Periocular Biometrics in the Visible Spectrum

Unsang Park, *Member, IEEE*, Raghavender Reddy Jillela, *Student Member, IEEE*, Arun Ross, *Senior Member, IEEE*, and Anil K. Jain, *Fellow, IEEE*

***Abstract*—The term periocular refers to the facial region in the immediate vicinity of the eye. Acquisition of the periocular biometric is expected to require less subject cooperation while permitting a larger depth of field compared to traditional ocular biometric traits (viz., iris, retina, and sclera). In this work, we study the feasibility of using the periocular region as a biometric trait. Global and local information are extracted from the periocular region using texture and point operators resulting in a feature set for representing and matching this region. A number of aspects are studied in this work, including the 1) effectiveness of incorporating the eyebrows, 2) use of side information (left or right) in matching, 3) manual versus automatic segmentation schemes, 4) local versus global feature extraction schemes, 5) fusion of face and periocular biometrics, 6) use of the periocular biometric in partially occluded face images, 7) effect of disguising the eyebrows, 8) effect of pose variation and occlusion, 9) effect of masking the iris and eye region, and 10) effect of template aging on matching performance. Experimental results show a rank-one recognition accuracy of 87.32% using 1136 probe and 1136 gallery periocular images taken from 568 different subjects (2 images/subject) in the Face Recognition Grand Challenge (version 2.0) database with the fusion of three different matchers.**

***Index Terms*—Biometrics, face, fusion, gradient orientation histogram, local binary patterns, periocular recognition, scale invariant feature transform.**

I. Introduction

BIOMETRICS is the science of establishing human identity based on the physical or behavioral traits of an individual [2], [3]. Several biometric traits such as face, iris, hand geometry, and fingerprint have been extensively studied in the literature and have been incorporated in both government and civilian identity management applications. Recent research in biometrics has explored the use of other human characteristics such as gait [4], conjunctival vasculature [5], knuckle joints [6], etc., as supplementary biometric evidence to enhance the performance of classical biometric systems.

Manuscript received April 19, 2010; revised October 11, 2010; accepted November 06, 2010. Date of publication December 03, 2010; date of current version February 16, 2011. An earlier version of this work appeared in the Proceedings of the International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2009. The work of R. R. Jillela and A. Ross was supported by IARPA BAA 09-02 through the Army Research Laboratory under Cooperative Agreement W911NF-10-2-0013. The work of A. K. Jain was supported in part by the World Class University (WCU) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (R31-10008). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing official policies, either expressed or implied, of IARPA, the Army Research Laboratory, or the U.S. Government. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Fabio Scotti.

U. Park is with the Computer Science and Engineering Department, Michigan State University, East Lansing, MI 48824 USA (e-mail: parkunsa@cse.msu.edu).

R. R. Jillela and A. Ross are with the Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26505 USA (e-mail: raghavender.jillela@mail.wvu.edu; arun.ross@mail.wvu.edu).

A. K. Jain is with the Computer Science and Engineering Department, Michigan State University, East Lansing, MI 48824 USA, and also with the Brain and Cognitive Engineering Department, Korea University, Seoul 136-713, Korea (e-mail: jain@cse.msu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2010.2096810

(a)[1] (b) (c) (d)

[1] http://jamblichus.files.wordpress.com/2009/10.

Fig. 1. Ocular biometric traits: (a) retina, (b) iris, (c) conjunctiva [10], and (d) periocular.

Ocular biometrics (see Fig. 1) has made rapid strides over the past few years primarily due to the significant progress made in iris recognition [7], [8]. The iris is the annular colored structure in the eye surrounding the pupil and its function is to regulate the size of the pupil thereby controlling the amount of light incident on the retina. The surface of the iris exhibits a very rich texture due to the numerous structures evident on its anterior layers. The random morphogenesis of the textural relief of the iris and its apparent stability over the lifetime of an individual (that has, however, been challenged recently), have made it a very popular biometric. Both technological and operational tests conducted under predominantly constrained conditions have demonstrated the uniqueness of the iris texture to an individual and its potential as a biometric in large-scale systems enrolling millions of individuals [7], [9]. Besides the iris, other ocular biometric traits such as retina and conjunctiva have been investigated for human recognition.

In spite of the tremendous progress made in ocular biometrics, there are significant challenges encountered by these systems:

1) The iris is a moving object with a small surface area that is located within the independently movable eyeball. The eyeball itself is located within another moving object—the head. Therefore, reliably localizing the iris in eye images obtained at a distance in unconstrained environments can be difficult [11]. Furthermore, since the iris is typically imaged in the near-infrared (NIR) portion (700–900 nm) of the electromagnetic (EM) spectrum, appropriate invisible lighting is required to illuminate it prior to image acquisition.
2) The size of an iris is very small compared to that of a face. Face images acquired with low resolution sensors or large standoff distances offer very little or no information about iris texture.

3) Even under ideal conditions characterized by favorable lighting conditions and an optimal standoff distance, if the subject blinks or closes his eye, the iris information cannot be reliably acquired.
4) Retinal vasculature cannot be easily imaged unless the subject is cooperative. In addition, the imaging device has to be in close proximity to the eye.
5) While conjunctival vasculature can be imaged at a distance, the curvature of the sclera, the specular reflections in the image, and the fineness of the vascular patterns can confound the feature extraction and matching modules of the biometric system [10].

In this work, we attempt to mitigate some of these concerns by considering a small region around the eye as an additional biometric. We refer to this region as the periocular region. We explore the potential of the periocular region as a biometric in color images pertaining to the visible spectral band. Some of the benefits in using the periocular biometric trait are as follows:

1) In images where the iris cannot be reliably obtained (or used), the surrounding skin region may be used to either confirm or refute an identity. Blinking or off-angle poses are common sources of noise during iris image acquisition.
2) The periocular region represents a good trade-off between using the entire face region or using only the iris texture for recognition. When the entire face is imaged from a distance, the iris information is typically of low resolution. On the other hand, when the iris is imaged at close quarters, the entire face may not be available thereby forcing the recognition system to rely only on the iris. However, the periocular biometric can be useful over a wide range of distances.
3) The periocular region can offer information about eye shape that may be useful as a soft biometric [12], [13].
4) When portions of the face pertaining to the mouth and nose are occluded, the periocular region may be used to determine the identity.
5) The design of a newer sensor is not necessary as both periocular and face regions can be obtained using a single sensor.

Only a few studies have been published on the use of the periocular region as a biometric. Park *et al.* [1] used both local and global image features to match periocular images acquired in the visible spectra and established its utility as a soft biometric trait. In their work, the authors also investigated the role of the eyebrow on the overall matching accuracy. Miller *et al.* [14] used scale and rotation invariant local binary pattern (LBP) to encode and match periocular images. They explicitly masked out the iris and sclera before the feature extraction process. In this work, our experiments are based on a significantly larger gallery and probe database than what was used by Miller *et al.* Further, we store only one image per eye in the gallery. We also automatically extract the periocular region from full face images.

