Periocular biometrics in the visible spectrum: A feasibility study

Conference Paper · October 2009

DOI: 10.1109/BTAS.2009.5339068 · Source: IEEE Xplore

CITATIONS
178

READS
262

3 authors, including:

Unsang Park
Sogang University
29 PUBLICATIONS **2,184** CITATIONS

SEE PROFILE

Arun Ross
Michigan State University
352 PUBLICATIONS **21,535** CITATIONS

SEE PROFILE

To appear in Biometrics: Theory, Applications and Systems (BTAS 09), Washington DC, September 2009

Periocular Biometrics in the Visible Spectrum: A Feasibility Study

Unsang Park, Arun Ross, and Anil K. Jain

***Abstract*— Periocular biometric refers to the facial region in the immediate vicinity of the eye. Acquisition of the periocular biometric does not require high user cooperation and close capture distance unlike other ocular biometrics (e.g., iris, retina, and sclera). We study the feasibility of using periocular images of an individual as a biometric trait. Global and local information are extracted from the periocular region using texture and point operators resulting in a feature set that can be used for matching. The effect of fusing these feature sets is also studied. The experimental results show a 77% rank-1 recognition accuracy using 958 images captured from 30 different subjects.**

I. Introduction

OCULAR biometrics has made rapid strides over the past few years primarily due to the significant progress made in iris recognition. 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 portion. The random morphogenesis of the textural relief of the iris and its *apparent* stability over the lifetime of an individual, have made it a very popular biometric. Both technological and operational tests conducted under predominantly constrained conditions have suggested the uniqueness of the iris texture across individuals and its potential as a biometric in large-scale systems enrolling millions of individuals [1, 2]. Indeed, even the two irises of an individual are observed to be different in their intricate textural content.

Besides the iris, other ocular traits have been investigated for human recognition, viz., the retinal and the conjunctival vasculature.

1. Retinal vasculature: The blood vessel pattern on the retina is believed to be unique across individuals [3]. Typically, a coherent light source is used to illuminate the vasculature pattern on the back of the eye and a CCD is used to image this pattern. However, a cooperative subject is assumed for procuring a good quality image that can be used during the matching phase.
2. Conjunctival vasculature: The vasculature pattern observed on the sclera of the eye has also been suggested as a potential biometric [4]. These blood vessels typically reside in the conjunctiva and the episclera layers of the sclera (although the term "conjunctival vasculature" is used to denote both sets of vessels), and are revealed when the iris is "off-axis" with respect to the imaging device. Thus, there is significant potential in utilizing these vasculature patterns along with the iris texture in a bimodal biometric system by employing a multispectral sensor for image acquisition.

Manuscript received June 7, 2009 and revised Aug 15, 2009.

Unsang Park and Anil K. Jain are with the Department of Computer Science and Electrical Engineering at Michigan State University (email: parkunsa@cse.msu.edu, jain@cse.msu.edu).

Arun Ross is with the Lane Department of Computer Science and Electrical Engineering at West Virginia University (email: arun.ross@mail.wvu.edu).

Fig. 1: Example periocular images from two different subjects: (a)(b) without eyebrows and (c)(d) with eyebrows.

In spite of the tremendous progress made in ocular biometrics (especially iris), 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 eye-ball. The eye-ball itself is located within another moving object – the head. Therefore, reliably localizing the iris in eye images obtained at a distance from unconstrained human subjects can be difficult [5]. Furthermore, since the iris is typically imaged in the near infrared portion (700 – 900nm) of the electromagnetic (EM) spectrum, appropriate invisible lighting is required to illuminate it prior to image acquisition.
2. 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.
3. 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 [6].

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. In this work we explore the potential of the periocular region as a biometric in **color images**. We do not use the near-IR spectrum in this paper, although the eventual goal is to use a multispectral acquisition device that can image the periocular region in both the visible and near-IR spectral bands [7]. This would ensure the possibility of combining the iris texture with the periocular texture. The use of the periocular region has several benefits:

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.
2. The use of the periocular region represents a good trade-off between using the entire face region or using only the iris for recognition. When the entire face is imaged from a distance, the iris information is typically of low resolution; this means the matching performance due to the iris modality will be poor. 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.
3. The periocular region can offer information about eye-shape that may be useful as a soft biometric.
4. The depth-of-field of iris systems can be increased if the surrounding ocular region were to be included as well.

