What is a 'Good' Periocular Region for Recognition?

Jonathon M. Smereka and B.V.K. Vijaya Kumar

Abstract

In challenging image acquisition settings where the performance of iris recognition algorithms degrades due to poor segmentation of the iris, image blur, specular reflections, and occlusions from eye lids and eye lashes, the periocular region has been shown to offer better recognition rates. However, the definition of a periocular region is subject to interpretation. This paper investigates the question of what is the best periocular region for recognition by identifying sub-regions of the ocular image when using near-infrared (NIR) or visible light (VL) sensors. To determine the best periocular region, we test two fundamentally different algorithms on challenging periocular datasets of contrasting build on four different periocular regions. Our results indicate that system performance does not necessarily improve as the ocular region becomes larger. Rather in NIR images the eye shape is more important than the brow or cheek as the image has little to no skin texture (leading to a smaller accepted region), while in VL images the brow is very important (requiring a larger region).

1. Introduction

Iris recognition is well known to provide high recognition rates in controlled acquisition conditions. However, in instances where the iris is occluded (by eye lashes, eye lids, non-frontal gaze, etc.) or where segmentation fails due to impairments such as harsh illumination, low image resolution, specular reflections, and/or blur from motion or defocus, the performance of the iris recognition algorithms can degrade considerably.

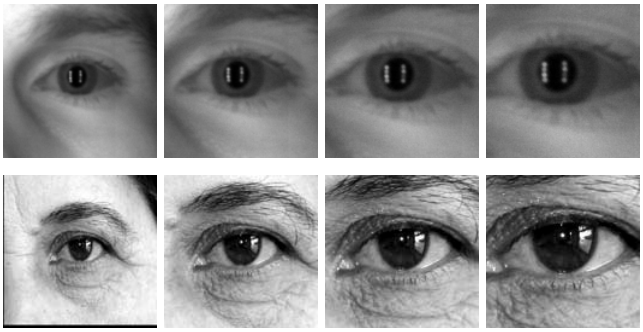
One potential solution is to use the ocular region of the face, sometimes referred to as periocular biometrics. Ocular modality benefits include avoiding iris segmentation [2] and an increased resistance to the effects of aging [4]. In addition, it has been shown that the periocular region can offer better recognition accuracy than face recognition at a great amount of deterioration [7]. This is due to the limited distortion that occurs within the periocular region compared to large deformation that accompanies a change in facial expression. Some examples of where the use of the ocular region can be beneficial are shown in Figure 1.

Figure 1: Examples of where periocular recognition can be beneficial (harsh illumination, occlusion, low resolution, etc) when attempting iris and/or face recognition.

One definition of the ocular region is the area of the face that includes the eyelids, eyelashes, eyebrow and the skin surrounding the eye. A few methods [19, 4] perform recognition using the bi-ocular region which include both eyes within a single image, however most ocular recognition approaches separate the left region from the right region where matching is performed individually. A survey of techniques for performing ocular recognition can be found in Ambika et al. [1].

To the authors' best knowledge, there is no commonly accepted definition of what exactly constitutes the best ocular region. Towards the goal of shedding further light on the effect of the ocular region on ocular recognition performance, this paper investigates the question of what is the best periocular region for recognition. As can be seen in Figure 2, the definition of a periocular region can correspond to several interpretations of how much information is available around (note that 'peri' is a Greek root for 'around' or 'surrounding') the eye. As we will show, recognition performance does not necessarily improve as more area around the eye becomes available in near-infrared (NIR) and visible light (VL) images. Specifically, the work presented identifies a distinct region of interest that provides the most support in matching for each sensor.

In determining the best periocular region, we test two fundamentally different algorithms on challenging periocular datasets of contrasting build. The two databases investigated are the Face and Ocular Challenge Series (FOCS) [8]



(a) Region 1 (b) Region 2 (c) Region 3 (d) Region 4

Figure 2: Examples of the normalized periocular regions for near-infrared (top) and visible light (bottom) images.

database captured in near-infrared (NIR) light from moving subjects in an unconstrained environment, and the University of Beira Interior Periocular (UBIPr) [11] database which is composed of visible light (VL) images of varying scale, pose, and illumination. The images from each dataset have had their eye corners labeled; we use the prepackaged labelings included with the UBIPr set and hand label the FOCS set as none are inherently included, to perform experiments over varying periocular regions. Figure 2 displays the differing regions of interest that are tested. The periocular region ranges from the case where the eye corners are at the edge of each image (referred to as 'Region 4') with very little cheek or brow information, to images ('Region 1') showing the complete area around the eye available in both datasets. In each region we refrain from masking the iris and/or eye itself as masking has been shown to be detrimental to system performance [12].