Since periocular biometrics is a relatively new area of research, it is essential to conduct a comprehensive study in order to understand the uniqueness and stability of this trait. Some of the most important issues that have to be addressed include the following:

1) Region definition: What constitutes the periocular region? Should the region include the eyebrows, iris, and the sclera, or should it exclude some of these components?
2) Feature Extraction: What are the best features for representing these regions? How can these features be reliably extracted?
3) Matching: How do we match the extracted features? Can a coarse classification be performed prior to matching in order to reduce the computational burden?
4) Image Acquisition: Which spectrum band (visible or NIR) is more beneficial for matching periocular biometrics?
5) Fusion: What other biometric traits are suitable to be fused with the periocular information? What fusion techniques can be used for this process?

In this work, we carefully address some of the above listed issues. The experiments conducted here discuss the performance of periocular matching techniques across different factors such as region segmentation, facial expression, and face occlusion. Experiments are conducted in the visible spectrum using images obtained from the Face Recognition Grand Challenge (FRGC 2.0) database [15]. The eventual goal would be to use a multispectral acquisition device to acquire periocular information in both visible and NIR spectral bands [16], [17]. This would facilitate combining the iris texture with the periocular region thereby improving the recognition performance.

II. Periocular Biometrics

The proposed periocular recognition process consists of a sequence of operations: image alignment (for the global matcher described in the next section), feature extraction, and matching. We adopt two different approaches to the problem: one based on global information and the other based on local information. The two approaches use different methods for feature extraction and matching. In the following section, the characteristics of these two approaches are described.

A. Global versus Local Matcher

Most image matching schemes can be categorized as being global or local based on whether the features are extracted from the entire image (or a region of interest) or from a set of local regions. Representative global image features include those based on color, shape, and texture [18]. Global features are typically represented as a fixed length vector, and the matching process simply compares these fixed length vectors, which is very time efficient. On the other hand, a local feature-based approach first detects a set of key points and encodes each of the key points using the surrounding pixel values, resulting in a local key point descriptor [19], [20]. Then, the number of matching key points between two images is used as the match score. Since the number of key points varies depending on the input image, two sets of key points from two different images cannot be directly compared. Therefore, the matching scheme has to compare each key point from one image against all the key points in the other image, thereby increasing the time for matching. There have been efforts to achieve constant time matching using the bag of words representation [21]. In terms of matching accuracy, local feature-based techniques have shown better performance [22]–[24].

When all the available pixel values are encoded into a feature vector (as is the case when global features are used), it becomes more susceptible to image variations especially with respect to

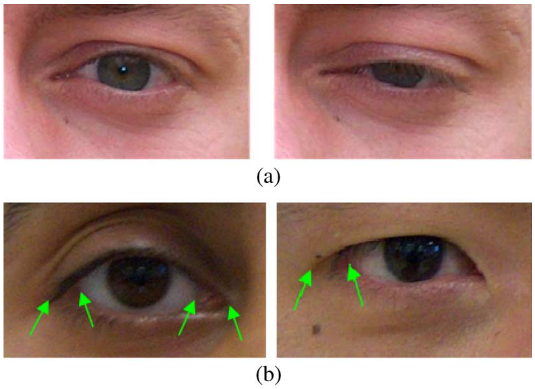


Fig. 2. Example images showing difficulties in periocular image alignment. (a) Illustrating eyelid movement; (b) presence of multiple corner candidates.

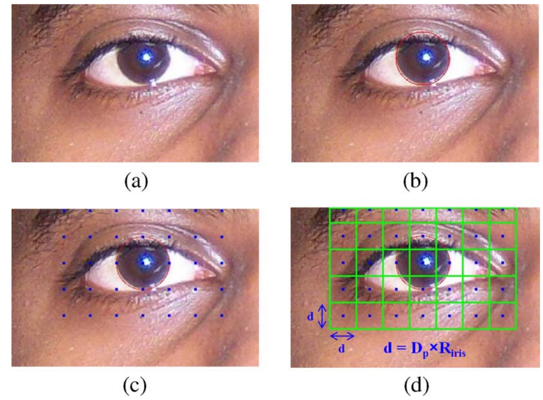


Fig. 3. Schematic of image alignment and feature extraction process. (a) Input image; (b) iris detection; (c) interest point sampling; (d) interest region sampling.

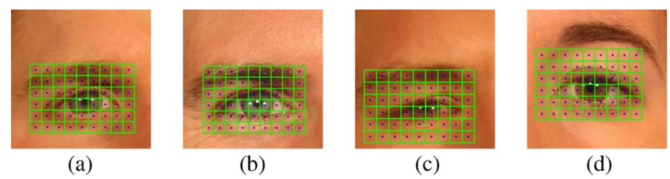


Fig. 4. Example images showing interest points used by the global matcher over the periocular region. Eyebrows are included in (a), (b), and (c), but not in (d).

geometric transformations and spatial occlusions. The local feature-based approach, on the other hand, is more robust to such variations because only a subset of distinctive regions is used to represent an image. This has made local feature-based approach to image retrieval very attractive.

B. Image Alignment

Periocular images across subjects contain some common components (e.g., iris, sclera, and eyelids) that can be represented in a common coordinate system. Once a common area of interest is localized, a global representation scheme can be used. The iris or eyelids are good candidates for the alignment process. Even though both the iris and eyelids exhibit motion, such variations are not significant in the periocular images used in this research. While frontal iris detection can be performed fairly well due to the approximately circular geometry of the iris and the clear contrast between the iris and sclera, accurate detection of the eyelids is more difficult. The inner and outer corners of the eye can also be considered as anchor points, but there can be multiple candidates as shown in Fig. 2. Therefore, we primarily use the iris for image alignment. A public domain iris detector based on the Hough transformation is used for localizing the iris [25]. The iris can be used for translation and scale normalization of the image, but not for rotation normalization. However, we overcome the small rotation variations using a rotation tolerant feature representation. The iris-based image alignment is only required by the global matching scheme. The local matcher does not require image alignment because the descriptors corresponding to the key points can be independently compared with each other.

C. Feature Extraction

We extract global features using all the pixel values in the detected region of interest that is defined with respect to the iris. The local features, on the other hand, are extracted from a set of characteristic regions. From the center C_{iris} and the radius R_{iris} of the iris, multiple $(= n_{pi})$ interest points $p_1, p_2, \ldots, p_{n_{pi}}$ are selected within a rectangular window defined around C_{iris} with a width of $6 \times R_{\text{iris}}$ and a height of $4 \times R_{\text{iris}}$, as shown in Fig. 3. The number of interest points is decided based on the sampling frequency $(1/D_p)$ which is inversely proportional to the distance between interest points, $D_p \times R_{\text{iris}}$. For each interest point p_i, a rectangular region r_i is defined. The dimension of each rectangle (r_i) in the ROI is of size $(D_p \times R_{\text{iris}})$ by $(D_p \times R_{\text{iris}})$. When $D_p = 1$, the size of the rectangle becomes $R_{\text{iris}} \times R_{\text{iris}}$ [see Fig. 3(d)]. The interest points used by the global matcher cover the eyebrows over 70% of the time as shown in Fig. 4. In a few cases, the region does not include the entire eyebrow. However, this does not affect the overall accuracy because the eyebrows are included in most cases and the SIFT uses the entire area of the image including the eyebrows. We construct the key point descriptors from r_i and generate a full feature vector by concatenating all the descriptors. Such a feature representation scheme using multiple image partitions is regarded as a local feature representation in some of the image retrieval literature [26], [27]. However, we consider this as a global representation scheme because all the pixel values are used in the representation without considering the local distinctiveness of each region.

