A. Brömme, C. Busch, A. Dantcheva, C. Rathgeb and A. Uhl (Eds.): BIOSIG 2017,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2017 41

Improvement of Iris Recognition based on Iris-Code Bit-Error Pattern Analysis

Christian Rathgeb[1], Christoph Busch[1]

Abstract: In this paper an advanced iris-biometric comparator is presented. In the proposed scheme an analysis of bit-error patterns produced by Hamming distance-based iris-code comparisons is performed. The lengths of sequences of horizontal consecutive mis-matching bits are measured and a frequency distribution is estimated. The difference of the extracted frequency distribution to that of an average genuine one obtained from a training set is used as a second comparison score. This score is then used together with the fractional Hamming distance in order to improve the recognition accuracy of an iris recognition system. In experimental evaluations relative improvements of approximately 45% and 10% in terms of false non-match rate at a false match rate of 0.01% are achieved on the CASIAv4-Interval and the BioSecure iris databases, respectively.

Keywords: Biometrics, iris recognition, iris-code, bit-error analysis, improved comparator.

1 Introduction

Generic iris recognition systems comprise four major components: (1) image acquisition, (2) segmentation, (3) feature extraction and (4) comparison. Based on Daugman's approach [Da04], the first three processing steps are performed on a reference iris image during enrolment to create a two-dimensional binary feature vector, i.e. iris-code. At the time of authentication an iris-code is extracted from a probe iris image and compared against a database of enrolled reference iris-codes. In the comparison stage Hamming distance (*HD*) scores between pairs of iris-codes and corresponding noise masks are estimated. Hence, the binary data representation of iris-codes enables a rapid comparison (and compact storage) achieving millions of comparisons per second per CPU core [Da04]. Circular bit shifts are applied to iris-codes and *HD* scores are estimated at different shifting positions, i.e. relative tilt angles caused by uncontrolled head poses. The minimal obtained *HD*, which corresponds to an optimal alignment, represents the final score.

Besides the Daugman de-facto standard for comparing iris-codes, different alternative comparators have been suggested in past years, see Sect. 2. The majority of proposed schemes aims at replacing the aforementioned *HD*-based algorithm by a modified comparator in order to improve the recognition performance. In most schemes findings obtained from a deeper analyses of the nature of the iris-code bits are utilized by those comparators. A prominent example for such an improvement is the assignment of weights to each bit position in an iris-code according to their expected *reliability*, e.g. in [ZD08, DST11].

[1] da/sec – Biometrics and Internet Security Research Group,
Hochschule Darmstadt, Germany, {christian.rathgeb,christoph.busch}@h-da.de

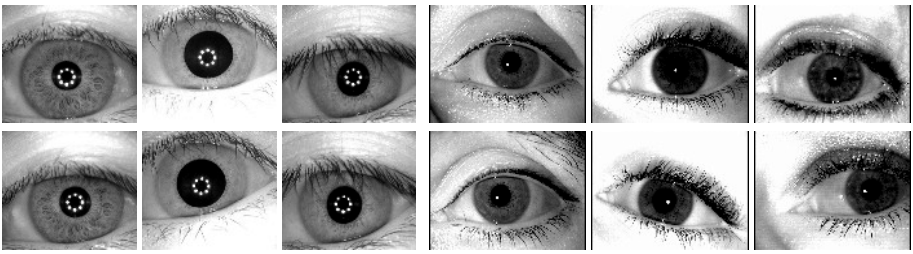
In this work we analyse entire bit-error patterns produced by *HD*-based iris-code comparisons, going beyond a local estimation of bit-errors. The presented approach measures the plausibility of an obtained bit-error pattern by comparing it to a pre-estimated model of genuine bit-error patterns. In particular, the frequency distribution of sequences of horizontal, i.e. circumferential, consecutive mis-matching bits is measured and its difference from the genuine model is used as secondary feature. This score can be interpreted as additional score, which can be estimated to achieve a more reliable decision for a distinct range of *HD* scores, e.g. $[0.35, 0.45]$. Hence, in contrast to most proposed comparators, our approach is designed to have negligible impact on comparison speed. For different iris databases it is shown that a weighted score-level fusion of the proposed score and the *HD* score improves the recognition accuracy of an iris recognition system, in particular at practical low false match rates.