The two techniques we use for recognition are Probabilistic Deformation Models (PDM) [2] and m-SIFT [13] as they perform matching in intrinsically different ways. PDM is based on building a graphical model to account for any relative distortion, or structured change, between a probe and gallery image, while m-SIFT imposes constraints on locations of matching keypoints to improve the authentication. In both cases sub-region scores over each of the periocular regions are examined to find which areas consistently provide the most support in matching, with equal error rate (EER) as a measure of overall system performance.

The rest of the paper is organized as follows. An overview of related work is provided in §2. The methods of recognition used in this work are briefly described in §3.1 and §3.2. Experimental details concerning dataset information and model parameters are provided in §4, followed by empirical results and observations in §5. Finally, §6 provides a summary along with concluding remarks.

2. Background

Though still in its infancy as a biometric modality, some effort has been devoted to the area of periocular recognition. Park et al. [12] were the first in exploring this task, making several fundamental recommendations. Their results demonstrate that masking the iris and/or eye, such as was done in early works [17, 18], will degrade system performance, and that it is preferable to include eyebrows (without any alterations). Merkow et al. [6] extended this result to soft-biometrics, using periocular data from Flickr images to determine gender, finding that the brow and eyes carry valuable information for discrimination. Oh et al. [10] also show that masking the eye/iris will degrade performance while the cheek region in visible light (VL) images does not contain significant discriminative textural information.

Hollingsworth et al. [3] analyzed human verification performance on periocular images under different lighting conditions. Results suggest that in NIR images, humans use the eyelashes, eyelid, tear duct, and eye shape as discriminating features when performing recognition, whereas in VL images, blood vessels and skin were slightly more helpful than eye shape and eyelashes. The authors suggest that periocular biometrics should focus on using VL images for best system performance. Miller et al. [7] show that only using the green channel in color images can increase performance.

Padole and Proenca [11] observed that the recognition performance is dependent on the camera distance at which the sample was captured (recommending a camera distance of ~7m) while also reporting the benefits of using eye corners as reference points to normalize the region. Uzair et al. [15] use this type of normalization to build image sets when doing recognition on the MBGC [9] dataset.

The goal of this research effort is to determine the best periocular region by identifying sub-regions of the ocular image that provide the best recognition performance. Previous efforts evaluated general system design decisions (e.g., whether masking the eye/iris is effective or not), and generally provide some evidence of how to improve performance (e.g., in VL images the eyebrow is effective while the cheek may not be), but overlook the effects of differing periocular regions in NIR and VL images at a sub-region level. The work presented identifies a region that consistently appears to be more useful in recognition tasks when matching VL and NIR images. We are able to identify the effective areas by examining the match scores over sub-regions of the ocular image. Testing is performed over 4 different definitions of the periocular region (Figure 2) with two different matching methods.

3. Recognition Models

The implementation details for each matching scheme tested are discussed in this section. They include methods

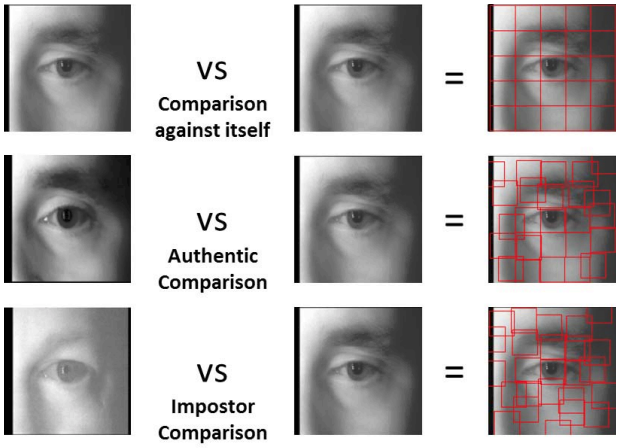


Figure 3: The red boxes are the size of each region of the probe that is individually compared against the gallery. The configuration of the boxes is based on the lateral shift required to center the patch over the area of best match.

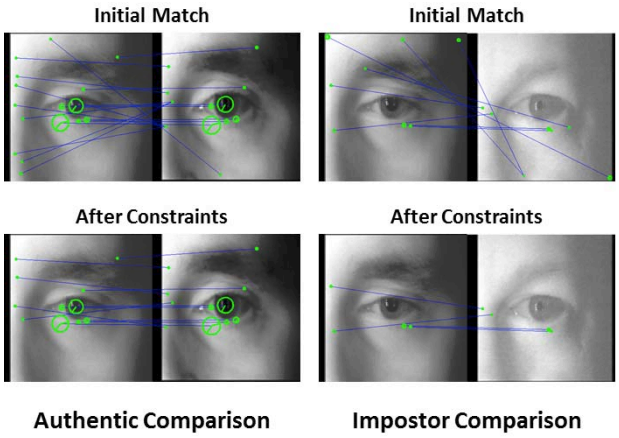


Figure 4: A set of matching interest points is identified with extra constraints applied to further reduce the number of keypoints used for score computation.

of measuring the match score at a detailed level to see which sub-regions of the ocular region are more useful for recognition. By testing two different methods (namely PDM and m-SIFT), we ensure that our observations about the sub-regions are not unique to a particular matching method. PDM cuts the images into patches, measuring similarity at a sub-region level as illustrated in Figure 3, while m-SIFT selects similar interest points based on their locations as shown in Figure 4.

