Alignment-Free Cancelable Iris Biometric Templates based on Adaptive Bloom Filters

C. Rathgeb, F. Breitinger and C. Busch
da/sec Biometrics and Internet Security Research Group
University of Applied Sciences Darmstadt, Germany
{christian.rathgeb,frank.breitinger,christoph.busch}@h-da.de

Abstract

Biometric characteristics are largely immutable, i.e. *unprotected storage of biometric data provokes serious privacy threats,* e.g. *identity theft, limited re-newability, or cross-matching. In accordance with the ISO/IEC 24745 standard, technologies of cancelable biometrics offer solutions to biometric information protection by obscuring biometric signal in a non-invertible manner, while biometric comparisons are still feasible in the transformed domain.*

In the presented work alignment-free cancelable iris biometrics based on adaptive Bloom filters are proposed. Bloom filter-based representations of binary biometric templates (iris-codes) enable an efficient alignment-invariant biometric comparison while a successive mapping of parts of a binary biometric template to a Bloom filter represents an irreversible transform. In experiments, which are carried out on the CASIA-v3 iris database, it is demonstrated that the proposed system maintains biometric performance for diverse iris recognition algorithms, protecting biometric templates at high security levels.

1. Introduction

Cancelable biometrics [15] consist of intentional, repeatable distortions of biometric signals based on transforms which provide a comparison of biometric templates in the transformed domain. Technologies of cancelable biometrics are commonly categorized as non-invertible transforms and biometric salting [16]. Transforms are designed to meet two major requirements of biometric information protection (ISO/IEC 24745). (1) Irreversibility: knowledge of the protected template can not be used to determine any information about the original biometric sample, while it should be easy to generate the protected template; (2) Unlinkability: different versions of protected biometric templates can be generated based on the same biometric data (renewability), while protected templates should not allow cross-matching (diversity). Meeting these requirements the vast majority of published approaches to cancelable biometrics report a significant decrease in recognition performance. This inevitable degradation of biometric performance is caused by two major issues: in most cases (1) local neighborhoods of feature elements are obscured and (2) transformed enrollment templates are not "seen", *i.e*. alignment can not be performed properly at the time of comparison [16].

The contribution of this work is the proposal of a generic approach to cancelable iris biometrics 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 storage and rapid processing of queries, Bloom filters convince by their wide field of applications, *e.g*. database applications [12] or network applications [3]. In the presented work Bloom filters are utilized to map binary biometric feature vectors to an alignment-free transformed domain. Eliminating the issue of feature alignment caused by cancelable transforms, a mapping of biometric feature elements to a Bloom filter is non-invertible, *i.e*. a re-construction of the original biometric template from stored Bloom filters is not feasible. In addition, unlinkability is provided by incorporating application-specific secrets, *i.e*. seed values. To the author's knowledge the proposed work represents the very first application of Bloom filter in order to achieve alignment-free cancelable biometric templates.

This paper is organized as follows: Sect. 2 reviews related works on cancelable biometrics in particular, based on iris. In Sect. 3 the concept of Bloom filters is introduced and the proposed system is described in detail. Experimental evaluations are presented in Sect. 4. In Sect. 5 conclusions are drawn and future work is discussed.

2. Related Work

Ratha *et al*. [15] 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 [20] the authors propose different techniques to generate cancelable iris biometrics based on non-invertible transforms and biometric salting, which are applied in image

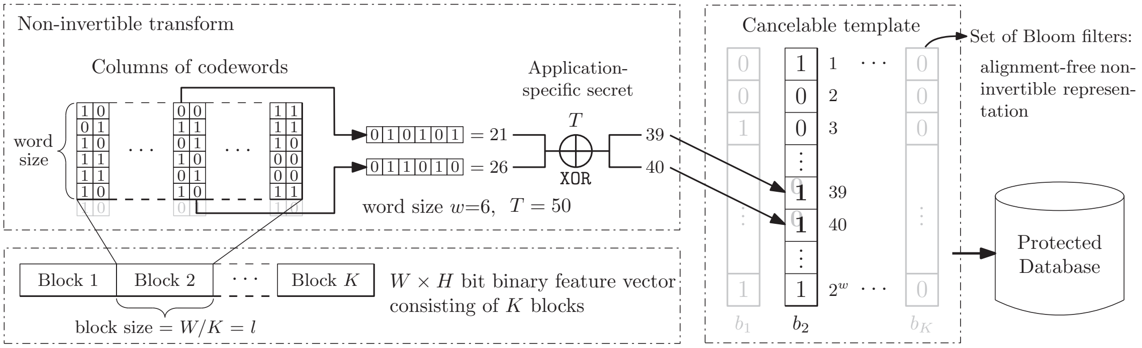


Figure 1. Basic operation mode of the proposed alignment-free cancelable iris biometric system based on adaptive Bloom filter.