The purpose of this work is to do a feasibility study on using periocular information as a biometric. Thus, images obtained in the visible spectrum are studied for this purpose.

II. Periocular Recognition

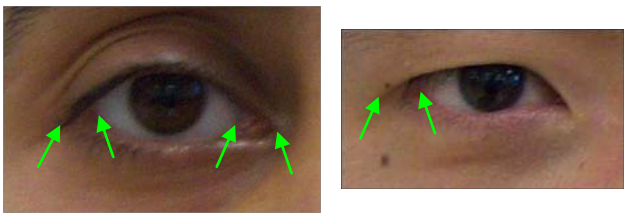
The proposed periocular recognition process consists of a sequence of operations: image alignment (for the global matcher described below), feature extraction, and matching. We adopt two different approaches to the problem: one based on global information and the other based on local information. The two approaches use different methods for feature extraction and matching. We will first review the characteristics of these two approaches, and describe each intermediate process.

A. Global vs. Local Matcher

Most image matching schemes can be categorized as global or local. The basic difference between global and local methods is 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 features are color, shape, and texture [8]. Global features are 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, the 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 descriptor) [9, 10]. Then, the number of matching key points between two images is calculated 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

(a) Example images showing eyelid movement



(b) Example images where multiple corner candidates are present

Fig. 2: Example images showing difficulties in periocular image alignment.

(a) Input image (b) Iris detection

(c) Interest point sampling (d) Interest region sampling

Fig. 3: Global descriptor construction process.

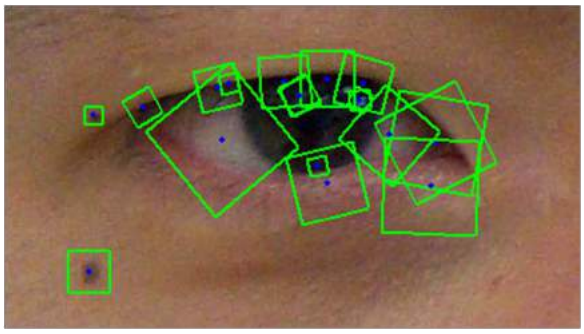


Fig. 4: Examples of local features and bounding boxes for descriptor construction in SIFT. Each bounding box is rotated with respect to the major orientation.

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 a constant time matching using local features through the bag of words representation [11].

In terms of the matching accuracy, local feature-based approaches have shown better performance. When all available pixel values are encoded into the feature vector (as is the case when global features are used), it becomes more susceptible to image variations especially with respect to 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 resulted in more active research on local feature based image retrieval schemes [12, 13, 14].

Face, iris, and hand mostly adopt a global representation scheme while fingerprint mostly adopts a local representation scheme. The basic criterion for determining different representations in image-based biometrics is whether the trait under consideration has a common morphology across all subjects. If we take the average of a hundred face, iris, or hand images after proper scaling and alignment, the output will still appear as a legitimate face, iris, or hand image. However, the average of a hundred fingerprint images will not look like a fingerprint image anymore. Therefore, the face, iris, or hand images can be aligned in a certain common coordinate space and encoded into a fixed length feature vector. However, fingerprint and other general images need to be represented by their local key points.

We use both global and local matching methods for periocular recognition in order to take advantage of the fixed length feature representation of the global scheme and the distinctiveness of the local scheme.

B. Image Alignment

Periocular images contain common components (i.e., 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, since the images were taken under similar operational conditions as traditional iris recognition systems, where variations due to the iris and eyelids are deliberately constrained. While frontal iris detection can be performed fairly well due to the approximately circular geometry of the iris and the clear contrast between iris and sclera, the accurate detection of 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 Hough transformation was used for localizing the iris [15]. 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 of 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 radius, R_{iris}, of the iris, multiple ($=n_{pi}$) interest points $p_1, p_2, \ldots, p_{npi}$ are selected within a rectangular window defined around C_{iris} with a width of $6\times R_{iris}$ and a height of $4\times R_{iris}$ 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, $D_p\times R_{iris}$.

For each interest point p_i, a rectangular region r_i is defined with a dimension of $D_p\times R_{iris}$ as an interest region. We construct the key point descriptors from r_i and generate a full feature vector by concatenating all the descriptors. The feature representation using partitioned image is regarded as a local feature representation in some image retrieval literature [16, 17]. However, we consider this as a global representation because all the pixel values are used in the representation without considering the local distinctiveness of each region.

