On the Application of Bloom Filters to Iris Biometrics

Christian Rathgeb, Frank Breitinger, Christoph Busch and Harald Baier

Abstract—In this work the application of adaptive Bloom filters to binary iris biometric feature vectors, i.e. iris-codes, is proposed. Bloom filters, which have been established as a powerful tool in various fields of computer science, are applied in order to transform iris-codes to a rotation-invariant feature representation. Properties of the proposed Bloom filter-based transform concurrently enable (1) biometric template protection, (2) compression of biometric data, and (3) acceleration of biometric identification, while at the same time no significant degradation of biometric performance is observed. According to these fields of application detailed investigations are presented.

Experiments are carried out on the CASIA-v3 iris database for different feature extraction algorithms. Confirming the soundness of the proposed approach, the application of adaptive Bloom filters achieves rotation-invariant cancelable templates maintaining biometric performance, a compression of templates down to 20–40% of original size, and a reduction of bit-comparisons to less than 5% leading to a substantial speed-up of the biometric system in identification mode.

Index Terms—Biometrics, iris recognition, Bloom filter, biometric template protection, biometric template compression, computationally efficient identification.

I. INTRODUCTION

IRIS BIOMETRIC RECOGNITION [8], [2], [40] is field-proven as a robust and reliable biometric technology. The iris's complex texture and its apparent stability hold tremendous promise for applying iris recognition in diverse application scenarios, such as border control, forensic investigations, as well as cryptosystems [12], [41]. Daugman's algorithm [10], forms the basis of the vast majority of today's iris recognition systems, which report (true positive) identification rates above 99% and equal error rates less than 1%: (1) at enrollment an image of a subject's eye is acquired; (2) in the pre-processing step the boundaries of the pupil and the outer iris are detected and the iris (in the approximated form of a ring) is "un-rolled" to obtain a normalized rectangular iris texture; (3) feature extraction is applied in order to generate a highly discriminative binary feature vector, i.e. iris-code; (4) at the time of authentication pairs of iris-codes are efficiently compared by calculating the Hamming distance between them, where template alignment is performed within a single dimension, applying a circular shift of iris-codes, to compensate against head tilts of a certain degree. Technologies of iris recognition are already deployed on national-sized databases, e.g. the Unique IDentification Authority of India (UIDAI) [44], which aims at registering all 1.2 billion Indian

citizens, is enrolling 1 million subjects on a daily basis. With about 300 million persons enrolled (status February, 2013), against which the daily intake has to be compared to check for duplicate identities the daily workflow of iris cross-comparisons results in 3×10^{14}, or 300 trillion (!). Resistance to false matches and comparison speed, which is achieved by various existing approaches [40], are vital for any large-scale biometric deployments. Nonetheless, the explosive impact and scale of iris recognition is accompanied by serious consequential issues, e.g. privacy concerns or computational limitations, which are still to be solved.

From a privacy perspective most concerns against the common use of biometrics arise from the storage and misuse of biometric data as well as the permanent tracking and observation of activities [7]. In addition, it has been shown that spoofed iris images can be re-constructed from stored iris-codes [45]. In accordance with the ISO/IEC IS 24745 [19] on biometric information protection, technologies of biometric template protection [21], [37] in particular, cancelable biometrics [34] meet the two major requirements of irreversibility and unlinkability. Cancelable biometrics which consist of intentional, repeatable distortions of biometric signals based on transforms that provide a comparison of biometric templates in the transformed domain, permanently protect biometric templates. However, the majority of approaches to cancelable biometrics report a significant decrease in biometric performance, which is caused by the fact that local neighborhoods of feature elements are often obscured and the transformed enrollment templates are not "seen" at the time of authentication, i.e. alignment can not be performed properly [37].

A binary representation of biometric data offers two major advantages, compact storage and rapid comparison [9]. Despite these benefits, it has been found that extracted iris-codes still suffer from low entropy [18], e.g. approximately 250 mutually independent bits out of 2048 in [10]. A compression of iris-codes enables an even more compact storage, e.g. in 2D bar codes, smart cards, or magnetic stripes [14]. Focusing on biometric identification, deployments of iris recognition perform brute force exhaustive searches which are accelerated in case the amount of required bit comparisons is significantly reduced, e.g. by utilizing compressed templates [13] or an alignment-free representation, which does not require circular bit-shifting [25].

The contribution of this work is the proposal of a generic approach to obtain a rotation-invariant representation of iris-codes based on adaptive Bloom filters. A Bloom filter [1] is a space-efficient probabilistic data structure representing a set in order to support membership queries. In addition to an efficient

C. Rathgeb, F. Breitinger, C. Busch and H. Baier are with the da/sec – Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany, http://www.dasec.h-da.de/.

storage and rapid processing of queries, Bloom filters convince by their wide field of applications, e.g. database applications [31] or network applications [4]. In [35] we have already demonstrated the applicability of Bloom filters in order to achieve template protection. In the presented work these ideas are extended and properties of Bloom filter-based transforms are utilized to tackle all of the aforementioned issues regarding (iris) biometrics:

1) *Template protection*: the successive mapping of parts of a binary biometric template to Bloom filters represents an irreversible transform achieving alignment-free protected biometric templates.

2) *Biometric data compression*: the proposed Bloom filter-based transform can be parameterized to obtain a desired template size, operating a trade-off between compression and biometric performance.

3) *Efficient identification*: a compact alignment-free representation of iris-codes enables a computaionally efficient biometric identification reducing the overall response time of the system.

According to these benefits, the proposed approach represents a secure template protection scheme which can be efficiently applied within an iris identification system.

The remainder of this work is organized as follows: related work with respect to iris biometric template protection, template compression, and computationally efficient iris biometric identification is summarized in Sect. II. In Sect. III the proposed approach is described in detail and applications of adaptive Bloom filters are proposed. Experimental evaluations are presented and obtained results are discussed in Sect. IV. Finally, conclusions are drawn in Sect. V.

II. RELATED WORK

Biometric template protection schemes [37] are commonly categorized as biometric cryptosystems and cancelable biometrics. Focusing on biometric cryptosystems the majority of existing approaches implement cryptographic primitives, e.g. fuzzy commitment scheme [23] or fuzzy vault scheme [22]. However, suggested approaches [16], [3], [26], have been exposed to be vulnerable to diverse attacks, e.g. based on statistical attacks [38] or via record multiplicity [42]. Complex key retrieval procedures, which are required at biometric comparison, prevent from a computationally efficient identification, representing another drawback of biometric cryptosystems. Ratha *et al.* [34] were the first introducing the concept of cancelable biometrics. In their work the authors apply image-based block permutations and surface-folding in order to obtain revocable biometric templates. In further work [47] the authors propose different techniques to generate cancelable iris biometrics based on non-invertible transforms and biometric salting, which are applied in image and feature domain. In order to preserve a computational efficient alignment of resulting iris-codes based on circular bit-shifting, iris textures and iris-codes are obscured in a row-wise manner, which means adjacency of pixels and bits is maintained along x-axis in image and feature domain, respectively. In [15] block re-mapping and image wraping have been applied

to normalized iris textures. For both types of transforms a proper alignment of resulting iris-codes is infeasible causing a significant decrease of biometric performance [37]. In [32] several enrollment templates are processed to obtain a vector of consistent bits. Revocability is provided by encoding the iris-code according to a subject-specific random seed. In case subject-specific transforms are applied in order to achieve cancelable biometrics, these transforms have to be considered compromised during inter-class comparisons [24]. Subject-specific secrets, be it transform parameters, random numbers, or any kind of passwords are easily compromised, i.e. performance evaluations have to be performed under the "stolen-secret scenario", where each impostor is in possession of valid secrets. In [33] cancelable iris templates are achieved by applying sector random projection to iris images. Again, recognition performance is only maintained if subject-specific random matrices are applied. In [5] non-invertible iris-codes are computed by thresholding inner products of the feature vector with randomly generated vectors. The random vectors are created by using a per-subject secret and a pseudo random number generator. Several normalized iris textures are multiplied with a random kernel in [6] to create concealed feature vectors. The vast majority of cancelable iris biometric systems only maintains biometric performance for settings which leave security doubtable, e.g. a row-wise permutation and shifting of iris texture stripes in [47] or a permutation of 32×32 pixel blocks within 512×64 pixel textures in [15]. Within approaches to biometric salting, e.g. in [5], [32], subject-specific secrets are incorporated while experiments are performed under the non-stolen-secret scenario omitting the actual biometric performance of the system.