The remainder of this paper is organized as follows: Sect. 2 briefly summarizes related works with respect to iris-biometric comparators. In Sect. 3 the proposed system is described in detail and evaluated. Finally, conclusions are drawn and potential future research directions are pointed out in Sect. 4.

2 Related Works

In the recent past numerous improved iris-biometric comparators have been proposed. Some of these require the processing of multiple reference samples during enrolment. In [ZD08] a weight map, which indicates the stability of iris-code bits, is obtained from several iris-codes by performing a weighted majority voting. Similar approaches based on personalized weight maps have been presented in [DST11, HSH17]. In these schemes comparison scores are estimated as a weighted sum of mis-matching bits. Note that for these modified comparators one can not expect that the comparison speed of a Hamming distance-based comparator is maintained. In [HBF09] so-called *fragile* bits, i.e. bits which exhibit a higher probability than others to flip their value during a genuine comparison, are detected by comparing several iris-codes obtained from a single eye instance. Since filters employed in the feature extraction stage set iris-code bits by the sign of obtained filter responses, these bits correspond to coefficients close to zero. That is, such bits can also be detected in a single iris-code [Da16]. It was shown that the recognition accuracy is improved, if detected fragile bits are incorporated into noise masks extracted in the iris segmentation stage. Moreover, masks encoding fragile bits can be employed as additional comparison sore to improve the performance of an iris recognition system [HBF11].

Further works utilize training sets to obtain statistics about iris-codes which are utilized by the comparator. In [RUW10] a static weight map indicating the reliability of each iris-code bit position, which is defined as the mean of discriminativity and stability, is estimated from a training set. During authentication most reliable bits are compared first to achieve a fast rejection of non-matching iris-codes in an identification scenario. A similar approach based on static masks has been presented in [Pr15]. Reported results suggest that static weight maps might vary depending on the used sensor or environmental conditions. In [RUW12] the progression of genuine comparison scores across all considered shifting



(a) CASIAv4-Interval (b) BioSecure

Fig. 1: Sample pairs of iris images of both datasets used in experimental evaluations.

Tab. 1: Overview of training and testing sets of employed datasets.

Database	Training set (left eye images)			Testing set (right eye images)		
	No. eyes	Gen. comp.	Imp. comp.	No. eyes	Gen. comp.	Imp. comp.
CASIAv4-Interval	198	4,454	19,503	197	4,343	19,306
BioSecure	210	1,260	21,945	210	1,260	21,945

positions are modelled by an inverse Gaussian of which the parameters are estimated from a training set. At authentication the deviation of comparison scores from the trained Gaussian is combined with the minimum *HD* score.

Given a single pair of iris-codes, in [RUW11] it is suggested to combine the minimum and the maximum *HD* score across shifting positions. Since genuine pairs of iris-codes can get out of phase in case of drastic mis-alignment exceptionally large *HD* scores become an indicator for a genuine comparison. More recently, a binary search technique which aims at accelerating the alignment process during iris-code comparisons was presented in [Ra16]. It is shown that, if the amount of considered shifting positions can be reduced, recognition accuracy is generally improved since *HD* scores of impostor comparisons remain higher.

3 Proposed System

3.1 Baseline System and Experimental Setup

In the employed iris recognition system, the iris of a given sample image is detected and transformed to a normalized rectangular texture of 512×64 pixels. The normalized iris texture is divided into texture stripes to obtain 10 one-dimensional signals, each one averaged from adjacent texture rows. A row-wise convolution with a Log-Gabor wavelet is performed on each signal and the real part of phase information is encoded to generate an iris-code consisting of 512×10 bits. Examples of generated iris-codes are depicted in Fig. 2. Implementations of the employed segmentation and feature extraction are available in [US17] and described in detail in [RUW13].