Mikilajczyk *et al.* [20] have categorized the descriptor types as distribution-based, spatial frequency-based, and differential-based. We use two well-known distribution-based descriptors: gradient orientation (GO) histogram [28] and local binary pattern (LBP) [29]. We quantize both GO and LBP into eight distinct values to build an eight bin histogram. The eight bin histogram is constructed from a partitioned subregion and concatenated across the various subregions to construct a full feature vector. A Gaussian blurring with a standard deviation σ is applied prior to extracting features using the GO and LBP methods in order to smooth variations across local pixel values. This subpartition-based histogram construction scheme has been successfully used in SIFT [22] for the object recognition problem. The local matcher first detects a set of salient key points in scale space. Features are extracted from the bounding boxes for each

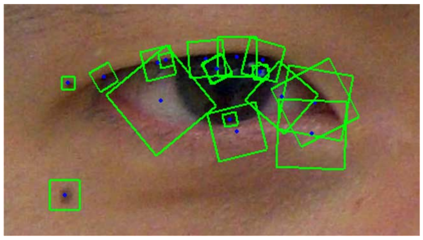


Fig. 5. Examples of local features and bounding boxes for descriptor construction in SIFT. Each bounding box is rotated with respect to the major orientation or gradient.

key point based on the gradient magnitude and orientation. The size of the bounding box is proportional to the scale (i.e., the standard deviation of the Gaussian kernel in scale space construction). Fig. 5 shows the detected key points and the surrounding boxes on a periocular image. While the global features are only collected around the eye, the local features are collected from all salient regions such as facial marks. Therefore, the local matcher is expected to provide more distinctiveness across subjects.

Once a set of key points are detected, these points can be used directly as a measure of image matching based on the goodness of geometrical alignment. However, such an approach does not take into consideration the rich information embedded in the region around each interest point. Moreover, when images are occluded or subjected to affine transformations, it will be beneficial to match individual interest points rather than relying on the entire set of interest points. We used a publicly available SIFT implementation [30] as the local matcher.

D. Match Score Generation

For the global descriptor, the Euclidean distance is used to calculate the matching scores. The distance ratio-based matching scheme [22] is used for the local matcher (SIFT).

E. Parameter Selection for Each Matcher

The global descriptor varies depending on the choice of σ and the frequency of sampling interest points $1/D_p$. SIFT has many parameters that affect its performance. Some of the representative parameters are the number of octaves (n_o), number of scales (n_s), and the cutoff threshold value t_{ex} related to the contrast of the extrema points. The absolute value of each extrema point in the Difference of Gaussian (DOG) space needs to be larger than t_{ex} to be selected as a key point. We construct a number of different descriptors for both the global and local schemes by choosing a set of values for σ, D_p, n_o, n_s, and t_{ex}. The set of parameters that results in the best performance in a training set is used on the test data for the global and local representations. We used a size of $6 \times R_{\text{iris}}$ by $4 \times R_{\text{iris}}$ (width $\times$ height) as the region for global feature extraction, 4 for σ, 0.7 (0.5) for D_p in GO (LBP), and 4, 4, 0.005 for n_0, n_s, and t_{ex}, respectively.

III. Experiments

A. Database

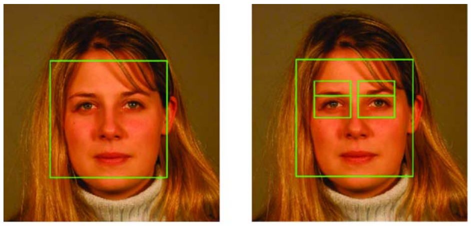
Two different databases were used in our experiments: DB1 and DB2. DB1 consists of 120 images (60 for probe and 60 for gallery) with two periocular images (left and right eye) per subject (30 subjects). Images in DB1 were captured in our laboratory using a NIKON COOLPIX P80 camera at a close distance, where a full image contains only the periocular region. The images in DB2 were taken from the FRGC (version 2.0) database [15]. FRGC 2.0 contains frontal images of subjects captured in a studio setting, with controlled illumination and background. A 4 Megapixel Canon PowerShot camera was used to capture the images [31], with a resolution of 1704×2272 pixels. The images are recorded in JPEG format with an approximate file size of 1.5 MB. The interpupillary distance, i.e., the distance between the centers of the two eyes of a subject in the FRGC images, is approximately 260 pixels. The FRGC database contains images with two different facial expressions for every subject: neutral and smiling. Fig. 6 shows two images of a subject with these two facial expressions. Three images (2 neutral and 1 smiling) of all the available 568 subjects in the FRGC database were used to form DB2, resulting in a total of 1704 face images. The FRGC database was assembled over a time period of 2 years with multiple samples of subjects captured in various sessions. However, the samples considered for the probe and gallery in this work belong to the same session, and do not have any time lapse between them. We used DB1 for parameter selection and then used these parameter values on DB2 for performance evaluation. We also constructed a small face image database including 40 different subjects collected at West Virginia University and Michigan State University to evaluate the perspective distortion effect on periocular biometrics.

(a) (b)

Fig. 6. Example images of a subject from the FRGC database [15] with (a) neutral and (b) smiling expressions.

B. Periocular Region Segmentation

It is necessary for the periocular regions to be segmented (cropped out) from full face images prior to feature extraction. Such a segmentation routine should be accurate, ensuring the presence of vital periocular information (eye, eyebrow, and the surrounding skin region) in the cropped image. Existing literature does not specify any guidelines for defining the periocular region. Therefore, segmentation can be performed to either include or discard the eyebrows from the periocular region. However, it can be hypothesized that the additional key points introduced by the inclusion of eyebrows can enhance recognition performance. To study the effect of the presence of eyebrows, periocular regions are segmented from the face images with and without eyebrows. The segmentation process was performed using the following techniques:



(a) (b)

Fig. 7. Example outputs of (a) face detection and (b) automatic periocular region segmentation. A set of heuristics is used to determine the periocular region based on the output of the face detector.

(a) (b) (c)

Fig. 8. Examples of incorrect outputs for face detection and periocular region segmentation.