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 [6] 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. In [13] 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 [8]. 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 [14] 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 [4] 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 [5] to create concealed feature vectors.

The vast majority of non-invertible transforms only maintains biometric performance for settings which leave security doubtable, *e.g.* a row-wise permutation and shifting of iris texture stripes in [20] or a permutation of 32×32 pixel blocks within 512×64 pixel textures in [6]. Within approaches to biometric salting, *e.g.* in [4, 13], subject-specific secrets are incorporated while experiments are performed under to non-stolen-secret scenario omitting the actual biometric performance of the system.

3. System Architecture

The key components of the proposed system which is depicted in Fig. 1 are described in detail as follows:

3.1. Bloom Filter-based Feature Transform

Basically, a Bloom filter b is a simple bit array of length n, where initially all bits are set to 0. In order to represent a set $S = \{x_1, x_2, ..., x_m\}$ 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 $h_i(x)$ of Bloom filter b are set to 1, for $1 \leq i \leq k$. An index 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 filter are suitable for any application where a distinct probability of false positive is acceptable.

In the proposed system the original concept of Bloom filters is adapted in order to achieve alignment-free cancelable iris biometrics. Generic iris recognition systems [2] extract binary feature vectors (iris-codes) which are generated from 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. 4 (e)-(f)). In the proposed scheme $W \times H$ iris-codes are divided into K blocks of equal size. Subsequently, the entire sequence of columns, where each column consists of w bits, of each block is successively transformed to according locations within Bloom filters, that is, a total number of K separate Bloom filters of length $n = 2^w$ form the protected 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. Thereby the original positions of codewords is concealed, *i.e.* given a Bloom filter b it is not clear from which column a distinct 1-bit in the protected template originated. In addition, it is most

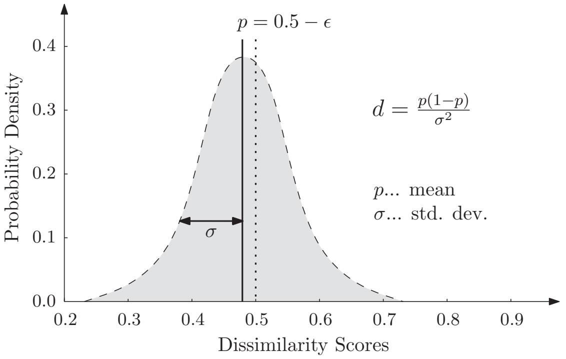


Figure 2. Binomial distribution of Hamming distance scores between different pairs of binary biometric feature vectors.

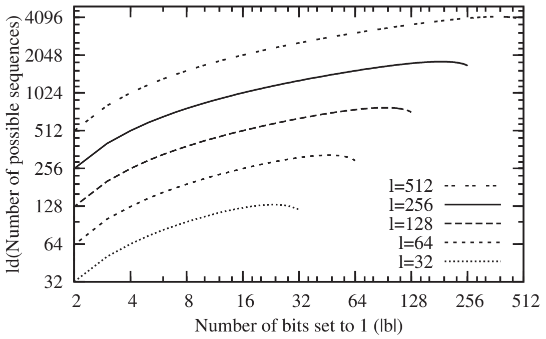


Figure 3. Amount of possible sequences (per block) for different block sizes and proportions of re-mapped codewords.

likely that diverse columns are mapped to a single index and the occurrence of distinct codewords can not be established from the protected template. 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 is incorporated, which is XORed with each codeword prior to mapping it to a Bloom filter, as depicted in Fig. 1. It is important to note that this secret is application-specific (and potentially subject specific) and does not serve any security purposes. Irreversibility and unlinkability will be discussed in detail in Sect. 3.3.

