



SIFT-based iris recognition revisited: prerequisites, advantages and improvements

C. Rathgeb¹ · J. Wagner¹ · C. Busch¹

Received: 16 August 2016 / Accepted: 8 May 2018 / Published online: 16 May 2018
© Springer-Verlag London Ltd., part of Springer Nature 2018

Abstract

Scale-invariant feature transform (SIFT), which represents a general purpose image descriptor, has been extensively used in the field of biometric recognition. Focusing on iris biometrics, numerous SIFT-based schemes have been presented in past years, offering an alternative approach to traditional iris recognition, which are designed to extract discriminative binary feature vectors based on an analysis of pre-processed iris textures. However, the majority of proposed SIFT-based systems fails to maintain the recognition accuracy provided by generic schemes. Moreover, traditional systems outperform SIFT-based approaches with respect to other key system factors, i.e. authentication speed and storage requirement. In this work, we propose a SIFT-based iris recognition system, which circumvents the drawbacks of previous proposals. Prerequisites, derived from an analysis of the nature of iris biometric data, are utilized to construct an improved SIFT-based baseline iris recognition scheme, which operates on normalized enhanced iris textures obtained from near-infrared iris images. Subsequently, different binarization techniques are introduced and combined to obtain binary SIFT-based feature vectors from detected keypoints and their descriptors. On the CASIAv1, CASIAv4-Interval and BioSecure iris database, the proposed scheme maintains the performance of different traditional systems in terms of recognition accuracy as well as authentication speed. In addition, we show that SIFT-based features complement those extracted by traditional schemes, such that a multi-algorithm fusion at score level yields a significant gain in recognition accuracy.

Keywords Scale-invariant feature transform · Biometrics · Iris recognition

1 Introduction

Biometrics refers to automated recognition of individuals based on their biological and behavioural characteristics [23, 25]. Due to its resistance to false matches, the iris represents one of the most powerful biometric characteristics [7]. In order to confirm an individual's identity accurately and reliably, iris recognition systems analyze the complex random texture that is visible in the iris of the eye. Iris recognition technologies [7, 16] are deployed in numerous large-scale nationwide projects and are currently entering the mobile market [43]. Following Daugman's approach [16], which is the core of most operational applications, four key modules constitute an iris recognition system: (1) *image acquisition*, where most current deployments require subjects to fully

cooperate with the system to capture images of sufficient quality; (2) *pre-processing*, which includes the detection of the pupil and the outer iris boundary, a normalization of the iris to a rectangular texture and enhancement of the textured image. Parts of the iris texture which are occluded by eyelids, eyelashes or reflections are stored in an according noise mask; (3) *feature extraction*, in which a binary feature vector, i.e. iris-code, is generated by applying appropriate filters to the pre-processed iris texture; (4) *comparison*, which is based on the estimation of Hamming distance (*HD*) scores between pairs of iris-codes and corresponding masks, can be performed rapidly achieving millions of comparisons per second per CPU core [16]. In the comparison stage, circular bit shifts are applied to iris-codes and *HD* scores are estimated at different shifting positions, i.e. relative tilt angles. The minimal obtained *HD*, which corresponds to an optimal alignment, represents the final score.

Focusing on the feature extraction stage, Daugman [15] proposed the use of two-dimensional Gabor filters, which are applied to a pre-processed iris image in a dimensionless

✉ C. Rathgeb
christian.rathgeb@h-da.de

¹ da/sec - Biometrics and Internet Security Research Group
Hochschule Darmstadt, Darmstadt, Germany

polar coordinate system at different scales. Each complex Gabor response is encoded into two bits by the signs of its real and imaginary value. Numerous alternative methods have been suggested for the purpose of iris feature extraction, e.g. one-dimensional Gabor filters [36], packets of Gabor wavelets [27], differences of discrete cosine transform (DCT) coefficients [39], custom-built filters [33, 49], including circular symmetric filters and tripole filters, or characterization of key local variations [34]. It has been shown that the vast majority of these methods resemble that of [16], in the sense that they quantize responses of other linear transforms or filters (replacing 2D Gabor filters), which are applied to iris textures in a similar manner [26].

Despite the above de facto standard approaches, some entirely different schemes have been proposed, which aim at extracting different types of iris features, e.g. salient key points [5, 57] or human-interpretable features [10, 50]. More recently, the use of deep neuronal networks has been introduced to iris recognition, too. With respect to the feature extraction stage, features obtained from convolutional neuronal networks have also revealed promising results, e.g. [30, 42]. In particular, several researchers have investigated the usefulness of SIFT [32] for iris recognition based on near-infrared (NIR) [1, 11, 37, 38, 56, 58] and visible wavelength (VW) iris images [5, 31, 37]. SIFT represents a generic method for image-based extraction and comparison of keypoints and corresponding descriptors, which is invariant to translations, rotations and scaling transformations in the image domain. However, compared to traditional approaches, the vast majority of proposed SIFT-based iris recognition schemes suffer from three major drawbacks, see Sect. 2: (1) *degraded biometric performance*: significant drops in recognition accuracy are reported, which might lead to unacceptable biometric performance, even in case of good-quality iris images; (2) *poor comparison speed*: the original comparison algorithm of SIFT requires a brute force comparison of extracted keypoint descriptors, which causes considerably longer response times compared to traditional HD-based iris comparators. (3) *extensive storage requirement*: SIFT-based reference data consist of a variable-sized set of keypoint descriptors, which requires significantly more storage compared to the compact binary iris-codes of traditional schemes.

In this work, we tackle all of the above-mentioned drawbacks of SIFT-based iris recognition systems. We show that, in case iris biometric input data are prepared properly and an adequate post-processing is applied, a SIFT-based iris recognition scheme is capable of maintaining the biometric performance obtained by different traditional systems. In order to achieve a rapid comparison, we extract SIFT-based binary codes from iris textures, which allow for a simultaneous matching of keypoints and an efficient retrieval of corresponding descriptors via look-up tables. Further, we

introduce a binarization of keypoint descriptors, which enables a fast comparison based on bit operations and, at the same time, significantly reduces storage requirements. The proposed system yields a substantial speed-up resulting in low transaction times comparable to those of traditional schemes. Moreover, the proposed system turns out to be a suitable candidate for biometric fusion, obtaining significant performance gains in a challenging multi-algorithm fusion with traditional schemes.

This article is organized as follows: Section 2 briefly summarizes the functionality of SIFT as well as existing applications of SIFT to iris recognition. Section 3 provides a detailed description of the constituting modules of the SIFT-based baseline iris recognition system. The generation of SIFT-based binary codes, the binarization of keypoint descriptors and the corresponding comparison techniques are introduced in Sect. 4. Experimental results are presented in Sect. 5. Finally, conclusions are drawn in Sect. 6.

2 Related work

The following subsections briefly review the fundamentals of SIFT (and its variants) and published approaches to SIFT-based iris recognition.

2.1 Scale-invariant feature transform

Given an image I , points of interest, i.e. keypoints, are extracted by detecting scale space extrema using a difference of Gaussian (*DoG*) function. The main idea behind scale space extrema detection is to identify stable features which are invariant to changes in scale and viewpoint. In the first step, a Gaussian filter $G(x, y, \sigma)$ of size σ is convolved with $I(x, y)$, $L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$. Subsequent scales are subtracted by employing a constant multiplicative factor k in order to obtain *DoG* images, $D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma)$. The set of Gaussian-smoothed images and corresponding *DoG* images form an octave, where different octaves are obtained by successively down sampling the original image by a factor of 2. In [32], it is suggested to use 3 octaves and $s = 3$ scales per octave with $\sigma = 1.6$ and $k = 2^{1/s}$. In order to cover a complete octave in the scale space extrema detection, $s + 3 = 6$ images have to be generated per octave. Keypoints are localized as minima or maxima by comparing each pixel in a *DoG* image against its eight neighbours in the same scale and corresponding nine neighbours in the scale above and below. Subsequently, the stability of detected feature points is verified by rejecting noise-sensitive points that exhibit low contrast and points which are along an edge. Therefore, additional parameters are introduced in [32], i.e. a contrast threshold of 0.04 and an edge threshold of 10. Figure 1 depicts SIFT keypoints

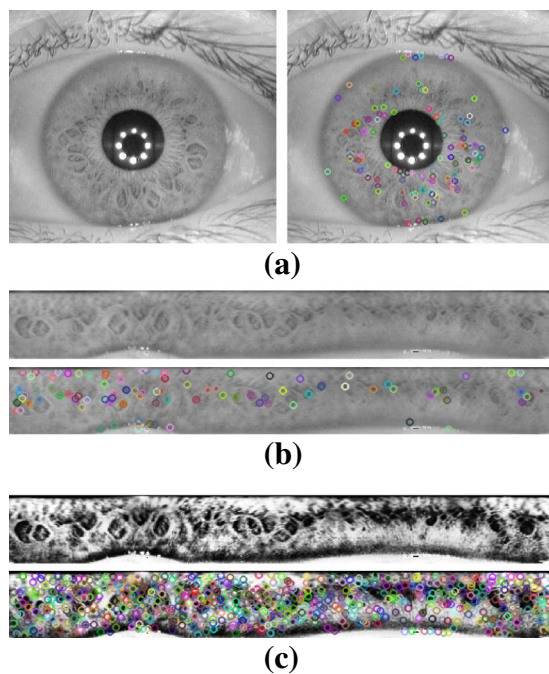


Fig. 1 Keypoint detection: SIFT keypoints extracted from **a** the iris, **b** the normalized iris texture and **c** the enhanced normalized iris texture for a sample image of the CASIAv4 iris database using the default parameters suggested in [32]. Keypoints are depicted using random colours for a better differentiation of nearby keypoints

extracted from a NIR iris image of size 320×280 pixels where keypoint detection is applied at different stages of the iris biometric processing chain.

In the next step, the dominant orientation of gradients within a 16×16 pixel window around each keypoint is estimated. Subsequently, a keypoint descriptor is extracted, relative to the rotation of the corresponding keypoint. Again, gradient magnitudes and orientations are sampled within a 16×16 pixel window and accumulated into orientation histograms summarizing the contents over 4×4 sub-regions where each orientation histogram consists of 8 bins. Hence, a total number of $D = 4 \times 4 \times 8 = 128$ bins form the descriptor of a keypoint.