- **Manual Segmentation**: The FRGC 2.0 database provides the coordinates of the centers of the two eyes and this was used to manually segment the periocular region. Such an approach was used to mitigate the effects of incorrect segmentation on the periocular matching performance.
- **Automatic Segmentation**: We used an automatic periocular segmentation scheme based on the OpenCV face detector [32] which is an implementation of the classical Viola-Jones algorithm [33]. Given an image, the OpenCV face detector outputs a set of spatial coordinates of a rectangular box surrounding the candidate face region. To automatically segment the periocular region, heuristic measurements are applied on the rectangular box specified by the face detector. These heuristic measurements are based on the anthropometry of the human face. Example outputs of the OpenCV face detector and the automatic periocular segmentation scheme are shown in Fig. 7.

It has to be noted that the success of periocular recognition directly depends on the segmentation accuracy. In the proposed automatic segmentation setup, the OpenCV face detector misclassified nonfacial regions as faces in 28 out of 1704 images in DB2 ($\approx$98.35% accuracy). Some of the wrongly classified outputs from the OpenCV face detector are shown in Fig. 8.

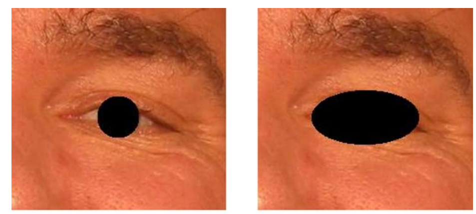
Based on the type of segmentation used (manual or automatic), and the decision to include or exclude the eyebrows from a periocular image, the following four datasets were generated from DB2:

- Dataset 1: Manually segmented, without eyebrows;
- Dataset 2: Manually segmented, with eyebrows;
- Dataset 3: Automatically segmented, without eyebrows;
- Dataset 4: Automatically segmented, with eyebrows.

The number of images obtained using the above-mentioned segmentation schemes and their corresponding sizes are listed in Table I. Note that manual segmentation generally crops the periocular region more tightly compared to automatic segmentation. Manual segmentation regions were normalized to a fixed size.

TABLE I
SIZE OF THE PERIOCULAR IMAGES OF THE DATABASES WITH RESPECT TO THE TYPE OF SEGMENTATION USED

DB & Segment. type	#img	Without eyebrows	With eyebrows
DB1, Manual	120	Varied from (438×198) to (940×453)	Varied from (438×290) to (940×787)
DB2, Manual	3408	241×121	241×226
DB2, Automatic	3408	Varied from (139×84) to (446×268)	Varied from (139×140) to (446×446)



(a) (b)

Fig. 9. Illustration of the mask on (a) iris and (b) entire eye region.

C. Masking Iris and Eye

As stated earlier, existing literature (both in the medical and biometric communities) does not offer a clear definition regarding the dimension of the periocular region. From an anatomical perspective, the term "peri-ocular" describes the surrounding regions of the eye. However, from a forensic/biometric application perspective, the goal is to improve the recognition accuracy by utilizing information from the shape of the eye, and the color and surface level texture of the iris. To study the effect of iris and sclera on the periocular recognition performance, we constructed two additional datasets by masking 1) the iris region only, and 2) the entire eye region of the images in Dataset 2 (see Fig. 9).

D. Recognition Accuracy

Using the aforementioned dataset configuration, the periocular recognition performance was studied. Each dataset is divided into a gallery containing 1 neutral image per subject, and a probe-set containing either a neutral or a smiling face image for each subject. Every probe image is compared against all the gallery images using the GO, LBP, and SIFT matching techniques. In this work, the periocular recognition performance is evaluated using 1) cumulative match characteristic (CMC) curves and rank-one accuracies, as well as 2) detection error trade-off (DET) curves and equal error rates (EERs).

Most biometric traits can be categorized into different classes, based on the nature (or type) of prominent patterns observed in their features. For example, fingerprints can be classified based on the pattern of ridges, while face images can be classified based on skin color. It is often desired to determine the class of the input probe image before the matching scheme is invoked.

TABLE II
RANK-ONE ACCURACIES FOR NEUTRAL–NEUTRAL MATCHING ON MANUALLY SEGMENTED DATASET (IN %) USING EYEBROWS AND L/R SIDE INFORMATION

	Without eyebrows		With eyebrows	
	Both sides	L-L and R-R	Both sides	L-L and R-R
GO	53.70%	53.78%	66.64%	66.64%
LBP	64.70%	64.88%	72.45%	72.45%
SIFT	62.06%	62.59%	78.79%	79.49%

*Number of probe and gallery images are both 1136.

TABLE III
RANK-ONE ACCURACIES FOR NEUTRAL–NEUTRAL MATCHING ON AUTOMATICALLY SEGMENTED DATASET (IN %) USING EYEBROWS AND L/R SIDE INFORMATION

	Without eyebrows		With eyebrows	
	Both sides	L-L and R-R	Both sides	L-L and R-R
GO	58.71%	59.10%	64.44%	64.43%
LBP	66.29%	66.46%	70.60%	70.51%
SIFT	59.07%	62.32%	77.82%	78.35%

*Number of probe and gallery images are both 1136.

This helps in reducing the number of matches required for identification by matching the probe image only with the gallery images of the corresponding class. This is also known as database indexing or filtering.

In the case of periocular recognition, the images can be broadly divided into two classes: left periocular region and the right periocular region. This classification is based on the location of the nose (left or right side) with respect to the inner corner of the eye in the periocular image. Periocular image classification can be potentially automated to enhance the recognition performance. However, in this work, this information is determined manually and used for observing the performance of the various matchers. Therefore, the following two different matching schemes were considered.

1) Retaining the side information: Left probe images are matched only against the left gallery images (L-L), and right probe images are matched only against right gallery images (R-R). The two recognition accuracies are averaged to summarize the performance of this setup.
2) Ignoring the side information: All probe periocular images are matched against all gallery images, irrespective of the side (L or R) they belong to.

This setup can also be understood as: (a) matching after performing classification and (b) matching without any classification.

For every dataset, all probe images containing a neutral expression are matched with their corresponding gallery images. Tables II and III indicate the rank-one accuracies obtained after employing the manual and automatic segmentation schemes, respectively.

From these results, it can be noticed that the recognition performance improves by incorporating the eyebrows in the periocular region. While the performance obtained using the automatic segmentation scheme is comparable to the manual segmentation scheme, slight degradation is observed due to incorrect face detection. The matching accuracies of GO and LBP are slightly better in automatically segmented images than those in the manually segmented images due to the partial inclusion of eyebrows during the automatic segmentation process. The best performance is observed when SIFT matching is used with periocular images containing eyebrows after manual segmentation (79.49%). The best performance under automatic segmentation is 78.35%.

TABLE IV
RANK-ONE ACCURACIES FOR NEUTRAL–SMILING MATCHING ON THE MANUALLY SEGMENTED DATASET (IN %) USING EYEBROWS AND L/R SIDE INFORMATION

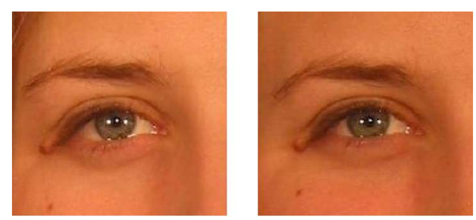
	Without eyebrows		With eyebrows	
	Both sides	L-L and R-R	Both sides	L-L and R-R
GO	46.21%	46.30%	61.09%	61.09%
LBP	58.63%	58.27%	67.52%	67.52%
SIFT	64.44%	64.79%	81.16%	81.60%

*Number of probe and gallery images are both 1136.