Focusing on iris biometric identification different mechanisms have been proposed in order to reduce the response time of the system. Biometric data does not have any natural sorting order, i.e. indexing databases represents a critical issue. In [17] a technique referred to as Beacon Guided Search is introduced. The algorithm is applied to a large-scale database of 632500 iris-codes enrolled in the United Arab Emirates (UAE), achieving a substantial improvement in search speed. However, computational efficiency comes at the cost of biometric performance, the same holds for other approaches [30], [36]. Based on the fact that entropy is not uniformly distributed across iris-codes [18], a compressed representation of the most reliable bits can be utilized to reduce the number of bit comparisons in a serial combination scenario. In [14] the generation of a short length iris-code (SLIC) is introduced which is applied in a two-stage identification [13], i.e. exhaustive $1 : n$ comparisons are performed based on the compressed template. By applying comparisons of original iris-codes only within a shortlist of most likely candidates the number of overall bit comparisons and the resulting response time is reduced. A pre-selection based on a rotation-invariant feature representation is presented in [25]. Since no circular bit-shifting is applied in the pre-selection step, the speed of identification is improved. In [39] an incremental comparison technique which successively compares the most reliable bits in iris-codes is applied to reduce bit comparisons.
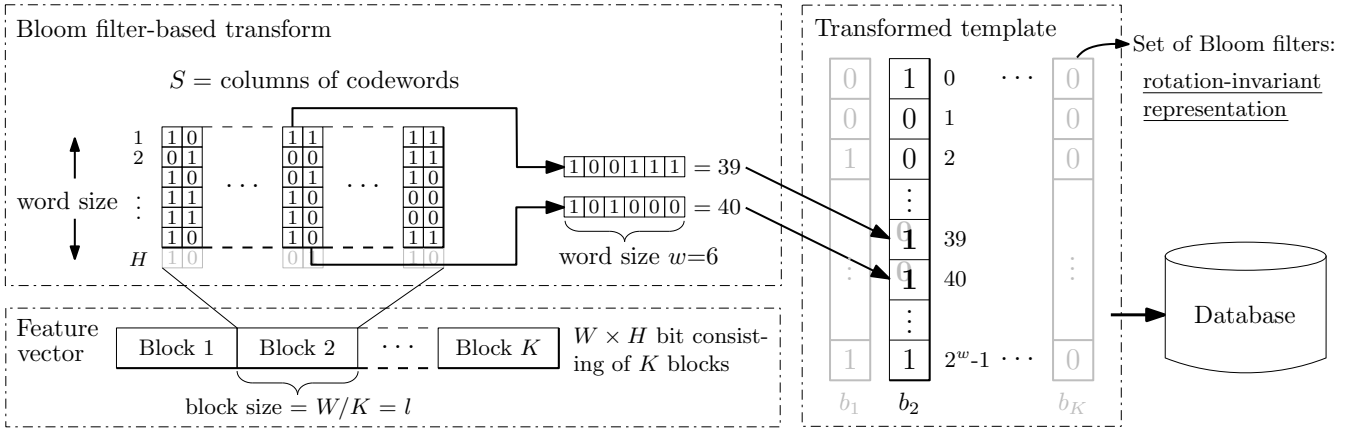
Fig. 1. Operation mode of the proposed rotation-invariant biometric templates applying Bloom filter-based transforms to feature vector columns. The highlighted codewords change in Bloom filter b_2 the element at index 39 (decimal representation of 100111) and also index 40 (decimal representation of 101000) to 1.

III. COMBINING BLOOM FILTERS AND IRIS RECOGNITION

Basically, a Bloom filter b is a bit array of length n, where initially all bits are set to 0. In order to represent a set S a Bloom filter traditionally utilizes k independent hash functions $h_1, h_2, ..., h_k$ with range $[0, n-1]$. For each element $x \in S$, bits at positions $h_i(x)$ of Bloom filter b are set to 1, for $1 \leq i \leq k$. A bit can be set to 1 multiple times, but only the first change has an effect. To test if an element y is in S, it has to be checked whether all position of $h_i(y)$ in b are set to 1. If this is the case, it is assumed that y is in S with a certain probability of false positive. If not, clearly y is not a member of S, hence, traditional Bloom filters are suitable for any application where a distinct probability of false positive is acceptable [1].

The original concept is adapted in different ways. Given a Bloom filter b of length n we restrict to inserting exactly l elements, where $l \leq n$. In case of uniformly distributed data the probability that a certain bit is set to 1 during the insertion of an element is $1/n$, i.e. the probability that a bit is still 0 is $1-1/n$. For inserting a total of l elements $1-(1-1/n)^l$ bits are expected to be set to 1. For $n = l \cdot c$ and $c \in \mathbb{N}$, i.e. n represents a multiple of l, $\lim_{n \to \infty}(1 - 1/n)^l = 1/e^{l/n}$. In addition, a trivial transform h is applied to each element $x \in S$ instead of multiple hash functions. Since feature elements are expected to be small the application of any hash function would not be resistant to brute force attacks.

In the following subsections the alignment-free adaptive Bloom filter-based transform and its properties with respect to template protection, biometric data compression and computationally efficient identification are described in detail.

A. Adaptive Bloom Filter-based Transform

In the proposed system adaptive Bloom filters are utilized in order to achieve an alignment-free representation of iris-codes. Generic iris recognition systems [2] extract binary feature vectors based on a row-wise analysis of normalized iris textures, i.e. iris-codes typically represent two-dimensional binary feature vectors of width W and height H (see Fig. 3 (e)-(f)). In the proposed scheme $W \times H$ iris-codes are divided into K blocks of equal size, where each column consists of $w \leq H$ bits. In case $w < H$ (e.g., for the purpose of compression), columns consist of the w upper most bits, i.e. features originating from outer iris bands, which are expected to contain less discriminitave information, are ignored. Subsequently, the entire sequence of columns of each block is successively transformed to according locations within adaptive Bloom filters, that is, a total number of K separate adaptive Bloom filters of length $n = 2^w$ form the template of size $K \cdot 2^w$. The transform is implemented by mapping each column within the 2D iris-code to the index of its decimal value, which is shown for two different codewords (=columns) as part of Fig. 1, for each column $x \in \{0, 1\}^w$, the mapping is defined as,

$$b[h(x)] = 1, \text{ with } h(x) = \sum_{j=0}^{w-1} x_j \cdot 2^j. \quad (1)$$

The very essence of the proposed transform is that it is alignment-free, i.e. generated templates (=sets of Bloom filters) do not need to be aligned at the time of comparison. Equal columns within certain blocks (=codewords) are mapped to identical indexes within adaptive Bloom filters, i.e. self-propagating errors caused by an inappropriate alignment of iris-codes are eliminated (radial neighborhoods persist). The rotation-compensating property of the proposed system comes at the cost of location information of iris-code columns. At block boundaries miss-alignment of iris-codes will distribute a certain amount of potentially matching codewords among different blocks, which would be mapped to neighbored Bloom filters. In experiments where ± 8 bit shifts are required to align iris-codes properly, miss-alignment did not affect biometric performance. In case larger rotation angles need to be anticipated, multiple columns of right and left neighbor-block can be mapped to the adaptive Bloom filter under construction in order to overcome this drawback.

B. Comparison in Transformed Domain

Typically, comparisons between binary biometric feature vectors are implemented by the simple XOR operator applied

to a pair of binary biometric feature vectors. The sum of all detected disagreements between any corresponding pairs of bits divided by the amount of compared bits yields the fractional Hamming distance (HD) as a measure of dissimilarity between pairs of binary biometric feature vectors [10]. Let $|b|$ denote the amount of bits within a Bloom filter b, which are set to 1. Then the dissimilarity DS between two Bloom filters b_i and b_j is defined as,

$$DS(b_i, b_j) = \frac{HD(b_i, b_j)}{|b_i| + |b_j|} \qquad |b_i| \neq 0, |b_j| \neq 0. \qquad (2)$$

If pairs of adaptive Bloom filters would be compared by merely estimating Hamming distances between these, mismatching bits between adaptive Bloom filters in which fewer bits are set to 1 would be weighted less and vice versa. Obviously, DS is computed as efficient as HD while DS does not have to be computed at numerous shifting positions. In order to incorporate masking bits obtained at the time of pre-processing, columns of iris-codes which are mostly affected by occlusions must not be mapped to adaptive Bloom filters, i.e. a seperate storage of bit masks is not required.

