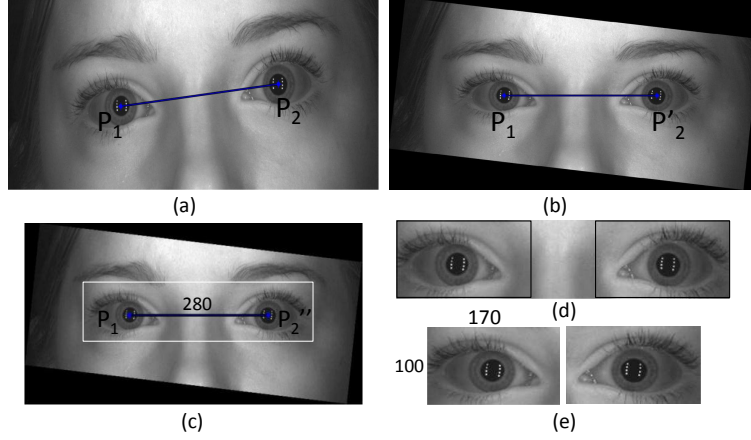
Figure 3: Periocular region normalization in MBGC NIR and visible spectrum videos. (a) Detected pupil centers (b) Rotation normalization (c) Scale normalization (d) Nose bridge removal (e) Normalized periocular regions

4. Automatic Periocular Region Extraction

Distance variations between the subject and the camera induces significant scale changes and head pose variations induces rotational effects in the video. These scale and rotation variations do not provide any discriminative information. We propose two algorithms for automatic periocular region extraction and normalization in videos and hyperspectral cubes.

4.1. Automatic Periocular Region Detection in Videos

In contrast to the previous approaches [13, 14] which selected only the best frames with the same scale and rotation for matching, we perform rotation and scale normalization. In each video frame (NIR or visible), we automatically detect the eye pair by using a Haar feature based ensemble of weak classifiers as proposed by Viola and Jones [36]. Then we crop a relatively large area containing both eyes and the nose bridge. For each eye, we automatically detect the pupil center using an improved version of Masek's algorithm [37, 38]. We first apply adaptive thresholds on the eye image to obtain a binary image. Following this, the largest local number of contour points is searched by contour processing the binary image. A circle is fitted to the detected largest contour points and its center is then selected as the pupil's center. The pupil centers in each eye are denoted by P_1 and P_2 in Fig. 3.

For in-plane rotation normalization, the cropped images are rotated such that the line through P_1 and P_2 becomes horizontal. For scale normalization, the dis-

9

tance between P_1 and P_2 is fixed and the full periocular region is equally scaled in both dimensions. By using fixed distance constraints, the normalized periocular region is divided into left and right periocular regions and the nose bridge is removed to simulate the situations when only one eye is visible. Thus we only utilize the eye region for recognition because in many practical situations, the other regions may not be available.

To enable single eye based recognition, we flip all the periocular regions to one side. Linear combinations of the original and flipped periocular regions can generate novel appearances and hence account for more intra-person variations. Note that the effect of flipping from left to right or right to left is the same.

4.2. Periocular Region Detection in Hyperspectral Images

We extend our analysis beyond the visible spectrum towards hyperspectral periocular biometric recognition. A hyperspectral image cube consists of images captured at multiple wavelengths of the electromagnetic spectrum. The signal to noise ratio of the bands near the blue wavelength region is very low and in general, no periocular region detector may ensure 100% detection rate in these images. However we assume that the location of the pupil centers in all bands are aligned. Thus we solve the problem of miss detections by estimating collective pupil center positions over all bands in a hyperspectral cube.

In each image of a hyperspectral cube, we automatically detect the eye pair using Viola and Jones detector [36] and for each eye, we automatically detect the pupil centers as follows. First, we apply thresholding on the eye image to obtain a segmented binary image. Next, we apply the morphological opening operation on the binary image. The opening operation enhances the dense region of the binary image corresponding to the pupil center and smooths the sparsely occupied regions (Fig. 4). We apply the opening operation n times and a value of $n = 3$ gave good results in our experiments. The location of the global minimum in the resultant opened image is considered as the pupil center. Here, we do not require circle fitting because of the availability of multiple bands to assist accurate pupil center localization. Figure 4 shows the proposed pupil center detection process.

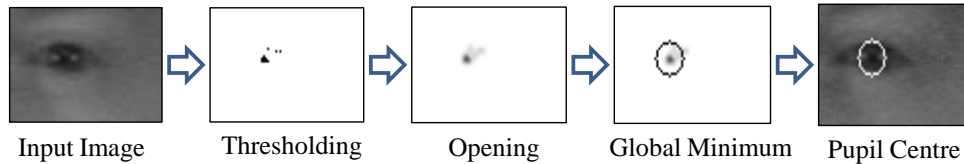| Input Image | Thresholding | Opening | Global Minimum | Pupil Centre |

Figure 4: Pupil center detection in hyperspectral databases. Detection errors are corrected by using the proposed algorithm.
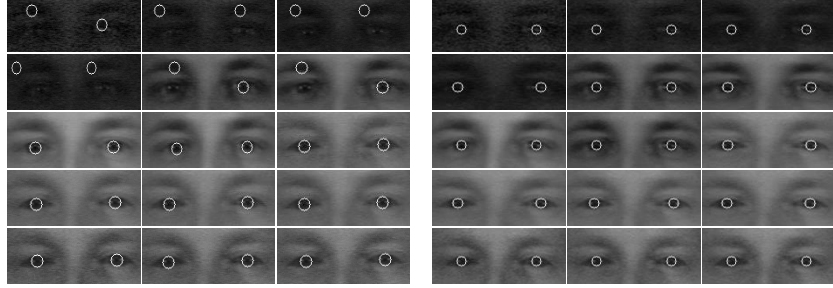
Figure 5: Detected pupil center (left) and corrected pupil centers (right) in some of the bands of CMU hyperspectral database.

We independently detect the pupil center in each band (Fig. 5). Then, each detection votes in a circle of radius r around the detected position, using Gaussian weights, in an accumulation array. The votes are accumulated over all bands of the same cube. The average of the coordinates in a small vicinity getting maximum votes over all bands is selected as the pupil center for the whole hyperspectral cube and used for rotation and scale normalization. The same rotation and scale normalization is applied to all bands of a cube. Cropping and flipping of the periocular regions is similar to the video databases.

5. Feature Extraction

We extract the following four types of features from the periocular regions.

5.1. Raw Pixel Values

To mimic typical surveillance videos, all normalized periocular images are scaled down to 20×30 pixels and the intensity values are used as features.