The fractional Hamming distance (*HD*) between a pair of iris-codes, *codeA*, *codeB*, and their according noise masks, *maskA*, *maskB* is defined as [Da04],

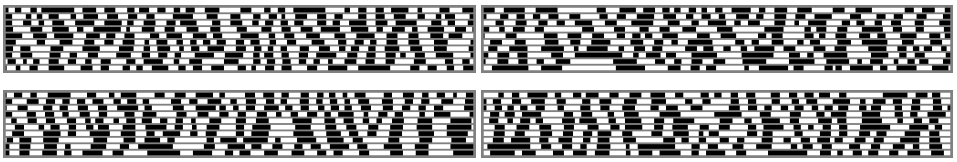
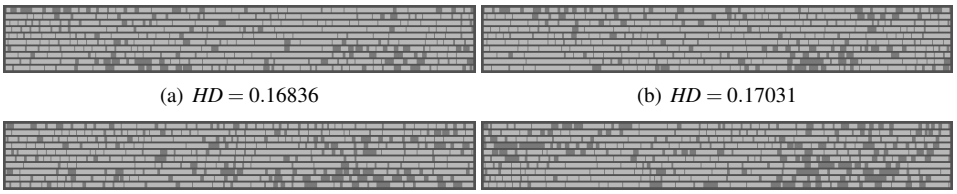


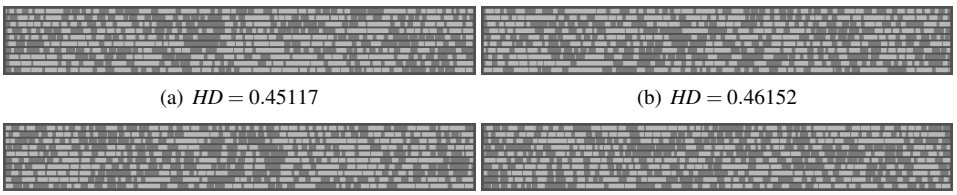
Fig. 2: Examples of iris-codes produced by four different iris images of used datasets.



(a) $HD = 0.16836$ (b) $HD = 0.17031$

(c) $HD = 0.20957$ (d) $HD = 0.24316$

Fig. 3: Examples of bit-error patterns produced by four genuine iris-code comparisons.



(a) $HD = 0.45117$ (b) $HD = 0.46152$

(c) $HD = 0.47461$ (d) $HD = 0.49004$

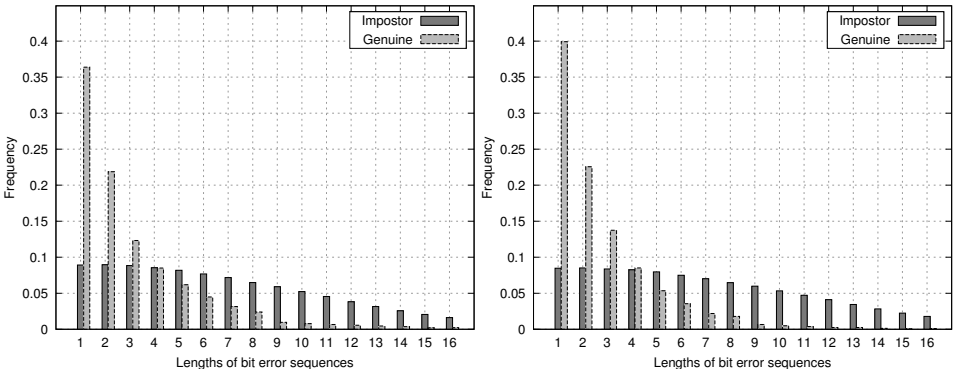
Fig. 4: Examples of bit-error patterns produced by four impostor iris-code comparisons.

$$HD = \frac{\|(codeA \oplus codeB) \cap maskA \cap maskB\|}{\|maskA \cap maskB\|}. \quad (1)$$

Experiments are conducted on the CASIAv4-Interval [CA17] and the BioSecure [Or10] iris database. Example images of both datasets are depicted in Fig. 1. An overview of the used training sets (left eye images) and testing sets (right eye images) is shown in Table 1. In experiments training and testing will be performed within and across both used databases.