Mikilajczyk et al. [10] 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 and local binary pattern (LBP) [18]. We quantize both GO and LBP into 8 distinct values to build an eight bin histogram. The eight bin histogram is constructed from a partitioned sub-region and concatenated to construct a full feature vector. A Gaussian blurring with a standard deviation σ is applied on both GO and LBP to smooth variations across local pixel values. This sub-partition based histogram construction scheme has been successfully used in SIFT [12] 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 key points 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. 4 shows the detected key points and 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, it is expected that the local matcher provides more distinctiveness.

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 points. Moreover, when there is affine transformation or occlusion it will be beneficial to match individual interest points rather than relying on the entire set of interest points. We used a publically available SIFT implementation [19] as the local matcher.

D. Matching Scheme

For the global descriptor, the simple Euclidean distance is used to calculate the matching distance. The distance ration based matching scheme [12] 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 of interest points, $1/D_p$.

SIFT has many parameters that affects its performance. Some of the representative parameters are the number of octaves (n_o), number of scales (n_s), and the cut-off 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.

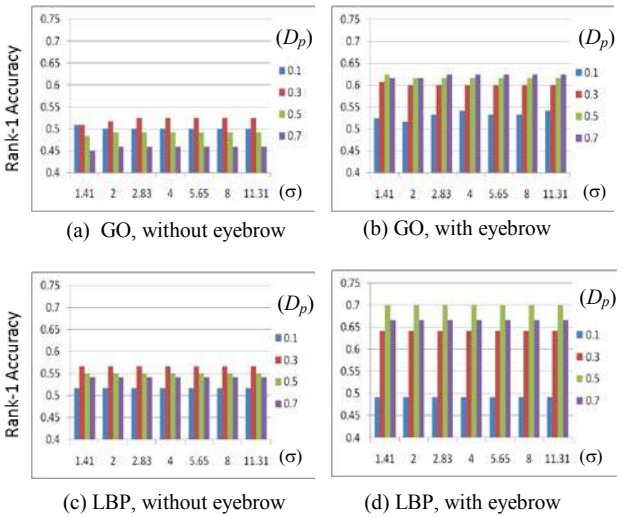
We construct a number of different descriptors for both the global and local schemes by choosing a set of different values for σ, D_p, n_o, n_s, and t_{ex}. The set of parameters that results in the best empirical performance is selected to be used for the global and local representations.

III. Experimental Results

A. Database

We collected 899 high-resolution face images from 30 different subjects in two different sessions (450 in session 1 and 449 in session 2, 14~15 images per subject in each session) using a Canon EOS 5D Mark II camera. The camera parameters were set to the following options: maximum resolution (21.1 Mega pixels - 5616×3744), Auto-Focus, Optical Vibration Reduction Image Stabilization, Portrait Mode, ISO AUTO, and JPEG format. Each subject was asked to sit ~4 feet away from the camera during data acquisition. We manually cropped the periocular region from each face image in two different ways: with and without eyebrows. Some example periocular images are shown in Fig. 1. The sizes of periocular images are in the range [419,892] for width and [182,400] for height with no eyebrow. The periocular images with eye brow shows height in the range of [265,713] with the same width range as those without eyebrow.

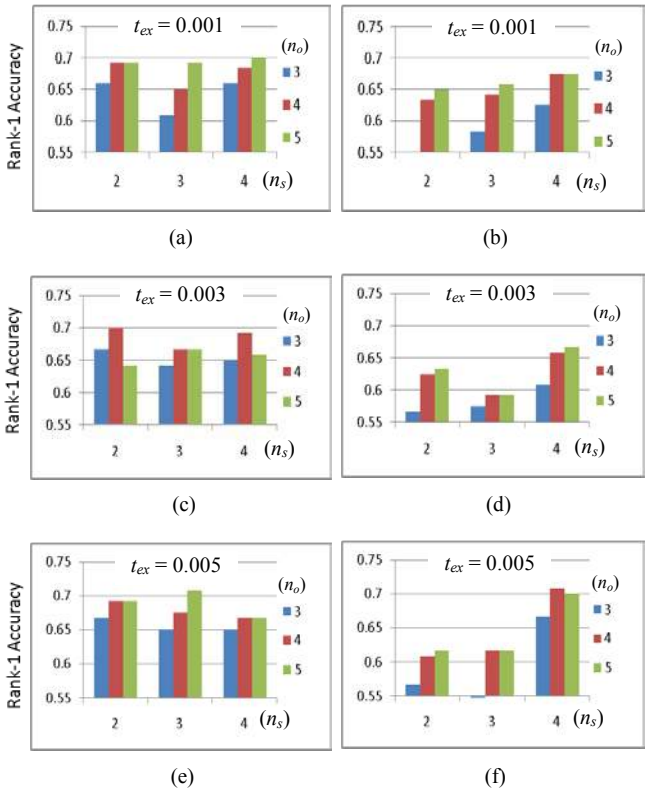
We assembled two different databases, DB1 and DB2, for the periocular recognition experiments. DB1 consists of 120 images with two (left and right eye) periocular images per subject per session. DB2 consists of 958 images with 898 probe and 60 gallery images. Probe dataset contains 28~30 periocular images per subject and gallery contains 2 periocular images per subject. DB1 is used for parameter selection and DB2 is used for evaluating the matching