The very essence of the proposed transform is that it is alignment-free, *i.e.* cancelable templates (=sets of Bloom filters) do not need to be aligned at the time of comparison. In traditional iris recognition systems, detected iris textures (in form of a ring) are transformed to polar coordinates. As a consequence a circular row-wise shifting of iris-codes corresponds to a rotation of the eye (see Fig. 4 (b)-(c)). In the presented approach equal columns within certain blocks (=codewords) are mapped to identical indexes within 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. A drastic initial miss-alignment of iris-codes would distribute a large amount of potentially matching codewords among different blocks which would be mapped to different Bloom filters. In experiments, which are performed on a database acquired in a rather constrained environment, this potential drawback did not affect biometric performance.

3.2. 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. 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,

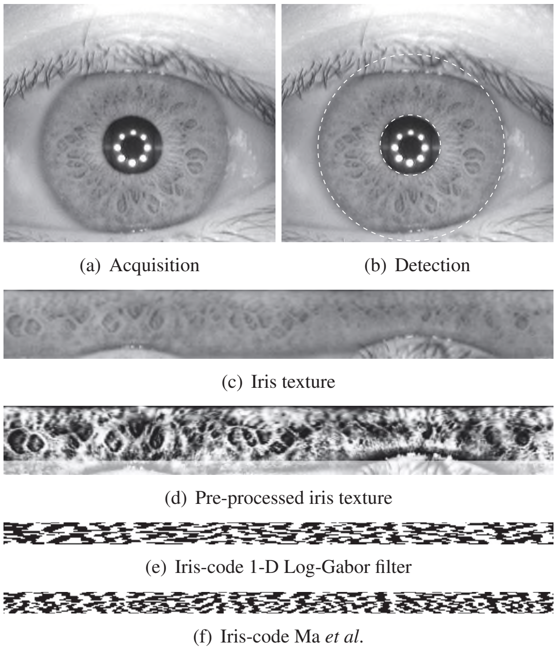
$$DS(b_i, b_j) = \frac{HD(b_i, b_j)}{|b_i| + |b_j|}. \tag{1}$$

If pairs of Bloom filters would be compared by merely estimating Hamming distances between these, mismatching bits between Bloom filters in which fewer bits are set to 1 would be weighted less and vice versa. Obviously, DSs are computed as efficient as HDs while DSs do not have to be computed at numerous shifting positions. Note that no masking bits are taken into account.

3.3. Irreversibility and Unlinkability

Within the presented scheme irreversibility is achieved by mapping column-wise codewords to Bloom filters. Given a Bloom filter b of length n we restrict to inserting only l codewords, where $l \leq n$ (blocks do not contain more than n columns). 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, $\lim_{n\to\infty}(1-1/n)^l = 1/\mathrm{e}^{l/n}$. Focusing on biometric data this theoretical expectation does not apply, since bits of binary biometric feature vectors must not be expected to be mutually independent, *i.e.* reasonable parts of feature vectors correlate.

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, shown in Fig. 2. In case all bits of each binary feature vector



(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.*

Figure 4. Preprocessing and applied feature extraction algorithms.

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, $\mathcal{B}(d, 0.5)$ [19].

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 l. Hence, the probability of re-mapping a bit to a certain position is $1 - |b|/l$. For a potential attacker the reconstruction of the original iris-code part involves an arranging of $|b|$ codewords to l positions. For $|b| \leq l$ the theoretical amount of possible sequences is recursively defined by $f(|b|, l)$ where each of the $|b|$ codewords have to appear at least once within l columns,

$$f(|b|,l) = \begin{cases} 1, & \text{if } |b| = 1\ , \\ |b|^l - \sum\limits_{i=1}^{|b|-1} \binom{|b|}{i} \cdot f(i,l) & \text{otherwise.} \end{cases} \tag{2}$$

In other words, all sequences where less than $|b|$ codewords appear are subtracted from the number of all possible sequences, $|b|^l$. Fig. 3 illustrates the rapid increase of possible sequences even for small values of $|b|$ (note the logarithmic scales of 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.

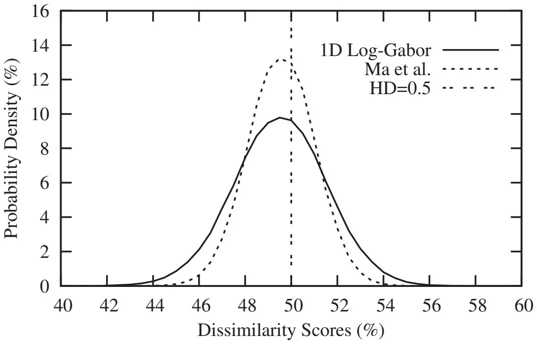


Figure 5. Binomial distribution of HDs of inter-class comparisons for both feature extraction algorithms.