Given two images, extracted sets of keypoints can be matched by estimating the Euclidean distance between corresponding keypoint descriptors. This means, a reference

descriptor is compared against all probe descriptors to determine the closest neighbour of distance d_1 . In order to decide whether the closest neighbour is a match, a ratio test is performed using the distance of the second closest neighbour d_2 . In case, $d_1/d_2 < t$ corresponding keypoints are considered to match, where $t \approx 0.8$ is suggested in [32]. An alternative to this ratio test is to cross-check matching pairs of keypoints. That is, only those matching pairs are retained where both keypoints are closest neighbours to each other. While the Euclidean distance between two points remains stable, the nearest neighbour of a specific point may change when changing the perspective of the comparison, which is illustrated in Fig. 2. By cross-checking whether pairs of descriptors are nearest neighbours, the number of outliers (false positives) can be minimized without employing a ratio test, which might require an application-specific threshold. On the contrary, cross-checking of obtained matching pairs of keypoints requires significantly more computational effort.

In order to provide a more efficient detection as well as matching of keypoints and corresponding descriptors, different alternative techniques have been proposed, most notably Speeded-Up Robust Features (SURF) [4]. For SURF, keypoint matching is accelerated by employing keypoint descriptors of reduced size, e.g. 64 or 32 bins. Moreover, since binary descriptors allow for a fast *HD*-based keypoint comparison, different methods of how to binarize keypoint descriptors [2, 28] or directly extract binary keypoint descriptors have been suggested, such as Binary Robust Independent Elementary Features (BRIEF) [8].

2.2 SIFT-based iris recognition

Numerous researchers have applied SIFT for the purpose of iris recognition. Table 1 summarizes most relevant works with respect to the input data from which SIFT keypoints and corresponding descriptors are extracted, major findings, employed databases and obtained biometric performance. It can be observed that, in comparison to conventional iris recognition systems, the vast majority of listed approaches do not reveal competitive biometric performance. Moreover, the complex procedure required to match sets of SIFT keypoints prevents from a rapid comparison of SIFT-based templates, which require significantly more storage than

Fig. 2 Example of cross-checking matching point pairs: Euclidean distances between two point sets, Green and Red, are estimated where **a** Green is matched against Red, **b** Red is matched against Green and **c** cross-checking is performed after matching Green against Red and vice versa

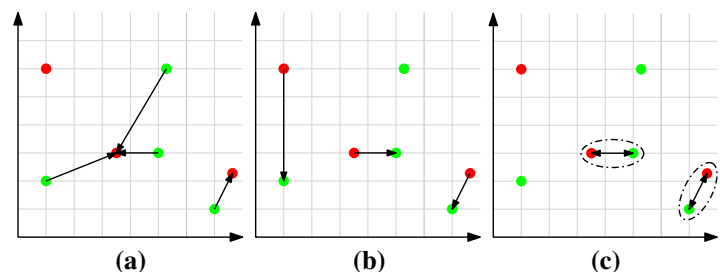


Table 1 Overview of related work: most relevant approaches of SIFT-based iris recognition, findings, employed databases, reported results and remarks

Method	Input data	Summary/findings	Database(s)	Results	Remarks
Zhu et al. [58]	Original/normalized NIR iris images	First application of SIFT to iris recognition; system based on original images outperforms the one based on normalized images	CASIAv1 [12]	FNMR = 8%, FMR = 0%	No treatment of false positive keypoint matches
Belcher and Du [5]	NIR iris images	Region-based SIFT features; for cooperative images, obtained performance is comparable to that of a Log-Gabor-based system; non-cooperative images lead to impractical results; evaluation for different quality classes	ICE [41], WVU [55]	EER = 5.57%, EER = 8.28%	No treatment of false positive keypoint matches
Alonso-Fernandez et al. [1]	NIR iris images	System based on traditional Log-Gabor features outperforms SIFT-based system; a fusion of both feature types is beneficial	BioSec [18]	EER = 11.52%	Database is considered challenging
Mehrotra et al. [37]	NIR/VW iris images	Indexing scheme using geometric hashing of SIFT keypoints	UBIRISv1 [46], BATH [53], CASIAv3 [13], IITK [19]	EER = 2.52%, EER = 5.18%, EER = 2.94%, EER = 3.97%	Results included for baseline system (without geometric hashing)
Mehrotra et al. [38]	NIR iris images	Combination of SIFT keypoints with Fourier-based feature descriptors (F-SIFT); phase-only correlation used for comparison	BATH [53], CASIAv3 [13]	FNMR = 5%, FMR = 0.72%, FNMR = 0%, FMR = 4.09%	–
Yang et al. [56]	Original/normalized/normalized enhanced NIR iris images	Application of SIFT to different types of input data: original images, normalized and normalized enhanced iris textures; best performance obtained for the latter	CASIAv1 [12]	EER = 3.01%	No treatment of false positive keypoint matches
Liu and Li [31]	Normalized VW iris images	Extraction of SIFT features from sub-images and keypoint matching for pair-wise sub-images; at comparison Bhattacharyya distance outperforms Euclidean distance	UBIRISv1 [46], UBIRISv2 [47]	EER = 7.66%, EER = 24.66%	–

Table 1 (continued)

Method	Input data	Summary/findings	Database(s)	Results	Remarks
Chen et al. [11]	NIR iris images	Discriminative SIFT feature selection based on orientation/magnitude probability distribution function; assignment of weights for iris texture sub-regions improves performance	CASIAv3 [13], MMU [40]	EER = 0.88%, EER = 0.9%	-

In case different variations of approaches are presented, best reported results were included

binary iris-codes. Early works on SIFT-based iris recognition [1, 5] were motivated by unconstrained acquisition scenarios where accurate segmentation of the iris region turns out to be challenging. By extracting SIFT keypoints directly from the original (coarsely segmented) iris image potential self-propagating errors in the normalization stage can be avoided. Reported performance rates of early approaches were not competitive compared to traditional iris recognition systems, while more recent proposals reveal acceptable biometric performance [11]. However, it has been shown that SIFT descriptors carry complement information which can be utilized to improve the performance of a traditional iris recognition system in a score-level fusion scenario [1]. While it has been shown that keypoint matching fails for challenging off-angle iris images [5], biometric performance tends to improve in case keypoints are extracted from (correctly) normalized and/or enhanced iris textures [31, 56]. Compared to original iris images, normalized textures of constant size yield a more stable amount of detected keypoints and contrast enhancement techniques increase the amount of detected keypoints, since less of them exhibit low contrast.

Another important issue is the detection of false positive keypoint matches. Since relative positions of detected keypoints are expected to change only slightly, geometrical constraints can be applied in order to detect false positive keypoint matches, which is referred to as *trimming of false matches* [45]. In [5, 11], it is suggested to divide the detected iris into regions and perform keypoint detection and matching per region. Similarly, a region-based matching of keypoints can be performed on normalized iris textures [31]. In [1], it is suggested to retain only matches within a distinct range of rotation and distance, which means that scale invariance and robustness to pupil dilation can no longer be guaranteed. Moreover, biometric performance might also be improved by employing alternative keypoint descriptors [38] or matching strategies [31]. SIFT can also be employed to detect and match macro-features, i.e. structures within the iris texture, such as furrows or crypts [50]. It is important to note that the estimation of a final comparison score between two given iris images has received only little attention. Within most approaches, it is suggested to count the number of matching keypoints [1, 5]. However, if compared sets of keypoints and corresponding descriptors are subject to strong variations in size, score normalization techniques play an important role.

In summary, it can be concluded that existing approaches to SIFT-based iris recognition fail to obtain competitive biometric performance rates. We also observe that diverse processing steps of SIFT-based iris recognition schemes, e.g. keypoint matching or score estimation, might allow for potential improvements to obtain a more reliable recognition system. Such a system would be of particular interest, since

proposed studies have demonstrated the potential of SIFT-based features to complement conventional iris-biometric feature vectors, i.e. iris-codes. In addition, comparison speed turns out to be vital to enable an operation of SIFT-based iris recognition in identification mode. Note, that all presented approaches require complex keypoint matching procedures, which do not allow for a rapid comparison. Finally, it is worth noting that so far there are no strategies for reducing the storage required within SIFT-based iris recognition schemes.

3 SIFT-based iris recognition system

In the subsequent subsections, we will summarize the pre-processing and feature extraction process the baseline and the proposed system will build upon. Further, the baseline SIFT-based comparator and its used score estimation are described.

3.1 Pre-processing and feature extraction

In the first step, the pupil and outer iris boundaries are detected and the iris is transformed to a normalized texture of $W \times H = 512 \times 64$ pixels according to the rubbersheet model [16], see Fig. 1a, b. The dimension of the normalized iris texture is derived from IS ISO/IEC 29794-6:2015 [20], where the iris radius, i.e. the radius of a circle approximating the iris–sclera boundary, is required to be at least 80 pixels. Subsequently, contrast-limited adaptive histogram equalization (CLAHE) [59] is applied to the texture resulting in an enhanced texture, as shown in Fig. 1c. CLAHE is performed using a block size of 40×64 pixels, 256 histogram bins and a clipping limit of 1. Similar image enhancement techniques are employed in conventional iris recognition systems relying on global feature extractors [7]. It can be observed that enhancing the contrast of the extracted iris texture significantly increases the amount of detected keypoints. By stretching the histogram of pixel intensities, detected keypoints exhibit higher contrast, i.e. numerous keypoints, which might have been discarded in the original image or the normalized texture due to low contrast, are now retained. Since CLAHE is performed locally, a rather uniform distribution of detected keypoints is obtained, which can not be achieved by operating a global contrast threshold in case images or image regions vary in contrast.

Let $\mathbf{p} = (x, y, \theta)$ be a detected keypoint at $(x, y) \in \mathbb{R}^2$ and $\theta \in [0, 360)$ its orientation, i.e. angle. The corresponding keypoint descriptor is denoted by $\mathbf{d} = (d_1, d_2, \dots, d_D)$, with $D = 128$ in case of SIFT. The biometric reference data \mathcal{R} for a single enhanced iris texture consist of a set of N detected keypoints and a corresponding set of their

keypoint descriptors, $\mathcal{R} = \{(\mathbf{p}_n, \mathbf{d}_n)\}, n = 1, \dots, N$. In contrast to a reliable feature selection procedure, which aims at reducing the extracted feature set to a discriminative subset thereof, the above-described pre-processing step increases N . Moreover, proposed feature selection methods mostly rely on global contrast thresholds [11] without analyzing whether feature descriptors of high contrast keypoints do carry the most discriminative information in case of iris biometrics. Note that, for the baseline system storage requirement, which increases with N , is considered a less relevant factor.

