Periocular Recognition Using Retinotopic Sampling and Gabor Decomposition

Fernando Alonso-Fernandez and Josef Bigun

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

Abstract. We present a new system for biometric recognition using periocular images based on retinotopic sampling grids and Gabor analysis of the local power spectrum. A number of aspects are studied, including: 1) grid adaptation to dimensions of the target eye vs. grids of constant size, 2) comparison between circular- and rectangular-shaped grids, 3) use of Gabor magnitude vs. phase vectors for recognition, 4) rotation compensation between query and test images, and 5) comparison with an iris machine expert. Results show that our system achieves competitive verification rates compared with other periocular recognition approaches. We also show that top verification rates can be obtained without rotation compensation, thus allowing to remove this step for computational efficiency. Also, the performance is not affected substantially if we use a grid of fixed dimensions, or it is even better in certain situations, avoiding the need of accurate detection of the iris region.

Keywords: Biometrics, periocular, eye, iris, Log-Polar mapping, Gabor decomposition.

1 Introduction and Related Work

Periocular recognition has gained attention recently in the biometrics field [1–9] due to demands for increased robustness of face or iris systems. Periocular refers to the face region in the immediate vicinity of the eye, including the eye, eyelids, lashes and eyebrows. Faces and irises have been extensively studied [10, 11], but periocular recognition has received revived attention recently, with suggestions that it may be as discriminative by itself as the face as a whole [5, 8]. Periocular region can be easily obtained with existing setups for face and iris, and the requirement of high user cooperation can be relaxed. An evident advantage is its availability over a wide range of acquisition distances even when the iris texture cannot be reliably obtained (low resolution, off-angle, etc.) [12] or under partial face occlusion (close distances). Most face systems use a holistic approach, requiring a full face image, so the performance is negatively affected in case of occlusion [10]. Also, the periocular region appears in iris images, so fusion with the iris texture has a potential to improve the overall recognition [9].

Most of the studies for periocular recognition have used Local Binary Patterns (LBP) [13] and, to a lesser extent, gradient orientation (GO) histograms [14] and

Scale-Invariant Feature Transform (SIFT) keypoints [15]. The best performance is consistently obtained with SIFT features (rank-one recognition accuracy: 81-94%, EER: 7%), followed by LBPs (rank-one: 74-87%, EER: 19%) and GO (rank-one: 67-90%, EER: 22%) [3, 7]. Comparison with face or iris is also done in some cases. For example, Park et al. [3] reported a rank-one accuracy of 99.77% using the whole face, but when the full face is not available (simulated by synthetically masking the face below the nose region), accuracy fell to 39.55%. This points out the strength of periocular recognition when only partial face images are available, for example in criminal scenarios with surveillance cameras, where it is likely that the perpetrator masks parts of his face. In the same direction, Miller et al. [5] found that, at extreme values of blur or down-sampling, periocular recognition performed significantly better than face. On the other hand, both face and periocular matching using LBPs under uncontrolled lighting were very poor, indicating that LBPs are not well suited for this scenario. Finally, Woodard et al. [9] fused periocular and iris information from near-infrared (NIR) portal data finding that periocular identification performed better than iris, and the fusion of the two modalities performed best. In most of these studies, periocular images were acquired in the visible range. Periocular on visible light works better than on NIR, because it shows melanin-related differences [7]. On the other hand, many iris systems work with NIR illumination due to higher reflectivity of the iris tissue in this range [16]. Unfortunately, the use of more relaxed scenarios will make NIR light unfeasible (e.g. distant acquisition, mobile devices, etc.) so there is a high pressure to the development of algorithms capable of working with visible light [17].

Here, we come up with a periocular recognition system based on retinotopic sampling grids positioned in the pupil center, followed by Gabor decomposition at different frequencies and orientations. This setup have been used in texture analysis [20], facial landmark detection and face recognition [2], and real-time face tracking and liveness assessment [1, 21], with high discriminative capabilities. We use the CASIA-IrisV3-Interval database [18] (2,655 eye images, 249 contributors), and the BioSec database [19] (3,200 images, 200 contributors). Although not directly comparable, our system achieves competitive verification rates in comparison with existing periocular recognition approaches [3, 7].