3.1. PDM

A technique proposed by Thornton et al. [14] to simultaneously estimate the non-linear deformation in the iris from the expansion and contraction of the pupil and match iris images using a Markov Random Field (MRF) showed a marked improvement in recognition rates in iris images where distortion was present. A similar method was adapted for ocular recognition in [2] and demonstrated the benefits of using the ocular region in environments where iris recognition can fail. This technique divides both the probe and gallery images into patches and the corresponding probe and gallery patches are correlated with each other to produce a set of correlation output arrays. For an authentic pair, the correlation output from a patch comparison from the probe and gallery images, should exhibit a peak at the location of the best match providing information about the (x, y) shift between the probe and gallery patches, while also providing a metric of similarity as peak height. A MRF is then built based on the shifts of neighboring image patches to compensate for the distortion that occurs between similar images. More details on PDM can be found within [2].

3.2. m-SIFT

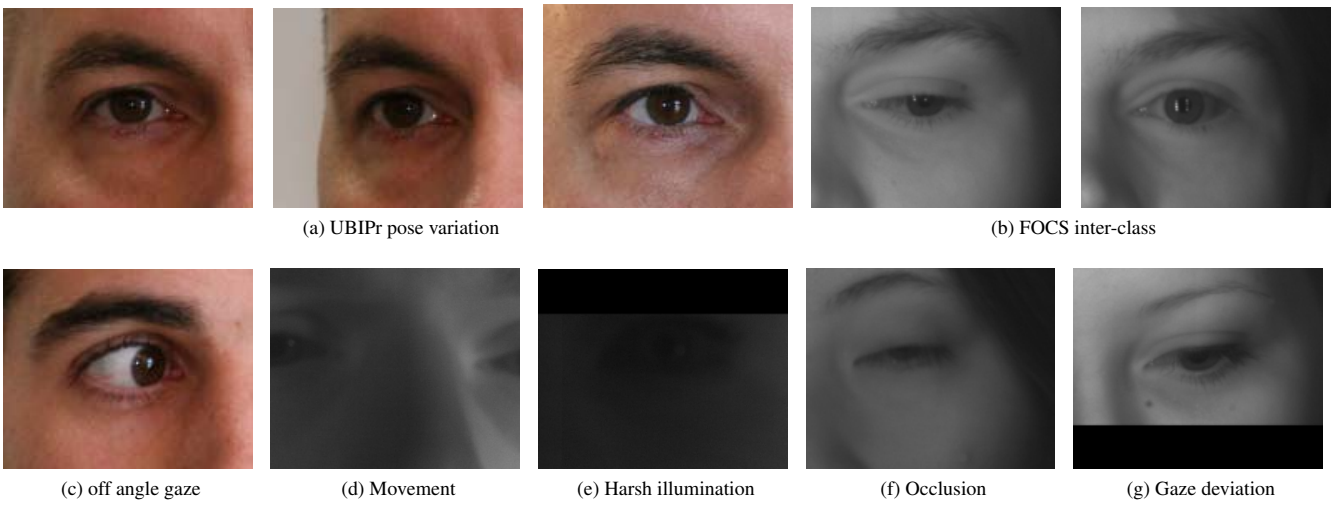
The Modified SIFT (m-SIFT) method described in [13], is a good method for basic point-to-point matching. Implementation involves the use of the Scale Invariant Feature Transform (SIFT) technique [5] (we used an open source implementation from the VLFeat library [16]) with additional matching constraints dependent on the location of similar keypoints, added to reduce the number of SIFT descriptors that need to be compared. The first constraint is looking at the angle of the matching keypoints between a query image and the template, while the second constraint involves calculating the distance. For a keypoint, p, located at a pixel location (x_1, y_1) in the query and (x_2, y_2) in the template, the magnitude of the angle (in degrees) and Euclidean distance are calculated between the two coordinates. In the original implementation of m-SIFT, to determine a match score, M, the number of matching keypoints are counted; however, we found a slight decrease in error rate by doing the following:

$$M = \begin{cases} \frac{-\text{mode}[\mathbf{E}]}{k} & k > 0 \\ \varepsilon & k = 0 \end{cases}$$

where ε is some large negative number, k is the number of matching keypoints (after applying the additional matching constraints), and $\mathbf{E}$ is a vector of the squared Euclidean distance between the the remaining keypoints (after applying the additional matching constraints). More details on m-SIFT can be found within [13].