5.2. PCA Coefficients of Pixels

The dimensionality of the normalized images is reduced by projecting them on 400 most significant PCA basis. The resulting 400 coefficients are used as feature vectors. Because the size of the normalized images is relatively large, they are resized to a smaller size for performing PCA. Note that for learning the PCA basis, only the gallery sets are used as training data. Using PCA coefficients as features has two main advantages. By dropping the least significant principal components, PCA suppresses the effect of noise such as partial and unaligned periocular regions present in the image-sets. By reducing the dimensionality of

11

15	16	23
21	23	17
25	25	24

0	0	1
0		0
1	1	1

	226	

LBP String = (01000111)
LBP code = 0+2+0+0+0+32+64+128 = 226
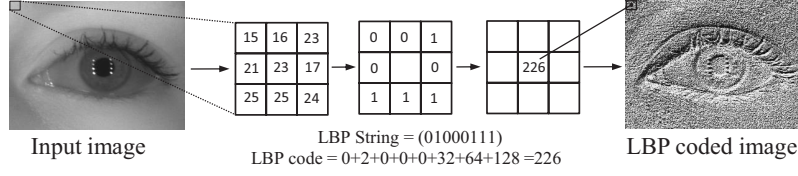
Input image   LBP coded image

Figure 6: LBP filtering process to generate feature type 2.

the normalized images to 400, the image-set classification algorithms achieved significant speedup.

5.3. LBP Features

From the 20×30 images, LBP features are extracted using circular (8,1) neighborhoods [24]. The LBP operator labels a pixel with a binary number computed by thresholding the gray-scale differences between the pixel and its neighborhood. The LBP of the pixel z in an image I can be represented as follows:

$$
\begin{aligned}
LBP_{P,R}(z) &= \{LBP_{P,R}^{(p)}(z)\}_{p=1,\dots,P}, \\
LBP_{P,R}^{(p)}(z) &= s(I^g(v_p) - I^g(z)), \\
s(z) &= \begin{cases} 1 & \text{if } z \geq 0, \\ 0 & \text{if } z < 0, \end{cases}
\end{aligned}
\tag{3}
$$

where $I^g(z)$ is the gray-scale value of the pixel z in the image I and $\{v_p\}_{p=1,\dots,P}$ as a set of P equally spaced pixels located on a circle of radius R and center z. We choose LBP features for their robustness to monotonic gray-scale changes resulting from global or local illumination variations [39]. Fig (6) shows the process of LBP filtering. The resulting LBP coded images are used as feature vectors.

5.4. PCA Coefficients of LBP Features

LBP features are computed over the normalized images and the dimensionality of the LBP codes is reduced to 400 using PCA.

6. Features and Classifiers Score Level Fusion

Considering periocular biometric as a set based classification problem, the label of the probe image-set p is to be predicted by computing its similarity with g gallery image-sets. There are c different classifiers, each generating a different similarity vector $\{S_k\}_{k=1}^c$ where $S_k = \{s_{k,j}\}_{j=1}^g$. These similarity vectors are
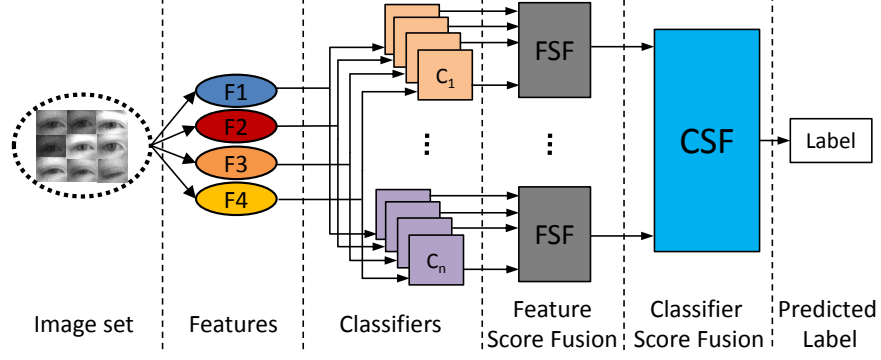
Figure 7: Block diagram of the two stage fusion: Feature Score Fusion (FSF) and Classifier Score Fusion (CSF). Different fusion strategies are used at both stages (Table 4).

fused into one vector $\mathcal{S} = \{\mathsf{s}_j\}_{j=1}^g$. For this purpose, we investigate the use of different classifier fusions techniques [40, 41] including the sum, product , min, max, and median rules, and majority voting.

We also propose a novel two stage fusion scheme using classifier error weighted summation technique (see Fig. 7). Let L_p^{Rule} be the predicted label by a specific rule, where Rule$\in$ { sum, prod, max, min, med, vote, SMF}.

$$
\begin{aligned}
\text{Sum rule: } L_p^{sum} &\triangleq \max_j(\sum_{k=1}^c s_{k,j}). \\
\text{Product rule: } L_p^{prod} &\triangleq \max_j(\prod_{k=1}^c s_{k,j}). \\
\text{Max rule: } L_p^{max} &\triangleq \max_j(\max_k(s_{k,j})). \\
\text{Min rule: } L_p^{min} &\triangleq \max_j(\min_k(s_{k,j})). \\
\text{Median rule: } L_p^{med} &\triangleq \max_j(Med_k(s_{k,j})). \\
\text{Majority voting rule: } L_p^{vote} &\triangleq Mode_k(\{l_k\}_{k=1}^c),
\end{aligned}
\tag{4}
$$

where $l_k \triangleq \max_j(s_{k,j})$ is the label predicted by the kth classifier. In these schemes, all classifiers are given equal weight. In contrast, we propose to weight each classifier inversely proportional to the error of that classifier on the validation dataset. We refer to this technique as *Error Weighted Fusion* (EWF). Specifically,
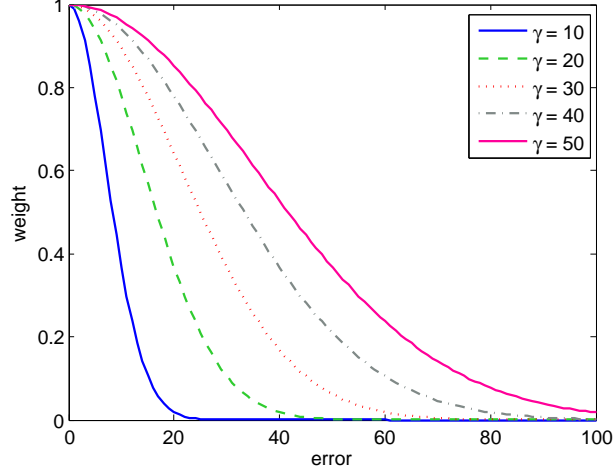
Figure 8: Plot of the proposed weight function w_k for different combinations of ξ_k and γ.

the following weight function is learned during the training phase:

$$w_k = \exp\left(-\frac{\xi_k^2}{\gamma^2}\right), \tag{5}$$

where ξ_k is the error rate of the kth classifier on the training data. The parameter γ controls the decay of the weight function with the increasing error rate: $1 \geq w_k \geq 0$. Figure 8 shows the variation of w_k for different combinations of ξ_k and γ. An appropriate value of γ is selected that minimizes the error on training data after fusion of classifiers. The fused similarity vector is computed as:

$$\mathcal{S} = \sum_{k=1}^{c} \exp\left\{\ln\left(\frac{S_k - \delta_k}{\lambda_k}\right) - \left(\frac{\xi_k^2}{\gamma^2}\right)\right\}, \tag{6}$$

where δ_k is used to shift the values in the S_k to start from zero and λ_k is the range of values in S_k. After fusion, the label L_p of the probe is assigned by the index containing maximum value over all gallery image-sets: $L_p^{SMF} = \max_j(\mathcal{S}_j)$. Note that if similarity is changed by distance, max function will be changed by min function. We perform score level fusion at two stages namely Feature Score Fusion (FSF) and Classifier Score Fusion (CSF) as shown in Fig. 7. We then experimentally study the performance of the fusion schemes defined above.