2   Recognition System

In our recognition system, input images are analyzed with a retinotopic sampling sensor, whose receptive fields consist in a set of modified Gabor filters designed in the log-polar frequency plane [1, 2]. The system is described next.

2.1   Sampling Grid

Our recognition strategy is based on a sparse retinotopic sampling grid obtained by log-polar mapping [1, 2], which is positioned in the pupil center (Figure 1, left). The grid has log-polar geometry, meaning that the density of sampling
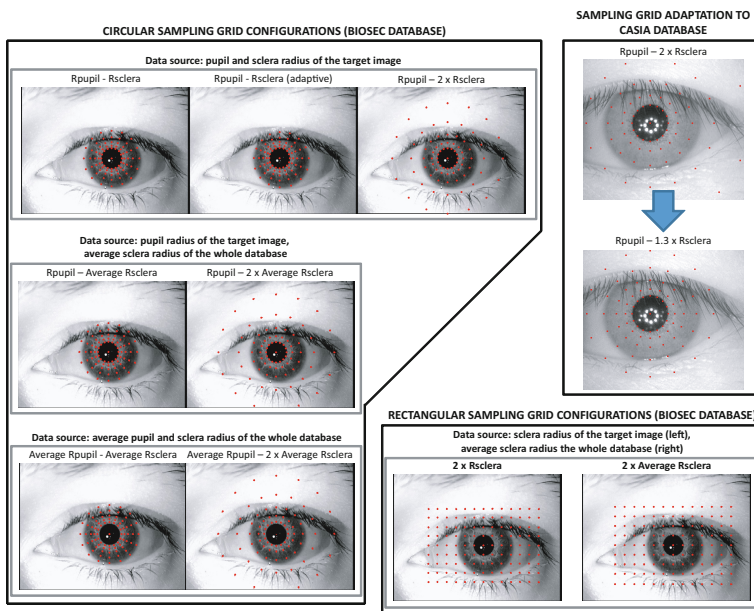
Fig. 1. Left: Circular sampling grid with different configuration of radius of the innermost and outermost circles. **Bottom right:** Uniform rectangular grids. Image is from the BioSec database. **Top right:** Sampling grid adaptation to the smaller images of CASIA database.

points decreases exponentially with the distance from the center. Such non-uniform sampling, with frequency decreasing from the center to the periphery, imitates the arrangement of photoreceptors in the human retina [22]. Each point of the grid is associated with a receptive field of the human eye. At each point, a Gabor decomposition of the image is performed to the effect that they mimic the simple cells of the primary visual cortex having the same receptive field but different spatial directions and frequencies [23]. The sparseness of the sampling grid allows direct filtering in the image domain without needing the Fourier transform, with significant computational savings [2] and even feasibility in real time [21]. In our experiments, we use a grid of 81 points arranged in 5 concentric circles, with 16 points per circle plus the point at the grid center. We also consider one case of non-concentricity between pupil and sclera circles (see the "adaptive" configuration in Figure 1, left, 1st row). For similitude with other previous works [3, 7], we also use a rectangular grid of 117 points, distributed uniformly in 9 rows and 13 columns (Figure 1, bottom right).

The circular grid is configured in several ways, employing different values of radius of the innermost and outermost circles: i) using pupil and sclera radius of the target image (Figure 1, left, 1st row), ii) using pupil radius of the target image only (2nd row), or iii) no usage of the pupil or sclera radius (3rd row). In cases ii) and iii), the average pupil and/or sclera radius of the whole database is used. This way, we evaluate the potential benefit of adapting the grid to the
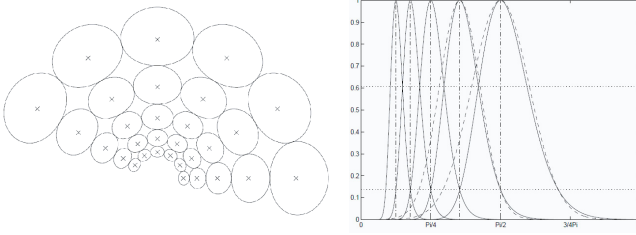
Fig. 2. Left: Iso-curves of the modified Gabor filters. Right: radial cross-sectional plot. Two standard Gabor filters are superimposed with dashed lines. Modified filters have a steeper cut-off on the low-frequency side, reducing overlap towards low frequencies. Image from [2].