4. Experiments

We investigate four different periocular regions as seen in Figure 2, referred to as 'Region 1', 'Region 2', 'Region 3', and 'Region 4'. Experiments over these regions



(a) UBIPr pose variation (b) FOCS inter-class

(c) off angle gaze (d) Movement (e) Harsh illumination (f) Occlusion (g) Gaze deviation

Figure 5: Examples of images from the UBIPr (color VL images) and FOCS (gray-scale NIR images) data sets.

allows for the results to be different from each other with respect to the amount of periocular information available while also considering the level of granularity in determining an appropriate periocular region for recognition. In each of the tests presented we use equal error rate (EER) as a measure of overall system performance displaying receiver operating characteristic (ROC) curves as well as cumulative match characteristic (CMC) curves for accuracy rates. Sub-regions effects are derived from the average score of each patch from PDM tests and the average of the weighted locations from m-SIFT tests. Due to space considerations sub-regions are only examined from authentic comparisons detailing the specific areas of the ocular image that are most important within an authentic match.

4.1. Challenging Ocular Images

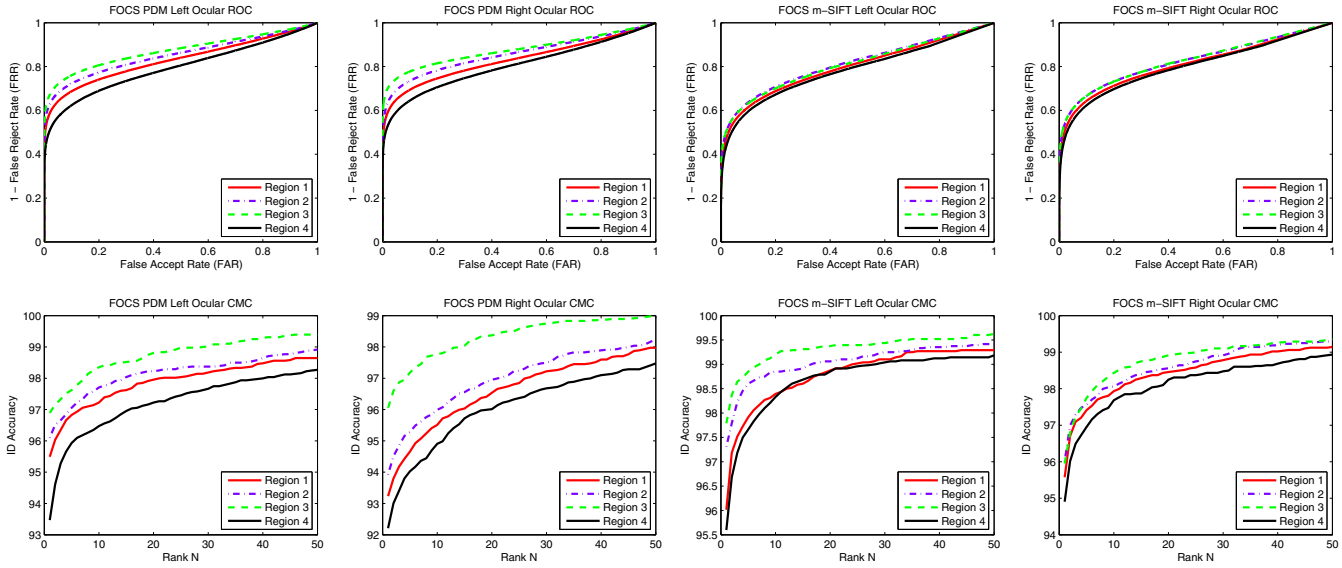
As shown in Figure 5, the FOCS and UBIPr images are challenging for both iris and ocular recognition. As a challenging data set for iris recognition, the FOCS images were captured from moving subjects (5d) with a near-infrared (NIR) sensor in an unconstrained environment showing drastic variations in sensor noise, illumination (5e), gaze (5g), and occlusion (5f). The UBIPr images captured with a visible light (VL) sensor, vary in pose (5a), gaze (5c), stand-off distance, and illumination. Pose variation images, which include frontal and both side views, can be seen in Figure 5a. The stand-off distance varies from 8 meters to 4 meters with varying resolution: 501×401 pixels for 8m, 561×541 pixels for 7m, 651×501 pixels for 6m, 801×651 pixels for 5m, and 1001×801 pixels for 4m.

An example of inter-class images found in the FOCS database, Figure 5b, shows that even at a resolution of 600×750 pixels, the small iris regions are difficult to extract from images for matching. Other issues such as specular reflections and out of focus blurring of the iris region seen in both the FOCS and UBIPr datasets favor the use of periocular recognition in each environment.

In total 19833 images are collected, 9581 (4792 left, 4789 right) from the FOCS set of 136 subjects and 10252 (5126 left, 5126 right) from the UBIPr set of 259 subjects. The number of samples per subject within the FOCS set vary between 2 ~ 236 samples/subject (however, most have at least 10 samples), while the UBIPr set is consistent with 30 samples/subject (with exception of two subjects which only have 29 samples and 27 samples total).