14

Table 3: Database details

Database	Subjects	Image-sets/subject	Images/set
UBIPr [3]	342	6	5
MGBC v2 Visible [1]	91	1-6	23-352
MGBC v2 NIR [1]	114	1-12	6-48
CMU Hyperspectral [2]	48	1-5	130

7. Experiments and Results

For periocular region based person identification, we use one periocular region image database, namely, the UBIPr database [3] and three public face databases. The face databases include the MBGC Visible spectrum and NIR video databases [1] and the CMU hyperspectral database [2]. These databases are summarized in Table 3 and their details are given below.

The *UBIPr Visible Spectrum Database*[3] consists of visible spectrum periocular images with variations in scale, illumination, eyeball movements, partial occlusion and head pose changes. The images were acquired at five different camera distances (4m to 8m), three poses ($0\,^\circ$, $30\,^\circ$, $-30\,^\circ$), three gazes ($0\,^\circ$, $30\,^\circ$, $-30\,^\circ$) and three levels of pigmentation (light, medium, heavy). The images in this database are converted to gray scale and normalized to reduce the effects of scale and pose variations.

The *MBGC v2 Visible Specturm* (VS) videos have significant scale and illumination variations (Fig. 9-a). Many frames have low signal to noise ratio and motion blur resulting in significant intra-person variations. The NIR illuminator used during *MBGC v2 NIR* video acquisition flashed for only a brief moment resulting in significant illumination variations, defocusing, motion blur, sensor noise, specular reflections, partially occluded faces and off-angle eyes making biometric recognition very challenging (Fig.9-b). Moreover, only part of the face is visible in most videos.

In each video frame, the periocular regions are extracted as detailed in Section 4.1. The distance between pupil centers is fixed to 280 pixels and a normalized image of size 430×100 containing both periocular regions is cropped. This normalized image is divided into left and right periocular regions by removing the central nose bridge portion of 90×100 pixels. The final image size of the left and right periocular regions is 170×100 (Fig. 3). The right side periocular regions are mirrored and combined with the left side regions to form a single set (Fig. 10).
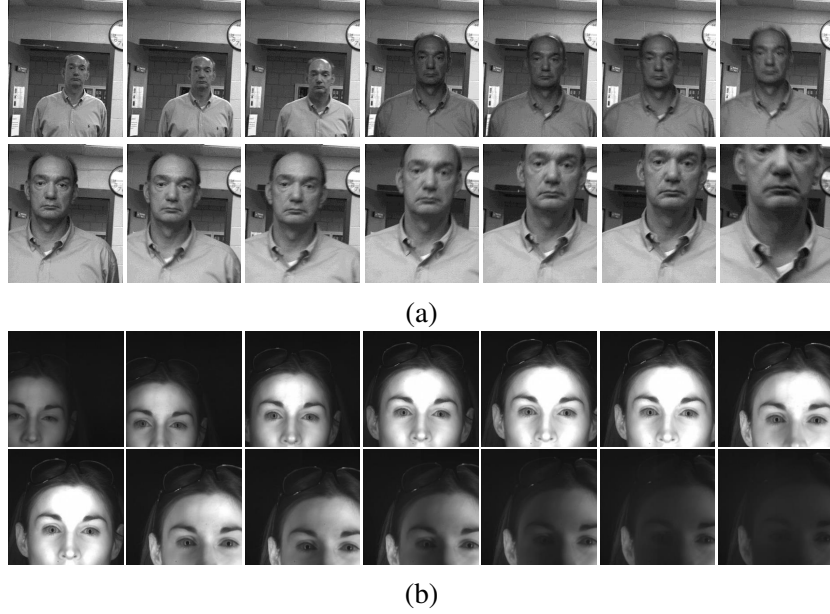
15

(a)

(b)

Figure 9: Sample video frames from the MBGC v2 (a) Visibe and (b) NIR databases.

Note that left to right mirroring gives the same results as right to left mirroring.

The *CMU Hyperspectral Database* [2] contains face images in 65 bands covering the spectral range of 450nm to 1100nm at 10nm step. The spatial resolution of the images is 640x480. In each image (band), we automatically detect the periocular region and normalize for rotation and scale variations. The distance between the eye centers is fixed to 90 pixels. A normalized image of size 160×40 is then cropped. The periocular region is divided into left and right regions by removing the nose bridge. The final image size of the periocular regions is 40×28 pixels (Fig. 10-c). The right periocular images are mirrored and combined with the left ones for form a set.

7.1. Experimental Setup

For each database, the gallery is constructed by randomly selecting one image-set per subject. The remaining image-sets are used as probes. Experiments are repeated 10-fold by randomly selecting different gallery and probe combinations each time. For image set classification, six algorithms are used including AHISD, CHISD [18], SANP [22], DCC [19], MMD [20] and MDA[21]. For every algorithm, we performed 10-fold experiments for each of the four feature vector types.

16

Figure 10: Sample images from one image-set of MBGC visible spectrum video (left) and CMU Hyperspectral database (right). See Fig. 2 for sample images from MBGC NIR and UBIPr databases.

For the sample based algorithms, the default input parameter values are used. For the structure based algorithms, the parameters are carefully tuned so that the best performance can be obtained on training data. For DCC, the subspace dimensions are set to 10 which preserves 90% energy and the corresponding 10 maximum canonical correlations are used to define set similarity. The embedding space is set to 100. For MMD and MDA, the ratio between Euclidean and Geodesic distance is varied from 2.0 to 0.01 with a step of 0.5. Maximum recognition rates were observed at 0.10 for MMD and 2.0 for MDA. The number of connected nearest neighbors for computing geodesic distance was tested at 14, 12 and 10. The best results were observed at the value of 10. The maximum canonical correlation was used in defining MMD. For MDA, the number of between-class NN local models was set to 5 and the dimension of MDA embedding space was set to 10 as recommended by Wang et al. [20, 21].

The error rate ξ_k in the weighted sum based fusion scheme is learned during training using a held-out validation set which is disjoint from the test set. For each database the validation set is generated by randomly choosing one image-set for each subject. The parameter ξ_k is found using the validation set. We observe that the accuracy of different algorithms on the validation set is almost the same as the test sets. After experimenting various values of γ, a fixed value $\gamma = 40$ is used for all the datasets.