TABLE V
RANK-ONE ACCURACIES FOR NEUTRAL–SMILING MATCHING ON THE AUTOMATICALLY SEGMENTED DATASET (IN %) USING EYEBROWS AND L/R SIDE INFORMATION

	Without eyebrows		With eyebrows	
	Both sides	L-L and R-R	Both sides	L-L and R-R
GO	51.50%	51.67%	58.54%	58.54%
LBP	59.95%	59.95%	64.08%	64.00%
SIFT	65.93%	66.99%	77.82%	78.34%

*Number of probe and gallery images are both 1136.



(a) (b)

Fig. 10. Right side periocular regions segmented from the face images in Fig. 6 containing neutral and smiling expressions, respectively. Note that the location of the mole under the eye varies in the two images due to the change in expression.

To compare the effect of varying facial expression on periocular recognition, the probe images in all the four datasets in DB2 containing the smiling expression are matched against their corresponding gallery images. Tables IV and V summarize the rank-one accuracies obtained using the manual and automatic segmentation schemes for this experiment.

The neutral–smiling matching results support the initial hypothesis that recognition performance can be improved by including the eyebrows in the periocular region. Also, neutral–smiling matching has lower performance than neutral–neutral matching for the GO and LBP methods. In contrast, there is no performance degradation for the SIFT matcher on

TABLE VI
RANK-ONE ACCURACIES AFTER MASKING OUT IRIS OR EYE REGION (NEUTRAL–NEUTRAL, MANUAL SEGMENTATION, WITH EYEBROWS)

	Without mask		Masking only the iris		Masking the entire eye	
	Both sides	L-L and R-R	Both sides	L-L and R-R	Both sides	L-L and R-R
GO	66.64%	66.64%	65.40%	65.49%	61.53%	61.53%
LBP	72.45%	72.45%	72.54%	72.53%	69.63%	69.63%
SIFT	78.79%	79.49%	68.75%	69.80%	37.59%	40.84%

*Number of probe and gallery images are both 1136.

the neutral–smiling experiments. In general, the SIFT matcher is more robust to geometric distortions than the other two methods [22]. Examples of such geometric distortions are shown in Fig. 10.

Tables II–V show that the performances obtained with and without classification (based on retaining or ignoring the L/R side information) are almost similar. This indicates that periocular images provide sufficient diversity between the two classes (left and right) and probably exhibit very little interclass similarity.

Table VI reports the recognition results after masking out the iris region or the entire eye region. It is observed that the use of the entire periocular image (i.e., no masking) yields higher recognition accuracy. The performance drop of the local matcher (SIFT) is significantly larger than those of the global matchers. This is due to the reduced number of SIFT key points which are mostly detected around the edges and corners of the eye, and are lost after masking.

E. Score Level Fusion

The results described above provide a scope to further improve the recognition performance. To enhance the recognition performance, score level fusion schemes can be invoked. In this work, score level fusion is implemented to combine the match scores obtained from multiple classes (left and right) and multiple algorithms (GO, LBP, and SIFT). The fusion experiments are described below.

1) Score level fusion using multiple instances: The match scores of dataset 4, obtained by matching left-left and right-right are fused together using the simple sum rule (equal weights without any score normalization). This process is repeated for each of the three matchers, individually.
2) Score level fusion using multiple algorithms: The fused scores obtained in the above process for each matcher are fused together by the weighted sum rule after using the minimum–maximum normalization.

Figs. 11 and 12 show the CMC curves obtained for the multi-instance and multialgorithm fusion schemes using the neutral–neutral match scores of dataset 4. The DET curves and EERs for GO, LBP, and SIFT matchers by score level fusion of multiple instances are shown in Fig. 13. Fig. 14 shows the normalized histograms of the match/nonmatch distributions for GO, LBP, and SIFT. A two-fold cross validation scheme is used to determine the appropriate weights for the fusion. From the figures, it can be noticed that the fusion of multiclass and

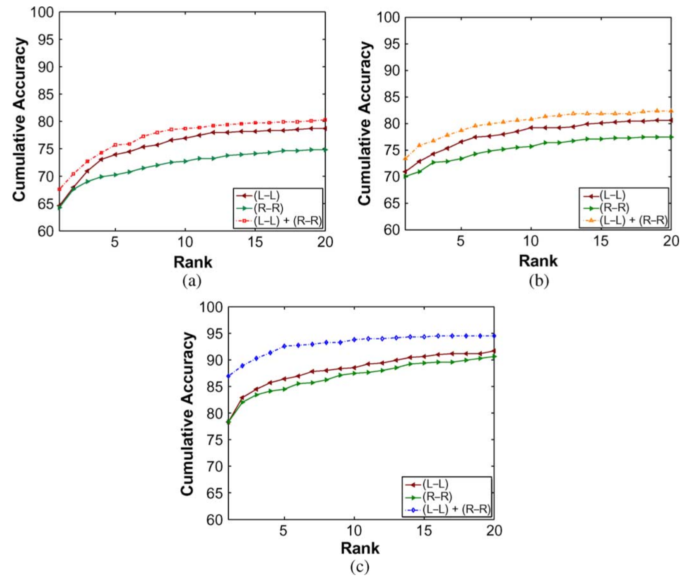


Fig. 11. CMC curves with fusion of (left-left) with (right-right) scores obtained from neutral–neutral matching for (a) GO, (b) LBP, and (c) SIFT matchers.

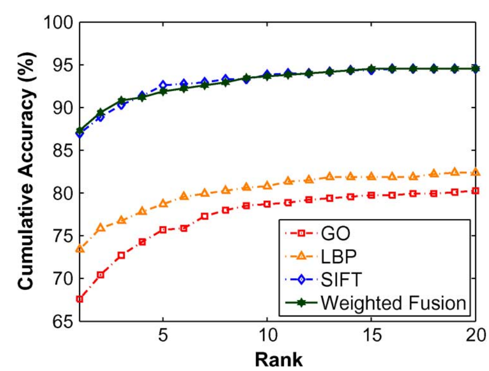


Fig. 12. CMC curves after fusing multiple classes (left and right eyes) and multiple algorithms (GO, LBP, and SIFT).

multialgorithm scores provides the best CMC performance. The fusion scheme did not result in any improvement in EER. We believe this is due to the noise in the genuine and imposter score distributions as shown in Fig. 14. The DET curves suggest the potential of using the periocular modality as a soft biometric cue.