3.2 Iris-Code Bit-Error Pattern Analysis

It is well known that bits in iris-codes are not mutually independent [Da04]. This is due to the internal spatial correlations within iris textures and the nature of employed filters [Da16]. Mis-matching bits between genuine iris-codes have been found to occur at boundaries of consecutive 0-bit or 1-bit sequences [HBF09, Da16]. That is, even for large *HD* scores lengths of sequences of consecutive mis-matching non-masked bits are expected to



(a) CASIAv4-Interval (b) BioSecure

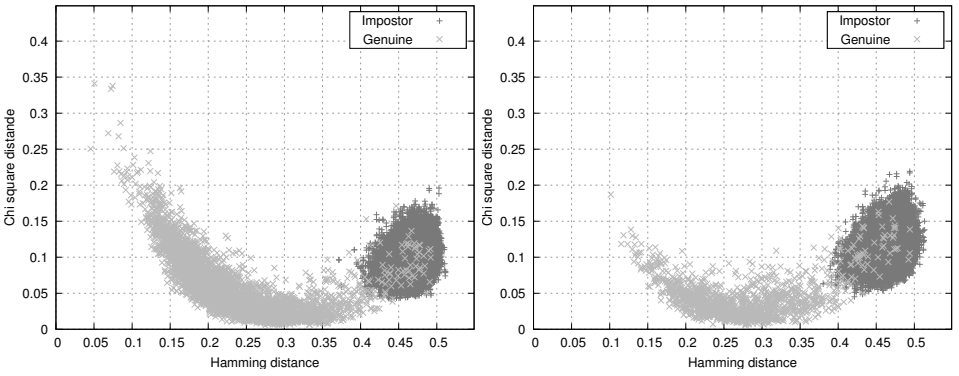
Fig. 5: Bit-error sequence lengths of bit-error patterns obtained from training sets.

be low. In contrast, for impostor comparisons these lengths tend to be higher. This is due to the facts that iris-codes of different eyes are uncorrelated and adjacent bits in iris-codes exhibit high correlation. Hence, the neighbouring bits of each non-matching bit have a high probability of being non-matching, too.

In our experiments left eye images of each database are processed in the training stage. Based on the training sets we perform all possible genuine comparisons and impostor comparisons based on the first image of each eye. Examples of bit-error patterns obtained by genuine and impostor comparisons are depicted in Fig. 3 and Fig. 4 (green pixels indicate matching bits; red pixels indicate non-matching bits). The lengths of horizontal sequences of consecutive mis-matching non-masked bits of genuine and impostor comparisons are counted and stored in separate histograms. For the training sets of the used datasets the obtained histograms are shown in Fig. 5. We observe that the frequency distributions for genuine and impostor comparisons are similar for both databases. Focusing on impostor distributions, in Fig. 5 it can be seen that, sequences of up to five consecutive mis-matching bits are almost equiprobable (also see Fig. 4). The similarity of distributions across both databases suggests that these mainly depend on the employed feature extractor (as will be shown in experimental evaluations).