4.2. Experimental Setup

The method of evaluation is set up such that each image is independently compared against another (left vs left and right vs right) to generate two score matrices for the left (4792×4792 for FOCS and 5126×5126 for UBIPr) and right (4789×4789 for FOCS and 5126×5126 for UBIPr) ocular images for each region selection respectively. Thus, when using a correlation filter for determining deformation and/or measuring the similarity between a probe and gallery with PDM, only a single gallery image is used with no other authentic or impostor images included. Furthermore, all ocular images are re-sized to 128×128 pixels for computation purposes. Though to preserve quality and address the slightly varying size of the eye within each image (which is measured from corner to corner), computing the normalized regions shown in Figure 2 requires down-sampling each image such that the region choice (‘Region 1’, etc.) is present when cropped (from the center of the image) at 128×128



(a) PDM Left (b) PDM Right (c) m-SIFT Left (d) m-SIFT Right

Figure 6: ROC and CMC curves for left and right ocular regions on the FOCS dataset.

pixels. Additionally, to perform a basic normalization for illumination each image's histogram is equalized.

Since training is required for PDM, each test takes a randomly chosen set consisting of half of the left (2396 FOCS and 2563 UBIPr) and right (2394 FOCS and 2563 UBIPr) ocular images to perform authentic comparisons for parameter estimation. Empirical testing with PDM has led to dividing the region into 36 non-overlapping patches (6×6 rectangular configuration), and based on the recommendation from [13] we set m-SIFT angle constraints to 20° and distance constraints to 35% of the image height.

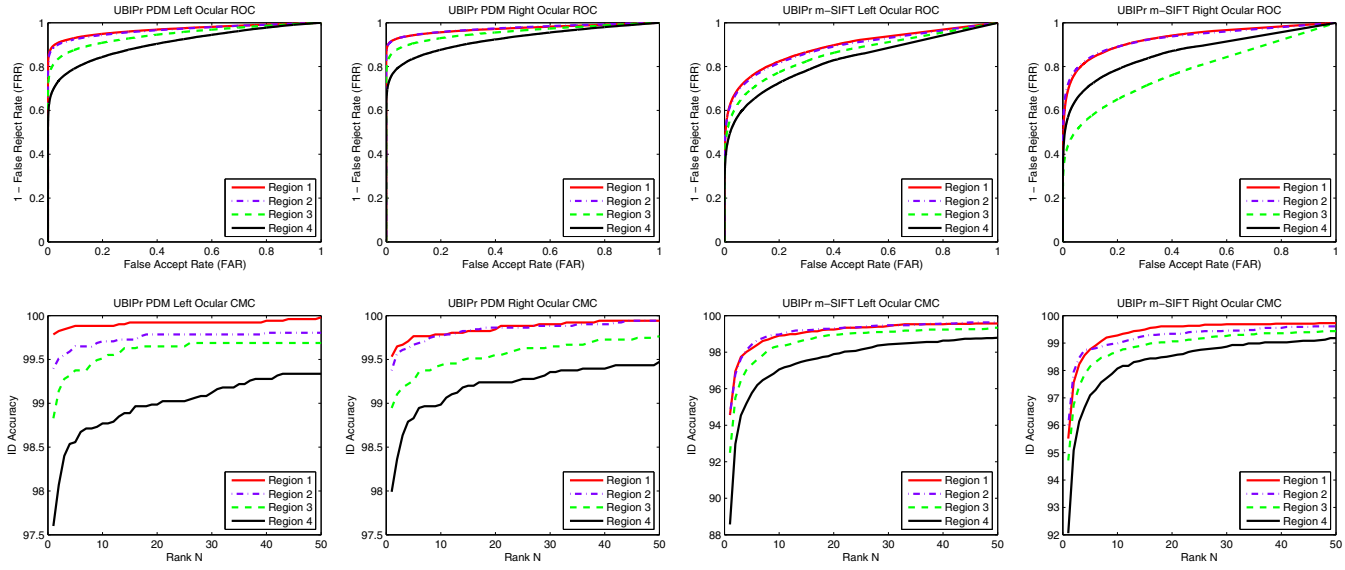
Visualization of effective periocular regions for matching is done for both PDM and m-SIFT experiments. For PDM results we examine average patch scores over the ocular region, while with m-SIFT we look at the sum of the keypoint locations weighted by their individual match score.

In each of the tests presented we use equal error rate (EER) as a measure of overall system performance. Within the FOCS database, each test on the left ocular region computes 22,963,264 comparisons and each test on the right ocular region computes 22,934,521 comparisons; of which 267,392 left (1.16%) and 267,273 right (1.17%) are authentic (262,600 left and 262,484 right, if not including self comparison). Thus, when moving along the diagonal of the receiver operating characteristic (ROC) curve (assuming an evenly 'bowed' ROC), a decrease in EER by 0.01% accounts to ~26 additional correctly classified authentic comparisons for both left and right regions, and ~2270 additional left and ~2267 right additional correctly classified impostor comparisons.