dimensions of the target eye, compared with placing a grid of constant size. Similarly, width of the rectangular grid is built with two different configurations (Figure 1, bottom right), one using the sclera radius of the target image and the second using the average sclera radius. Height of the rectangular grid is 4/6 of its width [3]. Figure 1, left and bottom right, shows grid configurations with the BioSec database. Due to smaller image size, there is less periocular region available in CASIA database, thus some configurations of the sampling grid has to be reduced accordingly to ensure that it is mostly contained in the image (Figure 1, top right).

2.2 Log-Polar Sampling in the Fourier Domain

The local power spectrum of the image is sampled at each point of the grid by a set of Gabor filters that constitute the associated receptive field. The simple cells are modeled with 30 Gabor filters, organized in 5 frequency channels and 6 equally spaced orientation channels (Figure 2, left). When only a small number of frequency channels is used, the (standard) Gaussian spectrum of Gabor filters produces a non-uniform frequency coverage. Given that the central frequency of the filters increases exponentially, a symmetric Gaussian shape results in excessive overlap towards the low (densely sampled) frequencies and poor coverage of high frequencies. To solve that, a set of modified Gabor filters is used [1] defined as Gaussians in the log-polar frequency plane. For a filter tuned to orientation φ_0 and angular frequency $\omega_0 = \exp(\xi_0)$:

$$G(\xi, \varphi) = A \exp\left(-\frac{(\xi - \xi_0)^2}{2\sigma_\xi^2}\right) \exp\left(-\frac{(\varphi - \varphi_0)^2}{2\sigma_\varphi^2}\right) \tag{1}$$

where A is a normalization constant and (ξ, φ) are the log-polar frequency coordinates $left((\xi, \varphi) = \left(\log|\overrightarrow{\omega}|, \tan^{-1}(\omega_x, \omega_y)\right)$. This way, filters are designed by arranging a set of identical Gaussians in a rectangular lattice in the log-polar frequency plane. When seen in the standard Fourier plane, the overlap in low

frequencies is significantly reduced (Figure 2, right). Filter wavelengths span the range from 4 to 16 pixels in half-octave intervals.

The Gabor responses are grouped into a single complex vector, which is used as identity model. Matching between two images is using either the magnitude of complex values or a binary vector obtained by phase binary quantization to 4 levels. Prior to matching with magnitude vectors, they are normalized to a probability distribution (PDF), and matching is done using the χ^2 distance [24]. Matching between binary vectors is done using the Hamming distance [16]. Rotation is accounted for by shifting the grid of the query image in counter- and clock-wise directions, and selecting the lowest distance, which corresponds to the best match between two templates.

Table 1. Verification results in terms of EER (CASIA database). The best case of each column is marked in bold. Results with rotation compensation: the relative EER variation with respect to no rotation compensation is given in brackets. R_p=pupil radius. R_s=sclera radius.