F. Periocular Recognition Under Nonideal Conditions

In this section, the periocular recognition performance is studied under various nonideal conditions:

1) Partial face images: To compare the performance of periocular recognition with face recognition, a commercial face recognition software, FaceVACS [34] was used to match the face images in DB2. A rank-one accuracy of 99.77% was achieved with only 4 nonmatches at rank-one and no enrollment failures using 1136 probe and 568 gallery images from the 568 different subjects (DB2). In such situations, it is quite logical to prefer face in lieu of periocular region. However, the strength of the periocular recognition lies in the fact that it can be used even in situations where only partial face images are available. Most face recognition systems use a holistic approach, which

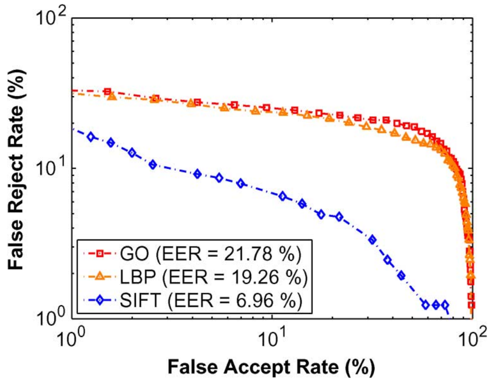


Fig. 13. DET curves for GO, LBP, and SIFT matchers obtained by the score level fusion of multiple classes.

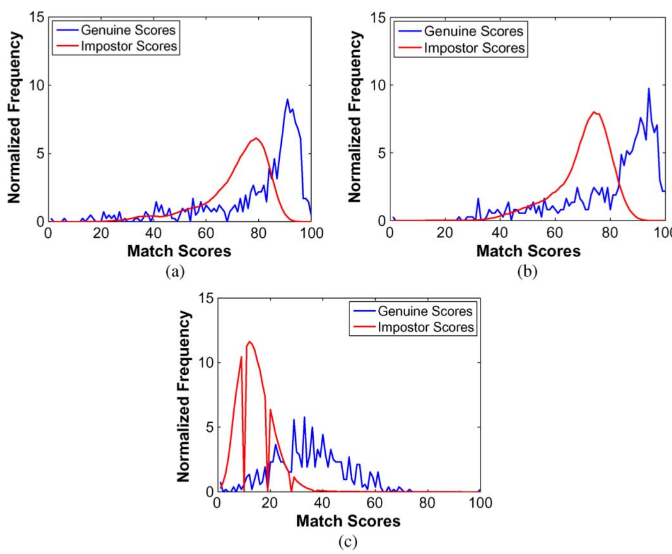


Fig. 14. Genuine and imposter matching score distributions for (a) GO, (b) LBP, and (c) SIFT, respectively.

(a) (b)

Fig. 15. Example of a partial face image. (a) Face image with mask applied under the nose region. (b) Detection of face and periocular regions.

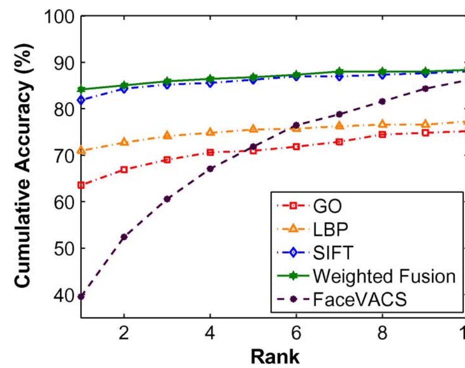


Fig. 16. CMC curves obtained on the partial face image dataset with the proposed periocular matcher and the FaceVACS face matcher.

(a) (b) (c) (d)

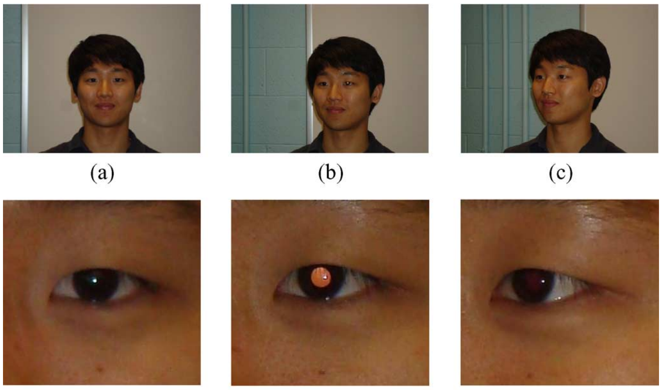
Fig. 17. Examples of periocular images with (a), (c) original and (b), (d) altered eyebrows using [35].

requires a full face image to perform recognition. In situations where a full face image is not available, it is quite likely that a face recognition system might not be successful. On the other hand, periocular region information could be potentially used to perform recognition. An example for such a scenario would be a bank robbery event where the perpetrator masks portions of the face to hide his identity.

To support the above stated argument, a dataset was synthetically constructed with partial face images. For every face image in DB2, a rectangular region of a specific size was used to mask the information below the nose region, as shown in Fig. 15(a), resulting in 1704 partial face images. The rank-one accuracy obtained on the partial face dataset using FaceVACS was observed to be 39.55%, much lower than the performance obtained with the full face dataset, DB2. For the periocular recognition, a total of 1663 faces out of the 1704 images (approximately 97.5%) were successfully detected using the OpenCV automatic face detector. Fig. 15(b) shows an example of a successfully detected partial face. The periocular regions with eyebrows were segmented again for the partial face dataset based on the same method used for the full face image. Fig. 16 shows the resulting performances of the matchers for neutral-versus-neutral matching. These results indicate the reliability of using periocular recognition in scenarios where face recognition may fail.

2) Cosmetic modifications: Considering the potential forensic applications, it is important to study the effect of cosmetic modifications to the shape of the eyebrow on periocular recognition performance. We used a web-based tool [35] to alter the eyebrows in 40 periocular images and conducted a matching experiment to determine its effect. Fig. 17 shows examples of the original periocular images along with their corresponding images with altered eyebrows. We have considered slight enlargement or shrinkage of the eyebrows. The average rank-one identification accuracies using the 40 altered (unaltered) images as probe and 568 images as gallery are 60% (70%), 65% (72.50%), and 82.50% (92.50%) using GO, LBP, and SIFT, respectively.

3) Perspective (or pose) variations: The periocular images considered in this work are cropped from facial images with frontal pose. However, the facial images might not always be in the frontal pose in a real operating environment. In this regard, a new dataset was collected with 40 different subjects under normal illumination conditions. A set of four face images with neutral expression were collected for each subject:



(a) (b) (c) (d) (e) (f)

Fig. 18. Examples of images with perspective variations. (a), (d) Frontal, (b), (e) 15° profile, and (c), (f) 30° profile.

TABLE VII
RANK-ONE ACCURACIES OBTAINED WITH POSE VARIATION DATA. ALL GALLERY IMAGES ARE FRONTAL BUT THE PROBE IMAGES ARE EITHER FRONTAL OR OFF-FRONTAL

	Frontal	15^o profile	30^o profile
GO	70.00%	52.50%	25.00%
LBP	70.00%	60.00%	50.00%
SIFT	85.00%	50.00%	40.00%

*Number of probe (gallery) images are 40 (608). Gallery image consists of 568 FRGC 2.0 images and 40 images collected at West Virginia University and Michigan State University.