The UBIPr database on the both ocular regions each have 26,275,876 comparisons; of which 76,840 (0.29%) from each side are authentic (71,714, if not including self comparisons). Similarly, when moving along the diagonal of the ROC curve (assuming an evenly 'bowed' ROC), a decrease in EER by 0.01% accounts to ~8 additional correctly classified authentic comparisons, and ~2620 additional correctly classified impostor comparisons for both left and right regions. Note that the labeled right periocular images found in the FOCS dataset are an equivalent user point of view to the images labeled as the left periocular region in the UBIPr dataset and vice-versa.

5. Results

The results from each test are in Table 1 as equal error rates (EER), while Figures 6 and 7 show the ROC and CMC curves on the FOCS and UBIPr datasets respectively. The average patch scores from PDM and normalized remaining keypoints (weighted at their respective locations) from each authentic comparison are displayed in Figure 8. Note that each set of sub-region scores for the FOCS dataset (Figure 8a) are scaled such that the minimum value is 0 (shown as dark blue regions, with black as 0) and the maximum value (dark red) is 0.001 and 0.032 for m-SIFT and PDM respectively. The scaling for the UBIPr database (Figure 8b) is set such that the minimum value is 0 (again shown as dark



(a) PDM Left (b) PDM Right (c) m-SIFT Left (d) m-SIFT Right

Figure 7: ROC and CMC curves for left and right ocular regions on the UBIPr dataset.

blue, with black as 0), and the maximum value (dark red) is 0.002 and 0.0245 for m-SIFT and PDM respectively. This adjusted scale setup allows for an easier visual comparison between neighboring regions. Next to each sub-region score distribution from each experiment is the average periocular region that was tested with the respective sub-region scores overlayed.

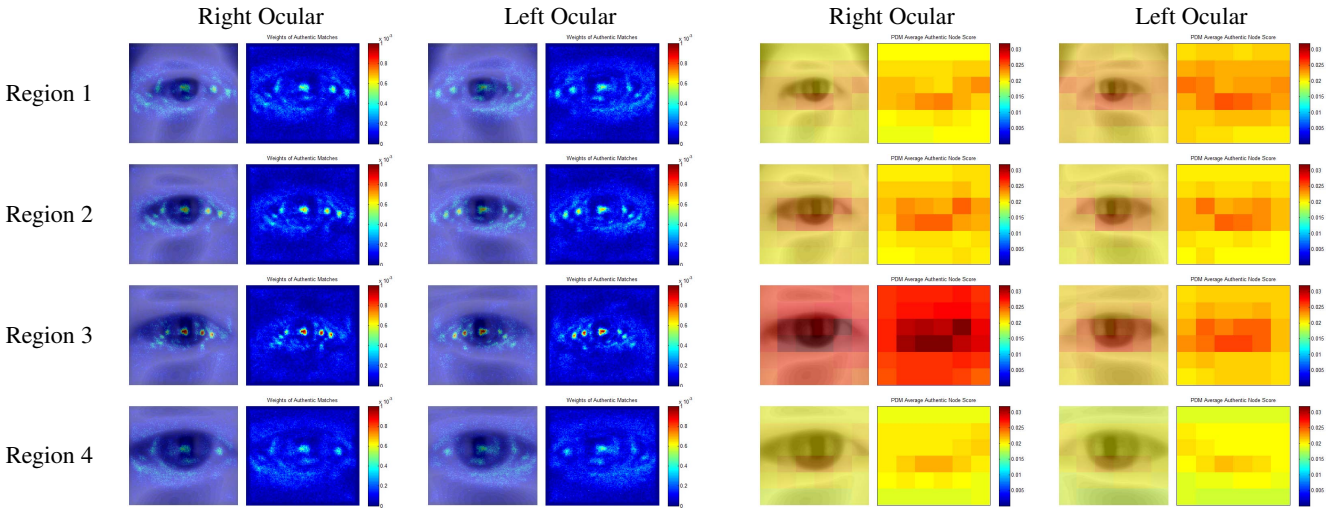
		FOCS		UBIPr	
		m-SIFT	PDM	m-SIFT	PDM
Region 1	Right	25.90%	23.95%	14.03%	6.43%
	Left	27.45%	24.23%	18.44%	7.62%
Region 2	Right	24.64%	21.24%	13.63%	6.43%
	Left	26.18%	21.93%	18.95%	8.09%
Region 3	Right	24.67%	18.85%	29.37%	9.57%
	Left	26.64%	19.57%	21.66%	11.62%
Region 4	Right	26.85%	26.60%	20.87%	14.31%
	Left	28.40%	27.55%	24.67%	16.89%

Table 1: EER for each tests on the FOCS and UBIPr datasets using m-SIFT and PDM methods.