3.2 Keypoint comparison and score estimation

Given a reference set \mathcal{R} and a probe set \mathcal{P} , keypoint descriptors are compared using the L_2 -norm. In the employed comparison step, each keypoint descriptor of \mathcal{R} is compared against all keypoint descriptors of \mathcal{P} and vice versa. After performing the previously described cross-checking step, a set of matching keypoints $\mathcal{M} = \{(\mathbf{p}_{rk}, \mathbf{p}_{pk})\}, k = 1, \dots, K$ is obtained, $\mathbf{p}_{rk} \in \mathcal{R}, \mathbf{p}_{pk} \in \mathcal{P}$. Since an increased amount of detected keypoints raises the probability of false positive keypoint matches, a detection and exclusion of such matches are performed based on geometric constraints. Given \mathcal{M} , only those matches are retained for which the corresponding keypoint coordinates lie within a distance defined by two thresholds, ϵ_x and ϵ_y , yielding a new set of matches \mathcal{M}' ,

$$\mathcal{M}' = \{(\mathbf{p}_{rl}, \mathbf{p}_{pl}) : \|x_{rl} - x_{pl}\| < \epsilon_x, \|y_{rl} - y_{pl}\| < \epsilon_y\}, \quad (1)$$

with $l = 1, \dots, L$ and $L \ll K$. These geometrical constraints are adjusted to detect false positive keypoint matches and to tolerate natural variance in keypoint locations. In particular, ϵ_x compensates for horizontal shifts of iris textures resulting from head tilts, while ϵ_y defines a vertical tolerance which might be caused by (non-accurate) normalization of iris images exhibiting large variation pupil dilation [51]. Figure 3 illustrates a SIFT-based comparison of a genuine pair of enhanced iris textures without and with employing geometrical constraints. Note that in contrast to [1], scale invariance is not affected since textures are normalized prior to keypoint detection. The proposed comparison is somewhat similar to a block or region-based comparison of keypoint descriptors [5, 31], in the sense that keypoints of matching descriptors have to be located within a $\epsilon_x \times \epsilon_y$ region. However, in the presented approach, these regions are, in transferred sense, dynamically stretched around each keypoint such that no matches are missed due to keypoints located at region boundaries. Moreover, matching pairs of descriptors are obtained from the entire sets of reference and probe descriptors so that the probability of false matches to occur within the defined region is further reduced.

Finally, a comparison score s between \mathcal{R} and \mathcal{P} is estimated. In diverse approaches [1, 5, 58], it is suggested to

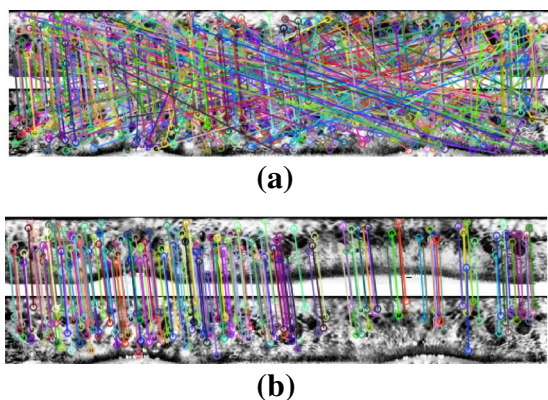


Fig. 3 Trimming of false matches: comparison of obtained keypoint matches from a genuine comparison **a** without and **b** with geometrical constraints. Keypoint matches are depicted using random colours for a better differentiation of nearby matches

return the number of obtained matching pairs of descriptors as final score, i.e. $s = \|\mathcal{M}'\|$. However, the amount of detected keypoints does significantly differ in case parts of the iris are occluded by eyelids, eyelashes or reflections. Moreover, potential attackers might perform presentation attacks [21] using textures which exhibit a large amount of keypoints in order to artificially increase the chance of a false accept. Hence, similar to the estimation of the fractional Hamming distance in relation to iris-code bits, comparison scores need to be normalized. Given \mathcal{R}, \mathcal{P} and obtained matches of keypoints \mathcal{M}' , a dissimilarity score s is estimated as,

$$s = 1 - \frac{\|\mathcal{M}'\|}{\min(\|\mathcal{R}\|, \|\mathcal{P}\|)}. \tag{2}$$

We suggest to define a lower limit for $\|\mathcal{R}\|$ as a quality check. Such a limit prevents from presentation attacks in which only a few descriptors are presented to the system and compared against the entire reference set, which might result (by chance) in a match and, hence, in a low dissimilarity score.

The presented baseline scheme reveals two major disadvantages: firstly, variable-sized sequences of real-valued keypoint triples and corresponding integer-valued descriptors cause an increased storage requirement, in contrast to conventional compact binary iris-codes; secondly, comparisons of keypoint descriptors are performed by estimating the L_2 -norm between corresponding histograms, which requires $\mathcal{O}(N^2)$ steps using the described comparator. Consequentially, the required computational effort is significantly compared to an efficient Hamming distance-based comparator, which might prevent the system to be operated in identification mode (on large databases). Note that the stated drawbacks apply to all proposals of SIFT-based

iris recognition. Hence, an acceleration of the presented baseline SIFT-based iris recognition system is of interest.

4 Proposed system

The proposed scheme builds upon the pre-processing and feature extraction of the previously described baseline scheme. Key components of our system, which are described in detail in the following subsections, include the generation of a keypoint-code and a look-up-table, a specific alignment procedure, a binarization of keypoint descriptors as well as a specific comparator.

4.1 Keypoint-code, look-up table and alignment

We assume that iris images are acquired from a direct frontal angle. This assumption is considered reasonable, since it has been shown that SIFT-based iris recognition becomes unreliable in case of off-angle image acquisition [5]. If iris images are taken frontally, the relative locations of keypoints detected in enhanced iris textures are expected to persist. Based on this assumption, a simultaneous pairing of keypoints can be achieved prior to a comparison of keypoint descriptors. For this purpose, we quantize a given keypoint triple $\mathbf{p} = (x, y, \theta)$ and map it to a two-dimensional binary map, in order to ease further processing steps. This two-dimensional binary map is referred to as *keypoint-code*. In the first step, function f maps a given keypoint triple to a quantized two-dimensional point, $f(\mathbf{p}, q_x, q_y, q_\theta) = \tilde{\mathbf{p}} = (\tilde{x}, \tilde{y}) \in \mathbb{N}^2$, where q_x, q_y and q_θ denote predefined quantisation factors and,

$$\begin{aligned} \tilde{x} &= \lfloor x/q_x \rfloor, \\ \tilde{y} &= \lfloor y/q_y \rfloor + \lfloor H/q_y \rfloor \cdot \lfloor \theta/q_\theta \rfloor. \end{aligned} \tag{3}$$

Hence, angle values only take influence on resulting y -coordinates. For instance, let $q_x = 4, q_y = 2, q_\theta = 90$, and according to the previously defined setting $H = 64$, a keypoint triple $\mathbf{p} = (15, 5, 100)$ is mapped to $(\tilde{x}, \tilde{y}) = (\lfloor 3.75 \rfloor, \lfloor 2.5 \rfloor + \lfloor 100/90 \rfloor \cdot \lfloor 64/4 \rfloor) = (3, 2 + 1 \cdot 16) = (3, 18)$. Note that, while this process reduces the dimensionality of the feature space, it does not cause any information loss and, hence, does not reduce storage requirement. However, in the following process of keypoint pairing, i.e. feature alignment, two-dimensional binary maps offer a more efficient internal representation.

In the second step, (\tilde{x}, \tilde{y}) is mapped into the binary keypoint-code \mathbf{C} of dimension $\tilde{W} \times \tilde{H} = \lfloor W/q_x \rfloor \times \lfloor H/q_y \rfloor \cdot \lfloor 360/q_\theta \rfloor$, such that 1s indicate keypoint coordinates, $\mathbf{C}[\tilde{x}][\tilde{y}] = 1$. In order to achieve a certain robustness against small changes in keypoint coordinates, neighbouring bit positions are set to 1

in case x/q_x or y/q_y are close to according boundaries. In the previous example, $\tilde{x} = \lfloor x/q_x \rfloor = \lfloor 15/4 \rfloor = 3$, while $x/q_x = 15/4 = 3.75$ is close to its right neighbour 4. Therefore, we define a threshold t , to which the distances to neighbouring positions are compared against,

$$C[\tilde{x} + i][\tilde{y}] = 1, C[\tilde{x}][\tilde{y} + j] = 1, C[\tilde{x} + i][\tilde{y} + j] = 1, \tag{4}$$

$$i = \begin{cases} 1, & \text{if } x \bmod q_x \geq q_x/2 + t \\ -1, & \text{if } x \bmod q_x < q_x/2 - t \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

$$j = \begin{cases} 1, & \text{if } y \bmod q_y \geq q_y/2 + t \\ -1, & \text{if } y \bmod q_y < q_y/2 - t \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

Obviously, $C[\tilde{x}][\tilde{y}] = C[\tilde{x} + i][\tilde{y}]$ or $C[\tilde{x}][\tilde{y}] = C[\tilde{x}][\tilde{y} + j]$ applies in case $i = 0$ or $j = 0$, respectively. For instance, let $t = 0.5$, then for $\mathbf{p} = (15, 5, 100)$ we get $x \bmod q_x = 15 \bmod 4 = 3 \geq 4/2 + 0.5$ and $y \bmod q_y = 5 \bmod 2 = 1$, resulting in $i = 1$ and $j = 0$. This means, in this example the right neighbour of $C[\tilde{x}][\tilde{y}]$ will be set to 1 as well, $C[\tilde{x} + i][\tilde{y}] = C[4][18] = 1$.