C. Template Protection

The Bloom filter-based transform conceals the original positions of codewords, i.e. given a Bloom filter b it is not clear from which column a distinct 1-bit in the generated protected template originated. In addition, it is most likely that diverse columns are mapped to a single index and the occurrence of distinct codewords can not be established from the stored template, i.e. the proposed transform achieves irreversible alignment-free templates, implementing cancelable biometrics. In order to provide unlinkability between multiple cancelable templates of a single subject an application specific secret T in form of a bit vector of length w, $T \in \{0, 1\}^w$, is incorporated. Each codeword is transformed applying this secret vector (of same length) by XORing both prior to mapping it to a Bloom filter. It is important to note that this secret is application-specific (and potentially subject specific) and is only incorporated as parameter in order to suffice the property of unlinkability required by the ISO/IEC IS 24745 [19]. Alternatively, different types of hash functions could be applied in different applications, or a single hash function could be parameterized based on an application specific seed (implementing MACs).

High correlation between codewords, especially neighboring ones, is expected. Consequently, a significant amount of codewords are mapped to identical positions in Bloom filters even for small values of l. Assume $|b|$ bits are set to 1 within a Bloom filter after inserting l codewords, i.e. $|b|$ different codewords occur in a block of length l. Hence, the amount of re-mapped bits is $1 - |b|/l$. For a potential attacker the reconstruction of the original iris-code block involves an arranging of $|b|$ codewords to l positions (K-times for the entire iris-code). For $|b| \leq l$ the theoretical amount of possible sequences is recursively defined by the function $f(|b|, l)$ where each of the $|b|$ codewords have to appear at least once within
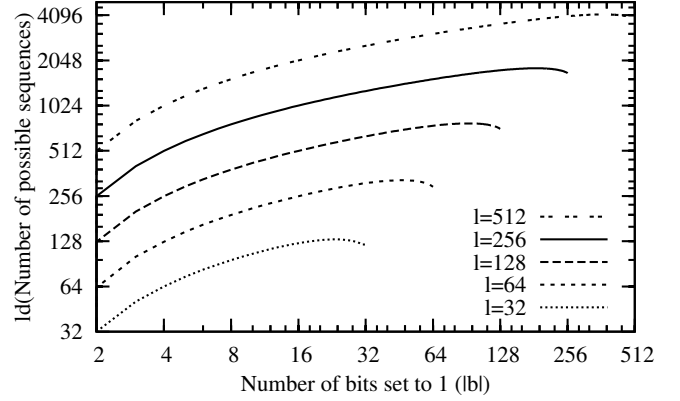

Fig. 2. Amount of possible sequences (per block) for different block sizes and proportions of re-mapped codewords.

l columns,

$$f(|b|, l) = \begin{cases} 1, & \text{if } |b| = 1, \\ |b|^l - \sum\limits_{i=1}^{|b|-1} \binom{|b|}{i} \cdot f(i, l) & \text{otherwise.} \end{cases} \qquad (3)$$

In other words, all sequences where less than $|b|$ codewords appear are subtracted from the number of all possible sequences, $|b|^l$. Fig. 2 illustrates the rapid increase of possible sequences even for small values of $|b|$ (note the logarithmic scales on both axis). Peaks are located around $3l/4$, in case of $l = |b|$ we get $f(l, l) = l!$ and $f(1, l) = 1$. For instance, for $l = 4$ and $|b| = 2$ we get $f(2, 4) = 2^4 - \binom{2}{1} \cdot f(1, 4) = 16 - 2 \cdot 1 = 14$ possible sequences, for $l = 4$ and $|b| = 3$ we get $f(3, 4) = 3^4 - \binom{3}{1} \cdot f(1, 4) - \binom{3}{2} \cdot f(2, 4) = 81 - 3 \cdot 1 - 3 \cdot 14 = 36$ possible sequences and for $l = 4$ and $|b| = 4$ we get $f(4, 4) = 4! = 24$ possible sequences and so forth. In experiments it will be demonstrated that for randomly generated bit vectors it is infeasible for potential attackers to cross-match pairs of protected templates extracted from a single subject.

D. Biometric Data Compression

The original template size is $W \times H$ bits. In the proposed scheme the template is divided into $W/l = K$ blocks of length l resulting in a template size of $2^w \cdot K = 2^w \cdot W/l$ where $w \leq H$. If we set $l = 2^q$ a compression is achieved if,

$$W/l \cdot 2^w < W \cdot H \Leftrightarrow 2^{w-q}/H < 1 \qquad (4)$$

applies, which is most likely the case as we will demonstrate in experiments. For instance, for an iris-code of size 2048 with $W = 256$ and $H = 8$, and the setting $l = 64$ and $w = 8$ we get $256/64 \cdot 2^8 = 1024 < 2048$, i.e. a compression down to 50% of the original size is achieved ($2^{8-6}/8 = 0.5$). Sizes of transformed templates are operated by setting parameters l and w. Both, increasing l and decreasing w reduces the overall size of the resulting template, see Eq. 4. Again the major advantage of the proposed transform is that compared to existing approaches to biometric template compression, e.g. [14], a comparison of compressed templates does not

require an optimal alignment within the presented scheme. It is important to note that algorithms may extract binary templates where distinct parts comprise features which should not be arranged in single columns, e.g. in [29] different parts of iris-codes represent real and complex values or minima and maxima extracted from different wavelet subbands.

E. Adaptive Bloom Filter-based Identification

Despite indexing techniques, original iris-codes have been combined with compressed and rotation-invariant templates in serial combination scenarios [13], [25]. For both types of attempts, compressed templates and alignment-free feature extractors have been found to exhibit unpractical biometric performance, requiring the application of a more sophisticated algorithm within a second stage. In contrast, as will be shown in experiments, the proposed Bloom filter-based transform generates rotation-invariant cancelable templates which maintain biometric performance.

If a biometric comparator is required to perform $\pm s$ bit shifts in each direction in order to compensate for head tilts the overall amount of bit comparisons increases to $W \cdot H \cdot (2s+1)$. This means for the proposed approach the number of required bit comparisons is reduced to,

$$100 \cdot 2^{w-q} / \left(H \cdot (2s+1) \right) \%. \tag{5}$$

For example, if a comparator performs ± 6 bit shifts and the proposed transform retains the template size (no compression) a reduction of bit comparisons down to $1/(12+1) \simeq 7.7\%$ is obtained, while no second algorithm is required. Again, the proposed system takes major advantage of its rotation-compensating property.

IV. Experimental Evaluations

Performance is estimated in terms of false non-match rate (FNMR) at a targeted false match rate (FMR), equal error rate (EER), and (true-positive) identification rate (IR). In accordance to the ISO/IEC IS 19795-1 [20] the FNMR of a biometric system defines the proportion of genuine attempt samples falsely declared not to match the template of the same characteristic from the same user supplying the sample. By analogy, the FMR defines the proportion of zero-effort impostor attempt samples falsely declared to match the compared non-self template. As score distributions overlap EERs are obtained, i.e. the system error rate where FNMR = FMR. The IR is the proportion of identification transactions by subjects enrolled in the system in which the subject's correct identifier is the one returned. In experiments identification is performed in the closed-set scenario returning the rank-1 candidate as identified subject (without applying a decision threshold).

A. Experimental Setup

Experiments are carried out using the CASIA-v3-Interval iris database[1] consisting of good quality NIR illuminated

(a) Acquisition   (b) Detection

(c) Iris texture

(d) Pre-processed iris texture

(e) Iris-code 1-D Log-Gabor filter
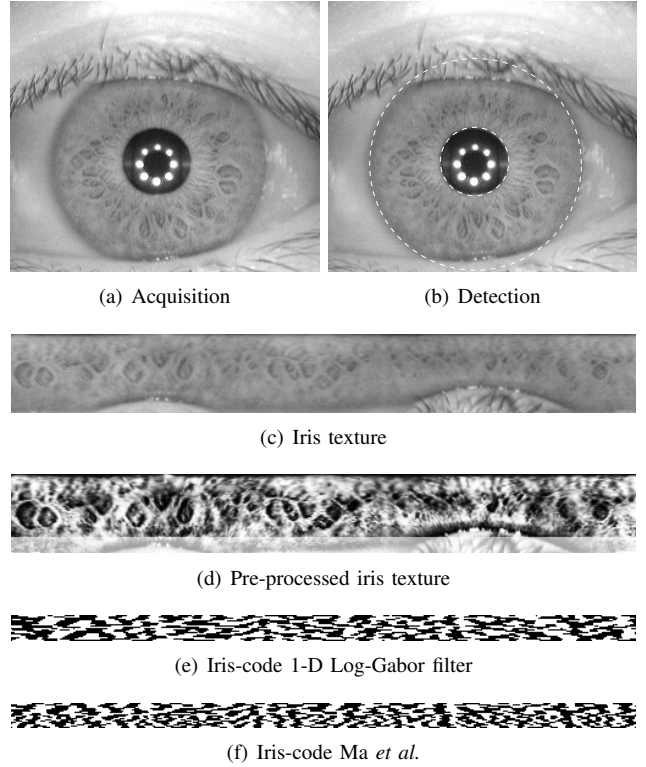

(f) Iris-code Ma *et al.*

Fig. 3. Preprocessing and both applied feature extraction algorithms.

