<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper presented at *IEEE International Conference on Image Processing, ICIP, Paris, France, 27-30 October, 2014*.

Citation for the original published paper:

Alonso-Fernandez, F., Bigun, J. (2014)

Best Regions for Periocular Recognition with NIR and Visible Images.

In: *2014 IEEE International Conference on Image Processing (ICIP)* (pp. 4987-4991). Piscataway, NJ: IEEE Press

<http://dx.doi.org/10.1109/ICIP.2014.7026010>

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:hh:diva-25468>

BEST REGIONS FOR PERIOCLAR RECOGNITION WITH NIR AND VISIBLE IMAGES

Fernando Alonso-Fernandez, Josef Bigun

Halmstad University. Box 823. SE 301-18 Halmstad, Sweden
{feralo, josef.bigun}@hh.se, <http://islabb.hh.se>

ABSTRACT

We evaluate the most useful regions for periocular recognition. For this purpose, we employ our periocular algorithm based on retinotopic sampling grids and Gabor analysis of the spectrum. We use both NIR and visible iris images. The best regions are selected via Sequential Forward Floating Selection (SFFS). The iris neighborhood (including sclera and eyelashes) is found as the best region with NIR data, while the surrounding skin texture (which is over-illuminated in NIR images) is the most discriminative region in visible range. To the best of our knowledge, only one work in the literature has evaluated the influence of different regions in the performance of periocular recognition algorithms. Our results are in the same line, despite the use of completely different matchers. We also evaluate an iris texture matcher, providing fusion results with our periocular system as well.

Index Terms— Biometrics, periocular, eye, Gabor filters

1. INTRODUCTION

Periocular recognition has gained attention recently as a promising trait for unconstrained biometrics [1]. It refers to the region in the immediate eye vicinity, including eyelids, lashes and eyebrows. This region can be easily obtained with existing face and iris setups, and the requirement of user cooperation can be relaxed. An evident advantage is its availability over a wide range of distances even when the iris texture cannot be reliably obtained (low resolution) or under partial face occlusion (close distances). Most face systems use a holistic approach, requiring a full face image, so occlusion affects performance dramatically [2]. Periocular region has also shown superior performance than face under extreme blur or down-sampling [3]. In addition, the periocular region appears in iris images, so fusion with the iris texture has potential to improve the overall recognition [4].

The study [5] identified which ocular elements humans find more useful for periocular recognition. With NIR im-

Fig. 1. Sampling grid configuration. Parameter d indicates the horizontal/vertical distance between adjacent points. Points with ‘X’ correspond to the iris texture region (case ‘ir’ of Figure 5). The remaining points are considered to mostly capture the region outside the iris ring (case ‘nir’ of Figure 5).

ages, eyelashes, tear ducts, eye shape and eyelids, were identified as the most useful, while skin was the less useful. But for visible data, blood vessels and skin were reported more helpful than eye shape and eyelashes. A similar study was done in [6], but with automatic algorithms (Probabilistic Deformation Models (PDM) [7] and m-SIFT [8]). This is, to the best of our knowledge, the only work evaluating the contribution of periocular regions to the performance of machine algorithms. Results were consistent with the study made with humans. With NIR images, regions around the iris (including the inner tear duct and lower eyelash) were the most useful, while cheek and skin texture were the less important. With visible images, on the other hand, the skin texture surrounding the eye was found very important, with the eyebrow/brow region (when present) also favored in visible range.

Fig. 2. Relative individual performance (EER) of each grid point. Magnitudes are re-scaled, so black color represents the minimum value, while white color represents the maximum. The row/column numbers correspond to those in Figure 1.

F. A.-F. thanks Swedish Research Council and EU Marie Curie program for funding his postdoctoral work. Authors acknowledge the CAISR program (Swedish Knowledge Foundation), the EU BBfor2 project and the EU COST Action IC1106. Authors also thank the Biometric Recognition Group (ATVS-UAM) for making the iris BioSec database available.