Keypoint-codes, which is merely constructed from sets of keypoint coordinates, will be used to determine correspondence between keypoints. Comparisons merely based on keypoint-codes, which do not carry information about descriptors of keypoints, are not expected to achieve reliable biometric recognition in particular, if the number of detected keypoints is specifically maximized. Hence, in addition to the keypoint-code, we extract a *look-up table*, denoted as \mathbf{L} , of same dimension, which consists of integer values defining the correspondence between 1s in the keypoint-code and keypoint descriptor indices. For the n -th keypoint $\mathbf{p}_n, n = 1, \dots, N, \mathbf{L}[\tilde{x}_n][\tilde{y}_n] = n$, where neighbouring positions might be set as well according to the above-described process. It is important to note that prior to generating the keypoint-code and the look-up table, keypoints are arranged in ascending order with respect to their response values. Thereby, keypoints with highest response values overlap those with lower response values in the keypoint-code and the look-up table, in case of identical quantized coordinates. Figure 4 illustrates two examples of keypoint-codes and visualizations of their look-up tables. From the look-up tables, we observe that cells originating from keypoints with highest response values (red cells) tend to overlap those originating from low-response keypoints (blue cells). From the look-up table in Fig. 4a, where $q_\theta = 360$ (angles are ignored), it can clearly be observed that keypoints with highest response values mostly occur in the upper regions of the enhanced iris

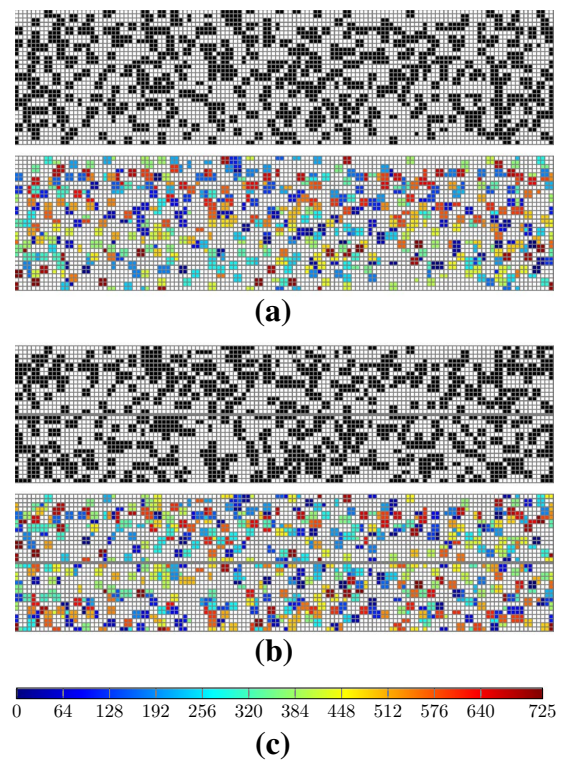


Fig. 4 Keypoint-code and look-up table: two examples of extracted binary codes and their corresponding visualized look-up tables extracted from the enhanced normalized texture of Fig. 1c. Chosen parameters of **a** and **b** yield identical dimensions for codes and look-up tables

texture, i.e. iris texture parts near the pupil which are hardly occluded by eyelids or eyelashes.

In order to efficiently retrieve correspondence between keypoints (and their descriptors) of a given reference and probe set, \mathcal{R} and \mathcal{P} , we align their keypoint-codes, \mathbf{C}_r and \mathbf{C}_p , respectively. In generic iris recognition schemes, the minimum *HD*-score obtained from various circular bit shifts of the probe iris-codes, which correspond to different relative tilt angles, represents an optimal alignment [16]. Similarly, we apply circular bit shifts to the probe keypoint-code, but seek for the maximum overlap of 1s between the reference and probe keypoint-code. At the optimal alignment, the intersection of the reference and probe keypoint-code is referred to as *alignment-code*. Let S be the maximum tolerated pixel-shift in both directions for enhanced iris textures, then the set of considered shifting positions is defined as $\mathbf{S} = \{-\lfloor S/q_x \rfloor, \dots, \lfloor S/q_x \rfloor\}$. Let $g(\mathbf{C}, s)$ denote a keypoint-code \mathbf{C} shifted by s bits, then the alignment-code \mathbf{A}_{rp} between \mathbf{C}_r and \mathbf{C}_p is estimated as,

$$\mathbf{A}_{rp} = \max_{s \in \mathbf{S}} (\mathbf{C}_r \cap g(\mathbf{C}_p, s)). \tag{7}$$

That is, the pairing of two sets of keypoints can now be performed based on efficient bit operations, comparable to HD-based comparators, where the amount of required bit comparisons depends on the dimension of keypoint-codes and the degree of rotation compensation. Note that, compared to the baseline system, the proposed alignment process implicitly trims false positive pairs of keypoints.

4.2 Binarization of descriptors and proposed comparator

In order to efficiently compare and store keypoints descriptors, we propose to binarize them, too. Given \mathcal{R} , the corresponding set of keypoint descriptors $\{\mathbf{d}_n\}$ is binarized to obtain $\{\mathbf{b}_n\}$ with $\mathbf{b}_n = (b_{1n}, \dots, b_{Dn})$ and $n = 1, \dots, N$, by defining a binarization threshold τ , such that,

$$b_{in} = \begin{cases} 1, & \text{if } d_{in} > \tau \\ 0, & \text{otherwise.} \end{cases}, \quad i = 1, \dots, D \quad (8)$$

where $D = 128$ in case of SIFT. With the use of this trivial, yet effective, quantization method, the most dominant bins of each descriptor histogram are set to 1. For the first 16 keypoint descriptors of keypoints detected in Fig. 1c, the resulting binarized descriptors are depicted in Fig. 5 for various values of τ . Obviously, the proposed binarization significantly reduces the required storage for extracted feature descriptors. In particular, the stored reference data consist of $\tilde{W} \times \tilde{H}$ bits and integers for the keypoint-code and the corresponding look-up table, respectively, and $N \times D$ bits for the set of binarized keypoint descriptors. In order to avoid arbitrary large sets of feature descriptors, the extraction of descriptors could be limited to N keypoints, which exhibit the highest response values.

Given a pair of feature descriptors, \mathbf{b}_r and \mathbf{b}_p , the distance function d is used for comparison,

$$d(\mathbf{b}_r, \mathbf{b}_p) = 1 - \frac{2\|\mathbf{b}_r \cap \mathbf{b}_p\|}{\|\mathbf{b}_r\| + \|\mathbf{b}_p\|}. \quad (9)$$

Hence, the amount of identical dominant descriptor bins, i.e. 1s, is counted and normalized by the total amount 1s occurring in both binarized descriptors. The proposed distance function operates on bit level and, thus, enables a highly efficient comparison of feature descriptors.

Let o be the shifting position which yields an optimal alignment, i.e. $\mathbf{A} = \mathbf{C}_r \cap g(\mathbf{C}_p, o)$. The final dissimilarity score between \mathcal{R} and \mathcal{P} is estimated as,

$$\frac{1}{\|\mathbf{A}_{rp}\|} \sum_{w=1}^{\tilde{W}} \sum_{h=1}^{\tilde{H}} \mathbf{A}_{rp}[w][h] \cdot d(\mathbf{b}_r[\mathbf{L}_r[w][h]], \mathbf{b}_p[\mathbf{L}_p[w+o][h]]). \quad (10)$$

In case, $\mathbf{A}_{rp}[w][h] = 1$, which indicates that at least one keypoint in \mathcal{R} and \mathcal{P} occur at identical coordinates, the corresponding binarized descriptors are obtained through the look-up table and compared. The amount of detected pairs $\|\mathbf{A}_{rp}\|$ is used to normalize the score.

5 Experiments

The subsequent subsections summarize the experimental setup and present evaluation with respect to biometric performance as well as time measurements.

5.1 Experimental setup

Experiments are carried out on CASIAv1 [12], CASIAv4-Interval [14] and the BioSecure iris database [44]. For the BioSecure database, we use all left eye images for performance evaluations and corresponding right eyes are for score normalization purposes in fusion scenarios (see Sect. 5.3). Number of subjects, images, image resolution, image type and resulting amounts of genuine and impostor authentication attempts are summarized in Table 2. Examples of iris images of each database are depicted in Fig. 6. In accordance with IS ISO/IEC 19795-1:2006 [22], biometric performance is evaluated in terms of genuine match rate (GMR) and false match rate (FMR) by plotting the detection error trade-off (DET) curve and the equal error rate (EER).

In Fig. 7, the frequency of locations of detected keypoints in enhanced iris textures across all datasets is depicted in a heatmap. In can be observed that less keypoints are detected in iris parts which are frequently occluded by eyelids (two-dome pattern), since these yield homogeneous regions in the enhanced texture. This implicit regional limitation of keypoint detection makes an additional storage of noise masks redundant, while these could obviously

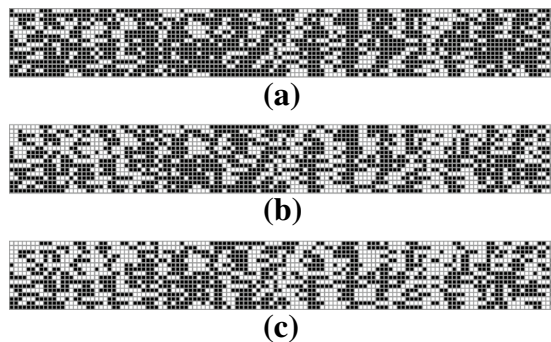


Fig. 5 Binarization of keypoint descriptors: examples for the first 16 binarized keypoint descriptors of the normalized enhanced texture of Fig. 1c according to three different thresholds. Each row represents one binarized keypoint descriptor. a $\tau = 10$. b $\tau = 20$. c $\tau = 30$

Table 2 Overview of employed databases w.r.t. image resolution, image type, corresponding number of subjects, number of iris images and the resulting number of genuine and impostor comparisons

Database	Resolution	Type	Subjects	Images	Gen. comp	Imp. comp
CASIAv1	320 × 280	NIR	108	756	8336	5778
CASIAv4	320 × 280	NIR	198	1332	4454	19,503
BioSecure	640 × 480	NIR	210	840	1260	21,945

Genuine and impostor scores have been derived following the recommendations in [35]

Fig. 6 Sample NIR iris images of the **a** CASIAv1, **b** CASIAv4-Interval and **c** BioSecure iris database

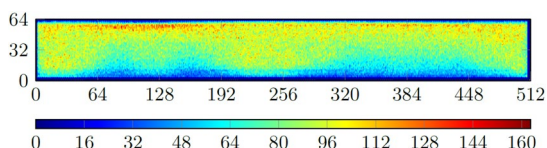
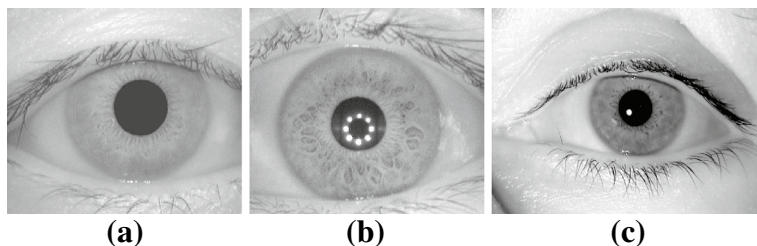


Fig. 7 Location of detected keypoints: frequencies of keypoint coordinates of SIFT keypoints detected in normalized enhanced textures extracted from all images of employed databases