(a) GO, without eyebrow (b) GO, with eyebrow

(c) LBP, without eyebrow (d) LBP, with eyebrow

Fig. 5: Rank-1 accuracies of the global matcher (GO and LBP) with different choices of parameter: (a)(c) without eyebrow and (b)(d) with eyebrow.



(a) (b)

(c) (d)

(e) (f)

Fig. 6: Rank-1 accuracies of the local matcher (SIFT) with different choices of parameter: (a)(c)(e) without eyebrow and (b)(d)(f) with eyebrow.

performance.

Table 1: Periocular recognition accuracy (%) with respect to the use of eyebrows and side information.

	Without eyebrow		With eyebrow	
	L or R eye *	Same eye **	L or R eye	Same eye
GO	52.5	49.2	62.5	60.8
LBP	56.7	50.8	70.0	66.7
SIFT	71.7	74.2	70.8	70.0
GO+SIFT	76.7	80.8	80.0	75.8
LBP+SIFT	76.7	80.8	80.0	78.3
GO+LBP+SIFT	73.3	77.5	80.0	79.2

* Left (Right) eyes can match with Right (Left) eyes
** Left (Right) eyes cannot match with Right (Left) eyes

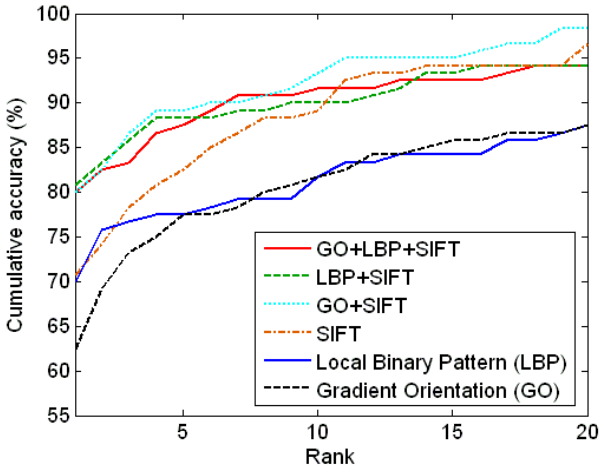
B. *Recognition Accuracy*

The recognition accuracy using the aforementioned periocular feature set is assessed using the Cumulative Match Characteristic (CMC) curve. For DB1, given N (=120) images $I_1, I_2, \ldots, I_N$, every image I_i is taken as the query and the rest of the images are used as the gallery. For DB2, separate set of probe and gallery images are used. Matching experiments on DB1 are performed with and without eyebrows, and with and without the Left/Right (eye side) information. When the Left/Right information is used, Left (Right) side periocular image can only match to the Left (Right) side. To take advantage of the characteristics of both global and local descriptors, we used a fusion scheme that combines the global and local information. We used a score level fusion based on weighted sum with min-max normalization. The weights are empirically selected for both the global and local matchers.