7.2. Results and Discussion

The proposed image-set based periocular biometric recognition has exhibited recognition rates much higher than the existing state-of-the-art techniques. In the following subsections, results of the six classification algorithms and different fusion schemes are discussed for each of the four databases. Results of existing
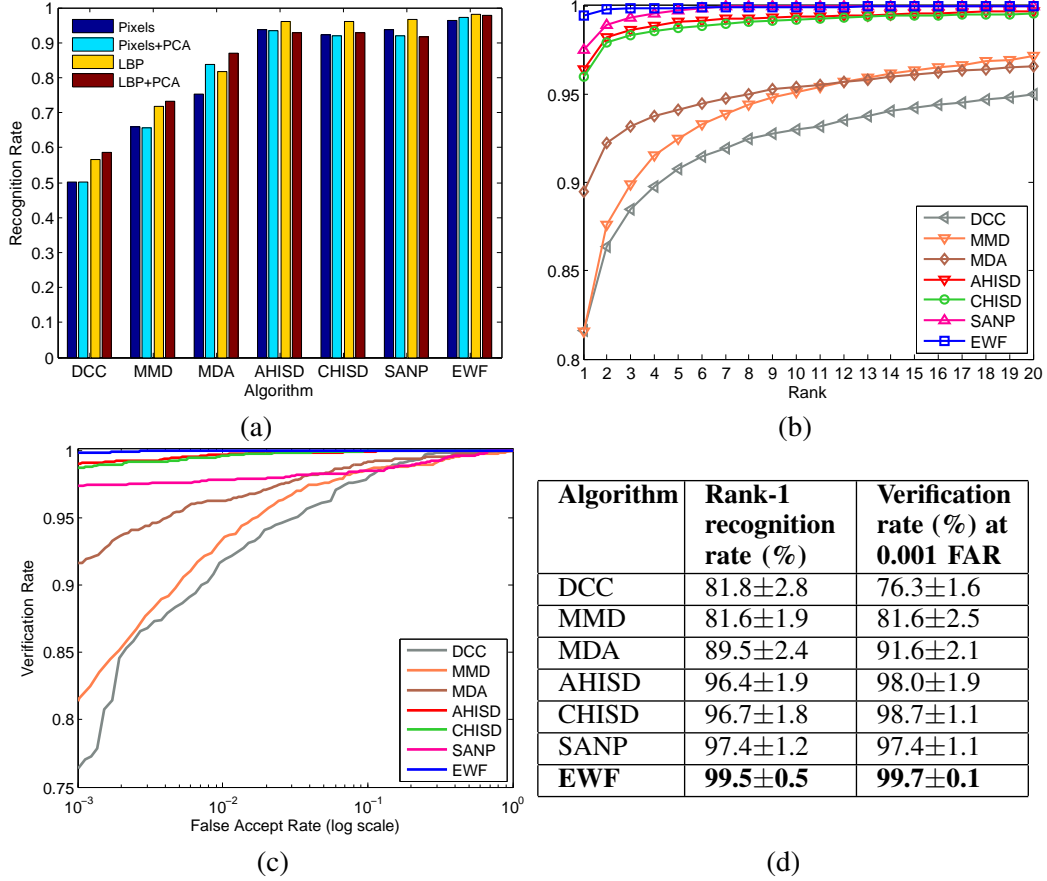
17

Figure 11: Results of 10-fold experiments on the UBIPr database. (a) Average recognition rates of individual algorithms on individual features. (b) CMC and (c) ROC curves of individual algorithms after feature score fusion. (d) Summarized results.

algorithms in the literature along with brief experimental settings are summarized in Table 2.

UBIPr database: Results on this database are reported in Fig. 11. The performance of each algorithm on individual feature types is shown in Fig. 11-a, where CHISD, AHISD and SANP achieve better performance than DCC, MMD and MDA. Thus sample based algorithms have outperformed structure based ones because only five images per set were not sufficient to estimate the structure. LBP features consistently achieve the highest performance among the sample based techniques. The last group of bars in Fig. 11-a show the performance of *Error Weighted Fusion* (EWF) of all algorithms on different features. Here, again LBP

features have achieved the highest recognition rate.

The performance of each algorithm is also studied after the feature score fusion (FSF) of all features using EWF. Figure 11-b and Fig. 11-c show the CMC and ROC curves respectively for each algorithm after EWF feature score fusion. The curves of the sample based algorithms are again above the structure based algorithms. The best performance in Fig. 11-b and Fig. 11-c correspond to the proposed two stage fusion algorithm (FSF + CSF as shown in Fig. 7). The proposed two stage fusion approach acheived the highest recognition rate of 99.5±0.5% and the highest verification rate of 99.9±0.1% at 0.001 FAR. The latter is significantly higher than the 38% verification rate at 0.001 FAR reported by Padole and Proenca in [3] using fusion of LBP+HOG+SIFT features. The consistent best performance of EWF demonstrates the effectiveness of the proposed image-set based classification as well as feature (FSF) and classifier score fusion (CSF) techniques.

MBGC visible range video database: Figure 12 shows the results on this database. The average recognition rates of each algorithm for each feature type are given in Fig. 12-a. Here, the structure based algorithms have outperformed the sample based ones because the number of images per set is significantly higher than the UBIPr dataset leading to improved structure estimation. The last group of bars are for feature score fusion of all algorithms. Feature type 4 (LBP+PCA) has shown better performance than the other features for structure based techniques and the feature score fusion. The large number of samples per set not only helps in better structure estimation but also facilitates better PCA subspace estimation after LBP filtering. Since the images in this database have more noise and blur, PCA and LBP together perform better noise filtering than either one alone.

The performance of each algorithm after feature score fusion (FSF) of all features using EWF is given in Fig. 12-b and 12-c and summarized in Table 12-d. Here, the plots of structure based techniques are above the sample based techniques. The proposed two stage fusion (FSF+CSF) using EWF has outperformed all other algorithms by achieving a rank-1 recognition rate of 98.5±1.0% and verification rate of 97.8±0.5% at 0.001 FAR.

MBGC NIR database: Results for this database are given in Fig. 13. The performance of individual algorithms on each feature type is shown in Fig. 13-a. Due to fewer images per set, sample based algorithms have again performed better than structure based ones with the exception of MDA. Note that in MDA, a discriminative function is learned such that the geometric distance between different set centers is maximized. The average of the NIR imagery of the periocular region, which also includes the iris, is more discriminative than the visual spectrum images. This fact is also reflected by the higher mean and lower standard deviation of

19

	(a)	(b)

Algorithm	Rank-1 recognition rate (%)	Verification rate (%) at 0.001 FAR
DCC	93.9±1.0	93.4±1.9
MMD	88.5±1.9	83.4±1.5
MDA	94.4±1.2	92.3±1.1
AHISD	69.4±2.9	66.5±3.3
CHISD	80.7±2.1	76.8±2.0
SANP	90.5±2.1	84.0±2.2
EWF	**98.5±1.0**	**97.8±0.5**