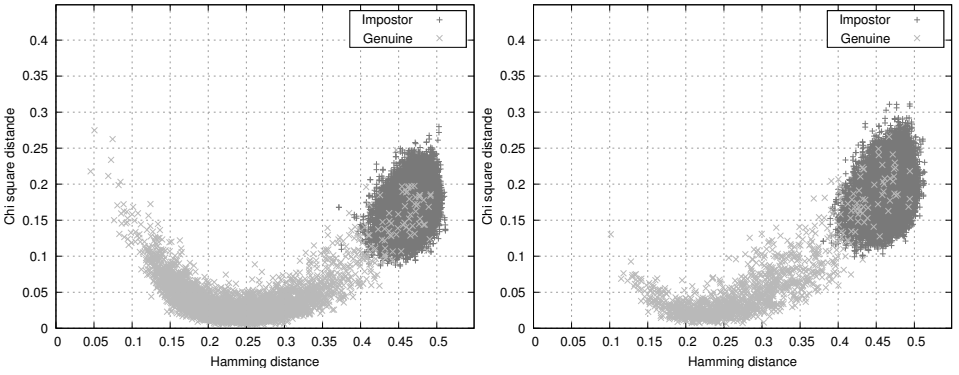
3.3 Improved Comparator

Given a pair of iris-codes, *codeA* and *codeB*, the *HD* score between them is estimated and the frequency distribution of sequences of consecutive mis-matching non-masked bits is stored in a histogram, *histAB*. This histogram is then compared against the average genuine model obtained during the training stage, *histGen*, by estimating the Chi square (χ^2) distance between both histograms,



(a) CASIAv4-Interval (b) BioSecure

Fig. 6: Scores obtained from testing sets with training performed on CASIAv4-Interval.



(a) CASIAv4-Interval (b) BioSecure

Fig. 7: Scores obtained from testing sets with training performed on BioSecure.

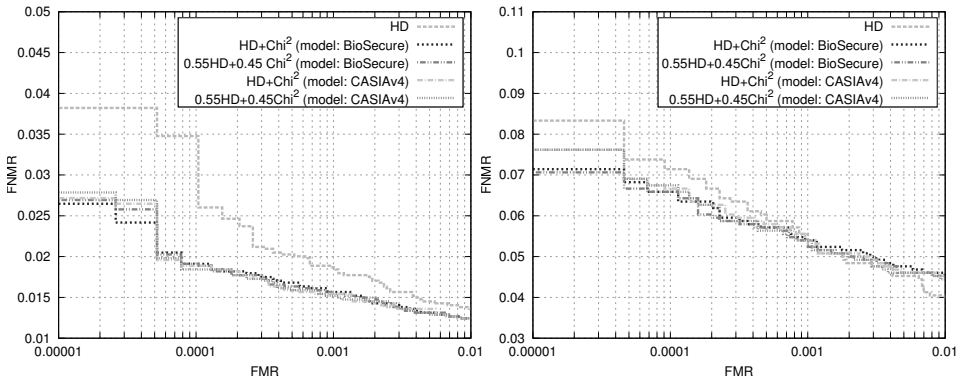
$$\chi^2(histAB, histGen) = 1/2k \sum_{i=1}^{k} (histAB_i - histGen_i)^2 / (histAB_i + histGen_i). \quad (2)$$

It has been found that the χ^2 distance is a suitable method for the proposed comparator. Alternatively, other similar methods could be employed to compare pairs of histograms, e.g. [PW10]. Note that only bit-error patterns obtained from genuine comparisons are used. No significant improvements were obtained for applying the proposed procedure to bit-error patterns produced by impostor comparisons.

Fig. 6 and Fig. 7 show scatter plots of *HD* scores and corresponding χ^2 distances for using different training sets. It can be observed that some large genuine *HD* scores still yield small χ^2 distances. Also, rather low genuine *HD* scores result in large χ^2 distance due to the small amount of bit-errors. However, as mentioned earlier, it is suggested to

Tab. 2: Performance rates (in %) obtained from the testing sets.