Unlinkability is provided by incorporating an application specific bit vector, denoted by $T \in \{0,1\}^w$, which is XORed with processed codewords prior to mapping these to Bloom filters. Alternatively, different types of hash functions could be applied in different applications, or a single hash function could be applied based on an application specific seed. 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.

4. Experimental Studies

Performance is estimated in terms of false non-match rate (FNMR) at a targeted false match rate (FMR) and equal error rate (EER). 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.

4.1. Experimental Setup

Experiments are carried out using the CASIA-v3-Interval iris database[1]. In experiments only left-eye images (1332 instances) are evaluated. At pre-processing the iris

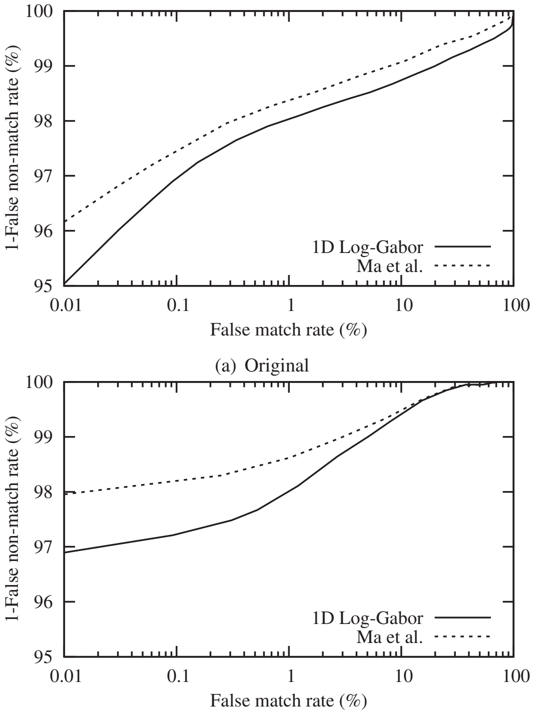
[1] The Center of Biometrics and Security Research, `http://www.idealtest.org`

Table 1. Performance rates (in %) for original feature extractors with and without shifting (FNMRs are obtained at FMR=0.01%).

Alignment	1-D Log Gabor		Ma *et al.*	
	1-FNMR	EER	1-FNMR	EER
± 8 bit shifts	95.03	1.58	96.16	1.19
No bit shift	81.48	8.35	72.17	17.41

of a given sample image is detected, un-wrapped to an enhanced rectangular texture of 512×64 pixel, shown in Fig. 4 (a)-(d) applying the weighted adaptive Hough algorithm proposed in [18]. 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 [11] 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.* [9]. 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. 4 (e)-(f).



(a) Original

(b) Bloom filter-based

Figure 6. ROC curves of the original biometric system and the best performing cancelable systems for both feature extractors.

The binomial distribution of Hamming distances between different pairs of binary biometric feature vectors for both algorithms is plotted in Fig. 5. 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. As previously mentioned feature alignment represents an essential task during comparison. Table 1 summarizes the biometric performance of both feature extractors in terms of 1-FNMRs at an FMR of 0.01% and EERs 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. It is expected that the proposed system will take major advantage of its rotation-compensating property. The according receiver operation characteristic (ROC) curves are depicted in Fig. 6 (a). For both methods practical performance rates are obtained while the iris-code extracted by the algorithm of Ma *et al.* exhibit twice as much degrees of freedom compared to the feature extraction of Masek.

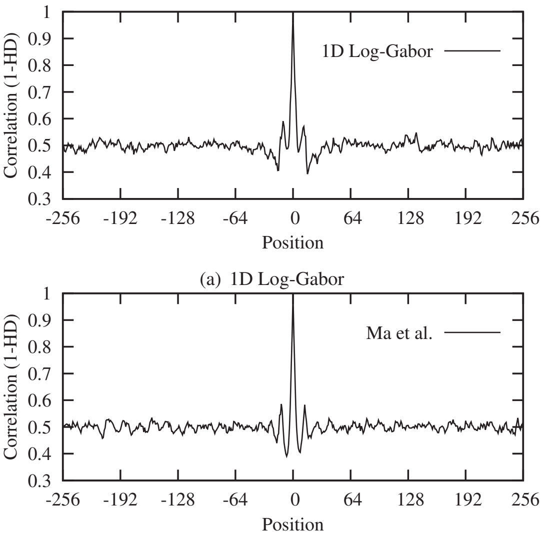
4.2. Performance Evaluation

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 dif-

[2] USIT – University of Salzburg Iris Toolkit v1.0, `http://www.wavelab.at/sources/`

Table 2. Performance rates (in %) for both feature extractor 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



(a) 1D Log-Gabor

(b) Ma *et al.*

Figure 7. Correlation (1-HD) between columns of iris-codes for both feature extraction algorithms.