TABLE I
Original systems: native performance rates (in %) for feature extractors with and without shifting (FNMRs are obtained at FMR=0.01%).

Alignment	1-D Log Gabor			Ma *et al.*		
	1-FNMR	EER	IR	1-FNMR	EER	IR
$\pm$ 8 bits	95.03	1.58	98.01	96.16	1.19	98.11
No shift	81.48	8.35	89.71	72.17	17.41	78.26

indoor images with 320×280 pixel resolution. The dataset comprises 2639 iris images of left and right eyes of 249 subjects resulting in a total number of 395 different classes. At pre-processing the iris of a given sample image is detected, un-wrapped to an enhanced rectangular texture of 512×64 pixel, shown in Fig. 3 (a)-(d) applying the weighted adaptive Hough algorithm proposed in [43]. The two-stage segmentation algorithm employs a weighted adaptive Hough transform iteratively refining a region of interest to find an initial center point, which is utilized to polar transform the image and extract polar and limbic boundary curves one after another from an (ellipso-)polar representation.

In the feature extraction stage custom implementations[2] of two different iris recognition algorithms are employed where normalized iris textures are divided into stripes to obtain 10 one-dimensional signals, each one averaged from the pixels of 5 adjacent rows (the upper 512×50 rows are analyzed). The first feature extraction method follows an implementation

by Masek [29] in which filters obtained from a Log-Gabor function are applied. Within this approach the texture is divided into 10 stripes to obtain 5 one-dimensional signals, each one averaged from the pixels of 5 adjacent rows, hence, the upper 512×50 pixel of preprocessed iris textures are analyzed. A row-wise convolution with a complex Log-Gabor filter is performed on the texture pixels. The phase angle of the resulting complex value for each pixel is discretized into 2 bits. The 2 bits of phase information are used to generate a binary code, which therefore is again $512 \times 20 = 10240$ bit. This algorithm is somewhat similar to Daugman's use of Log-Gabor filters, but it works only on rows as opposed to the 2-dimensional filters used by Daugman. The second feature extraction algorithm was proposed by Ma *et al.* [27]. Within this algorithm a dyadic wavelet transform is performed on 10 signals obtained from the according texture stripes, and two fixed subbands are selected from each transform resulting in a total number of 20 subbands. In each subband all local minima and maxima above an adequate threshold are located, and a bit-code alternating between 0 and 1 at each extreme point is extracted. Using 512 bits per signal, the final code is $512 \times 20 = 10240$ bit. Sample iris-codes generate by both feature extraction methods are shown in Fig. 3 (e)-(f).

A common way to estimate the average entropy ($\simeq$ amount of mutually independent bits) of biometric feature vectors is to measure the provided "degrees-of-freedom" which are defined by $d = p(1-p)/\sigma^2$, where p is the mean HD and σ^2 the corresponding variance between comparisons of different pairs of binary feature vectors. In case all bits of each binary feature vector of length z would be mutually independent, comparisons of pairs of different feature vectors would yield a binomial distribution, $\mathcal{B}(z,k) = \binom{z}{k}p^k(1-p)^{z-k} = \binom{z}{k}0.5^z$ and the expectation of the Hamming distance would be $1/z \cdot \mathbb{E}(X \oplus Y) = zp \cdot 1/z = p = 0.5$, where X and Y are two independent random variables in $\{0,1\}$. In reality p decreases to $0.5 - \epsilon$ while Hamming distances remain binomially distributed with a reduction in z in particular, $\mathcal{B}(d, 0.5)$ [46]. The 1D Log-Gabor feature extractor achieves a total of 592 degrees of freedom for a mean of 0.493 and an according standard deviation of 0.021. The algorithm of Ma *et al.* yields 1291 degrees of freedom for a mean of 0.498 and a standard deviation of 0.013.

Feature alignment represents an essential task at comparison. Table I summarizes the biometric performance of both feature extractors for ± 8 circular bit shifts and no bit shifting. Obviously, biometric performance is significantly improved if templates are aligned properly, where 8 circular bit shifts in each direction was found to be an adequate choice. As expected improved identification rates are obtained in the closed-set evaluation returning rank-1 candidates whitout considering any decision threshold. For both methods practical performance rates are obtained while the iris-code extracted by the algorithm of Ma *et al.* exhibits twice as much degrees of freedom compared to the feature extraction of Masek.

High correlation appears between neighboring columns of iris-codes, for both algorithms correlations in terms of 1-HD are plotted in Fig. 4 for more than 10 000 randomly chosen iris-code columns. As expected, directly neighboring columns
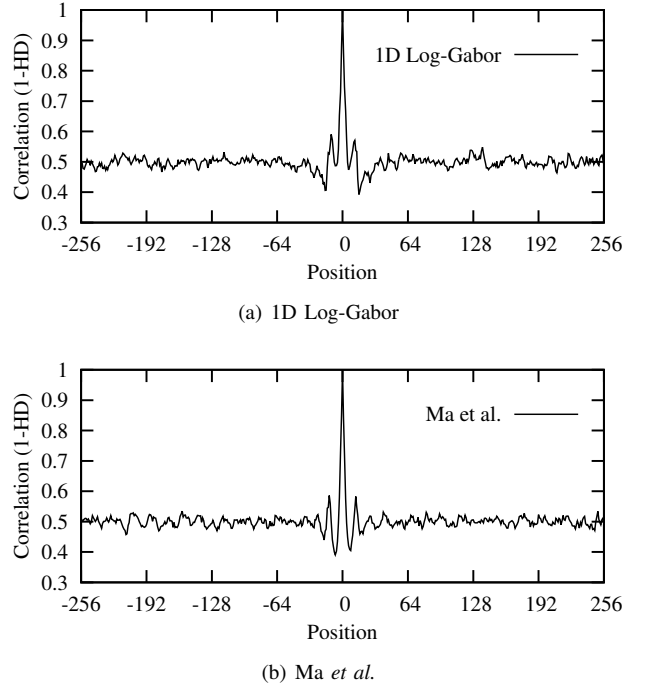

(a) 1D Log-Gabor

(b) Ma *et al.*

Fig. 4. Correlation (1-HD) between columns of iris-codes for both feature extraction algorithms.

exhibit high correlation since they originate from neighbouring pixel blocks in the iris texture which are not mutually independent. Columns with high correlation are surrounded by columns exhibiting rather low correlation, that is, from the estimated degrees-of-freedom an average iris-code extracted by the algorithm of Ma *et al.* corresponds to 1291 Bernoulli trials which means concatenated sequences of 0s and 1s exhibit an average length of $\simeq 8$ (10240/1291) bit. By analogy, for the 1D Log-Gabor feature extractor according sequences exhibit an average length of $\simeq 17$ (10240/592) bit (see Fig. 3(e)-(f)).

B. Performance Evaluation

Focusing on the applied feature extraction algorithms extracted iris-codes are divided in an upper 512×10 bit half and a lower 512×10 bit half as these represent real and complex values or minima and maxima extracted from different wavelet subbands, respectively. In case HDs are estimated based on column-wise codewords, i.e. a single error between two codewords defines a mis-match, FNMRs slightly increase with the size of codewords while EERs increase rather fast. 1-FNMRs and EERs for codeword sizes of $w = 8$ to $w = 10$ bits are summarized in Table II, where codewords start at the top of the upper and lower half (the original biometric performance corresponds to a word size of $w = 1$).

Table III - V summarize obtained 1-FNMRs, EERs, and IRs for different word sizes w and block sizes l for both feature extraction algorithms. From the obtained results it is clear that rotations of $\pm$ 8 bits, which significantly affect original systems, are compensated. The according receiver operation characteristic (ROC) curves are depicted in Fig. 5. Biometric performance is maintained or even improved for

TABLE II
ORIGINAL SYSTEMS: PERFORMANCE RATES (IN %) FOR BOTH FEATURE EXTRACTOR FOR HD-BASED COMPARISONS FOR DIFFERENT WORD SIZES (FNMRs ARE OBTAINED AT FMR=0.01%).