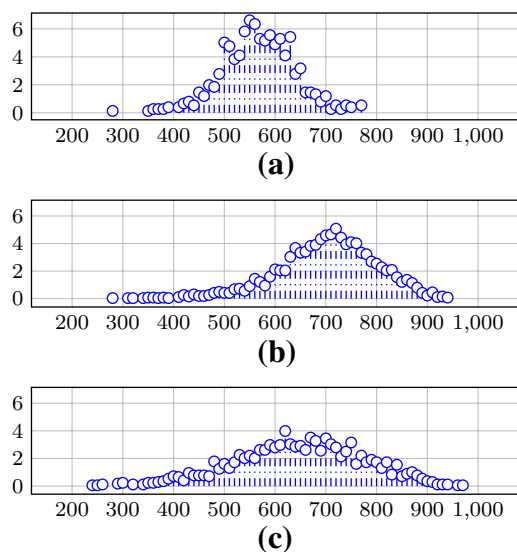


Fig. 8 Number of detected keypoints: distributions of the amount of detected keypoints in normalized enhanced textures for the three employed databases. **a** CASIAv1. **b** CASIAv4-Interval. **c** BioSecure

be easily incorporated in order to exclude unreliable keypoints. Figure 8 shows the frequency distribution of the amount of detected keypoints in enhanced iris textures on each database. We observe that the number of detected keypoints might vary considerably, e.g. if large parts of iris textures are affected by occlusions. Hence, that the number of detected keypoints could serve as an input for the quality check module of an iris recognition system. Across all used databases, an average number of keypoints extracted from enhanced iris textures is $N=645$. Hence, the look-up table has to consist of at least 10 bits per entry in order to index all keypoints. Together with the keypoint-code and the binarized descriptors, the stored template consists of $\tilde{W} \cdot \tilde{H} + 10 \cdot \tilde{W} \cdot \tilde{H} + N \cdot D$ bits. In contrast, assuming that a keypoint triple and its response value can be stored in 32 bits and that each feature descriptor bin consists of 8 bits, the average template size of the baseline system is estimated as $N \cdot (32 + 128 \cdot 8)$ bits. Depending on the used quantisation factors, $q_x, q_y,$ and q_θ , a significant reduction in storage requirement can be achieved in the proposed system. For instance, let $q_x = 4, q_y = 4, q_\theta = 90$, we get $\tilde{W} = 512/4 = 128$ and $\tilde{H} = 64/4 \cdot 360/90 = 64$, such that the average size of stored template results in $128 \cdot 64 + 10 \cdot 128 \cdot 64 + 645 \cdot 128 \approx 170$ kB. In contrast, the baseline system would on average require $645 \cdot (32 + 128 \cdot 8) \approx 665$ kB.

5.2 Performance evaluation

In a first experiment, we measure the biometric performance of the baseline system of Sect. 3. Table 3 compares obtained performance rates using SIFT, SURF and BRIEF keypoint descriptors across employed datasets for different reasonable values of ϵ_x and ϵ_y . As can be seen, for best configurations,

Table 3 Biometric performance in terms of EER (%) on employed databases for SIFT, SURF and BRIEF keypoint detectors and corresponding descriptors using an all-against-all comparison of keypoint matches and different parameters for geometrical constraints

ϵ_x	ϵ_y	CASIAv1			CASIAv4-Interval			BioSecure		
		SIFT	SURF	BRIEF	SIFT	SURF	BRIEF	SIFT	SURF	BRIEF
12	2	0.485	1.511	7.207	0.916	1.972	4.317	3.382	5.145	7.292
12	3	0.503	1.311	6.840	0.931	1.752	3.964	3.138	4.356	5.907
12	4	0.516	1.385	7.206	0.943	1.685	3.755	2.971	3.874	5.580
12	5	0.525	1.511	7.067	0.943	1.753	3.835	2.866	3.720	5.474
12	6	0.533	1.686	7.367	0.891	1.757	4.039	2.781	3.720	5.495
14	2	0.263	1.455	7.664	0.605	1.451	4.487	3.274	4.856	6.968
14	3	0.262	1.328	7.238	0.530	1.302	4.001	3.023	4.040	5.948
14	4	0.262	1.359	7.234	0.475	1.369	4.044	2.664	3.888	5.739
14	5	0.311	1.599	7.708	0.493	1.468	4.062	2.608	3.621	5.096
14	6	0.350	1.660	7.888	0.488	1.409	4.152	2.555	3.612	5.331
16	2	0.048	1.398	7.809	0.445	1.371	4.651	3.166	4.721	6.989
16	3	0.049	1.367	7.618	0.427	1.122	4.148	2.850	3.987	5.825
16	4	0.049	1.407	7.709	0.332	1.125	4.320	2.779	3.701	5.532
16	5	0.049	1.630	7.683	0.334	1.278	4.306	2.488	3.740	5.271
16	6	0.079	1.621	7.966	0.332	1.213	4.336	2.398	3.653	5.235
18	2	0.049	1.455	8.081	0.412	1.404	4.746	3.341	5.085	7.569
18	3	0.049	1.319	7.945	0.335	1.136	4.431	2.841	4.049	6.414
18	4	0.079	1.494	8.172	0.297	1.189	4.450	2.590	3.772	6.097
18	5	0.049	1.700	8.425	0.292	1.278	4.561	2.451	3.586	5.606
18	6	0.088	1.629	8.648	0.292	1.184	4.713	2.384	3.720	5.523
20	2	0.079	1.424	8.659	0.399	1.351	4.807	3.341	5.145	7.651
20	3	0.088	1.415	8.269	0.316	1.149	4.649	3.016	3.992	6.503
20	4	0.088	1.551	8.556	0.249	1.208	4.640	2.696	4.052	6.205
20	5	0.048	1.717	8.578	0.246	1.264	4.793	2.675	3.612	5.669
20	6	0.127	1.743	8.687	0.249	1.256	5.015	2.631	3.729	5.829
22	2	0.088	1.629	8.875	0.402	1.383	5.217	3.461	4.845	8.105
22	3	0.088	1.533	8.896	0.311	1.165	4.943	3.063	4.045	6.323
22	4	0.088	1.494	9.006	0.246	1.206	4.961	2.703	3.871	6.344
22	5	0.088	1.669	8.845	0.265	1.275	5.269	2.788	3.687	5.897
22	6	0.136	1.726	9.237	0.302	1.302	5.205	2.698	3.745	6.154

Bold values indicate the best results (lowest error rates)

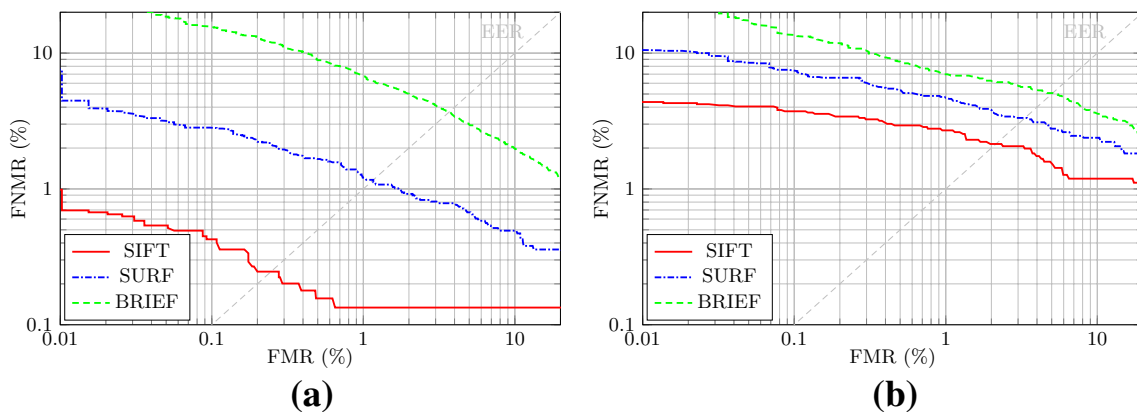


Fig. 9 DET curves for the best parameter configurations for SIFT, SURF and BRIEF keypoint detectors and corresponding keypoint descriptors using all-against-all comparison of keypoint descriptors,

cross-checking of keypoint matches and geometrical constraints. **a** CASIAv4-Interval. **b** BioSecure

ϵ_x values are significantly larger than ϵ_y values, since rotations of the eye (head tilts) cause significant variations in x -coordinates of detected keypoints. The DET curves for the best configurations on the CASIAv4-Interval and the BioSecure database are plotted in Fig. 9. While SURF and BRIEF extract more compact descriptors, which enable a faster comparison of keypoints, recognition accuracy drastically drops for these approaches compared to SIFT, which obtains practical performance rates across all datasets. These drastic performance drops, which have been observed across all datasets, stress the importance of discriminative keypoint

descriptors. If no cross-checking is applied, the EER of the best configuration on the CASIAv4-Interval database is almost doubled, from 0.246 to 0.431%. From Table 1, we observe that the proposed SIFT-based baseline system, which operates on enhanced iris textures and utilizes geometrical constraints as well as cross-checking of detected keypoint matches, significantly outperforms the vast majority of proposed approaches.