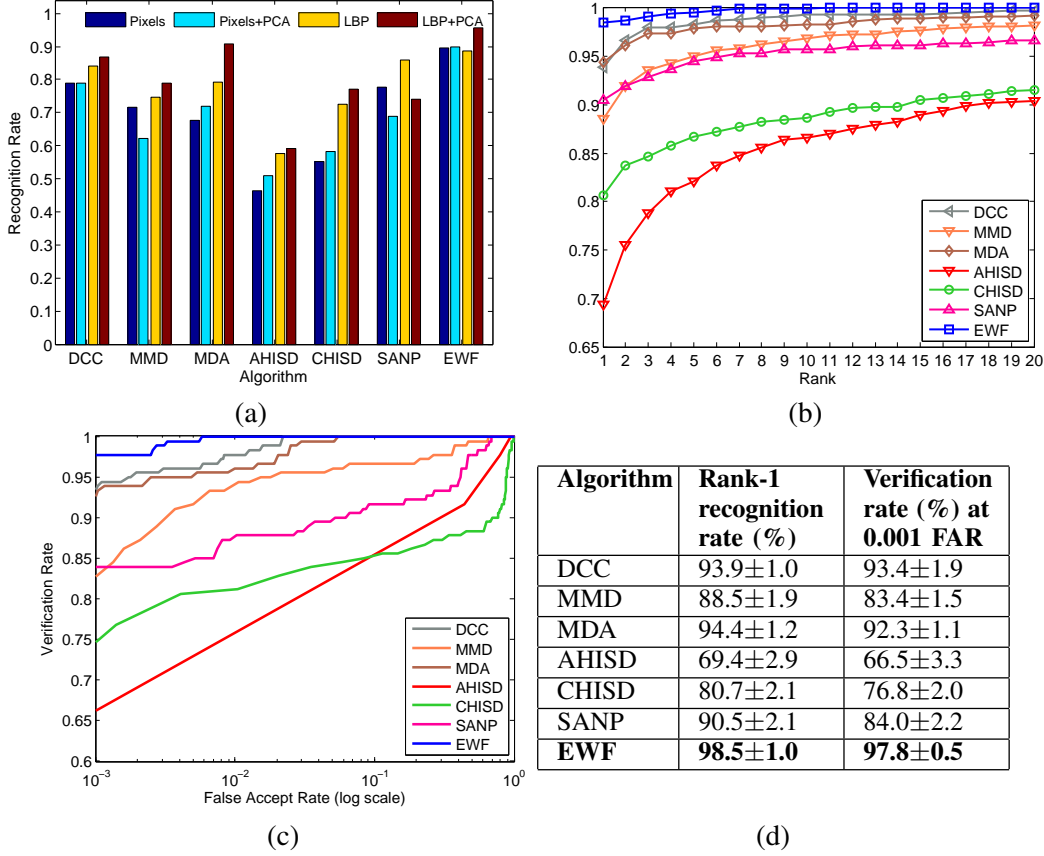(c)   (d)

Figure 12: Results of 10-fold experiments on the visible range MBGC video database. (a) Average recognition rates of individual algorithms on individual features. (b) CMC and (c) ROC curves of individual algorithms after feature score fusion. (d) Summarized results.

the NIR recognition rate of 99.9±0.1%, compared to visual spectrum recognition 98.5±1.0%. One may argue that the higher performance of NIR images is due to better image quality but this is balanced by the higher number of images per set in the visible spectrum image sets. The NIR illuminator gives better illumination, but at the same time induces non-linear specularities. LBP filtering has remained more successful in removing the uneven illumination variations therefore, LBP features have shown better performance than raw pixels and raw pixels + PCA. The reason why LBP+PCA resulted in lower performance is because an accurate PCA subspace could not be estimated from fewer images.

The performance of each algorithm after feature score fusion using EWF fusion is shown in Fig. 13-b and Fig. 13-c. The CMC and ROC curves of the
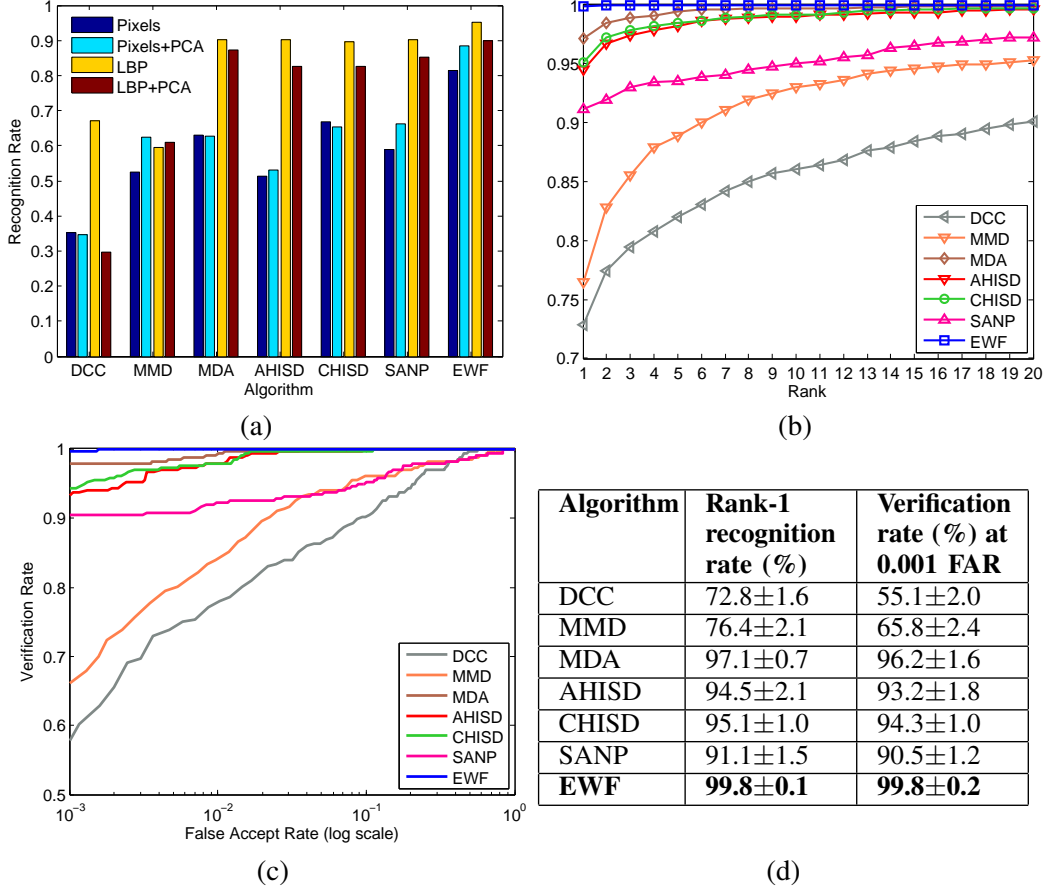
20

Figure 13: Results of 10-fold experiments on MBGC NIR video database: (a) Average recognition rates of individual algorithms on individual features. (b) CMC and (c) ROC curves of individual algorithms after feature score fusion. (d) Summarized results.