Table 3. Performance rates (1-FNMRs at FMR=0.01 in %) for both feature extractors and different configurations of the system.

Algorithm	Word size	Block size l (bits)				
	w (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 4. EERs (in %) for both feature extractors and different configuration of the cancelable system.

Algorithm	Word size	Block size l (bits)				
	w (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	–

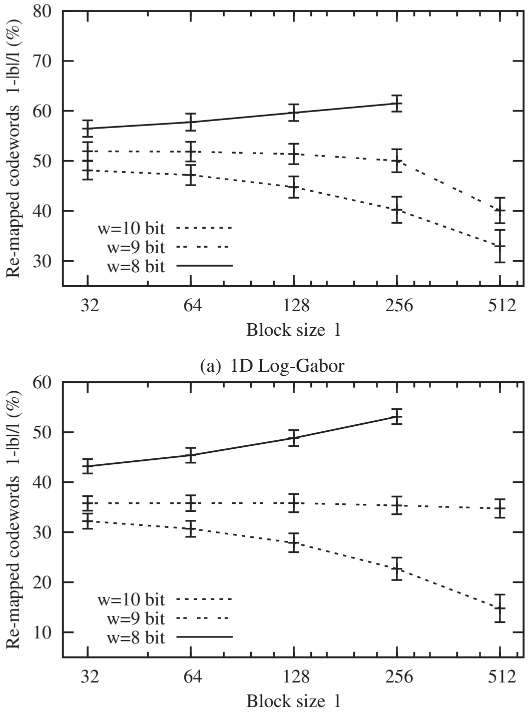
ferent wavelet subbands, respectively. In case HDs are estimated based on column-wise codewords, *i.e.* a single error between two codewords defines a mis-match, FNMRs slightly increase with the size of codewords while EERs increase rather fast. 1-FNMRs and EERs for codeword sizes of $w = 8$ to $w = 10$ bits are summarized in Table 2, where codewords start at the top of the upper and lower half and the original biometric performance corresponds to a word size of $w = 1$.

As already stated high correlation appears between neighboring columns of iris-codes. For both algorithms correlations in terms of 1-HD are plotted in Fig. 7 for more than 10 000 randomly chosen iris-code columns. As expected directly neighboring columns exhibit high correlation since bits in iris-codes 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 concatenated sequences of 0s and 1s exhibit an average length of $\simeq$17 (10240/592) bit (cf. Fig. 4(e)-(f)).

Table 3 and Table 4 summarize obtained 1-FNMRs and according EERs for different word sizes w and block sizes l for both feature extraction algorithms. Biometric performance is maintained or even improved for small block sizes, which support the previous claim that initial miss-alignments do not cause a drastic decrease in biometric performance. Throughout experiments best result were achieved for the maximum word size of 10 bit. Fig. 6 (b) depicts the ROC curve for the best configuration of $K = 10240/(32 \cdot 10) = 32$ 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 while the original biometric template is protected. For greater block sizes (e.g. $l = 2^8$) biometric performance decreases drastically, since too much local information is lost. Regarding resulting template sizes for most configurations $K \cdot 2^w < 10 \cdot 2^{10}$, which means a compression of the original template is achieved.