Table 4 summarizes obtained performance rates for the proposed system for appropriate parameter settings. For all evaluations, we use $S = 18$, which corresponds to a relative

Table 4 Biometric performance in terms of EER (%) on employed databases for the proposed system using different quantization parameters and thresholds for the binarization of keypoint descriptors

q_x	q_y	q_θ	CASIAv1			CASIAv4-Interval			BioSecure		
			$\tau = 10$	$\tau = 15$	$\tau = 20$	$\tau = 10$	$\tau = 15$	$\tau = 20$	$\tau = 10$	$\tau = 15$	$\tau = 20$
2	2	60	0.350	0.310	0.319	0.289	0.292	0.313	2.857	3.104	3.415
2	2	45	0.232	0.302	0.302	0.359	0.383	0.428	2.788	2.947	3.337
2	2	36	0.302	0.350	0.406	0.249	0.225	0.316	2.961	3.318	3.749
2	3	60	0.358	0.350	0.406	0.313	0.311	0.359	2.550	2.815	3.178
2	3	45	0.310	0.310	0.310	0.359	0.337	0.375	2.691	2.859	3.019
2	3	36	0.262	0.262	0.310	0.340	0.359	0.426	3.092	3.263	3.420
2	4	60	0.350	0.358	0.358	0.297	0.268	0.270	2.693	2.931	3.178
2	4	45	0.302	0.341	0.319	0.289	0.313	0.313	2.592	2.700	3.092
2	4	36	0.398	0.398	0.398	0.268	0.268	0.294	2.640	3.021	3.184
2	5	60	0.533	0.524	0.516	0.359	0.287	0.356	2.864	2.767	3.009
2	5	45	0.398	0.446	0.476	0.359	0.359	0.359	2.377	2.608	2.610
2	5	36	0.319	0.350	0.271	0.268	0.287	0.356	2.651	2.613	2.696
3	2	60	0.524	0.454	0.446	0.313	0.294	0.354	3.092	3.175	3.433
3	2	45	0.476	0.350	0.350	0.354	0.364	0.340	2.597	3.095	3.251
3	2	36	0.398	0.437	0.533	0.289	0.263	0.375	2.693	3.021	3.018
3	3	60	0.485	0.485	0.564	0.313	0.292	0.356	2.767	3.014	3.305
3	3	45	0.485	0.437	0.437	0.292	0.361	0.423	2.532	2.624	2.956
3	3	36	0.437	0.350	0.485	0.289	0.337	0.380	2.705	2.774	3.092
3	4	60	0.437	0.398	0.485	0.340	0.246	0.313	2.622	2.830	3.245
3	4	45	0.310	0.350	0.350	0.337	0.292	0.337	2.147	2.601	2.785
3	4	36	0.350	0.437	0.446	0.273	0.268	0.313	2.610	2.862	3.150
3	5	60	0.533	0.524	0.603	0.359	0.362	0.359	2.945	2.945	3.101
3	5	45	0.476	0.358	0.350	0.380	0.426	0.404	2.259	2.223	2.451
3	5	36	0.446	0.437	0.428	0.268	0.292	0.356	2.772	2.615	2.709
4	2	60	0.398	0.350	0.446	0.340	0.313	0.340	2.788	3.092	3.325
4	2	45	0.406	0.446	0.446	0.383	0.359	0.359	3.014	3.362	3.487
4	2	36	0.406	0.494	0.485	0.268	0.268	0.294	2.942	2.871	3.168
4	3	60	0.485	0.485	0.485	0.359	0.316	0.356	2.866	3.083	3.099
4	3	45	0.350	0.398	0.454	0.388	0.428	0.404	2.629	2.776	3.003
4	3	36	0.310	0.350	0.398	0.311	0.316	0.364	2.853	2.862	3.099
4	4	60	0.485	0.446	0.516	0.356	0.292	0.337	2.375	2.714	3.097
4	4	45	0.310	0.319	0.398	0.335	0.335	0.356	2.541	3.005	3.332
4	4	36	0.463	0.485	0.485	0.402	0.356	0.378	2.594	2.772	3.090
4	5	60	0.572	0.620	0.572	0.429	0.402	0.471	2.687	2.465	2.848
4	5	45	0.468	0.446	0.612	0.302	0.319	0.416	2.543	2.732	2.933
4	5	36	0.494	0.572	0.590	0.287	0.292	0.380	2.227	2.375	2.696

Bold values indicate the best results (lowest error rates)

rotation of approximately $\pm 10^\circ$, and $t = 1$ in order to achieve rotation invariance and a more robust pairing of keypoint coordinates, respectively. In case only the overlap of probe and reference keypoint-codes, $\|\mathbf{A}_{pp}\|$, would be employed as comparison score, i.e. not taking into account corresponding keypoint descriptors, biometric performance significantly decreases. For instance, for the best configuration on the CASIAv4-Interval database the obtained EER increases from 0.225 to 5.021%. For the binarization of keypoint descriptors thresholds of $\tau = 10, 20$, and 30 have been found to reveal best biometric performance. As can be observed, on the more challenging CASIAv4-Interval and BioSecure database, the presented scheme maintains the recognition accuracy of the SIFT-based baseline iris recognition system. Best performance rates, which indicate a robust pairing of keypoints, are achieved for rather small quantisation values of q_x and q_y . Focusing on keypoint angles, best recognition accuracies are obtained for q_θ values of 45 or less. In order to further assess the goodness of the presented binarization method, we evaluate the proposed system employing two alternative descriptor binarization schemes referred to as Binarized Scale-Invariant Feature Transform (B-SIFT) [29] and Binarization of Gradient Orientation Histograms (BIG-OH) [3]. In the B-SIFT approach, the median bin value of each descriptor is used as threshold based on which all bins are binarized. In the BIG-OH scheme, consecutive bin values are compared and greater/smaller-or-equal relations are binary encoded. Both approaches employ *HD*-scores to determine the dissimilarity between keypoint descriptors. Moreover, in case keypoint descriptors are not binarized, the L_2 -norm can be employed in order to compare them. Obtained performance rates for applying the L_2 -norm on original keypoint descriptors and the B-SIFT as well as the BIG-OH scheme are shown in Table 5 for various parameter configurations. The proposed binarization technique outperforms the B-SIFT as well as the BIG-OH approach on all datasets. Performance rates achieved when applying the L_2 -norm on original keypoint descriptors underline that information loss is negligible in the presented binarization method. DET curves which compare the best configurations of the proposed system to the aforementioned schemes are depicted in Fig. 10 for the CASIAv4-Interval and the BioSecure database. The corresponding probability density functions of the proposed system for both databases are depicted in Fig. 11.

5.3 Comparison and fusion with traditional systems

Further, we compare and fuse the proposed system with two traditional approaches. The first feature extraction method follows the Daugman-like 1D-Log Gabor feature extraction algorithm of Masek [36] (LG), and the second follows the algorithm proposed by Ma et al. [34] (QSW) based on a quadratic spline wavelet transform. Both methods divide

the pre-processed iris texture into stripes to obtain 10 one-dimensional signals, each one averaged from the pixels of 5 adjacent rows, and extract iris-codes consisting of 512×20 bits. In the comparison stage, *HD*-based scores are estimated performing ± 16 bit shifts to compensate misalignments. Implementations of the employed feature extractors are freely available in the University of Salzburg Iris Toolkit (USIT) [52]. Table 6 compares the biometric performance of LG and QSW to the best configurations of the previously evaluated schemes. We observe that the proposed system reveals competitive biometric performance compared to the traditional systems across all databases achieving the lowest EERs on the CASIAv4-Interval and the BioSecure database.

In order to investigate whether features extracted by the proposed system complement those extracted by traditional schemes, as reported in [1], we perform a biometric fusion on the BioSecure database, which has been found to be the most challenging one. In the context of biometric fusion, score-level fusion using the sum-rule with proper normalization has been observed to result in competitive performance [24]. To fuse scores obtained by the traditional and the proposed system, we perform a *z*-score normalization prior to performing a score-level fusion based on the sum-rule. Normalization parameters are estimated employing score distributions obtained from right eye comparisons. From the DET curves in Fig. 12, we observe that a fusion of the traditional schemes does not yield a significant gain in biometric performance. In contrast, a fusion of the proposed system with LG or QSW does reveal a clear improvement in recognition accuracy resulting in EERs of 1.58 and 1.74%, respectively.

5.4 Time measurements

Experiments were performed on a single core of an Intel Core i7-3610QM CPU with 3.2G Hz on a standard work station with sufficient RAM. Keypoint-codes, binarized descriptors and iris-codes are internally represented as two-dimensional arrays consisting of bytes (chars). Comparisons are performed by using according bit operations where the number of bits set, i.e. 1s, within resulting bytes is derived from a 8-bit look-up table. The implementation can be considered lightweight and fully portable as it does not make use of larger look-up tables which require a significant amount of RAM nor of CPU-specific PopCnt (population count) functions. That is, our implementation yields an upper bound for the time required for a single comparison. Figure 13 depicts average execution time and corresponding boxplots for a single iris biometric comparison required by all considered systems. As can be seen, the proposed system clearly outperforms the baseline systems based on SIFT and SURF revealing an average execution time comparable to that of the BRIEF-based baseline system, which obtained unpractical recognition accuracy. Using B-SIFT or BIG-OH in order to binarize descriptors slightly reduces the

Table 5 Biometric performance in terms of EER (%) on employed databases using different quantization parameters with L_2 -norm and fractional HD , after binarization with BIG-OH and B-SIFT, for the comparison of feature descriptors

q_x	q_y	q_θ	CASIAv1			CASIAv4-Interval			BioSecure		
			L_2	BIG-OH	B-SIFT	L_2	BIG-OH	B-SIFT	L_2	BIG-OH	B-SIFT
2	2	60	0.249	0.970	0.905	0.350	1.198	0.766	3.195	5.746	4.077
2	2	45	0.127	0.708	0.668	0.459	1.085	0.737	3.068	5.091	4.278
2	2	36	0.214	0.786	0.834	0.374	1.136	0.721	3.630	6.185	4.492
2	3	60	0.327	0.970	0.874	0.420	1.412	0.838	2.801	6.115	4.125
2	3	45	0.214	0.747	0.620	0.359	1.058	0.764	2.921	5.318	3.978
2	3	36	0.209	0.699	0.795	0.473	1.257	0.788	3.224	6.086	4.273
2	4	60	0.314	1.066	0.874	0.353	1.465	0.795	3.017	5.878	4.045
2	4	45	0.305	0.795	0.620	0.348	1.120	0.694	2.601	5.165	4.132
2	4	36	0.253	0.786	0.747	0.380	1.257	0.812	3.311	6.037	3.964
2	5	60	0.380	1.319	1.193	0.389	1.679	0.938	2.549	5.158	3.653
2	5	45	0.358	1.018	0.882	0.321	1.154	0.881	2.474	4.910	3.729
2	5	36	0.257	0.978	0.826	0.398	1.369	0.780	2.515	5.401	4.047
3	2	60	0.367	0.930	0.922	0.386	1.482	0.900	3.272	5.882	4.300
3	2	45	0.349	0.708	0.572	0.416	1.163	0.852	3.214	5.315	4.342
3	2	36	0.305	0.804	0.852	0.432	1.311	0.855	3.076	5.940	3.955
3	3	60	0.445	1.057	0.970	0.412	1.675	1.008	3.113	5.721	4.195
3	3	45	0.327	0.843	0.826	0.424	1.326	0.911	2.753	5.488	4.206
3	3	36	0.394	0.834	0.834	0.441	1.353	0.876	2.916	5.892	3.931
3	4	60	0.353	1.000	0.922	0.390	1.760	1.034	2.921	6.196	4.358
3	4	45	0.197	0.769	0.660	0.396	1.312	0.890	2.596	5.230	4.033
3	4	36	0.319	0.930	0.795	0.403	1.573	1.008	2.928	6.115	4.116
3	5	60	0.389	1.249	1.097	0.470	1.934	1.013	2.710	5.841	3.969
3	5	45	0.288	1.026	0.708	0.401	1.501	0.967	2.580	4.849	3.493
3	5	36	0.363	0.970	0.865	0.434	1.800	1.010	2.579	5.313	3.731
4	2	60	0.275	0.843	0.930	0.383	1.693	0.927	3.162	6.021	4.213
4	2	45	0.358	0.882	0.795	0.478	1.340	0.900	3.083	5.152	4.365
4	2	36	0.349	0.922	0.882	0.392	1.580	0.879	2.963	5.894	4.121
4	3	60	0.384	1.018	1.057	0.381	1.817	1.136	3.034	6.189	4.054
4	3	45	0.349	0.970	0.843	0.431	1.415	1.021	2.606	5.345	4.040
4	3	36	0.305	0.891	0.891	0.428	1.651	0.975	3.008	6.026	4.211
4	4	60	0.358	1.040	1.018	0.430	2.053	1.026	2.710	5.993	4.042
4	4	45	0.284	0.852	0.786	0.382	1.479	0.933	3.090	5.523	4.287
4	4	36	0.415	0.961	1.018	0.465	1.774	1.082	3.203	6.046	4.061
4	5	60	0.502	1.249	1.280	0.596	2.294	1.106	2.428	5.472	3.749
4	5	45	0.341	1.066	0.930	0.453	1.641	1.083	2.597	5.029	3.793
4	5	36	0.411	1.074	1.049	0.456	2.032	1.112	2.483	5.297	3.830