Word size	1-D Log Gabor		Ma et al.	
w (bits)	1-FNMR	EER	1-FNMR	EER
10	93.65	2.61	95.44	1.75
9	96.11	2.31	95.70	1.69
8	96.16	1.54	96.93	1.63

TABLE III
1-FNMRs (IN %) AT FMR=0.01% FOR DIFFERENT CONFIGURATIONS OF THE ADAPTIVE BLOOM FILTER-BASED TRANSFORM.

Algorithm	Word size w (bits)	Block size l (bits)				
		2^5	2^6	2^7	2^8	2^9
1-D Log Gabor	10	96.36	93.45	84.75	60.19	41.48
	9	95.90	92.07	81.73	60.15	41.22
	8	94.78	90.89	79.33	50.89	–
Ma et al.	10	97.95	95.08	86.70	75.29	52.68
	9	97.49	93.50	84.04	66.08	27.46
	8	96.52	92.17	75.39	31.40	–

TABLE IV
EERs (IN %) FOR DIFFERENT CONFIGURATIONS OF THE ADAPTIVE BLOOM FILTER-BASED TRANSFORM.

Algorithm	Word size w (bits)	Block size l (bits)				
		2^5	2^6	2^7	2^8	2^9
1-D Log Gabor	10	1.49	2.12	3.17	5.04	9.04
	9	1.67	2.15	3.32	6.10	12.17
	8	1.83	2.24	4.74	11.55	–
Ma et al.	10	1.14	1.72	2.64	5.08	8.99
	9	1.44	1.95	3.74	7.57	13.90
	8	1.47	2.51	4.79	11.97	–

TABLE V
IRs (IN %) FOR DIFFERENT CONFIGURATIONS OF THE ADAPTIVE BLOOM FILTER-BASED TRANSFORM.

Algorithm	Word size w (bits)	Block size l (bits)				
		2^5	2^6	2^7	2^8	2^9
1-D Log Gabor	10	98.01	97.34	95.14	90.53	78.61
	9	97.89	96.93	93.96	86.44	70.23
	8	97.45	95.93	92.12	75.60	–
Ma et al.	10	98.87	98.05	96.52	91.35	82.14
	9	98.56	97.54	94.62	85.93	68.08
	8	98.05	96.72	92.73	70.74	–

small block sizes, which support the previous claim that initial miss-alignments do not cause a drastic decrease in biometric performance. Again, performance is improved in identification mode. Throughout experiments best result were achieved for the maximum word size of 10 bit, i.e. $K = 10240/(32 \cdot 10) = 32$ blocks of $l = 32$ codewords which are mapped to 32 Bloom filters of size $n = 2^w = 2^{10}$. By applying the DS metric, which represents an improved biometric comparator, to pairs of Bloom filters accuracy is gained. For greater block sizes (e.g. $l = 2^8$) biometric performance decreases. While an increase of block sizes provides rotations-invariance for higher degrees of miss-alignment, it increases the chance that identical codewords occur within blocks, i.e. local information is lost leading to a greater overlap of intra- and inter-class score distributions.

C. Cancelable Templates

The security of the entire approach relies on the non-invertible mapping of codewords to a Bloom filter. W.l.o.g. this transform obscures the original position of the codeword as well as the number a codeword occurs, hence, for different configurations certain amounts of codewords are mapped to an identical position within according adaptive Bloom filters. Fig. 6 depicts the average percentage of re-mapped codewords and according standard deviations for both algorithms and applied configurations, according to Fig. 2 remapping $1 - |b|/l \simeq 1 - 3/4 = 25\%$ would be optimal in terms of security. In contrast to uniformly distributed data for configurations of $w = 9$ and $w = 10$ bit more codewords are re-mapped for smaller block sizes which is caused by correlation between iris-code columns. For $w = 8$ bit, the amount of re-mapped codewords increases with the block size l, i.e. more information is lost compared to larger values of

w. As a result, in general, biometric performance decreases with the word size w (see Table III–V), as opposed to the case where no information about codeword positions is lost (see Table II).

For best performing configurations (w.r.t. accuracy), mapping $l = 32$ codewords of length $w = 10$ to a $n = 2^{10}$ bit adaptive Bloom filters, for the 1D Log-Gabor feature extraction $\sim 48\%$ of codewords are re-mapped, $1 - |b|/l \simeq 0.48$ (see Fig. 6(a)). By analogy, for the algorithm of Ma et al. on average $\sim 32\%$ of codewords are re-mapped (see Fig. 6(b)). Focusing on the 1D Log-Gabor feature extractor, according to the previously estimated amount of possible sequences (see Fig. 2, $l = 32$) a potential attacker would have to try $\sim 2^{126}$ different sequences, $|b| = 32 \cdot (1 - 0.48) = 16.64$, for each of the $K = 32$ blocks. For the algorithm of Ma et al. the average amount of re-mapped codewords is even lower resulting in $\sim 2^{131}$ different sequences for $|b| = 32 \cdot (1 - 0.32) = 21.76$. By increasing block sizes security is significantly increased, e.g. for the feature extractor of Ma et al. a total number of $\sim 2^{283}$ possible sequences have to be tried per block in order to guess the original iris-code for $w = 10$ and $l = 64$, with $|b| = 32 \cdot (1 - 0.31) = 22.08$, while the system still reveals a practical EER of 1.72%. Obviously, a cancelable biometric system is operated through a natural trade-off between security and biometric performance.

Unlinkability, i.e. the infeasibility of cross-matching different protected templates of a single subject, represents a major issue of biometric template protection, however, experimental studies on unlinkability are commonly neglected [28]. According to the ISO/IEC IS 24745 on biometric information protection, unlinkability can be implemented by
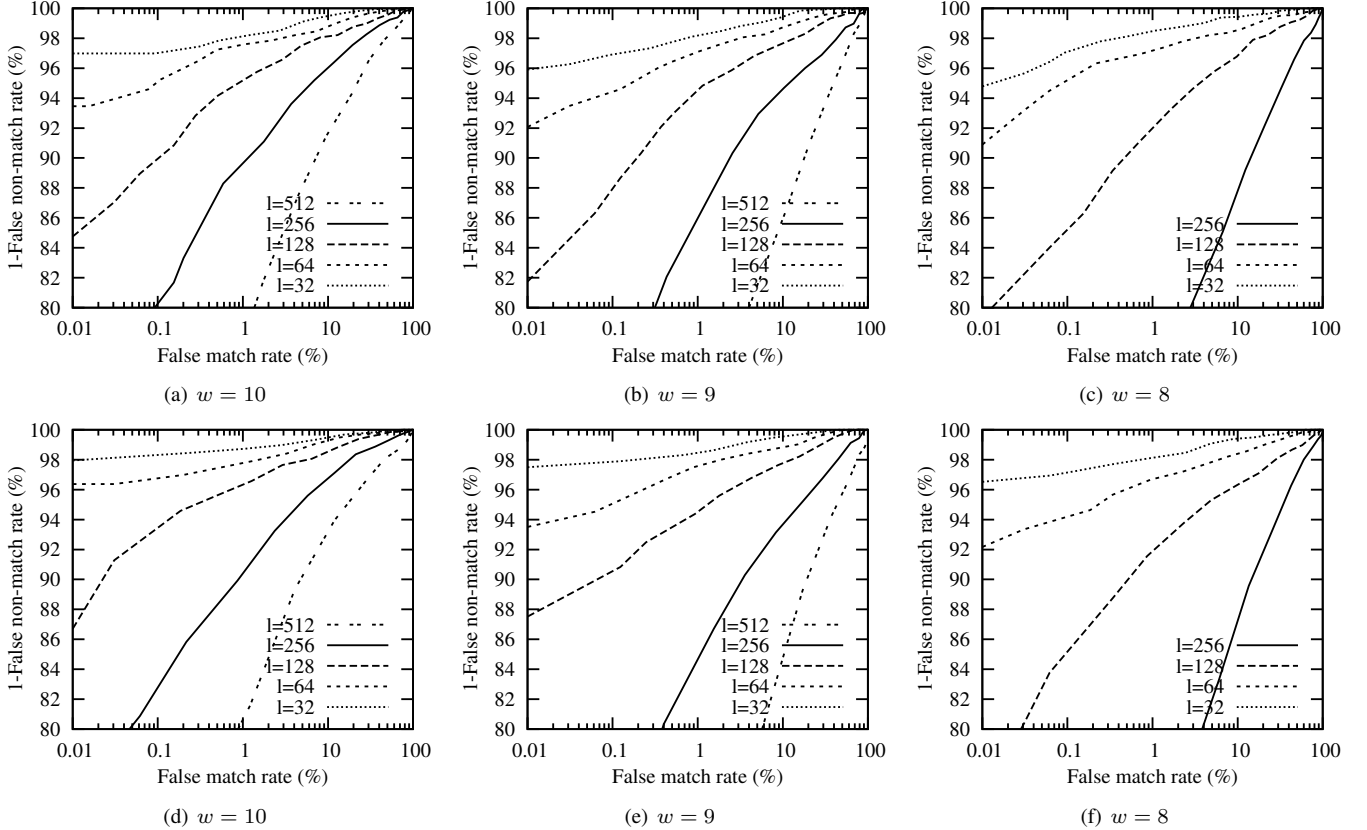
Fig. 5. ROC curves for the 1D Log-Gabor feature extractor (a)-(c) and the algorithm of Ma *et al.* (d)-(f) for different settings of block sizes and word sizes.