sample based techniques are above the structure based techniques. The proposed two stage fusion (FSF+CSF) using EWF achieved a rank-1 identification rate of 99.9±0.1% and a verification rate of 99.8±0.2% at 0.001 FAR (Table 13-d). The previous best identification rate on this database was reported by Woodard *et al.* [14] which is 96.5% using only 85 subjects of the database (Table 2). We achieved higher identification rate while using 114 subjects from the database which is a more challenging scenario. We excluded 35 subjects because only four periocular regions per subject could be automatically detected. Moreover, Woodard *et al.* [14] manually extracted the periocular regions whereas we performed this task automatically.
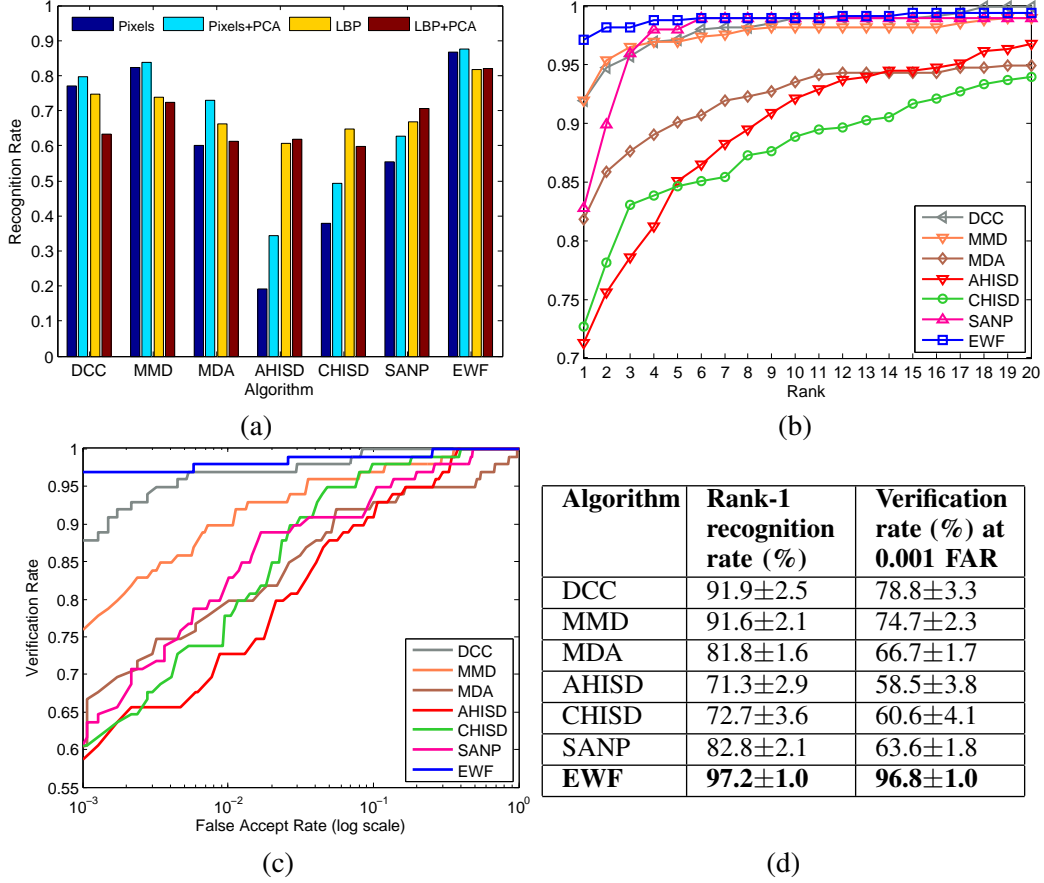
Figure 14: Results of 10-fold experiments on CMU Hyperspectral database: (a) Average recognition rates of individual algorithms on individual features. (b) CMC and (c) ROC curves of individual algorithms after feature score fusion. (d) Summarized results.

Algorithm	Rank-1 recognition rate (%)	Verification rate (%) at 0.001 FAR
DCC	91.9±2.5	78.8±3.3
MMD	91.6±2.1	74.7±2.3
MDA	81.8±1.6	66.7±1.7
AHISD	71.3±2.9	58.5±3.8
CHISD	72.7±3.6	60.6±4.1
SANP	82.8±2.1	63.6±1.8
EWF	**97.2±1.0**	**96.8±1.0**

CMU Hyperspectral Face database: For thoroughness, we also performed experiments on the hyperspectral face cubes. To the best of our knowledge, this is the first work on hyperspectral periocular biometrics that includes as many as 65 bands covering the visible and infra-red range (from 450nm to 1100nm at 10nm step). The results on this database are compiled in Fig. 14. On individual features (Fig. 14-a), structure based algorithms have performed better than sample based ones (Fig. 14-a) because the number of images per set are 130 leading to a good structure estimation. Pixels+PCA gives the best performance for structure based techniques and when feature score fusion is performed (the last set of bars in 14-a). Figure 14-b and Fig. 14-c show the CMC and ROC curves of individual

Table 4: Average identification Rates and std (%) of 10-fold experiments using different combinations of Feature Score Fusion (FSF) and Classifier Score Fusion (CSF).

MBGC Visible range video database							
CSF / **FSF** — Prod	Max	Median	Voting	Sum	SWV	WMV	EWF
Prod	33.6±1.2	94.7±1.1	97.5±0.6	78.5±1.8	95.6±1.0	96.3±1.1	97.5±0.6	97.1±1.1
Max	63.7±1.1	77.7±2.2	94.8±0.9	93.0±0.7	95.4±1.5	95.1±1.2	96.3±1.1	97.1±0.7
Median	73.7±3.7	91.7±2.5	97.5±0.8	96.5±0.7	97.1±1.1	97.0±1.8	97.2±1.4	97.2±1.4
Voting	-	-	-	97.8±0.4	-	-	-	
Sum	69.9±3.5	91.5±1.6	97.5±0.8	97.0±0.7	97.3±1.2	97.4±1.2	97.3±1.5	97.3±1.5
SWV	68.7±3.6	90.3±1.9	97.0±0.8	96.7±1.5	96.7±1.5	94.3±1.8	96.7±1.5	97.2±1.4
EWF	74.1±3.6	91.5±2.5	96.9±1.6	97.7±0.1	97.4±1.4	97.4±1.4	97.7±1.5	**98.5±1.0**
MBGC NIR video database							
Prod	18.4±2.5	91.8±1.3	97.5±0.9	73.6±2.0	97.4±1.0	97.4±1.0	99.1±0.5	99.0±0.7
Ma	65.2±1.7	83.3±2.5	96.9±1.3	97.2±1.0	97.9±0.8	97.7±0.8	98.7±0.8	99.0±0.5
Median	75.5±2.2	93.3±1.7	98.0±0.7	98.2±0.9	98.2±0.5	98.7±0.8	99.1±0.5	99.1±0.5
Voting	-	-	-	97.6±0.9	-	-	-	
Sum	73.0±2.0	91.5±2.0	98.1±0.7	98.2±0.9	98.1±0.5	97.7±0.8	99.6±0.4	99.6±0.4
SWV	71.8±1.7	90.7±1.3	97.9±0.8	97.7±0.8	98.4±0.4	96.8±0.9	98.4±0.4	97.7±0.8
WMV	73.2±2.1	93.1±1.7	98.4±0.6	97.7±0.8	98.1±0.5	98.7±0.8	98.4±0.4	99.6±0.4
EWF	73.9±2.9	93.7±2.0	98.4±0.6	98.8±0.8	98.3±0.4	98.7±0.8	99.6±0.4	**99.9±0.1**

algorithms after feature score fusion (FSF) of all features. The proposed two stage fusion (FSF+CSF) using feature and classifier score fusion with EWF achieves the highest rank-1 recognition rate of 97.2±1.0% and verification rate of 96.8±1.0% at 0.001 FAR. Figure 14-d summarizes the results.