In previous research [9], we proposed a periocular system based on uniform sampling grids positioned in the pupil

center (Figure 1), followed by Gabor decomposition of the spectrum. The system achieved competitive error rates with respect to other periocular approaches [1], with results reported using data acquired with near-infrared (NIR) illumination. This paper evaluates the influence of different periocular regions on the performance of our system by considering Gabor responses from selected grid points only. We look for the best combination of grid points by Sequential Forward Floating Selection (SFFS) [10]. A second novelty with respect to [9] is the inclusion of data acquired in visible range. With NIR data, the eye region is found to be the best in our analysis. On the other hand, the best region with visible data is the surrounding skin texture. Our results are in the same line as those in [6], despite we use a different machine algorithm. This is not to affirm that these are universally the best regions for periocular recognition, but that the features of the two studies behave equally under the same conditions. Other popular features also proposed for periocular recognition, such as Local Binary Patterns or Histograms of Gradient Orientation [1] (not tested in this study) may lead to a different result, and would need of additional studies. We also evaluate an iris texture matcher based on 1D Log-Gabor wavelets. The performance of this matcher is considerably worse under visible data, but it is able to complement our periocular system, obtaining better performance with the fusion of the two systems. This complementarity is not observed with NIR images.

2. PERIOULAR RECOGNITION SYSTEM

The system used is described in [9], which is based on the face detection and recognition system of [11, 12]. It uses a sparse retinotopic sampling grid positioned in the eye center. The grid has rectangular geometry, with uniform sampling points (Figure 1). We use a relatively low dense grid, as we have observed that more dense grids do not necessarily give better performance [13]. The local power spectrum is sampled at each cell of the grid by applying a set of Gabor filters organized in 5 frequency channels and 6 equally spaced orientation channels. For each grid point, we group the magnitude of its Gabor responses into a vector $\mathbf{v} = (v_1, \dots, v_N)$ of $N=5 \times 6=30$ elements. Matching between two images is done by computing the χ^2 distance [14] between corresponding points of the grid. According to Figure 1, this results in $7 \times 9=63$ sub-distances (BioSec database) and $5 \times 7=35$ sub-distances (MobBIO) between any two given images, one sub-distance per grid point. Prior to matching with magnitude vectors, they are normalized to a probability distribution (PDF). For this purpose, each vector element v_i ($i = 1, \dots, N$) is divided by the sum of all vector elements $\sum_{j=1}^N v_j$. We also carry out experiments without separating the responses of each grid point. For this purpose, all Gabor responses of all grid points are grouped into a single vector, resulting in one distance between any two images.

Fig. 3. Verification performance for an increasing number of grid points selected. Performance without selection (using all grid points) is also given for reference.

3. DATABASES AND EXPERIMENTAL PROTOCOL

We use the BioSec [15] and MobBIO [16] databases. From BioSec, we select 1,200 images from 75 individuals acquired in 2 sessions (4 images of each eye per person, per session). Images are of 480×640 pixels, acquired with a LG IrisAccess EOU3000 close-up infrared iris camera. MobBIO has been captured with the Asus Eee Pad Transformer TE300T Tablet (a webcam in visible light) in one session. Images in MobBIO were captured in two different lightning conditions, with variable eye orientations and occlusion levels. Distance to the camera was kept constant, however. Here, we use the training dataset, with 800 iris images of 200×240 pixels from 100 individuals (4 images of each eye per person). We have manually annotated the two databases [17], computing the radius and center of the pupil and sclera circles, which are used as input for the experiments. Similarly, we have also modeled eyelids as circles, computing the radius and center of those circles too. Due to different image size, Gabor filter wavelengths span from 4 to 16 pixels with MobBIO and 16 to 60 with BioSec. This covers approximately the range of pupil radius of each database, as given by the groundtruth.

We carry out verification experiments. We consider each eye as a different user (200 available users in MobBIO, 150 in BioSec). Experiments with MobBIO are as follows. Genuine matches are done by comparing each image of a user to his/her remaining images, avoiding symmetric matches. Impostor matches are obtained by comparing the 1st image of a user to the 2nd image of the remaining users. We then get $200 \times 6=1,200$ genuine and $200 \times 199=39,800$ impostor matchings. With BioSec, genuine matches for a given user are obtained by comparing all images of the 1st session to all images of the 2nd session. Impostor matches are obtained by comparing the 2nd image of the 1st session of a user to the 2nd image of the 2nd session of the remaining users. We then obtain $150 \times 4 \times 4=2,400$ genuine and $150 \times 149=22,359$ impostor matchings. Note that experiments with BioSec are made by matching images of different sessions, but these inter-session experiments are not possible with MobBIO.