(a) (b) (c)

Fig. 19. Examples of images showing occlusions pertaining to (a) 10%, (b) 20%, and (c) 30% of the periocular image area.

TABLE VIII
RANK-ONE ACCURACIES OBTAINED USING OCCLUSION DATA

	10% occlusion	20% occlusion	30% occlusion
GO	27.02%	13.38%	6.60%
LBP	18.66%	9.59%	3.78%
SIFT	25.97%	20.51%	10.12%

*Number of probe and gallery images are both 140.

TABLE IX
EFFECT OF TEMPLATE AGING ON THE RANK-ONE ACCURACIES

	Same session	Different session (up to 3 months, Fall 2003)	Different session (up to 9 months, from Fall 2003 to Spring 2004)
GO	86.43%	77.86%	56.43%
LBP	84.29%	78.57%	66.43%
SIFT	97.14%	82.14%	70.00%

*Number of probe and gallery images are both 140.

two frontal, one 15° left profile, and one 30° left profile. While one frontal image per subject was used to construct the gallery, the other three images were used as probe. An additional 568 images from Dataset 2 were added to the gallery. The periocular regions from the gallery and probe face images were segmented using the manual segmentation scheme described in Section III-B. Fig. 18 shows some example facial images along with their corresponding periocular regions. Table VII lists the rank-one accuracies of periocular recognition obtained with perspective variations.

It is noticed that variations in the perspective (profile) view can significantly reduce the recognition accuracy.

4) Occlusions: In a real operating environment, the periocular region could sometimes be occluded due to the presence of structural components such as long hair or glasses. To study the effect of occlusion on periocular recognition performance, three datasets were generated by randomly occluding 10%, 20%, and 30% of the periocular images in Dataset 2. Fig. 19 shows example images for each case. The recognition results are summarized in Table VIII. It is observed that the performance significantly drops with increasing amount of occlusion in the periocular region.

5) Template Aging: The periocular images used in all the earlier experiments were collected in the same data acquisition session. To evaluate the effect of time-lapse on the identification performance of periocular biometric, we conducted an additional experiment using data collected over multiple sessions. We used the face images of 70 subjects in the FRGC 2.0 database collected in Fall 2003 and Spring 2004. Three face images were selected for each subject from Fall 2003. The first image was used as the gallery image; the second image, where the subject was wearing the same clothes as the first one, was used as the same-session probe image; the third image, where the subject was wearing different clothes, was used as the different-session probe image. Further, the image of the corresponding subject from Spring 2004 was also used as a different-session probe image (with larger time-lapse).

Table IX shows the rank-1 identification accuracy in these experiments. As expected, the performance decreases as the time lapse increases. Template aging is a challenging problem in many biometric traits (e.g., facial aging). Further efforts are required to address the template aging problem in periocular biometrics.

IV. CONCLUSIONS AND FUTURE WORK

In this paper, we investigated the use of the periocular region for biometric recognition and evaluated its matching performance using three different matchers based on global and local feature extractors, viz., GO, LBP, and SIFT. The effects of various factors such as segmentation, facial expression, and eyebrows on periocular biometric recognition performance were discussed. A comparison between face recognition and periocular recognition performance under simulated nonideal conditions (occlusion) was also presented. Additionally, the effects of pose variation, occlusion, cosmetic modifications, and template

TABLE X
AVERAGE DIFFERENCE IN RANK-ONE ACCURACIES OF PERIOCULAR RECOGNITION UNDER VARIOUS SOURCES OF DEGRADATION

Factors	Average difference in rank-one accuracy
With eyebrows vs. no eyebrows	10.66%
Manual vs. automatic	0.06%
Neutral-neutral vs. neutral-smiling	3.36%
L-L and R-R vs. both sides	0.33%
Unmasked vs. masked iris	3.65%
Unmasked vs. masked eye	15.95%
No occlusion vs. 10% occlusion	48.67%
Same session vs. different session	25.00%

aging on periocular recognition were presented. Experiments indicate that it is preferable to include eyebrows and use neutral facial expression for accurate periocular recognition. Matching the left and right side of periocular images individually and then combining the results helped in improving recognition accuracy. The combination of both global and local matcher improve the accuracy marginally, which may be further improved by using more robust global matchers. Manually segmented periocular images showed slightly better recognition performance than automatically segmented images. Removing the iris or eye region, and partially occluding the periocular region degraded the recognition performance. Altering the eyebrows and template aging also degraded the matching accuracy. Table X reports the average difference in rank-one accuracies of periocular recognition under various scenarios.

On an average, the feature extraction using GO, LBP, and SIFT takes 4.68, 4.32, and 0.21 seconds, respectively, while matching takes 0.14, 0.45, and 0.10 seconds, respectively, on a 2.99-GHz CPU and 3.23-GB RAM PC in a Matlab environment with periocular images of size 241×226 (width$\times$height). The performance of periocular recognition could be further enhanced by incorporating the information related to the eye shape and size. Fusion of periocular (either in NIR or visible spectrum) with iris is another topic that we plan to study.

ACKNOWLEDGMENT

Anil K. Jain is the corresponding author of this paper.

REFERENCES

[1] U. Park, A. Ross, and A. K. Jain, "Periocular biometrics in the visible spectrum: A feasibility study," in *Proc. Biometrics: Theory, Applications and Systems (BTAS)*, 2009, pp. 153–158.

[2] *Handbook of Biometrics*, A. K. Jain, P. Flynn, and A. Ross, Eds. New York: Springer, 2007.

[3] R. Clarke, "Human identification in information systems: Management challenges and public policy issues," *Inf. Technol. People*, vol. 7, no. 4, pp. 6–37, 1994.

[4] J. B. Hayfron-Acquah, M. S. Nixon, and J. N. Carter, "Automatic gait recognition by symmetry analysis," in *Proc. Audio-and-Video-Based Biometric Person Authentication (AVBPA)*, 2001, pp. 272–277.

[5] R. Derakhshani and A. Ross, "A texture-based neural network classifier for biometric identification using ocular surface vasculature," in *Proc. Int. Joint Conf. Neural Networks (IJCNN)*, 2007, pp. 2982–2987.

[6] A. Kumar and Y. Zhou, "Human identification using knucklecodes," in *Proc. Biometrics: Theory, Applications and Systems (BTAS)*, 2009, pp. 147–152.

[7] J. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 11, pp. 1148–1161, Nov. 1993.

[8] A. Ross, "Iris recognition: The path forward," *IEEE Computer*, vol. 43, no. 2, pp. 30–35, Feb. 2010.

[9] K. W. Bowyer, K. Hollingsworth, and P. J. Flynn, "Image understanding for iris biometrics: A survey," *Comput. Vis. Image Understanding*, vol. 110, no. 2, pp. 281–307, 2008.

[10] S. Crihalmeanu, A. Ross, and R. Derakhshani, "Enhancement and registration schemes for matching conjunctival vasculature," in *Proc. Int. Conf. Biometrics (ICB)*, 2009, pp. 1240–1249.