encrypting biometric references employing different (secret) keys, provided that the secret keys are managed appropriately to avoid collusion [19]. An application-specific secret bit vector is XORed with each iris-code column before applying adaptive Bloom filter-based transforms. To further enhance security with respect to unlinkability, this XOR-encryption based on a secret key could be substituted by an application of non-linear functions. In order to investigate the unlinkability of the presented approach we focus on the best performing configuration in terms of accuracy, i.e. $l = 32$ columns comprising $w = 10$ bits are successively mapped to according adaptive Bloom filters of size $n = 2^{10}$. Subsequently, obtained inter-class distributions are compared to distributions yielded by comparing adaptive Bloom filters originating from a single iris-code which are obscured by different bit vectors. Obtained score distributions are depicted in Fig. 7 for both algorithms, where unlinkability studies have been obtained from more than 10 000 genuine comparison with randomly chosen bit vectors. The comparison of different cancelable templates generated from a single iris-code does not allow cross-matching since resulting dissimilarity scores are generally higher than that of impostor comparisons within a single application.

D. Compressed Templates

Regarding resulting template sizes, which are depicted in Fig. 8 for most configurations $K \cdot 2^w < W \cdot H = 20 \cdot 2^9$ applies, which means a compression of the original template is achieved. Again, a trade-off is observed, between template

size and biometric performance. Smallest template sizes (10% of original size), for the configuration $w = 8$ and $l = 2^8$, result in rather un-practical performance rates of EERs >10%, while compressions down to 20% or 40% of the original size almost maintains accuracy, see Table III–V. Fig. 9 shows examples of resulting codes for both feature extraction algorithms. As can be seen for the algoritm of Ma *et al.* which provides more degrees of freedom, more bits are set to 1. Extracted codes appear suitable to be used in aforementioned application scenarios.

E. Identification Speed

In Fig. 10 the number of bit comparison of different configurations and the resulting IRs are compared to the original systems (requiring ±8 bit shifts). A significant reduction of bit comparisons (at least <25% of original system) is obtained for all settings of w and l while biometric performance is maintained for decreasing the bit comparisons down to ∼5%, which corresponds to a comparison of 2^{13} bits.

Identification is performed on a 2.30 GHz system. In order to minimize computational overhead caused by file access operations, all enrolled iris-codes are loaded a priori. A $1 : n$ comparison of a single iris-code of 10240 bits applying ±8 bit shifts takes on average 607 ms, which is defined as the baseline (=100%) of computational effort. Optimized C-based programs are able to compare more than one million iris-codes per second [11] which has been confirmed by published tests (e.g. NIST IREX-3). However, this difference is irrelevant
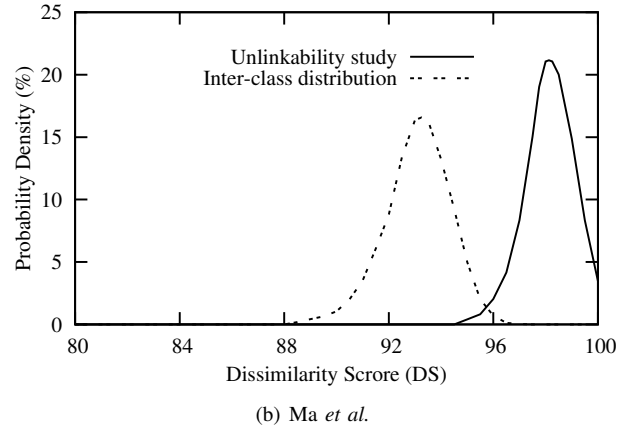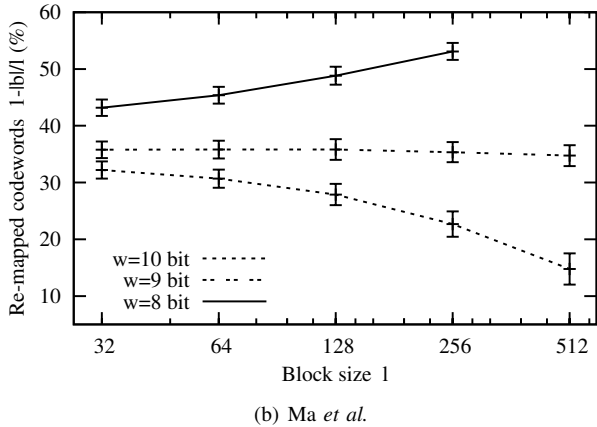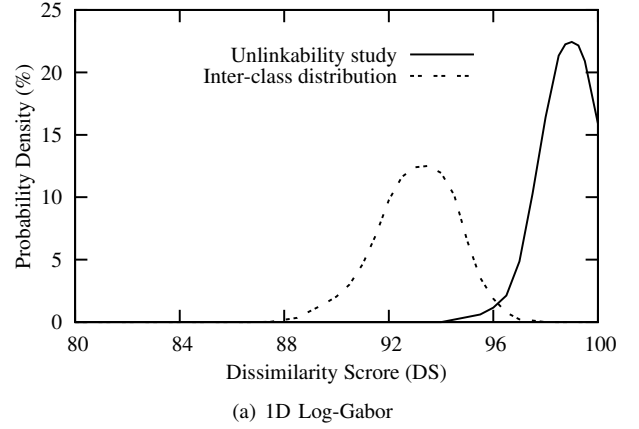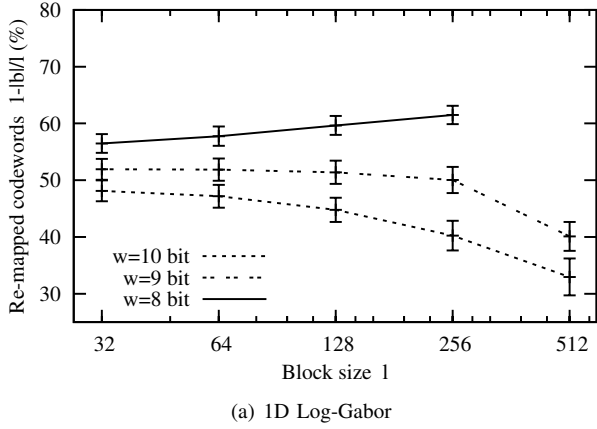
(a) 1D Log-Gabor

(a) 1D Log-Gabor

(b) Ma *et al.*

(b) Ma *et al.*

Fig. 6. Proportion of re-mapped codewords, $1\text{-}|b|/l$, for different block sizes l and word sizes w for both feature extractors.

Fig. 7. Score distributions for inter-class comparisons and according unlinkability tests for both feature extractors.

since we aim at comparing the two types of techniques based on the same platform and report speed-up in percent. Experimental results according to the average relative time required to identify a single subject compared to the original algorithms are summarized in Table VI. As expected, due to remaining inevitable computational overhead the obtained speed-up does not precisely relate to according template sizes, still, speed-up is substantial. For the best configuration with respect to biometric performance, $w = 10$ and $l = 2^5$, a four-fold speed-up is achieved ($607 \cdot 0.2357 = 143.07$ms). Up to a ten-fold speed-up comparable biometric performance is maintained, see Fig. 10. Furthermore, it is important to note that, while the applied database consits only of a few hundred subjects, in contrast to an indexing approach, the size of the applied dataset is irrelevant as well, since the proposed approach aims at achieving linear speed up requiring a $1 : n$ comparison.