We also conduct matching experiments of the iris texture [18]. The iris ring-shaped region is unwrapped to a normalized rectangle [19] and next, a 1D log-Gabor wavelet is ap-

plied plus phase binary quantization to 4 levels. Matching between binary vectors is done with the normalized Hamming distance [19], which incorporates the noise mask (given here by the eyelids groundtruth). Rotation is accounted for by shifting the query image in counter- and clock-wise directions, and selecting the lowest distance, which corresponds to the best match between two templates. Some fusion experiments are also done between the periocular and the iris matchers. The fused distance is computed as the mean value of the distances given by the two matchers, which are first normalized to the same range using tanh-estimators [20].

Fig. 4. Grid points chosen by the SFFS algorithm. Two particular cases are highlighted: *i*) when the performance with SFFS reaches that obtained by using all grid points and *ii*) when the best performance with SFFS is obtained. Red arrows indicate the range around the minimum EER where the performance does not vary considerably (see Figure 3).

4. RESULTS

Figure 2 shows the individual performance of each grid point, computed using its corresponding sub-distance. The best performance with BioSec is given the iris region (note the black ‘U’ around the center). Contrarily, these points are among those performing worse in MobBIO. Other ‘good’ (black-ish) regions with BioSec are around the iris neighborhood, which correspond mostly to the sclera and lower eyelids. On the other hand, skin regions (the first two rows) have the worst performance. In MobBIO, on the contrary, skin regions (those far away from the grid center) have the best performance.

We then evaluate the combination of grid points. The best grid points are found by Sequential Forward Floating Selection (SFFS) [10]. Given n grid points to combine, the resulting matching distance between two images is obtained by averaging its n sub-distances. As criterion value of the SFFS algorithm, we use the EER given by the n averaged sub-distances. Figure 3 gives the performance results as we increase the number of grid points selected with SFFS, while Figure 4 depicts the actual grid points selected (superimposed

on an eye image for better assessment of the selected regions). We also give in Figure 3 results of grid points combination selected on the basis of its individual performance (the best n performing ones). The performance without selection (using all grid points, see Section 2) is also given for reference.

A substantial performance improvement is observed with an appropriate combination of grid points (the EER is reduced by more than 25% with SFFS in the two databases). Also, SFFS selection leads to better performance than combining the best individual grid points. This is expected, since the combination of top individual classifiers may be outperformed by other appropriate combination of weaker classifiers [21, 22]. The best performance is obtained with a combination of 26 grid points (BioSec) and 21 (MobBIO). DET curves of these cases are given in Figure 5, top (‘26c’ and ‘21c’). In addition, the performance with BioSec remains somewhat constant in the range of 22-28 grid points, and 20-25 with MobBIO. It is worth noting that the optimum number of grid points is quite similar for the two databases, despite different image size, acquisition conditions, sampling grid configuration, or wavelength span of the Gabor filters.

When the grid points are properly chosen, the performance obtained by using all grid points can be improved by combining a small number of points only (6 with BioSec, 5 with MobBIO). DET curves of these are also given in Figure 5, top (‘6c’ and ‘5c’). It is also observed an initial sharp improvement in performance, which stabilizes at around 10 grid points. Once that the best performance is reached, the addition of more points actually degrades the EER. Again, it is worth highlighting that despite the obvious difference between the two databases, the number of grid points at which relevant events happen are very similar in both databases.

We observe interesting phenomena by analyzing the grid points selected with SFFS (Figure 4). With BioSec, points of the iris region are chosen from the beginning. When the best performance is obtained (given by the red arrow), selected points include the iris, sclera and eyelashes. The surrounding skin, on the contrary, is mostly discarded. This can be due to the NIR illumination, which reveals the details of the iris texture [23], but over-illuminates the surrounding skin region (hiding its texture details). With MobBIO, on the other hand, the best performance is obtained with points mostly pertaining to the skin region, sclera and eyelashes. Contrarily to BioSec, when the number of selected points is low, the iris region is never chosen. Also worth noting, the optimal configuration in both databases never includes the center of the grid. The black region of the pupil (captured by this point) is not expected to be very different between individuals, thus it does not provide useful, discriminative information.