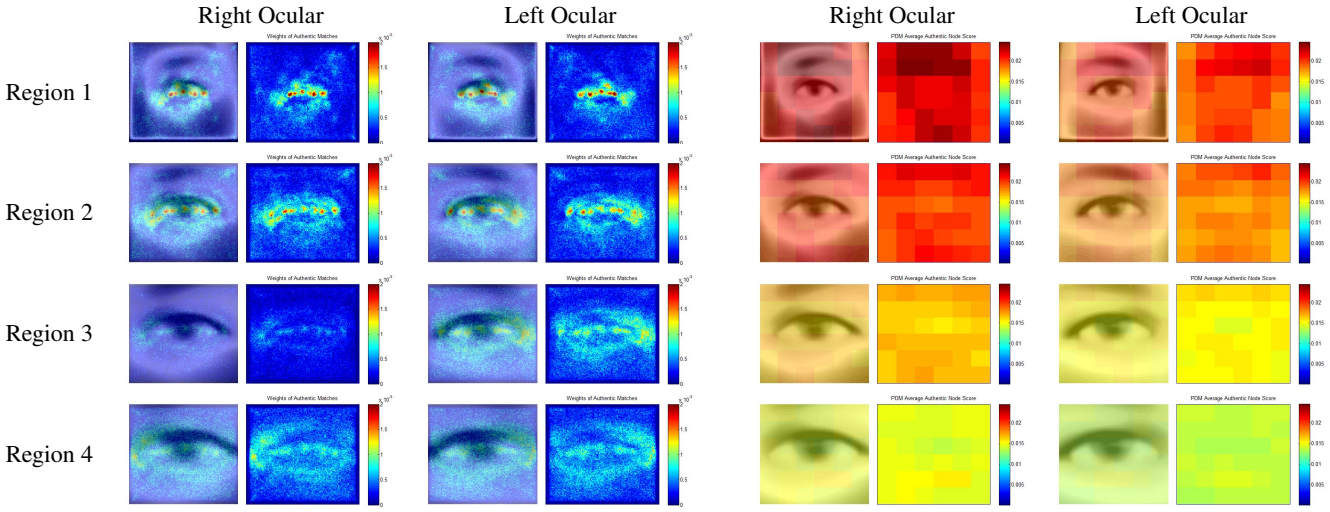
The EERs for the FOCS dataset using m-SIFT are relatively close across all periocular regions, varying by ~1-2% EER, with the best found at Region 2 (though Region 3 is within ~0.03-0.46% EER). PDM shows its best EERs at Region 3 with significantly more variation between the other regions. Referring to the sub-region scores for the FOCS database in Figure 8a, both left and right ocular images for the PDM method shows the highest scores across the eye itself with consistently smaller scores on the cheek and brow areas when present. The m-SIFT keypoint weights reflect a similar result with scores being much larger on the eye itself than in surrounding areas. Specifically in both sub-region results the area near the inner tear duct of the eye and lower boundary between the eye and cheek appear to contribute the most to the final match scores; which may help explain why Region 3 outperforms Region 4 in NIR images. Region 4 appears to cut off part of the inner tear duct while also completely removing any possible brow information, which while scoring low, does contribute to the final match score more than any skin texture.

Based on high sub-region score locations and the EER performance of the regions tested the best cropping for NIR images requires a tight cut around the eye itself with a small amount area just outside of eye corners. Including some brow information has shown to help, though cheek and skin texture may not be necessary.

The results on the UBIPr dataset of VL images show a large degradation in recognition performance for both m-SIFT and PDM as the amount of area surrounding the eye decreases. From the EERs found in Table 1, we see that for both methods Region 1 and Region 2 are roughly equivalent, while there is a decrease in performance from using Region 3 or Region 4. Referring to Figure 8b, the sub-region scores for both methods are able to display and help explain the value of using the surrounding skin texture when



(a) FOCS



(b) UBIPr

Figure 8: Average authentic patch scores from PDM and keypoint weights (at their respective locations) from m-SIFT on the FOCS (a) and UBIPr (b) datasets. Note that the labeled right periocular images found in the FOCS dataset are an equivalent user point of view to the images labeled as the left periocular region in the UBIPr dataset and vice-versa.

performing periocular recognition with VL images. Viewing the m-SIFT sub-region scores we see that the largely weighted keypoints are not only centered on the eye itself but are also abundantly spread outside of the eye. This is distinctly different than what is shown in the results on the FOCS dataset in Figure 8a where the larger keypoint weights are located very close to the eye boundaries with little noticeable clustering throughout the area around the eye. In addition, the UBIPr sub-region scores from PDM favor the eyebrow/brow region of the periocular image with consistently larger patch scores in areas where the brow is present. Comparatively higher PDM patch scores are also seen in on the skin/cheek areas under the eye (which are largest in Regions 1 & 2) as opposed to at the boundaries or tear ducts.