V. CONCLUSIONS

The wide use of (iris) biometrics raises the need for privacy protection. Technologies of cancelable biometrics are designed to permanently secure biometric data, preventing from identity fraud and privacy violation. In addition, while a binary representation of biometric features enable a rapid comparison computational limits are reached deploying national-
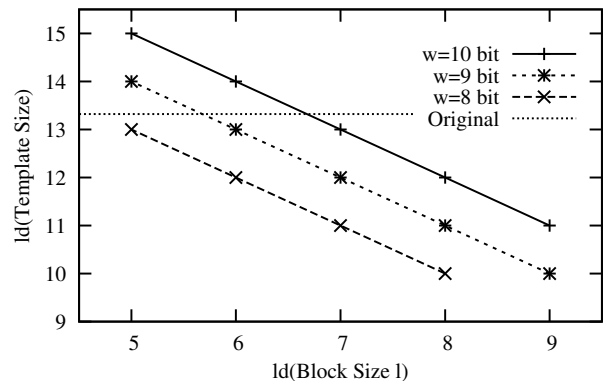

Fig. 8. Resulting template sizes for different block sizes l of the proposed system compared to the original algorithms.

sized biometric databases in identification mode and public deployments of iris recognition are still based on a brute force exhaustive search through a database. While the majority of approaches to biometric database indexing suffer from a significant decrease in biometric performance, indexing protected biometric templates represents an even greater challenge.

In this article alignment-free cancelable iris biometric templates based on adaptive Bloom filters are introduced. The generic adaptive Bloom filter-based transform which is applied
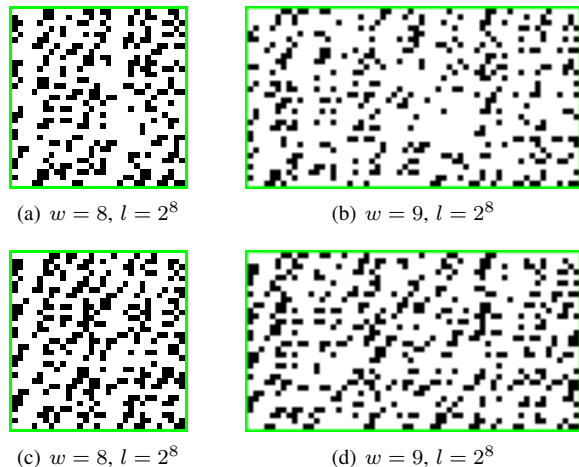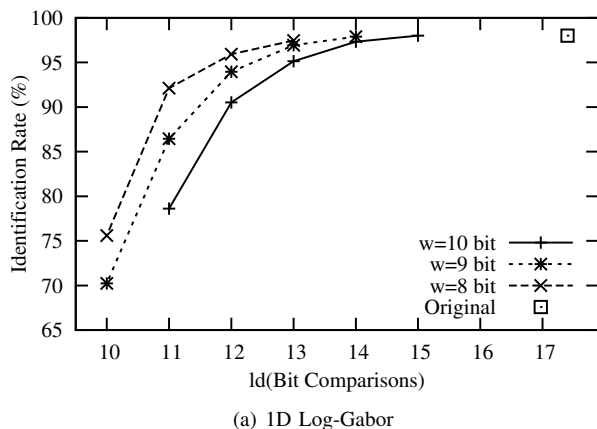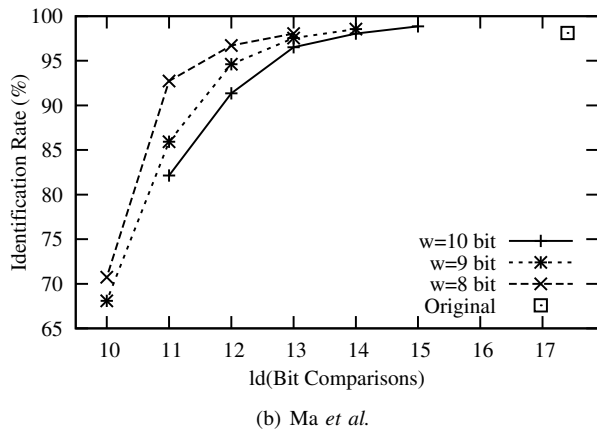
(a) $w = 8, l = 2^8$ (b) $w = 9, l = 2^8$

(c) $w = 8, l = 2^8$ (d) $w = 9, l = 2^8$

Fig. 9. Sample compressions for the 1D Log-Gabor (a)-(b) and the Ma *et al.* (c)-(d) feature vector of sample iris-codes in Fig. 3 (e)-(f) (256×4 and 256×8 codes have been rearranged for visualization).

(a) 1D Log-Gabor

(b) Ma *et al.*

Fig. 10. Amount of required bit comparisons and resulting IRs for different configurations of the adaptive Bloom filter-based transform.

TABLE VI
RELATIVE TIME (IN %) COMPARED TO THE TRADITIONAL APPROACH FOR DIFFERENT CONFIGURATIONS OF THE ADAPTIVE BLOOM FILTER-BASED TRANSFORM.

Word size w (bits)	Block size l (bits)				
	2^5	2^6	2^7	2^8	2^9
10	23.57	13.07	8.51	5.51	4.03
9	13.09	8.50	5.51	4.01	3.48
8	8.50	5.55	4.03	3.47	–

to binary feature vectors of different iris recognition algorithms enables (1) template protection, (2) a compression of biometric data, and (3) computational efficient biometric identification. Existing approaches to iris biometric template protection suffer from low biometric performance or utilize rather insecure alignment-preserving transforms. In contrast, the proposed rotation-invariant Bloom filter-based transform provides a high level of security while recognition accuracy is maintained. In addition, the presented scheme can be parameterized in order to highly compress biometric templates (down to 10% of original size). Furthermore, since bit-shifting is obsolete at the time of biometric comparison (in transformed domain) a substantial speed-up of biometric identification is achieved. Finally, it is important to note that the proposed approach can be utilized in order to generate a fixed-length protected template based on a variable-length binary biometric feature vector which may be the case for other biometric characteristics, e.g. fingerprints. To the authors' knowledge the proposed approach represents the very first template protection scheme which enables compression and computationally efficient identification.

ACKNOWLEDGMENT

REFERENCES

[1] B. Bloom. Space/time tradeoffs in hash coding with allowable errors. *Communications of the ACM*, 13(7):422–426, 1970.
[2] K. Bowyer, K. Hollingsworth, and P. Flynn. Image understanding for iris biometrics: A survey. *Computer Vision and Image Understanding*, 110(2):281–307, 2007.
[3] J. Bringer, H. Chabanne, G. Cohen, B. Kindarji, and G. Zemor. Theoretical and practical boundaries of binary secure sketches. *IEEE Transactions on Information Forensics and Security*, 3:673–683, 2008.
[4] A. Broder and M. Mitzenmacher. Network Applications of Bloom Filters: A Survey. *Internet Mathematics*, 1(4):485–509, 2005.
[5] S. C. Chong, A. T. B. Jin, and D. N. C. Ling. High security iris verification system based on random secret integration. *Computer Vision and Image Understanding*, 102(2):169–177, 2006.
[6] S. C. Chong, A. T. B. Jin, and D. N. C. Ling. Iris authentication using privatized advanced correlation filter. In D. Zhang and A. Jain, editors, *Proc. 1st Int'l Conf. on Biometrics*, volume 3832 of *LNCS*, pages 382–388. Springer, 2006.
[7] S. Cimato, M. Gamassi, V. Piuri, R. Sassi, and F. Scotti. Privacy in biometrics. In *Biometrics: fundamentals, theory, and systems*. Wiley, 2009.
[8] J. Daugman. High confidence visual recognition of persons by a test of statistical independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1148–1161, 1993.
[9] J. Daugman. The importance of being random: statistical principles of iris recognition. *Pattern Recognition*, 36(2):279–291, 2003.

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

[11] J. Daugman. Probing the uniqueness and randomness of iriscodes: Results from 200 billion iris pair comparisons. *Proc. of the IEEE*, 94(11):1927–1935, 2006.

[12] J. Daugman. Iris recognition at airports and border-crossings. In *Encyclopedia of Biometrics*. Springer, 2009.

[13] J. E. Gentile, N. Ratha, and J. Connell. An efficient, two-stage iris recognition system. In *Proc. IEEE 3rd Int'l Conf. on Biometrics: Theory, Applications, and Systems*, pages 1–5. IEEE, 2009.

[14] J. E. Gentile, N. Ratha, and J. Connell. Slic: Short-length iris codes. In *Proc. IEEE 3rd Int'l Conf. on Biometrics: Theory, Applications, and Systems*, pages 1–5. IEEE, 2009.