7.3. Recognition Rate Comparison of Different Fusion Schemes

We observe that score fusion at two stages produces better results than a single stage fusion. We have explored all combinations of different score fusion schemes discussed in Section 6 at both stages (feature score and classifier score). We also compare the proposed EWF fusion algorithm to the closely related classifier fusion methods, Simple Weighted Vote (SWV) and the Weighted Majority Vote (WMV), presented by Seco et al. in [42]. Table 4 summarizes the accuracies for different combinations of Feature Score Fusion (FSF) and Classifier Score Fusion (CSF). The proposed inverse Error Weighted Fusion (EWF) outperforms the other fusion methods when used at both stages. The accuracy of the product rule and min rule when used at both levels is the minimum because these rules are affected by the lowest performers. The weights of different features and classifiers are learned in the training stage according to the performance on the training data.
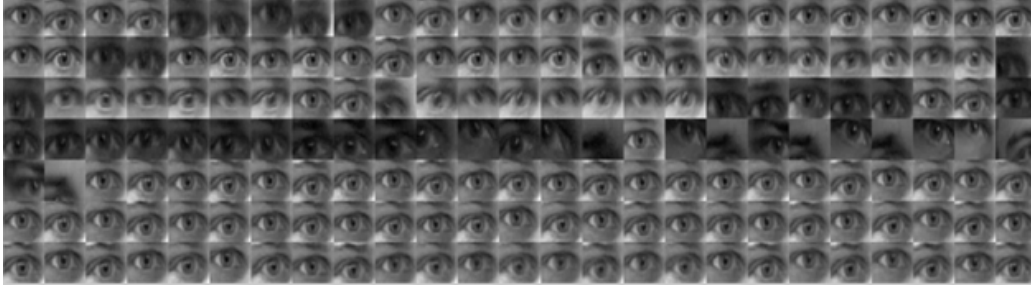
23

Figure 15: Sample images from an image set of the visible range MBGC video dataset. The image set contains noise in the form of misdetected periocular regions.

Therefore, the poor performers get lower weights in the test stage and the overall accuracy significantly improves over the sum rule.

7.4. Effect of Periocular Region Misdetection on Accuracy

The accuracy of periocular region detection may affect the performance of periocular recognition. The accuracy of the proposed algorithm will depend on the total images in a set and the ratio of misdetections to the correct detection. Figure 15 shows samples from an image set of the visible range MBGC video database. Alongside correct detections, the set also contains around 14% noisy samples in the form of misdetected periocular regions. In our experiments, this set was misclassified by the AHISD algorithm because the affine hull of the set was not able to accommodate for the outlier samples. On the other hand, the same image set was classified correctly by the DCC algorithm as DCC models the image set structure using the most dominant subspace basis. However, if most of the set samples contain large detection errors and there are not enough correct samples available for accurate set structure estimation, then the accuracy of the set based periocular recognition may slightly degrade, We observe that the proposed framework is robust to approximately 17% outlier samples in image sets without significant change in accuracy.

7.5. Execution Time

We compare the execution times of the six image set classification algorithms on the five databases. Table 5 shows the average execution time for matching one probe set with 114 gallery sets, on a 3.4GHz CPU with 8GB RAM for the MBGC NIR database. The relative execution times for the other databases are similar. We observe that the sample based techniques are computationally more expensive

24

Table 5: Execution time in seconds for matching a single probe image-set with 114 gallery image sets in the MBGC NIR database using 600-d LBP features.

Algorithm	DCC	MMD	MDA	AHISD	CHISD	SANP	FSF+CSF
Train	13.11	6.16	21.95	N/A	N/A	N/A	164.91
Test	0.14	0.11	0.09	0.15	2.31	5.23	32.12

than the structure based techniques. The overall execution time of the proposed two stage fusion approach is the sum of execution time for all algorithms. Note that the execution time for SANP is lower than reported by the original authors [22] and in our preliminary work [23]. This was achieved by normalizing all features to unit norm. This significantly improves the convergence speed of the Accelerated Proximal Gradient method used to find the sparse approximated nearest points in the SANP algorithm without affecting the accuracy.

8. Conclusion and Future Work

This paper presented a fully automatic algorithm for periocular biometric based person identification. An extensive study of the periocular biometric in the visible, NIR and the hyperspectral images was carried out to find the feasibility of this emerging biometric. Periocular biometric recognition was formulated as an image set classification problem. Such a formulation provides robustness to intra person variations such as pose, eyelid and eyeball movements. Periocular image sets were automatically extracted and represented by four feature types and six image set classification algorithms were used for classification. Fusion was performed at two stages namely feature scores and classifier scores. Different fusion schemes were studied and the inverse Error Weighted Fusion (EWF) was found to yield the best performance. Experiments were performed on four public databases, including MBGC Visible and NIR, CMU Hyperspectral and UBIPr. Recognition rates significantly higher than the current state of the art were achieved. Our results demonstrate the feasibility of image-set based periocular biometrics for real world person identification in scenarios where the full face is not visible or may have changed. Currently we report the computational time on 600 dimensional LBP features. The computational time can be significantly reduced by reducing the dimensionality of the features. In future we plan to carry out a thorough comparison of periocular biometrics with the full face biometrics.

Acknowledgements

References

[1] P. Phillips, P. Flynn, T. Scruggs, K. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, W. Worek, Overview of the face recognition grand challenge, in: IEEE International Conference on Computer Vision and Pattern Recognition, 2005, volume 1, pp. 947–954.

[2] L. Denes, P. Metes, Y. Liu, Hyperspectral Face Database, Technical Report CMU-RI-TR-02-25, Robotics Inst., Pittsburgh, PA, 2002.

[3] C. Padole, H. Proenca, Periocular recognition: Analysis of performance degradation factors, in: International Conference on Biometrics, 2012, pp. 439–445.

[4] F. Juefei-Xu, K. Luu, M. Savvides, T. Bui, C. Suen, Investigating age invariant face recognition based on periocular biometrics, in: International Joint Conference on Biometrics, 2011, pp. 1–7.

[5] J.-X. Du, C.-M. Zhai, Y.-Q. Ye, Face aging simulation and recognition based on NMF algorithm with sparseness constraints, Neurocomputing 116 (2013) 250 – 259.

[6] K. W. Bowyer, K. Hollingsworth, P. J. Flynn, Image understanding for iris biometrics: A survey, Computer Vision and Image Understanding 110 (2008) 281–307.