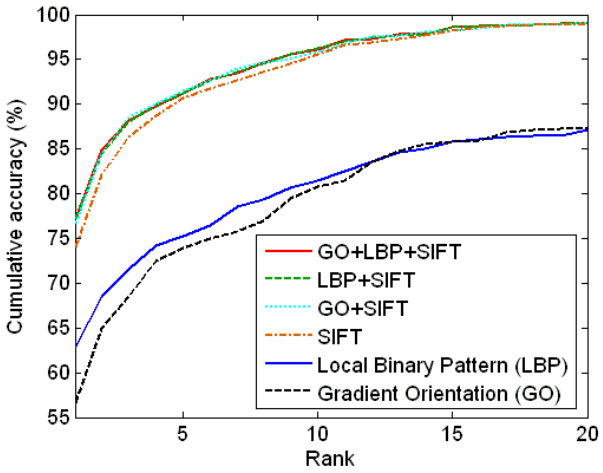
Fig. 5 shows the rank-1 accuracy of the global matcher using GO and LBP descriptors based on different configuration of the parameters. The best performance was observed to be 62.5% and 70.0% for the GO and LBP descriptors, respectively. The performance of both GO and LBP shows a dependency on D_p rather than σ. The use of eyebrow showed better recognition accuracy for both the GO and LBP descriptors.

Fig. 6 shows the rank-1 accuracy of the SIFT matcher. Larger values of t_{ex} and n_s than those shown in Fig. 6 resulted in lower accuracy. With a large value of n_s, the standard deviation of the Gaussian kernel increases by a small amount when constructing the scale space, resulting in smaller values across the DOG space. This has a similar effect as increasing t_{ex}, which also decreases the matching accuracy. Larger n_o helps in improving the matching accuracy, in general. The accuracy decreases with the use of eyebrows. We believe this is due to the noisy keypoints detected around the eyebrow, which results in false matches thereby inflating the imposter matching scores. The best rank-1 performance is obtained as 74.2% with no eyebrow and using information about the location of the periocular region (i.e., left or right eye).

The matching accuracy of the best global matchers, local matcher, and the resulting fusion schemes on DB1 and DB2 are shown in Fig. 7. The best performance is observed to be 80.8% and 77.3% by fusing the LBP based global matcher with SIFT on DB1 and DB2, respectively.



(a)



(b)

Fig. 7: CMC curve of the global, local, and fusion matchers on (a) DB1 and (b) DB2.

The first five rows of Fig. 8 show examples of image pairs that were not correctly matched at rank-1 by the LBP and SIFT schemes but were correctly matched after fusion for the first five rows. The last two rows show failure cases both before and after fusion. Periocular images from different subjects appear similar in the last two rows, resulting in the false matches.

IV. Conclusions and Future Work

We have proposed a method for using periocular images as a biometric trait. Both global and local descriptors have been explored for feature extraction and matching. Further, a score-level fusion scheme was employed to enhance the recognition accuracy. Based on the evaluation of a total of 958

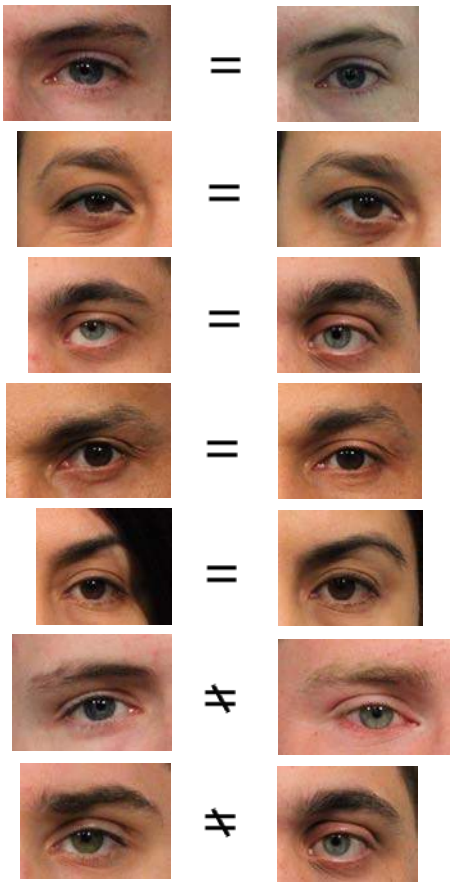


Fig. 8: Example image pairs that were not correctly matched at rank-1 by neither the LBP nor the SIFT methods, but were correctly matched at rank-1 after fusion for the first five rows. Last two rows show failure cases in LBP, SIFT, and fusion

.

images, taken from 30 different subjects, the proposed method demonstrates the feasibility of using periocular images as a biometric trait. The feature extraction time is about 6 sec for both LBP and SIFT, and the matching time is negligible on a 2.4 GHz, 4GB RAM PC.