Based on the facts noted in the previous paragraph, we have also conducted experiments where we manually select: *i*) grid points capturing the iris region, and *ii*) grid points capturing the region outside the iris (see Figure 1). DET curves of these are given in Figure 5, bottom (curves ‘ir’ and ‘nir’, re-

spectively). With Biosec, these two selections perform worse than the whole grid (curve ‘ns’) and much worse than the optimal SFFS selection (‘26c’). This suggests that both regions (iris texture and the surrounding neighborhood) are important for periocular recognition in this database, at least with our matcher. On the other hand, with MobBIO, points outside the iris ring (‘nir’) exhibit a performance similar to the optimal SFFS selection (‘21c’), while performance with points of the iris region (‘ir’) is substantially degraded even. One may suggest that the smaller resolution of MobBIO results in an unusable iris texture. However, the fusion experiments with an iris matcher (analyzed next), shows that the iris texture can be very complementary to our periocular matcher.

Figure 5 also gives performance of the iris matcher and the fusion with our periocular system. The iris matcher works much better with BioSec, which is reasonable since iris systems usually work better in NIR range [5]. An additional factor could be the difference in image size between the two databases, and the more adverse acquisition conditions of MobBIO. It is relevant also that the periocular system works better than the iris matcher in MobBIO. The small image size makes more difficult to reliably extract identity information from the iris texture. When it comes to complementarity, however the fusion of the iris and periocular systems does improve performance with MobBIO, which is not the case with BioSec (see the fusion cases of Figure 5, top).

It is worth highlighting that, with BioSec, the fusion of the iris matcher (‘it’) with the periocular system working on the iris region (‘ir’) does not improve the performance (Figure 5, bottom). However, this same fusion case with MobBIO does improve the recognition performance of the best matcher. The two systems are shown to be complementary in visible light, even extracting features from the same eye region. The fusion of the iris matcher (‘it’) with the periocular system working outside the iris region (‘nir’) is even better with MobBIO, since features are extracted from different image regions. It is interesting that with BioSec, this fusion case (‘it+nir’) pushes the DET towards the best individual matcher (‘it’), even outperforming it for low FRR values. In some sense, this suggest that the two systems can also be complementary in visible light. The big difference in performance between the iris and periocular systems in BioSec can be one reason of the absence of improvement in most of the fusion cases analyzed here. Working towards the improvement of our periocular system, specially in visible images, can be one avenue to overcome this issue.

5. CONCLUSION

We study the influence of different periocular regions on the recognition performance. We use a periocular algorithm based on retinotopic sampling grids and Gabor analysis of the spectrum [9], which is evaluated with both NIR and visible iris images. The best periocular regions for each dataset are

Fig. 5. Verification results of the periocular system. Top: automatic selection using SFFS. Bottom: manual selection of grid points pertaining/not pertaining to the iris texture region. Fusion results with the iris matcher are also provided.

selected with SFFS [10]. We find the iris neighborhood as the best region with NIR data, while the surrounding skin texture is the most important area with visible images. These results are similar to others found in the literature with different matchers [6], and also consistent with the most useful features identified by human observers [5].

The size of the optimum periocular region is very similar for both databases (measured as the number of sampling grid points giving the best performance). This is very interesting, considering that the selected regions are different in each case, as well as other important differences between the two databases (image size, acquisition conditions, configuration of the sampling grid, or wavelength span of the Gabor filters). We also evaluate an iris texture matcher based on 1D Log-Gabor wavelets. Despite the poorer performance of the iris matcher with webcam data, its fusion with the periocular system results in improved performance. This complementarity is not observed with NIR images.

Currently, we are working on analyzing the effects of specific perturbations on periocular recognition, specially scale changes, and how the sampling grid and Gabor filters can be adapted to these conditions. This includes making use of periocular databases of higher resolution [1]. We are also working on the accurate localization of the eye center, with the aim of incorporating such stage to our periocular system [13].