Comparator	Training	CASIAv4-Interval			BioSecure		
		$FNMR_{0.01}$	$FNMR_{0.001}$	$FNMR_0$	$FNMR_{0.01}$	$FNMR_{0.001}$	$FNMR_0$
HD	–	3.48	3.83	3.85	7.38	8.26	8.34
$HD+\chi^2$	CASIAv4-Interval	1.98	2.69	2.79	6.89	7.54	7.62
$0.55HD+0.45\chi^2$		1.96	2.65	2.72	6.75	7.39	7.62
$HD+\chi^2$	BioSecure	1.94	2.69	2.70	6.59	7.16	7.17
$0.55HD+0.45\chi^2$		1.92	2.63	2.65	6.56	6.99	7.14



(a) CASIAv4-Interval (b) BioSecure

Fig. 8: Detection error trade-off curves obtained from the testing sets.

estimate the χ^2 distance only for a distinct range of obtained *HD* scores, e.g. $[0.35, 0.45]$. As can be seen in Fig. 6 and Fig. 7, in such a range a diagonal line would achieve the best separation of genuine and impostor scores. That is, for a pre-defined interval the χ^2 distance is estimated as an assisting score and combined with the *HD* scores employing a weighted score-level fusion using the sum-rule. Further, we observe χ^2 distances of impostors are generally larger if the model obtained from the BioSecure training set is employed. This is because in the histogram of the BioSecure training set sequences of small lengths are weighted higher compared to the histogram of the CASIAv4-Interval database (see Fig. 5). Also, it can be seen that χ^2 distances of genuine as well as impostors are slightly larger on the BioSecure testing set. This might suggest that this database is more noisy than the CASIAv4-Interval database, which is also reflected by the obtained performance rates.

In accordance to the ISO/IEC IS 19795-1 [Int11] biometric performance is estimated in terms of false non-match rate (*FNMR*) at a targeted false match rate (*FMR*), denoted by $FNMR_{FMR}$. Obtained *FNMR*s at *FMR*s of 0.01%, 0.001% and 0% are listed in Table 2. The resulting detection error trade-off (DET) curves are shown in Fig. 8. Across considered *FMR*s the recognition accuracy is generally enhanced by the fusion of *HD* scores and χ^2 distances, which is performed within the *HD* score interval of $[0.35, 0.45]$. Due to the

fact that the histograms of bit-error sequences are similar for both databases, no significant performance drops are observed if the training is be performed on a different dataset. When using a weighted fusion only small improvements can be achieved. As an alternative to the simple (weighted) sum-rule fusion support vector machines (SVMs) could be trained to separate genuine from impostor scores.

4 Conclusions and Future Work

In this work we presented an advanced iris-biometric comparator to improve the biometric performance in an iris recognition system. In contrast to many published works, we propose an analysis of bit-error patterns produced by iris-code comparisons. In particular, we construct a model for the expected frequency distribution resulting from a genuine comparison based on a training set of iris-codes. The difference of an obtained bit-error pattern to that of the pre-trained one can be used as a second comparison score in combination with the fractional Hamming distance. At practical false match rates the recognition accuracy has be significantly improved on different databases. Reported preliminary improvements motivate further investigations of bit-error patterns of iris-code comparisons. Models of bit-errors could be, (1) constructed for different intervals of *HD* scores to improve the robustness of the proposed comparator, (2) extended to also analyse vertical, i.e. radial, correlations of bit-errors, (3) constructed for different regions of iris textures, since entropy has been found to vary significantly across iris texture regions.

Building a model for genuine bit-error patterns might be of interest for other research fields. In particular, models of bit-error patterns produced by iris-code pairs could be employed in presentation attack detection techniques [GGB16]. Moreover, machine learning techniques, e.g. convolutional neuronal networks, could be used to reliably identify error patterns produced by genuine iris-code comparisons.

Acknowledgements

This work was partially supported by the German Federal Ministry of Education and Research (BMBF) as well as by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within the Center for Research in Security and Privacy (CRISP).

References

[CA17] CASIA: , Chinese Academy of Sciences' Institute of Automation – Iris Image Database V4.0 – Interval. `http://biometrics.idealtest.org`, 2017.

[Da04] Daugman, J.: How iris recognition works. Trans. on Circuits and Systems for Video Technology, 14(1):21–30, 2004.

[Da16] Daugman, J.: Information Theory and the IrisCode. Trans. on Information Forensics and Security, 11(2):400–409, Feb 2016.