			Circular geometry - EER (%)			
			No rotation compensation		Rotation compensation	
	innermost circle	outermost circle	Gabor magnitude	Gabor phase	Gabor magnitude	Gabor phase
1	R_p	R_s	6.13	19.75	**5.85** (-4.59%)	6.88 (-65.17%)
2	R_p	R_s (adaptive)	**5.96**	**18.91**	**5.68** (-4.66%)	**5.66** (-70.04%)
3	R_p	$1.3 \times R_s$	6.95	21.55	6.73 (-3.05%)	10.45 (-51.48%)
4	R_p	avg(R_s)	6.33	20.25	6.18 (-2.38%)	7.52 (-62.85%)
5	R_p	$1.3 \times$avg(R_s)	7.22	22.13	6.97 (-3.48%)	10.90 (-50.75%)
6	avg(R_p)	avg(R_s)	7.53	25.40	7.32 (-2.79%)	19.57 (-22.95%)
7	avg(R_p)	$1.3 \times$avg(R_s)	7.99	26.28	7.68 (-3.96%)	22.18 (-15.60%)

			Rectangular geometry - EER (%)			
			No rotation compensation		Rotation compensation	
	height	width	Gabor magnitude	Gabor phase	Gabor magnitude	Gabor phase
1	$1.3 \times (4/6) \times R_s$	$1.3 \times R_s$	**7.89**	23.40	**7.81** (-1.11%)	15.83 (-32.37%)
2	$1.3 \times (4/6) \times$avg(R_s)	$1.3 \times$avg(R_s)	8.37	**23.23**	8.20 (-2.05%)	**15.64** (-32.66%)

3   Experiments

3.1   Databases and Protocol

We use the CASIA-IrisV3-Interval [18] and the BioSec baseline [19] databases. CASIA has 2,655 NIR images of 280×320 pixels (height×width) from 249 contributors in 2 sessions, with 396 different eyes (the number of images per contributor and per session is not constant, and not all the individuals have images of the two eyes). The BioSec database has 3,200 NIR images of 480×640 pixels from 200 individuals in 2 sessions. Each person contributes with 4 images of the two eyes per session (thus, 400 different eyes). We have manually annotated all images of the database, computing the radius and the center of the iris and sclera circles, which are used as input for the experiments.

Table 2. Verification results in terms of EER (BioSec database). The best case of each column is marked in bold. Results with rotation compensation: the relative EER variation with respect to no rotation compensation is given in brackets. R_p=pupil radius. R_s=sclera radius.

	innermost circle	outermost circle	Circular geometry - EER (%)			
			No rotation compensation		Rotation compensation	
			Gabor magnitude	Gabor phase	Gabor magnitude	Gabor phase
1	R_p	R_s	18.71	28.38	18.47 (-1.30%)	15.78 (-44.39%)
2	R_p	R_s (adaptive)	18.69	**27.52**	18.52 (-0.91%)	**13.88** (-49.56%)
3	R_p	$2 \times R_s$	16.67	32.76	16.86 (+1.11%)	25.36 (-22.58%)
4	R_p	avg(R_s)	19.09	29.39	18.66 (-2.23%)	16.34 (-44.37%)
5	R_p	$2 \times$avg(R_s)	16.84	32.94	16.59 (-1.49%)	25.46 (-22.73%)
6	avg(R_p)	avg(R_s)	17.46	38.93	17.49 (+0.20%)	31.10 (-20.12%)
7	avg(R_p)	$2 \times$avg(R_s)	**15.25**	37.88	**15.08** (-1.14%)	33.66 (-11.16%)

	height	width	Rectangular geometry - EER (%)			
			No rotation compensation		Rotation compensation	
			Gabor magnitude	Gabor phase	Gabor magnitude	Gabor phase
1	$2 \times (4/6) \times R_s$	$2 \times R_s$	**15.25**	**35.76**	**15.18** (-0.46%)	**29.97** (-16.19%)
2	$2 \times (4/6) \times$avg(R_s)	$2 \times$avg(R_s)	15.56	38.34	15.41 (-0.94%)	32.95 (-14.06%)

We consider each eye as a different user. Verification performance experiments with the CASIA database are as follows. Genuine matches are obtained by comparing each image of a user to the remaining images of the same user, avoiding symmetric matches. Impostor matches are obtained by comparing the 1st image of a user to the 1st image of the next 100 users. With this procedure, we obtain 9,018 genuine and 31,477 impostor scores. With the BioSec database, genuine matches for a given user are obtained by comparing the 4 images of the 1st session to the 4 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 all the remaining users. With this, we obtain $400 \times 4 \times 4$=6,400 genuine and 400×399=159,600 impostor scores. Note that experiments with the BioSec database are made by matching images of different sessions, but these inter-session experiments are not possible with CASIA-IrisV3-Interval, since it does not contain session information.