We also performed a face recognition experiment on the full-face images using 449 images in session 2 as probes and 30 images in session 1 as gallery images. A commercial face recognition engine, FaceVACS [20], was used for this purpose. A 100% accuracy was obtained in the face recognition test. This implies the following; i) periocular biometric should be used as a secondary method supporting the primary biometric or as an alternative when the primary biometric is not available and ii) periocular region contains ~80% of the identity information in associated with the face. This complies with the results in an earlier study in [21], where the periocular region was shown to be the most important region in identifying a face.

Future work will involve utilizing **multispectral** information for feature extraction; using more robust image alignment and matching methods; combining the periocular matcher with iris matcher; and developing more robust feature encoding schemes. We would also like to study the impact of cosmetics on the texture of the periocular region and the ensuing recognition capability.

V. Acknowledgements

The authors are grateful to Dr. Thirimachos Bourlai and Brian DeCann of WVU for coordinating the data collection activity.

References

[1] J. Daugman, "High confidence visual recognition of persons by a test of statistical independence," IEEE TPAMI, 15(11), pp. 1148-1161, 1993.

[2] K. W. Bowyer, K. Hollingsworth, and P. J. Flynn, "Image understanding for iris biometrics: A survey," Computer Vision and Image Understanding, 110(2), pp. 281-307, 2008.

[3] D. Usher, Y. Tosa, and M. Friedman, "Ocular biometrics: simultaneous capture and analysis of the retina and iris," Advances in Biometrics: Sensors, Algorithms and Systems, Springer Publishers, pp. 133-155, 2008.

[4] R. Derakhshani and A. Ross, "A texture-based neural network classifier for biometric identification using ocular surface vasculature," Proc. of International Joint Conference on Neural Networks (IJCNN), pp. 2982-2987, 2007.

[5] J. Matey, D. Ackerman, J. Bergen, and M. Tinker, "Iris recognition in less constrained environments," Advances in Biometrics: Sensors, Algorithms and Systems, Springer Publishers, 2008.

[6] S. Crihalmeanu, A. Ross and R. Derakhshani, "Enhancement and registration schemes for matching conjunctival vasculature," Proc. of the 3rd IAPR/IEEE International Conference on Biometrics (ICB), pp. 1240-1249, 2009.

[7] C. Boyce, A. Ross, M. Monaco, L. Hornak, and X. Li, "Multispectral iris analysis: a preliminary study," Proc. of IEEE Computer Society Workshop on Biometrics at CVPR, pp. 51-59, 2006.

[8] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," 22(12), pp. 1349-1380, 2000.

[9] C. Schmid and R. Mohr, "Local Grayvalue Invariants for Image Retrieval," IEEE TPAMI, 19(5), pp. 530-535, 1997.

[10] Krystian Mikolajczyk and Cordelia Schmid, "A performance evaluation of local descriptors," IEEE TPAMI, 27(10), pp. 1615-1630, 2005.

[11] R. Fergus, P. Perona, and A. Zisserman, "Object class recognition by unsupervised scale-invariant learning," Proc. of CVPR (2), pp. 264-271, 2003.

[12] D. Lowe, "Distinctive image features from scale-invariant key points," International Journal of Computer Vision, 60(2), pp. 91-110, 2004.

[13] K. Mikolajczyk and C. Schmid. "An affine invariant interest point detector," Proc of ECCV, pp. 128-142, 2002.

[14] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "SURF: Speeded Up Robust Features," Computer Vision and Image Understanding, 110(3), pp. 346--359, 2008.

[15] Libor Masek and Peter 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.

[16] S. Rudinac, M. Uscumlic, M. Rudinac, G. Zajic, and B. Reljin, "Global image search vs. regional search in CBIR systems," WIAMIS '07, pp. 14, 2007.

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

[18] Matti Pietikäinen, "Image analysis with local binary patterns," Image Analysis, pp. 115-118, 2005.

[19] http://www.vlfeat.org/~vedaldi/code/sift.html.

[20] FaceVACS FaceVACS Software Developer Kit, Cognitec, http://www.cognitec-systems.de

[21] Behrooz Kamgar-Parsi, Behzad Kamgar-Parsi, and A. K. Jain: Synthetic Eyes. Proc. of Audio and Video-Based Person Authentication (AVBPA), pp. 412-420, 2003