4.3. Security Analysis

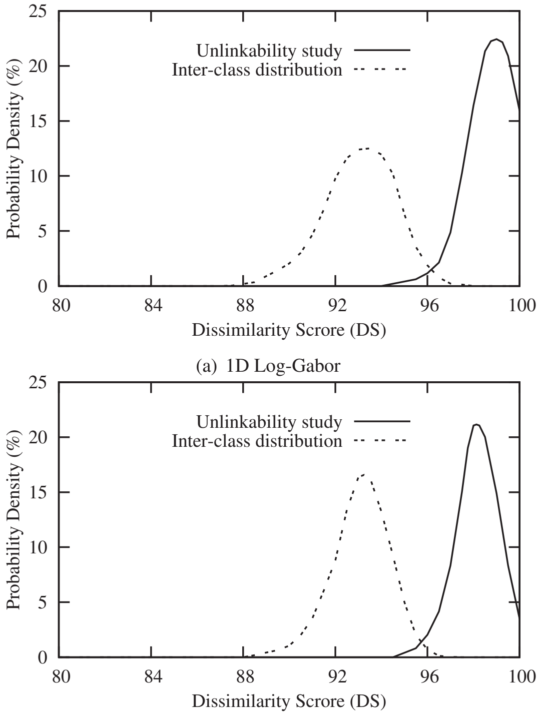
The security of the entire approach relies on the non-invertible mapping of codewords to a Bloom filter. W.l.o.g. this transform obscures the original position of the codeword as well as the number a codeword occurs, hence, for different configurations certain amounts of codewords are mapped to an identical position within according Bloom fil-



(a) 1D Log-Gabor

(b) Ma *et al.*

Figure 8. Proportion of re-mapped codewords, 1-$|b|/l$, for different block sizes l and word sizes w for both feature extractors.



(a) 1D Log-Gabor

(b) Ma *et al.*

Figure 9. Score distributions for inter-class comparisons and according unlinkability tests for both feature extractors.

ters. Fig. 8 depicts the average percentage of re-mapped codewords and according standard deviations for both algorithms and applied configurations, according to Fig. 3 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 3 and Table 4), as opposed to the case where no information about codeword positions is lost (see Table 2).

For best performing configurations (w.r.t. accuracy), mapping $l = 32$ codewords of length $w = 10$ to a $n = 2^{10}$ bit Bloom filters, for the 1D Log-Gabor feature extraction $\sim$ 48% of codewords are re-mapped, $1-|b|/l \simeq 0.48$ (see Fig. 8(a)). By analogy, for the algorithm of Ma *et al.* on average $\sim 32\%$ of codewords are re-mapped (see Fig. 8(b)). Focusing on the 1D Log-Gabor feature extractor, according to the previously estimated amount of possible sequences (see Fig. 3, $l = 32$) 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, the cancelable biometric system is operated through a natural trade-off between security and biometric performance.

4.4. Unlinkability Study

Unlinkability, *i.e.* the infeasibility of cross-matching different protected templates of a single subject, represents a major issue of biometric template protection. Experimental studies on unlinkability are commonly neglected [10] while several template protection schemes, *e.g.* fuzzy vault [7], have been exposed to be highly vulnerable to attacks via record multiplicity [17].

In the proposed cancelable scheme unlinkability is achieved by incorporating an application-specific bit vector which is XORed with each iris-code column prior to transforms. 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 Bloom filters of size $n = 2^{10}$. Subsequently, obtained inter-class distributions are compared to distributions yielded by comparing Bloom filters originating from a single iris-code which are obscured by different bit vectors. Obtained score distributions are depicted in Fig. 9 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.

5. Conclusion and Future Work

In this work alignment-free cancelable iris biometrics based on adaptive Bloom filter are proposed. Mapping biometric feature elements to Bloom filter represents an efficient non-invertible transform which provides a rapid alignment-free biometric comparison in transformed domain. A comprehensive experimental evaluation based on different iris biometric feature extractors confirms the soundness of the presented approach, providing protected templates and maintaining biometric performance of original recognition systems.

Due to the fact that a Bloom filter-based representation of biometric features does not require template alignment it could be applied in an efficient serial combination of feature extractors in order to accelerate biometric identification. In addition, if it is possible to establish an ordering of Bloom filter these could serve as a fuzzy hash in order to index biometric databases.

Acknowledgements

This work has been partially funded by the European FP7 FIDELITY project (SEC-2011-284862) and the Center of Applied Security Research Darmstadt (CASED).

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] A. Broder and M. Mitzenmacher. Network applications of bloom filters: A survey. *Internet Mathematics*, 1(4):485–509, 2005.

[4] 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.

[5] 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.

[6] 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.

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

[8] 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.

[9] 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.

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

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

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

[13] 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.

[14] J. K. Pillai, V. M. Patel, R. Chellappa, and N. K. Ratha. Sectored random projections for cancelable iris biometrics. In *Proc. IEEE Int'l Conf. on Acoustics Speech and Signal Processing*, pages 1838–1841. IEEE, 2010.

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

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

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

[18] 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.

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

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