[15] J. Hämmerle-Uhl, E. Pschernig, and A. Uhl. Cancelable iris biometrics using block re-mapping and image warping. In P. Samarati, M. Yung, F. Martinelli, and C. Ardagna, editors, *Proc. 12th Int'l Information Security Conf.*, volume 5735 of *LNCS*, pages 135–142. Springer, 2009.

[16] F. Hao, R. Anderson, and J. Daugman. Combining Cryptography with Biometrics Effectively. *IEEE Transactions on Computers*, 55(9):1081–1088, 2006.

[17] F. Hao, J. Daugman, and P. Zielinski. A fast search algorithm for a large fuzzy database. *IEEE Transactions on Information Forensics and Security*, 3(2):203–212, 2008.

[18] K. P. Hollingsworth, K. W. Bowyer, and P. J. Flynn. The best bits in an iris code. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(6):964–973, 2009.

[19] ISO/IEC JTC1 SC27 Security Techniques. *ISO/IEC 24745:2011. Information Technology - Security Techniques - Biometric Information Protection*. International Organization for Standardization, 2011.

[20] ISO/IEC TC JTC1 SC37 Biometrics. *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, Mar. 2006.

[21] A. K. Jain, K. Nandakumar, and A. Nagar. Biometric template security. *EURASIP Journal Adv. Signal Process*, 2008:1–17, 2008.

[22] A. Juels and M. Sudan. A fuzzy vault scheme. In *Proc. IEEE Int'l Symp. on Information Theory*, page 408. IEEE, 2002.

[23] A. Juels and M. Wattenberg. A fuzzy commitment scheme. In *Proc. 6th ACM Conf. on Computer and Communications Security*, pages 28–36. ACM, 1999.

[24] A. Kong, K.-H. Cheunga, D. Zhanga, M. Kamelb, and J. Youa. An analysis of BioHashing and its variants. *Pattern Recognition*, 39(7):1359–1368, 2006.

[25] M. Konrad, H. Stögner, A. Uhl, and P. Wild. Computationally efficient serial combination of rotation-invariant and rotation compensating iris recognition algorithms. In *Proc. 5th Int'l Conf. on Computer Vision Theory and Applications*, volume 1, pages 85–90. INSTICC, 2010.

[26] C. Lee, J. Choi, K. Toh, S. Lee, and J. Kim. Alignment-free cancelable fingerprint templates based on local minutiae information. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 37(4):980–992, 2007.

[27] L. Ma, T. Tan, Y. Wang, and D. Zhang. Efficient iris recognition by characterizing key local variations. *IEEE Transactions on Image Processing*, 13(6):739–750, 2004.

[28] E. Maiorana. Biometric cryptosystem using function based on-line signature recognition. *Expert Systems with Applications*, 37(4):3454–3461, 2010.

[29] L. Masek. Recognition of human iris patterns for biometric identification. Master's thesis, University of Western Australia, 2003.

[30] R. Mukherjee and A. Ross. Indexing iris images. In *Proc. 19th Int'l Conf. on Pattern Recognition (ICPR'08)*, pages 1–4, 2008.

[31] J. Mullin. Optimal semijoins for distributed database systems. *IEEE Transactions on Software Engineering*, 16(5):558 –560, may 1990.

[32] O. Ouda, N. Tsumura, and T. Nakaguchi. Tokenless cancelable biometrics scheme for protecting iris codes. In *Proc. 20th Int'l Conf. on Pattern Recognition*, pages 882–885. IEEE, 2010.

[33] J. K. Pillai, V. M. Patel, R. Chellappa, and N. K. Ratha. Secure and robust iris recognition using random projections and sparse representations. *IEEE Transactions on Pattern Analysis and Machine Intellegence*, 33(9):1877–1893, 2011.

[34] N. Ratha, J. Connell, and R. Bolle. Enhancing security and privacy in biometrics-based authentication systems. *IBM Systems Journal*, 40(3):614–634, 2001.

[35] C. Rathgeb, F. Breitinger, and C. Busch. Alignment-free cancelable iris biometric templates based on adaptive bloom filters. In *In Proc. of the 6th IAPR Int'l Conf. on Biometrics (ICB'13)*, pages 1–8, 2013.

[36] C. Rathgeb and A. Uhl. Iris-biometric hash generation for biometric database indexing. In *Proc. 20th Int'l Conf. on Pattern Recognition*, pages 2848–2851. IEEE, 2010.

[37] C. Rathgeb and A. Uhl. A survey on biometric cryptosystems and cancelable biometrics. *EURASIP Journal on Information Security*, 2011(3), 2011.

[38] C. Rathgeb and A. Uhl. Statistical attack against fuzzy commitment scheme. *IET Biometrics*, 1(2), 2012.

[39] C. Rathgeb, A. Uhl, and P. Wild. Incremental iris recognition: A single-algorithm serial fusion strategy to optimize time complexity. In *Proc. IEEE 4th Int'l Conf. on Biometrics: Theory, Applications, and Systems*, pages 1–6. IEEE, 2010.

[40] C. Rathgeb, A. Uhl, and P. Wild. *Iris Biometrics: From Segmentation to Template Security*. Number 59 in Advances in Information Security. Springer, 2012.

[41] A. Ross. Iris recognition: The path forward. *Computer*, 43:30–35, 2010.

[42] W. Scheirer and T. Boult. Cracking Fuzzy Vaults and Biometric Encryption. In *Proc. Biometrics Symposium*, pages 1–6. IEEE, 2007.

[43] A. Uhl and P. Wild. Weighted adaptive hough and ellipsopolar transforms for real-time iris segmentation. In *Proc. 5th Int'l Conf. on Biometrics*, pages 1–8, 2012.

[44] Unique Identification Authority of India. Aadhaar: http://uidai.gov.in/. retrieved March, 2013.

[45] S. Venugopalan and M. Savvides. How to generate spoofed irises from an iris code template. *IEEE Transactions on Information Forensics and Security*, 6(2):385–395, 2011.

[46] R. Viveros, K. Balasubramanian, and N. Balakrishnan. Binomial and negative binomial analogues under correlated bernoulli trials. *The American Statistician*, 48(3):243–247, 1984.

[47] J. Zuo, N. K. Ratha, and J. H. Connel. Cancelable iris biometric. *Proc. 19th Int'l Conf. on Pattern Recognition*, pages 1–4, 2008.

APPENDIX

Proof *(by induction).* for all $|b|, l \in \mathbb{N}$, $l \geq |b| > 1$, the theoretical amount of possible sequences is $f(|b|, l)$, where $|b|$ codewords have to appear at least once within l columns,

$$f(|b|, l) = |b|^l - \sum_{i=1}^{|b|-1} \binom{|b|}{i} \cdot f(i, l). \tag{6}$$

Base case: $f(1, l) = 1$, and for $|b| = 2$, the number of possible sequences is $2^l - 2$, i.e. all possible sequences minus the two sequences where only one codeword occurs,

$$f(2, l) = 2^l - \sum_{i=1}^{1} \binom{2}{i} \cdot f(i, l) = 2^l - \binom{2}{1} \cdot f(1, l) = 2^l - 2.$$

(6) is true for the base case, $|b| = 2$.

Induction step: $|b| \to |b| + 1$, suppose (6) is true for $|b|$. For $|b| + 1$ the number of all possible sequences is $(|b| + 1)^l$, the subtracted number of possible i-element subsets are now of a set containing $|b| + 1$ elements, and sequences comprising $|b|$ codewords are subtracted. We get,

$$
\begin{aligned}
f(|b| + 1, l) = & \ (|b| + 1)^l - \binom{|b| + 1}{|b|} \cdot f(|b|, l) \\
& - \sum_{i=1}^{|b|-1} \binom{|b| + 1}{i} \cdot f(i, l) \\
= & \ (|b| + 1)^l - \binom{|b| + 1}{|b| + 1 - |b|} \cdot f(|b|, l) \\
& - \sum_{i=1}^{|b|-1} \binom{|b| + 1}{i} \cdot f(i, l) \\
= & \ (|b| + 1)^l - (|b| + 1) \cdot f(|b|, l) \\
& - \sum_{i=1}^{|b|-1} \binom{|b| + 1}{i} \cdot f(i, l) \\
= & \ (|b| + 1)^l - \sum_{i=1}^{|b|} \binom{|b| + 1}{i} \cdot f(i, l)
\end{aligned}
$$

Conclusion: by the principle of induction, (6) is true for all $|b|, l \in \mathbb{N}$, $l \geq |b| > 1$.

$\square$