Based on high sub-region score locations and the EER performance of the regions tested the best cropping for VL images requires a large cut around the eye including the eyebrow and surrounding skin texture (cheek) for necessary discrimininative information.

6. Conclusion

In this paper we examined the question of what is the best periocular region for recognition when using near-infrared (NIR) or visible light (VL) sensors. Two fundamentally different algorithms were tested on challenging periocular datasets of contrasting build and with the effects of each evaluated at a sub-region level. Our results for VL images indicate a slight contradiction to previous work [10], where we found that the cheek region can be beneficial to recognition performance, however the eyebrow/brow region is the most discriminative, with eye shape having the smallest contribution. This suggests using a larger area surrounding the eye (Region 1) for periocular recognition with VL images. While the opposite was seen with NIR images; finding that the shape of the eye is more important, specifically the inner tear duct and lower boundary between the eye and cheek, with the brow contributing the least. This allows for better recognition performance with a smaller periocular region (Region 3) when using NIR images.

References

[1] D.R. Ambika, K. R. Radhika, and D. Seshachalam. The eye says it all: Periocular region methodologies. In *Multimedia Computing and Systems (ICMCS), 2012 International Conference on*, pages 180–185, May 2012.

[2] Vishnu Naresh Boddeti, Jonathon M Smereka, and B.V.K. Vijaya Kumar. A comparative evaluation of iris and ocular recognition methods on challenging ocular images. In *Biometrics (IJCB), 2011 International Joint Conference on*, pages 1 –8, oct. 2011.

[3] K.P. Hollingsworth, S.S. Darnell, P.E. Miller, D.L. Woodard, E. Ortiz, and K.W. Bowyer. Human and machine performance on periocular biometrics under near-infrared light and visible light. *Information Forensics and Security, IEEE Transactions on*, 7(2):588–601, April 2012.

[4] Felix Juefei-Xu, Khoa Luu, Marios Savvides, Tien D. Bui, and Ching Y. Suen. Investigating age invariant face recognition based on periocular biometrics. In *Biometrics (IJCB), 2011 International Joint Conference on*, pages 1 –7, oct. 2011.

[5] David G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60(2):91–110, November 2004.

[6] J. Merkow, B. Jou, and M. Savvides. An exploration of gender identification using only the periocular region. In *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, pages 1–5, Sept. 2010.

[7] P.E. Miller, J.R. Lyle, S.J. Pundlik, and D.L. Woodard. Performance evaluation of local appearance based periocular recognition. In *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, pages 1–6, Sept. 2010.

[8] NIST. Face and ocular challenge series (focs).

[9] NIST. Multiple biometric grand challenge (mbgc).

[10] Beom-Seok Oh, Kangrok Oh, and Kar-Ann Toh. On projection-based methods for periocular identity verification. In *Industrial Electronics and Applications (ICIEA), 2012 7th IEEE Conference on*, pages 871–876, July 2012.

[11] C.N. Padole and H. Proenca. Periocular recognition: Analysis of performance degradation factors. In *Biometrics (ICB), 2012 5th IAPR International Conference on*, pages 439–445, April 2012.

[12] Unsang Park, R.R. Jillela, A. Ross, and A.K. Jain. Periocular biometrics in the visible spectrum. *Information Forensics and Security, IEEE Transactions on*, 6(1):96 –106, 2011.

[13] A. Ross, R. Jillella, V.N. Boddeti, J. Smereka, B.V.K. Vijaya Kumar, R. Barnard P. Pauca, and R. Plemmons. Matching highly non-ideal ocular images: An information fusion approach. *Intl. Conf. on Biometrics (ICB)*, March 2012.

[14] J. Thornton, M. Savvides, and V. Kumar. A bayesian approach to deformed pattern matching of iris images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(4):596 –606, 2007.

[15] Muhammad Uzair, Arif Mahmood, Ajmal Mian, and Chris McDonald. Periocular biometric recognition using image sets. In *Applications of Computer Vision (WACV), 2013 IEEE Workshop on*, pages 246–251, Jan. 2013.

[16] A. Vedaldi and B. Fulkerson. Vlfeat: An open and portable library of computer vision algorithms, 2008.

[17] D.L. Woodard, S. Pundlik, P. Miller, R. Jillela, and A. Ross. On the fusion of periocular and iris biometrics in non-ideal imagery. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 201–204, Aug. 2010.

[18] D.L. Woodard, S.J. Pundlik, J.R. Lyle, and P.E. Miller. Periocular region appearance cues for biometric identification. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 162 –169, 2010.

[19] Juefei Xu, M. Cha, J.L. Heyman, S. Venugopalan, R. Abiantun, and M. Savvides. Robust local binary pattern feature sets for periocular biometric identification. In *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, pages 1–8, Sept. 2010.