We conduct matching experiments of iris texture using 1D log-Gabor filters [25]. The iris region is unwrapped to a normalized rectangle using the Daugman's rubber sheet model [16] and next, a 1D Log-Gabor wavelet is applied plus phase binary quantization to 4 levels. Matching is done using the Hamming distance. Rotation is accounted for with the same procedure of Section 2.2.

Some fusion experiments are also done. Prior to fusion, scores of the individual systems are normalized to log-likelihood ratios. Given a score s_i, normalization is achieved as $f_i = a_0 + a_1 \cdot s_i$. Weights a_0, a_1 are trained via logistic regression as in [26]. This is done separately for each system. Normalized scores from each system are then summed for the fusion. This trained fusion approach it has shown superior performance than other fusion rules [26, 28].
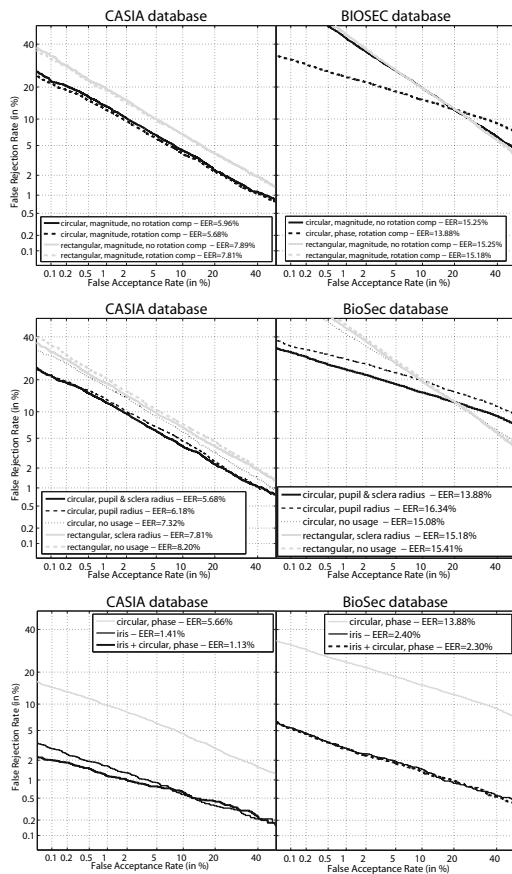
Fig. 3. First row: Circular vs. rectangular grid (best cases of Tables 1 and 2, with and w/o rotation compensation). **Second row**: Grid adaptation to the target image (best cases of Tables 1 and 2 depending whether the sampling grid makes use of the pupil and/or sclera radius values of the target image, with rotation compensation). **Third row**: Comparison of iris and periocular recognition (best cases of Tables 1 and 2, with rotation compensation). The best fusion combination found between our periocular system and the iris matcher is also shown.

3.2   Results

EER results with the different sampling grid configurations are given in Tables 1 and 2. It is observed that rotation compensation does not have appreciable effects in the Gabor magnitude (EER is reduced up to 4.7% in the best case), but there is a substantial improvement with phase vectors (up to 50% with BioSec, 70% with CASIA). The positive result, however, is that Gabor magnitude without rotation compensation performs similar to phase vectors after rotation compensation with circular grids, and even much better if we use square grids (EER is

about 50% less). This would allow to save computational time by suppressing rotation compensation.

Comparing circular and rectangular grids (Figure 3, 1st row), there is no clear winner. With CASIA, circular grids perform slightly better, but with BioSec, the best geometry depends on the DET region. Also remarkably, BioSec database has higher error rates than CASIA. Apart from differences given by the sensor (which cannot be assessed with the information available), one reason could be that experiments with BioSec are inter-session. On the contrary, one can think that BioSec images are bigger (480×640 vs. 280×320 pixels) with more periocular region available, so error rates should be lower. One way to assess the latter would be to increase the density and number of grid points with Biosec, a direction that we are currently exploring.