6. REFERENCES

- [1] G. Santos and H. Proenca, "Periocular biometrics: An emerging technology for unconstrained scenarios," in *Proc. IEEE CIBIM*, April 2013, pp. 14–21.
- [2] Unsang Park, Raghavender R. Jillela, Arun Ross, and Anil K. Jain, "Periocular biometrics in the visible spectrum," *IEEE TIFS*, vol. 6, no. 1, pp. 96–106, 2011.
- [3] P. E. Miller, J. R. Lyle, S. J. Pundlik, and D. L. Woodard, "Performance evaluation of local appearance based periocular recognition," *Proc. IEEE BTAS*, 2010.
- [4] D. Woodard, S. Pundlik, P. Miller, R. Jillela, and A. Ross, "On the fusion of periocular and iris biometrics in non-ideal imagery," *Proc. ICPR*, 2010.
- [5] Karen Hollingsworth, Shelby Solomon Darnell, Philip E. Miller, Damon L. Woodard, Kevin W. Bowyer, and Patrick J. Flynn, "Human and machine performance on periocular biometrics under near-infrared light and visible light," *IEEE TIFS*, vol. 7, no. 2, pp. 588–601, 2012.
- [6] J.M. Smereka and B.V.K.V. Kumar, "What is a good periocular region for recognition?," in *Proc. IEEE CVPRW*, June 2013, pp. 117–124.
- [7] V.N. Boddeti, J.M. Smereka, and B.V.K.V. Kumar, "A comparative evaluation of iris and ocular recognition methods on challenging ocular images," in *Proc. IJCB*, Oct 2011, pp. 1–8.
- [8] A. Ross, R. Jillela, J.M. Smereka, V.N. Boddeti, B.V.K.V. Kumar, R. Barnard, Xiaofei Hu, P. Pauca, and R. Plemmons, "Matching highly non-ideal ocular images: An information fusion approach," in *Proc. ICB*, March 2012, pp. 446–453.
- [9] F. Alonso-Fernandez and J. Bigun, "Periocular recognition using retinotopic sampling and gabor decomposition," *Proc. WIAF, in conjunction ECCV*, vol. Springer LNCS-7584, pp. 309–318, 2012.
- [10] P. Pudil, J. Novovicova, and J. Kittler, "Flotating search methods in feature selection," *Pattern Recognition Letters*, vol. 15, pp. 1119–1125, 1994.
- [11] Fabrizio Smeraldi, O. Carmona, and Josef Bigün, "Saccadic search with gabor features applied to eye detection and real-time head tracking," *Image Vision Comput.*, vol. 18, no. 4, pp. 323–329, 2000.
- [12] Fabrizio Smeraldi and Josef Bigün, "Retinal vision applied to facial features detection and face authentication," *Pattern Recognition Letters*, vol. 23, no. 4, pp. 463–475, 2002.
- [13] F. Alonso-Fernandez and J. Bigun, "Eye detection by complex filtering for periocular recognition," *Proc. IWBF*, 2014.
- [14] A. Gilperez, F. Alonso-Fernandez, S. Pecharroman, J. Fierrez, and J. Ortega-Garcia, "Off-line signature verification using contour features," *Proc. ICFHR*, 2008.
- [15] J. Fierrez, J. Ortega-Garcia, D. Torre-Toledano, and J. Gonzalez-Rodriguez, "BioSec baseline corpus: A multimodal biometric database," *Pattern Recognition*, vol. 40, no. 4, pp. 1389–1392, April 2007.
- [16] Ana F. Sequeira, Jo ao C. Monteiro, Ana Rebelo, and Hélder P. Oliveira, "Mobbio: a multimodal database captured with a portable handheld device," *Proc. VIS-APP*, vol. 3, pp. 133–139, 2014.
- [17] H. Hofbauer, F. Alonso-Fernandez, P. Wild, J. Bigun, and A. Uhl, "A ground truth for iris segmentation," *Proc. International Conference on Pattern Recognition, ICPR*, 2014.
- [18] Libor Masek, "Recognition of human iris patterns for biometric identification," M.S. thesis, School of Computer Science and Software Engineering, University of Western Australia, 2003.
- [19] J. Daugman, "How iris recognition works," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 14, pp. 21–30, 2004.
- [20] A.K. Jain, K. Nandakumar, and A. Ross, "Score Normalization in Multimodal Biometric Systems," *Pattern Recognition*, vol. 38, no. 12, pp. 2270–2285, December 2005.
- [21] E.S. Bigun, J. Bigun, B. Duc, and S. Fischer, "Expert Conciliation for Multi Modal Person Authentication Systems by Bayesian Statistics," *Proc. AVBPA*, vol. Springer LNCS-1206, pp. 291–300, 1997.
- [22] J. Kittler, M. Hatef, R. Duin, and J. Matas, "On Combining Classifiers," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226–239, March 1998.
- [23] K.W. Bowyer, K. Hollingsworth, and P.J. Flynn, "Image understanding for iris biometrics: a survey," *Computer Vision and Image Understanding*, vol. 110, pp. 281–307, 2007.