Bold values indicate the best results (lowest error rates)

execution time, since these schemes do not require a normalization of intermediate scores. However, it has been shown that these schemes fail to maintain the recognition accuracy of the SIFT-based baseline system. Hence, focusing on SIFT, the proposed system yields the best trade-off between biometric performance and authentication speed. The traditional systems

still outperform the proposed system in terms of execution time as these require less bit comparisons. Nevertheless, as previously mentioned, a limitation of the template construction to a fixed number of N keypoints, which exhibit highest response values, could be used to achieve further speed-ups as well as more constant execution times.

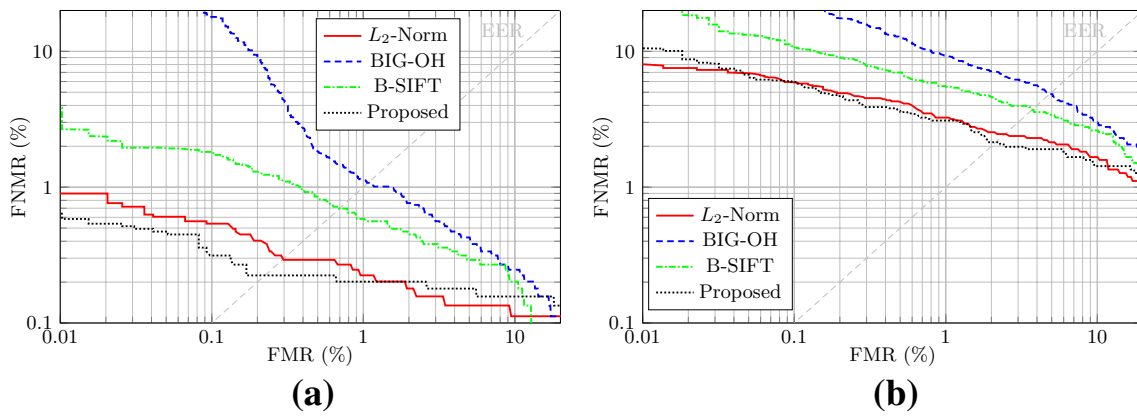


Fig. 10 DET curves for the best parameter configurations for L_2 -norm and HD -based comparison of binarized keypoint descriptors BIG-OH, B-SIFT and the proposed method. **a** CASIAv4-Interval. **b** BioSecure

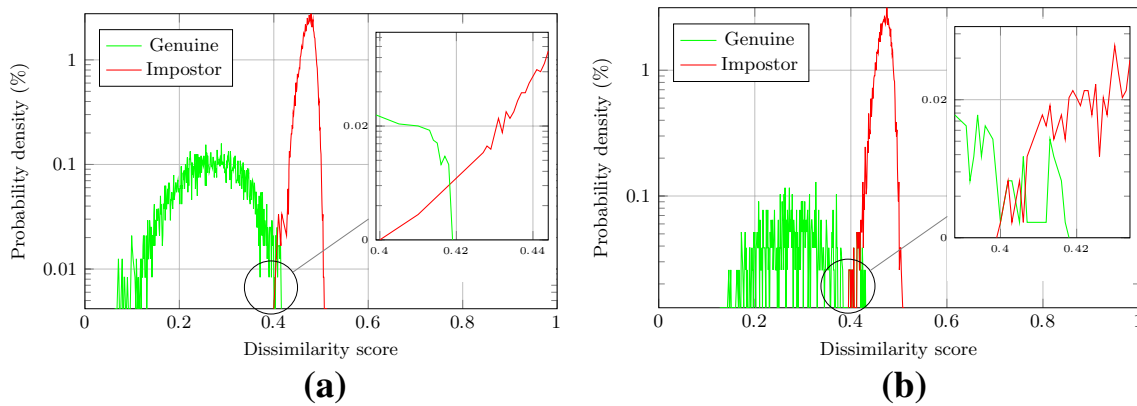


Fig. 11 Probability density functions for the best parameter configurations of the proposed method. **a** CASIAv4-Interval. **b** BioSecure

Table 6 Summary of biometric performance in terms of EER (%) for all databases obtained by the baseline systems and best parameter configurations all keypoint-based approaches

Method	CASIAv1	CASIAv4-Interval	BioSecure
QSW \pm 16 bit shift	0.485	0.356	2.930
LG \pm 16 bit shift	1.144	0.447	2.301
SIFT with geom. constr.	0.048	0.246	2.384
SURF with geom. constr.	1.311	1.122	3.612
BRIEF with geom. constr.	6.840	3.755	5.096
Keypoint-code with L_2 -norm	0.127	0.321	2.474
Keypoint-code with BIG-OH	0.708	1.058	4.849
Keypoint-code with B-SIFT	0.620	0.694	3.729
Proposed	0.232	0.225	2.147

Bold values indicate the best results (lowest error rates)

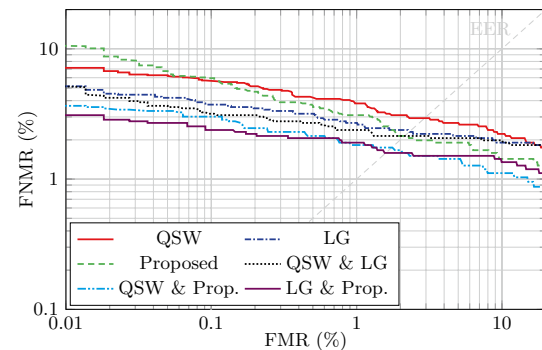
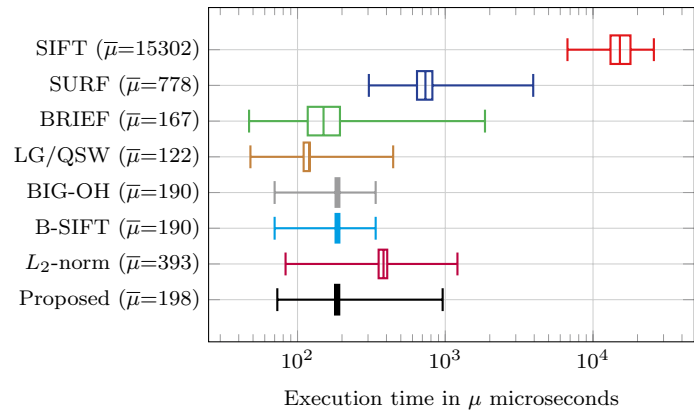


Fig. 12 DET curves on the BioSecure database for the baseline systems, the best parameter configuration of the proposed method and pair-wise score-level fusions thereof

Fig. 13 Comparison of the average execution time (in microseconds) required for a single pair-wise biometric comparison, boxplots depict the corresponding median, upper/lower quartile and upper/lower whisker for different methods (note the logarithmic scale on x -axis)



6 Conclusion

In contrast to proposed studies, we have shown that the discriminative power of SIFT-based features is comparable to that of traditional NIR iris recognition systems. In order to obliterate the additional shortcomings regarding authentication speed and storage requirement, we proposed an improved SIFT-based iris recognition system achieving an approximately 75-fold speed-up (without the use of intrinsic functions), compared to an according baseline system. The authentication speed provided by the proposed approach is similar to that of traditional scheme. Furthermore, employed binarization methods significantly reduce storage requirements. Hence, the proposed approach represents the first practical applicable SIFT-based iris recognition system in particular, in scenarios where reliable recognition and comparison speed represent crucial factors such as large-scale identification systems. Finally, we confirmed that SIFT-based features complement those extracted by traditional schemes, such that a multi-algorithm yields a significant gain in recognition accuracy, as reported in [1]. It is important to note that generic improvement which have been the proposed for traditional iris recognition system, e.g. by using generalized or personalized weight maps [17], can be easily integrated in order to further improve recognition accuracy.

An application of the proposed system to VW iris images is subject to future work. Since VW iris recognition schemes tend to exhibit inferior recognition accuracy compared to systems based on NIR images [7], a computationally efficient multi-algorithm fusion would be especially appealing in order to improve performance rates. Computational efficiency might as well play an important role in mobile iris recognition systems [7, 48], where computing and storage capacities are limited. Further, the presented concept could be applied to speed-up fingerprint recognition systems, which perform a brute-force

comparison of vicinity-based descriptors of minutiae triples, e.g. [6, 9]. In contrast to iris recognition, where alignment involves a one-dimensional shifting, in case of a fingerprint recognition system proper pre-alignment methods would have to be employed. In order to facilitate reproducible research [54], an implementation of the improved SIFT-based baseline iris recognition system is made available together with [52] and the proposed system will be included upon acceptance.

Acknowledgements Funding was partially provided 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 Center for Research in Security and Privacy (CRISP).