[11] J. Matey, D. Ackerman, J. Bergen, and M. Tinker, "Iris recognition in less constrained environments," *Advances in Biometrics: Sensors, Algorithms and Systems*, pp. 107–131, 2008.

[12] S. Bhat and M. Savvides, "Evaluating active shape models for eyeshape classification," in *Proc. ICASSP*, 2008, pp. 5228–5231.

[13] A. Jain, S. Dass, and K. Nandakumar, "Soft biometric traits for personal recognition systems," in *Proc. Int. Conf. Biometric Authentication (LNCS 3072)*, 2004, pp. 731–738.

[14] P. E. Miller, A. W. Rawls, S. J. Pundlik, and D. L. Woodard, "Personal identification using periocular skin texture," in *Proc. ACM 25th Symp. Applied Computing*, 2010, pp. 1496–1500, ACM Press.

[15] NIST, Face Recognition Grand Challenge Database [Online]. Available: http://www.frvt.org/FRGC/

[16] C. Boyce, A. Ross, M. Monaco, L. Hornak, and X. Li, "Multispectral iris analysis: A preliminary study," in *Proc. IEEE Workshop on Biometrics at CVPR*, 2006, pp. 51–59.

[17] D. Woodard, S. Pundlik, P. Miller, R. Jillela, and A. Ross, "On the use of periocular and iris biometrics in non-ideal imagery," in *Proc. Int. Conf. Pattern Recognition (ICPR)*, 2010, pp. 201–204.

[18] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Contentbased image retrieval at the end of the early years," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 12, pp. 1349–1380, Dec. 2000.

[19] C. Schmid and R. Mohr, "Local grayvalue invariants for image retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 5, pp. 530–535, May 1997.

[20] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.

[21] R. Fergus, P. Perona, and A. Zisserman, "Object class recognition by unsupervised scale-invariant learning," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2003, pp. 264–271.

[22] D. Lowe, "Distinctive image features from scale-invariant key points," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[23] K. Mikolajczyk and C. Schmid, "An affine invariant interest point detector," in *Proc. Eur. Conf. Computer Vision (ECCV)*, 2002, pp. 128–142.

[24] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Surf: Speeded up robust features," *Comput. Vis. Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[25] L. Masek and P. Kovesi, MATLAB Source Code for a Biometric Identification System Based on Iris Patterns The School of Computer Science and Software Engineering, University of Western Australia, 2003.

[26] S. Rudinac, M. Uscumlic, M. Rudinac, G. Zajic, and B. Reljin, "Global image search vs. regional search in CBIR systems," in *Int. Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)*, 2007, pp. 14–17.

[27] K. Chang, X. Xiong, F. Liu, and R. Purnomo, "Content-based image retrieval using regional representation," *Multi-Image Analysis*, vol. 2032, pp. 238–250, 2001.

[28] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2005, pp. 886–893.

[29] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.

[30] SIFT Implementation [Online]. Available: http://www.vlfeat.org/vedaldi/code/sift.html

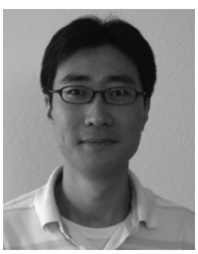
[31] P. Phillips, P. Flynn, T. Scruggs, K. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Jun. 2005, vol. 1, pp. 947–954.

[32] OpenCV: Open Source Computer Vision Library [Online]. Available: http://sourceforge.net/projects/opencvlibrary/

[33] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2001, pp. 511–518.

[34] FaceVACS Software Developer Kit Cognitec Systems GmbH [Online]. Available: http://www.cognitec-systems.de

[35] TAAZ, Free Virtual Make Over Tool [Online]. Available: http://www.taaz.com/



Unsang Park (S'06–M'07) received the B.S. and M.S. degrees from the Department of Materials Engineering, Hanyang University, South Korea, in 1998 and 2000, respectively. He received the second M.S. and Ph.D. degrees from the Department of Computer Science and Engineering, Michigan State University, in 2004 and 2009, respectively.

From 2009, he was a Postdoctoral Researcher in the Pattern Recognition and Image Processing Laboratory, Michigan State University. His research interests include biometrics, video surveillance, image processing, computer vision, and machine learning.

Raghavender Reddy Jillela (S'09) received the B.Tech. degree in electrical and electronics engineering from Jawaharlal Nehru Technological University, India, in May 2006. He received the M.S. degree in electrical engineering from West Virginia University, in December 2008. He is currently working toward the Ph.D. degree in the Lane Department of Computer Science and Electrical Engineering, West Virginia University.

His current research interests are image processing, computer vision, and biometrics.

Arun Ross (S'00–M'03–SM'10) received the B.E. (Hons.) degree in computer science from the Birla Institute of Technology and Science, Pilani, India, in 1996, and the M.S. and Ph.D. degrees in computer science and engineering from Michigan State University, East Lansing, in 1999 and 2003, respectively.

Between 1996 and 1997, he was with the Design and Development Group of Tata Elxsi (India) Ltd., Bangalore, India. He also spent three summers (2000–2002) with the Imaging and Visualization Group of Siemens Corporate Research, Inc., Princeton, NJ, working on fingerprint recognition algorithms. He is currently a Robert C. Byrd Associate Professor in the Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown. His research interests include pattern recognition, classifier fusion, machine learning, computer vision, and biometrics. He is actively involved in the development of biometrics and pattern recognition curricula at West Virginia University. He is the coauthor of *Handbook of Multibiometrics* and coeditor of *Handbook of Biometrics*.

Dr. Ross is a recipient of NSF's CAREER Award and was designated a Kavli Frontier Fellow by the National Academy of Sciences in 2006. He is an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING and the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY.

Anil K. Jain (S'70–M'72–SM'86–F'91) is a university distinguished professor in the Department of Computer Science and Engineering, Michigan State University, East Lansing. His research interests include pattern recognition and biometric authentication. He received the 1996 IEEE TRANSACTIONS ON NEURAL NETWORKS Outstanding Paper Award and the Pattern Recognition Society best paper awards in 1987, 1991, and 2005. He served as the editor-in-chief of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE (1991–1994).

Dr. Jain is a fellow of the AAAS, ACM, IAPR, and SPIE. He has received Fulbright, Guggenheim, Alexander von Humboldt, IEEE Computer Society Technical Achievement, IEEE Wallace McDowell, ICDM Research Contributions, and IAPR King-Sun Fu awards. The holder of six patents in the area of fingerprints, he is the author of a number of books, including *Handbook of Fingerprint Recognition* (2009), *Handbook of Biometrics* (2007), *Handbook of Multibiometrics* (2006), *Handbook of Face Recognition* (2005), *BIOMETRICS: Personal Identification in Networked Society* (1999), and *Algorithms for Clustering Data* (1988). ISI has designated him a highly cited researcher. According to Citeseer, his book *Algorithms for Clustering Data* (Prentice-Hall, 1988) is ranked #93 in most cited articles in computer science. He served as a member of the Defense Science Board and The National Academies committees on Whither Biometrics and Improvised Explosive Devices.