[DST11] Dong, W.; Sun, Z.; Tan, T.: Iris Matching Based on Personalized Weight Map. IEEE Trans. on Pattern Analysis and Machine Intelligence, 33(9):1744–1757, 2011.

[GGB16] Galbally, J.; Gomez-Barrero, M.: A review of iris anti-spoofing. In: Proc. Int'l Workshop on Biometrics and Forensics (IWBF'16). pp. 1–6, 2016.

[HBF09] Hollingsworth, K. P.; Bowyer, K. W.; Flynn, P. J.: The Best Bits in an Iris Code. IEEE Trans. on Pattern Analysis and Machine Intelligence, 31(6):964–973, 2009.

[HBF11] Hollingsworth, K. P.; Bowyer, K. W.; Flynn, P. J.: Improved Iris Recognition through Fusion of Hamming Distance and Fragile Bit Distance. IEEE Trans. on Pattern Analysis and Machine Intelligence, 33(12):2465–2476, Dec 2011.

[HSH17] Hu, Y.; Sirlantzis, K.; Howells, G.: A novel iris weight map method for less constrained iris recognition based on bit stability and discriminability. Image and Vision Computing, 58:168 – 180, 2017.

[Int11] International Organization for Standardization. ISO/IEC 19795-1:2006. Information Technology - Biometric performance testing and reporting – Part 1: Principles and framework, 2011.

[Or10] Ortega-Garcia, J.; Fierrez, J.; Alonso-Fernandez, F.; Galbally, J.; Freire, M. R. et al.: The Multiscenario Multienvironment BioSecure Multimodal Database (BMDB). IEEE Trans. on Pattern Analysis and Machine Intelligence, 32(6):1097–1111, 2010.

[Pr15] Proença, H.: Iris Recognition: What Is Beyond Bit Fragility? IEEE Trans. on Information Forensics and Security, 10(2):321–332, 2015.

[PW10] Pele, O.; Werman, M.: The Quadratic-Chi Histogram Distance Family. In: 11th European Conf. on Computer Vision (ECCV'10). pp. 749–762, 2010.

[Ra16] Rathgeb, C.; Hofbauer, H.; Uhl, A.; Busch, C.: TripleA: Accelerated Accuracy-preserving Alignment for Iris-Codes. In: Proc. of the 9th IAPR/IEEE Int'l Conf. on Biometrics (ICB'16). pp. 1–8, 2016.

[RUW10] Rathgeb, C.; Uhl, A.; Wild, P.: Incremental Iris Recognition: A Single-algorithm Serial Fusion Strategy to Optimize Time Complexity. In: Proc. of the 4th IEEE Int'l Conf. on Biometrics: Theory, Application, and Systems 2010 (IEEE BTAS'10). pp. 1–6, 2010.

[RUW11] Rathgeb, C.; Uhl, A.; Wild, P.: Shifting Score Fusion: On Exploiting Shifting Variation in Iris Recognition. In: Proc. of the 26th ACM Symposium on Applied Computing (SAC'11). pp. 1–5, 2011.

[RUW12] Rathgeb, C.; Uhl, A.; Wild, P.: Iris-Biometric Comparators: Exploiting Comparison Scores towards an Optimal Alignment under Gaussian Assumption. In: Proc. of the 5th IAPR/IEEE Int'l Conf. on Biometrics (ICB'12). pp. 1–6, 2012.

[RUW13] Rathgeb, C.; Uhl, A.; Wild, P.: Iris Recognition: From Segmentation to Template Security, volume 59 of Advances in Information Security. Springer Verlag, 2013.

[US17] USIT: , University of Salzburg Iris Toolkit. `http://www.wavelab.at/sources/Rathgeb16a` Version 2.0.x, 2017.

[ZD08] Ziauddin, S.; Dailey, M. N.: Iris recognition performance enhancement using weighted majority voting. In: 15th Int'l Conference on Image Processing (ICIP'08). pp. 277–280, 2008.