References

1. Alonso-Fernandez F, Tome-Gonzalez P, Ruiz-Albacete V, Ortega-Garcia J (2009) Iris recognition based on SIFT features. In: International conference on biometrics, identity and security (BIDS'09), pp 1–8
2. Baber J, Dailey MN, Satoh S, Afzulpurkar N, Bakhtyar M (2014) BIG-OH: binarization of gradient orientation histograms. *Image Vis Comput* 32(11):940–953
3. Baber J, Dailey MN, Satoh S, Afzulpurkar N, Bakhtyar M (2014) Big-oh: Binarization of gradient orientation histograms. *Image Vis Comput* 32(11):940–953
4. Bay H, Ess A, Tuytelaars T, Gool LV (2008) Speeded-up robust features (SURF). *Comput Vis Image Underst* 110(3):346–359
5. Belcher C, Du Y (2009) Region-based SIFT approach to iris recognition. *Opt Lasers Eng* 47(1):139–147
6. Bringer J, Despiegel V (2010) Binary feature vector fingerprint representation from minutiae vicinities. In: 4th IEEE International conference on biometrics: theory applications and systems (BTAS'10), pp 1–6
7. Burge MJ, Bowyer K (2016) Handbook of iris recognition, 2nd edn. Springer, New York
8. Calonder M, Lepetit V, Strecha C, Fua P (2010) BRIEF: binary robust independent elementary features. In: Proceedings of the 11th European conference on computer vision (ECCV'10), pp 778–792

9. Cappelli R, Ferrara M, Maltoni D (2010) Minutia cylinder-code: a new representation and matching technique for fingerprint recognition. *IEEE Trans Pattern Anal Mach Intell* 32(12):2128–2141
10. Chen J, Shen F, Chen D, Flynn P (2016) Iris recognition based on human-interpretable features. *IEEE Trans Inf Forensics Secur* 99:1–10
11. Chen Y, Liu Y, Zhu X, He F, Wang H, Deng N (2014) Efficient iris recognition based on optimal subfeature selection and weighted subregion fusion. *Sci World J* 2014:1–19
12. Chinese Academy of Sciences' Institute of Automation (2002) CASIA Iris Image Database V1.0. <http://biometrics.idealtest.org>
13. Chinese Academy of Sciences' Institute of Automation (2002) CASIA Iris Image Database V3.0. <http://biometrics.idealtest.org>
14. Chinese Academy of Sciences' Institute of Automation (2012) CASIA Iris Image Database V4.0. <http://biometrics.idealtest.org>
15. Daugman J (1993) High confidence visual recognition of persons by a test of statistical independence. *IEEE Trans Pattern Anal Mach Intell* 15(11):1148–1161
16. Daugman J (2004) How iris recognition works. *IEEE Trans Circuits Syst Video Technol* 14(1):21–30. <https://doi.org/10.1109/ICIP.2002.1037952>
17. Dong W, Sun Z, Tan T (2011) Iris matching based on personalized weight map. *IEEE Trans Pattern Anal Mach Intell* 33(9):1744–1757
18. Fierrez J, Ortega-Garcia J, Toledano DT, Gonzalez-Rodriguez J (2007) Biosec baseline corpus: a multimodal biometric database. *Pattern Recognit* 40(4):1389–1392
19. Indian Institute of Technology Kanpur (2009) IITK Iris Image Database. <http://www.cse.iitk.ac.in/users/biometrics/>
20. ISO/IEC JTC1 SC27 Security Techniques (2015) ISO/IEC 29794:2015. Information technology—biometric sample quality—part 6: iris image data. International Organization for Standardization
21. ISO/IEC JTC1 SC27 Security Techniques (2016) ISO/IEC 30107-1:2016. Information technology—biometric presentation attack detection—part 1: framework. International Organization for Standardization
22. ISO/IEC TC JTC1 SC37 Biometrics (2006) ISO/IEC 19795-1:2006. Information technology—biometric performance testing and reporting—part 1: principles and framework. International Organization for Standardization and International Electrotechnical Committee
23. Jain AK, Flynn P, Ross AA (2008) *Handbook of biometrics*. Springer, New York
24. Jain AK, Klare B, Ross AA (2015) Guidelines for best practices in biometrics research. In: *Proceedings of the 8th international conference on biometrics 2015 (ICB'15)*, pp 1–5
25. Jain AK, Ross A, Prabhakar S (2004) An introduction to biometric recognition. *IEEE Trans Circuits Syst Video Technol* 14:4–20. <https://doi.org/10.1109/TCSVT.2003.818349>
26. Kong AWK, Zhang D, Kamel MS (2010) An analysis of iriscode. *IEEE Trans Image Process* 19(2):522–532
27. Krichen E, Mellakh MA, Garcia-Salicetti S, Dorizzi B (2004) Iris identification using wavelet packets. In: *Proceedings of the 17th international conference on pattern recognition (ICPR'04)*, vol 4, pp 335–338
28. Li J, Lu Z (2011) B-SIFT: a highly efficient binary SIFT descriptor for invariant feature correspondence. In: *Workshop on intelligent science and intelligent data engineering*, pp 426–433
29. Li J, Lu Z (2012) Intelligent science and intelligent data engineering: second sino-foreign-interchange workshop, IScIDE 2011, chap. B-SIFT: a highly efficient binary SIFT descriptor for invariant feature correspondence, pp 426–433
30. Liu N, Zhang M, Li H, Sun Z, Tan T (2016) Deepiris: learning pairwise filter bank for heterogeneous iris verification. *Pattern Recognit Lett* 82:154–161 An insight on eye biometrics
31. Liu X, Li P (2012) An iris recognition approach with SIFT descriptors. In: *7th International conference on advanced intelligent computing theories and applications. With aspects of artificial intelligence*, pp 427–434
32. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60(2):91–110
33. Ma L, Tan T, Wang Y, Zhang D (2003) Personal identification based on iris texture analysis. *IEEE Trans Pattern Anal Mach Intell* 25(12):1519–1533
34. Ma L, Tan T, Wang Y, Zhang D (2004) Efficient iris recognition by characterizing key local variations. *IEEE Trans Image Process* 13(6):739–750. <https://doi.org/10.1109/TIP.2004.827237>
35. Mansfield A, Wayman J (2002) Best practices in testing and reporting performance of biometric devices. Tech. rep., U.K. Government Biometrics Working Group
36. Masek L (2003) Recognition of human iris patterns for biometric identification. Master's thesis, University of Western Australia
37. Mehrotra H, Majhi B, Gupta P (2010) Robust iris indexing scheme using geometric hashing of SIFT keypoints. *J Netw Comput Appl* 33(3):300–313 Recent advances and future directions in biometrics personal identification
38. Mehrotra H, Majhi B, Sa P (2011) Unconstrained iris recognition using F-SIFT. In: *8th International conference on information, communications and signal processing (ICICS'11)*, pp 1–5
39. Monro DM, Rakshit S, Zhang D (2007) Dct-based iris recognition. *IEEE Trans Pattern Anal Mach Intell* 29(4):586–595
40. Multimedia University (2004) MMU iris image database. <http://pesona.mmu.edu.my/ccteo>
41. National Institute of Standards and Technology (NIST) (2005) Iris challenge evaluation (ICE). <http://www.nist.gov/itl/iad/ig/ice.cfm>
42. Nguyen K, Fookes C, Ross A, Sridharan S (2017) Iris recognition with off-the-shelf cnn features: a deep learning perspective. *IEEE Access* 99:1–1
43. Nigam I, Vatsa M, Singh R (2015) Ocular biometrics: a survey of modalities and fusion approaches. *Inf Fusion* 26:1–35
44. Ortega-Garcia J, Fierrez J, Alonso-Fernandez F, Galbally J, Freire MR, Gonzalez-Rodriguez J, Garcia-Mateo C, Alba-Castro JL, Gonzalez-Agulla E, Otero-Muras E, Garcia-Salicetti S, Allano L, Ly-Van B, Dorizzi B, Kittler J, Bourlai T, Poh N, Deravi F, Ng MW, Fairhurst M, Hennebert J, Humm A, Tistarelli M, Brodo L, Richiardi J, Drygajlo A, Ganster H, Sukno FM, Pavani SK, Frangi A, Akarun L, Savran A (2010) The multisenario multienvironment biosecure multimodal database (BMDB). *IEEE Trans Pattern Anal Mach Intell* 32(6):1097–1111
45. Park U, Pankanti S, Jain AK (2008) Fingerprint verification using sift features
46. Proença H, Alexandre LA (2005) Ubiiris: a noisy iris image database. In: *Proceedings of the 13th international conference on image analysis and processing (ICIAP'05)*, pp 970–977
47. Proença H, Filipe S, Santos R, Oliveira J, Alexandre L (2010) The UBIRIS.v2: a database of visible wavelength iris images captured on-the-move and at-a-distance. *IEEE Trans Pattern Anal Mach Intell* 32(8):1529–1535
48. Raja KB, Raghavendra R, Stokkenes M, Busch C (2015) Multimodal authentication system for smartphones using face, iris and periocular. In: *Proceedings of the international conference on biometrics (ICB'15)*, pp 143–150
49. Sun Z, Tan T (2009) Ordinal measures for iris recognition. *IEEE Trans Pattern Anal Mach Intell* 31(12):2211–2226

50. Sunder MS, Ross A (2010) Iris image retrieval based on macro-features. In: Proceedings of the 20th international conference on pattern recognition (ICPR'10), pp 1318–1321
51. Tomeo-Reyes I, Ross A, Clark A, Chandran V (2015) A biomechanical approach to iris normalization. In: International conference on biometrics (ICB'15), pp 9–16
52. USIT (2016) University of Salzburg iris toolkit. <http://www.wavelab.at/sources/Rathgeb16a>. Version 2.0
53. University of Bath (2004) University of Bath iris image database. <http://www.bath.ac.uk/>
54. Vandewalle P, Kovacevic J, Vetterli M (2009) Reproducible research in signal processing—what, why, and how. *IEEE Signal Process Mag* 26(3):37–47
55. West Virginia University (2005) WVU iris image database. <http://www.csee.wvu.edu/>
56. Yang G, Pang S, Yin Y, Li Y, Li X (2012) SIFT based iris recognition with normalization and enhancement. *Int J Mach Learn Cybern* 4(4):401–407
57. Yu L, Zhang D, Wang K (2007) The relative distance of key point based iris recognition. *Elsevier Pattern Recognit* 40(2):423–430
58. Zhu R, Yang J, Wu R (2006) Iris recognition based on local feature point matching. In: International symposium on communications and information technologies (ISCIT'06), pp 451–454
59. Zuiderveld K (1994) *Graphics gems iv*. chap. Contrast limited adaptive histogram equalization, pp 474–485