We now evaluate the effect of adapting the grid to the dimensions of the target eye (Figure 3, 2nd row). With CASIA and circular grids, adaptation both to the pupil and sclera radius is always best. However, if the outer dimension of the grid is fixed, performance is not too much affected (EER from 5.68% to 6.18%). With rectangular grids on CASIA, using a grid of constant dimensions does not have a dramatic impact neither (EER from 7.81% to 8.20%). Concerning BioSec database, the best case depends on the region of the DET. For low FAR, adaptation both to the pupil and sclera radius is always best. For low FRR however, both circular and rectangular grids with fixed dimensions are able to give the best performance.

Finally, we compare the performance of our periocular system with the iris expert (Figure 3, 3rd row). We also assess its fusion. We have tested all the fusion possibilities between the iris expert and the cases (rows) of Tables 1 and 2 (with rotation compensation only), reporting in Figure 3, 3rd row, the combination with the lowest EER. The best fusion on CASIA is with case 4 (phase) of Table 1. With BioSec, the best fusion is with case 6 (phase). As it can be observed, the dedicated iris machine expert works much better than our periocular approach. In addition, the fusion results in a relatively appreciable improvement on CASIA at low FAR, but no improvement is observed on BioSec.

4   Conclusion

We propose the use of retinotopic sampling grids positioned in the pupil center for recognition using periocular images. The local power spectrum is sampled at each grid point by a set of Gabor filters tuned to different frequencies and orientations [1, 2]. The system is evaluated with two databases acquired using iris sensors. One advantage of periocular systems is that existing setups for face and iris can be used for recognition purposes.

We compare the use of Gabor magnitude and phase vectors. We also carry out rotation compensation experiments by shifting the grid of the query image to find the best match with the test image. Results show that rotation compensation does not have appreciable effects in the Gabor magnitude, but produces a significative performance improvement with phase vectors. In any case, Gabor

magnitude without rotation compensation already performs at similar or better levels than phase vectors after rotation compensation.

Depending on the database, it is better to use circular or rectangular grids. We also evaluate the effect of adapting the grid to the dimensions of the target eye. Although adaptation is the optimal solution in most cases, the performance is not too much affected if we use a grid of fixed dimensions, or it is even better in certain situations. The only requirement of our system is the availability of the center of the eye and, in some cases, of the pupil radius. These are available even when the iris texture is difficult to extract [12]. Finally, we compare our system with a dedicated iris expert using 1D Log-Gabor wavelets [25]. We observe that the iris expert works much better than our periocular approach, and the fusion of the two results in a slight improvement with one of the databases.

Future work includes comparison and fusion of our approach with other existing periocular recognition algorithms [7] and evaluation on images in the visual spectrum [3]. Also, existing works do not focus on detection of the periocular region (it is manually extracted), but on texture analysis only. Only Park et al. [3] used a Viola-Jones face detector plus heuristics measurements to extract the periocular region, so successful extraction relies on an accurate detection of the (whole) face. We will explore the use of eye detectors that does not need the whole face. In this sense, the sampling grid and Gabor decomposition used in this work has been already used for the task of facial landmark detection [1, 2, 21], and it will be the source of future work.

References

1. Smeraldi, F., Carmona, O., Bigün, J.: Saccadic search with gabor features applied to eye detection and real-time head tracking. IVC 18 (2000)
2. Smeraldi, F., Bigün, J.: Retinal vision applied to facial features detection and face authentication. Pattern Recognition Letters 23 (2002)
3. Park, U., Jillela, R.R., Ross, A., Jain, A.K.: Periocular biometrics in the visible spectrum. IEEE TIFS 6 (2011)
4. Miller, P.E., Rawls, A.W., Pundlik, S.J., Woodard, D.L.: Personal identification using periocular skin texture. In: Proc. ACM SAC (2010)
5. Miller, P.E., Lyle, J.R., Pundlik, S.J., Woodard, D.L.: Performance evaluation of local appearance based periocular recognition. In: Proc. IEEE BTAS (2010)
6. Bharadwaj, S., Bhatt, H.S., Vatsa, M., Singh, R.: Periocular biometrics: When iris recognition fails. In: Proc. IEEE BTAS (2010)