[7] Y. Song, W. Cao, Z. He, Robust iris recognition using sparse error correction model and discriminative dictionary learning, Neurocomputing 137 (2014) 198 – 204.

[8] J. Huang, X. You, Y. Yuan, F. Yang, L. Lin, Rotation invariant iris feature extraction using gaussian markov random fields with non-separable wavelet, Neurocomputing 73 (2010) 883–894.

[9] U. Park, R. Jillela, A. Ross, A. Jain, Periocular biometrics in the visible spectrum, IEEE Trans. on Info. Forensics and Sec. 6 (2011) 96–106.

[10] U. Park, A. Ross, A. Jain, Periocular biometrics in the visible spectrum: A feasibility study, in: IEEE International Conference on Biometrics: Theory, Applications, and Systems, 2009, pp. 1–6.

[11] K. Oh, B.-S. Oh, K.-A. Toh, W.-Y. Yau, H.-L. Eng, Combining sclera and periocular features for multi-modal identity verification, Neurocomputing 128 (2014) 185–198.

[12] K. Hollingsworth, S. Darnell, P. Miller, D. Woodard, K. Bowyer, P. Flynn, Human and machine performance on periocular biometrics under near-infrared light and visible light, IEEE Trans. on Info. Forensics and Sec. 7 (2012) 588–601.

[13] D. Woodard, S. Pundlik, J. Lyle, P. Miller, Periocular region appearance cues for biometric identification, in: IEEE International Conference on Computer Vision and Pattern Recognition Workshops, 2010, pp. 162–169.

[14] D. Woodard, S. Pundlik, P. Miller, R. Jillela, A. Ross, On the fusion of periocular and iris biometrics in non-ideal imagery, in: International Conference on Pattern Recognition, 2010, pp. 201–204.

[15] J. Adams, D. Woodard, G. Dozier, P. Miller, K. Bryant, G. Glenn, Genetic-based type ii feature extraction for periocular biometric recognition: Less is more, in: International Conference on Pattern Recognition, 2010, pp. 205–208.

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

[17] Z. Cui, H. Chang, S. Shan, B. Ma, X. Chen, Joint sparse representation for video-based face recognition, Neurocomputing 135 (2014) 306–312.

[18] H. Cevikalp, B. Triggs, Face recognition based on image sets, in: IEEE International Conference on Computer Vision and Pattern Recognition, 2010, pp. 2567–2573.

[19] T.-K. Kim, J. Kittler, R. Cipolla, Discriminative learning and recognition of image set classes using canonical correlations, IEEE Trans. Pattern Anal. Mach. Intell. 29 (2007) 1005–1018.

[20] R. Wang, S. Shan, X. Chen, W. Gao, Manifold-manifold distance with application to face recognition based on image set, in: IEEE International Conference on Computer Vsion and Pattern Recognition, 2008, pp. 1–8.

[21] R. Wang, X. Chen, Manifold discriminant analysis, in: IEEE International Conference on Computer Vision and Pattern Recognition, 2009, pp. 429–436.

[22] Y. Hu, A. Mian, R. Owens, Face recognition using sparse approximated nearest points between image sets, IEEE Trans. Pattern Anal. Mach. Intell. 34 (2012) 1992–2004.

[23] M. Uzair, A. Mahmood, A. Mian, C. McDonald, Periocular biometric recognition using image sets, in: IEEE Workshop on the Applications of Computer Vision, 2013, pp. 246–251.

[24] A. H. T. Ahonen, M. Pietikainen, Face Description with Local Binary Patterns: Application to Face Recognition, IEEE Trans. Pattern Anal. Mach. Intell. 28 (2006) 2037–2041.

[25] P. E. Miller, A. W. Rawls, S. J. Pundlik, D. L. Woodard, Personal identification using periocular skin texture, in: Proceedings of the ACM Symposium on Applied Computing, 2010, pp. 1496–1500.

[26] NIST, Multiple Biometric Grand Challenge (MBGC) dataset http://face.nist.gov/mbgc/, 2008.

[27] S. Bharadwaj, H. Bhatt, M. Vatsa, R. Singh, Periocular biometrics: When iris recognition fails, in: IEEE International Conference on Biometrics: Theory, Applications, and Systems, 2010, pp. 1–6.

[28] H. Proenca, S. Filipe, R. Santos, J. Oliveira, L. Alexandre, The ubiris.v2: A database of visible wavelength iris images captured on-the-move and at-a-distance, IEEE Trans. Pattern Anal. Mach. Intell. 32 (2010) 1529–1535.

[29] V. Boddeti, J. Smereka, B. Kumar, A comparative evaluation of iris and ocular recognition methods on challenging ocular images, in: International Joint Conference on Biometrics, 2011, pp. 1–8.

[30] NIST, Face and Ocular Challenge Series (FOCS) dataset http://www.nist.gov/itl/iad/ig/focs.cfm, 2010.

[31] F. Juefei-Xu, M. Savvides, Unconstrained periocular biometric acquisition and recognition using cots ptz camera for uncooperative and non-cooperative subjects, in: IEEE Workshop on the Applications of Computer Vision, 2012, pp. 201–208.

[32] V. P. Pauca, M. Forkin, X. Xu, R. Plemmons, A. A. Ross, Challenging ocular image recognition, BTHI, SPIE 8029 (2011) 80291V–80291V–13.

[33] F. Alonso-Fernandez, J. Bigün, Periocular recognition using retinotopic sampling and gabor decomposition, in: European Conference on Computer Vision, Workshops (2), 2012, pp. 309–318.

[34] BIT, CASIA Iris Image Database http://biometrics.idealtest.org, 2013.

[35] J. Fierrez-aguilar, J. Ortega-garcia, D. Torre-toledano, J. Gonzalez-rodriguez, Biosec baseline corpus: A multimodal biometric database, Pattern Recognition (2007) 1389–1392.

[36] P. Viola, M. Jones, Robust real-time face detection, International Journal of Computer Vision 57 (2004) 137–154.

[37] L. Masek, Recognition of Human Iris Patterns for Biometric Identification, Technical Report, The University of Western Australia, 2003.

[38] Y. Lee, R. Micheals, P. Phillips, Improvements in video-based automated system for iris recognition, in: IEEE Workshop on Motion and Video Computing (WMVC), 2009, pp. 1–8.

[39] B. Yang, S. Chen, A comparative study on local binary pattern LBP based face recognition: LBP histogram versus LBP image, Neurocomputing 120 (2013) 365 – 379.

[40] J. Kittler, M. Hatef, R. P. W. Duin, J. Matas, On combining classifiers, IEEE Trans. Pattern Anal. Mach. Intell. 20 (1998) 226–239.

[41] A. Sinha, H. Chen, D. Danu, T. Kirubarajan, M. Farooq, Estimation and decision fusion: A survey, Neurocomputing 71 (2008) 2650–2656.

[42] F. Moreno Seco, J. Inesta, P. Ponce de Leon, L. Mico, Comparison of classifier fusion methods for classification in pattern recognition tasks, in: International Workshops on Structural and Syntactic Pattern Recognition 2006, pp. 705–713.