7. Hollingsworth, K., Darnell, S.S., Miller, P.E., Woodard, D.L., Bowyer, K.W., Flynn, P.J.: Human and machine performance on periocular biometrics under near-infrared light and visible light. IEEE TPAMI 7 (2012)
8. Woodard, D.L., Pundlik, S.J., Lyle, J.R., Miller, P.E.: Periocular region appearance cues for biometric identification. In: Proc. IEEE CVPR Biometrics Workshop (2010)
9. Woodard, D.L., Pundlik, S.J., Miller, P., Jillela, R., Ross, A.: On the fusion of periocular and iris biometrics in non-ideal imagery. In: Proc. ICPR (2010)
10. Li, S., Jain, A. (eds.): Handbook of Face Recognition. Springer (2004)
11. Bowyer, K., Hollingsworth, K., Flynn, P.: Image understanding for iris biometrics: a survey. Computer Vision and Image Understanding 110 (2007)
12. Matey, J., Ackerman, D., Bergen, J., Tinker, M.: Iris recognition in less constrained environments. In: Advances in Biometrics: Sensors, Algorithms and Systems (2008)
13. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE TPAMI 24 (2002)
14. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Proc. IEEE CVPR (2005)
15. Lowe, D.: Distinctive image features from scale-invariant key points. International Journal of Computer Vision 60 (2004)
16. Daugman, J.: How iris recognition works. IEEE TCSVT 14 (2004)
17. NICE II. Noisy Iris Challenge Evaluation, Part II (2010), http://nice2.di.ubi.pt/
18. CASIA Iris Image Database, http://biometrics.idealtest.org
19. Fierrez, J., Ortega-Garcia, J., Torre-Toledano, D., Gonzalez-Rodriguez, J.: BioSec baseline corpus: A multimodal biometric database. Patt. Recogn. 40 (2007)
20. Bigün, J., du Buf, J.M.H.: N-folded symmetries by complex moments in gabor space and their application to unsupervised texture segmentation. IEEE TPAMI 16 (1994)
21. Bigun, J., Fronthaler, H., Kollreider, K.: Assuring liveness in biometric identity authentication by real-time face tracking. In: Proc. CIHSPS (2004)
22. Bigun, J.: Vision with Direction. Springer (2006)
23. Hubel, D.H.: Eye, brain, and vision. Scientific American Library. Distributed by W.H. Freeman, New York (1988)
24. Gilperez, A., Alonso-Fernandez, F., Pecharroman, S., Fierrez, J., Ortega-Garcia, J.: Off-line signature verification using contour features. In: Proc. ICFHR (2008)
25. Masek, L.: Recognition of human iris patterns for biometric identification. Master's thesis, School of Computer Science and Software Engineering, Univ. Western Australia (2003)
26. Alonso-Fernandez, F., Fierrez, J., Ramos, D., Gonzalez-Rodriguez, J.: Quality-based conditional processing in multi-biometrics: Application to sensor interoperability. IEEE TSMC-A 40(6) (2010)
27. Duda, R., Hart, P., Stork, D.: Pattern Classification, 2nd edn. (2004)
28. Poh, N., Bourlai, T., Kittler, J., Allano, L., Alonso-Fernandez, F., Ambekar, O., Baker, J., Dorizzi, B., Fatukasi, O., Fierrez, J., Ganster, H., Ortega-Garcia, J., Maurer, D., Salah, A., Scheidat, T., Vielhauer, C.: Benchmarking Quality-dependent and Cost-sensitive Score-level Multimodal Biometric Fusion Algorithms. IEEE TIFS 4(4) (2009)