TOWARDS CANCELABLE MULTI-BIOMETRICS BASED ON BLOOM FILTERS:
A CASE STUDY ON FEATURE LEVEL FUSION OF FACE AND IRIS

Christian Rathgeb*, Marta Gomez-Barrero†, Christoph Busch*, Javier Galbally† and Julian Fierrez†

*da/sec - Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany
Biometric Recognition Group - ATVS, EPS, Universidad Autonoma de Madrid, Spain
{christian.rathgeb,christoph.busch}@h-da.de, {marta.barrero,javier.galbally,julian.fierrez}@uam.es

ABSTRACT

In this work we propose a generic framework for generating an irreversible representation of multiple biometric templates based on adaptive Bloom filters. The presented technique enables a feature level fusion of different biometrics (face and iris) to a single protected template, improving privacy protection compared to the corresponding systems based on a single biometric trait. At the same time, a significant gain in biometric performance is achieved, confirming the soundness of the proposed technique.

Index Terms— Template protection, biometric fusion, Bloom filter, face, iris

1. INTRODUCTION

Biometric template protection schemes [1], commonly categorized as biometric cryptosystems and cancelable biometrics, offer solutions to privacy preserving biometric authentication. Cancelable biometrics consist of intentional, repeatable distortions of biometric signals that provide a comparison of biometric templates in the transformed domain. In accordance with the ISO/IEC IS 24745 [2] on biometric information protection, technologies of cancelable biometrics meet the two major requirements of irreversibility, i.e. 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, and unlinkability.

Ratha et al. [3] were the first to introduce the concept of cancelable biometrics applying non-invertible transforms to biometric data at enrolment, using application-dependent parameters. During authentication, biometric inputs are transformed and compared to the enrolled protected templates. In past years several concepts implementing cancelable biometrics based on face or iris have been proposed. For a detailed survey on cancelable biometrics, the reader is referred to [1].

Most existing approaches to cancelable biometrics report a significant decrease in biometric performance [1]. This restriction can be overcome by introducing multi-biometric template protection schemes [5], since a combination of different biometric characteristics generally leads to higher accuracy [6]. Within a conventional biometric system, a fusion of different biometric information can be performed at feature, score and decision level [7], as defined in the ISO/IEC TR 24722 on multimodal and other multi-biometric fusion [8]. Since both score and decision levels require a separate storage of (protected) templates, with respect to template protection, feature level has been identified as the preferable approach [9]. Performing such a fusion represents a great challenge: it requires a generic framework in order to establish a common representation of biometric features [5]. In addition, feature alignment turns out to be a critical issue since protected templates comprising information of more than one biometric instance are expected to require a complex alignment process.

In [10] face and ear features are fused using random projections and transformation-based feature extraction, reducing dimensionality with PCA and clustering the features. The final protected templates improve the performance of both unimodal traits. The spiral and the continuous components belonging to two different fingerprints from the same subject are mixed in one cancelable template in [11]. In [12], voice and iris data are fused applying different cancelable transformations for each modality.

In 2013, Rathgeb et al. [13] introduced cancelable biometrics based on Bloom filters, which has been applied to different biometric characteristics, e.g. iris [13] or face [14]. It has been shown that the concept can be used to generate irreversible protected templates which are capable of maintaining biometric performance, e.g. EERs of approximately 1% and 5% for iris and face, respectively [13, 14]. Furthermore, it has been shown that the concept can be employed in a multi-biometric scenario, in particular a multi-instance single-algorithm scenario protecting iris-codes at feature level [15]. Recently, Hermans et al. [16] provided a security analysis of this scheme, in particular of the original proposal [13]. While the authors confirm the irreversibility property of the system, it is shown that the scheme is vulnerable to cross-matching attacks. In particular, they prove that, given two cancelable templates generated from a single biometric sample, it is possible to cross-match the templates.
In this work, we (1) propose a generic feature level fusion of Bloom filter-based templates based on face and iris, (2) restricting our aim to the irreversibility of the stored reference. On the one hand, compared to [15], we adress a multi-instance multi-algorithm scenario, which represents a more challenging issue w.r.t. multi-biometric fusion at feature level. On the other hand, the design of transforms based on secret tokens in order to achieve unlinkability is subject of future work. Unlinkability can be achieved by transforming the feature vector prior to the Bloom filter-based transform or by extending the transform to a non-linear function, as suggested in [16]. In addition, in [17] it is shown that Bloom filter-based representations of biometric templates can be directly mapped to a set of unique features suitable as input for cross-matching resistant cryptographic primitives.

We have selected these characteristics as they can be captured with a single device that may have a sensor for 2D-face image acquisition and an NIR-sensor for iris acquisition. Such a single device has clear advantages regarding the transaction time, besides being more user-friendly than using multiple sensors for a single recognition system. It is shown that the proposed technique significantly improves the privacy protection and maintains the performance, i.e. $\sim 0.4\%$ EER, of a conventional score level fusion on the BioSecure face corpus and the IITD iris database.

This paper is organized as follows: In Sect. 2 the proposed feature level fusion based on Bloom filters is described in detail. Experimental evaluations are presented in Sect. 3 and conclusions are drawn in Sect. 4.

2. PROTECTED FEATURE LEVEL FUSION

In this section, the Bloom filter-based feature level fusion, the comparison of protected templates, and the resulting privacy protection in terms of irreversibility are described.

2.1. Bloom Filter-based Transform and Fusion

In the proposed multi-biometric system depicted in Fig. 1, Bloom filters, i.e. bit arrays of length $2^w$ where initially all bits are set to 0, are utilized in order to obtain an irreversible representation of binary face and iris features. For both characteristics, binary feature vectors are arranged in a two-dimensional matrix of width $W_F (W_I)$ and height $H_F (H_I)$. Each two-dimensional binary code is then divided into blocks of equal size where each column consists of $w_F (w_I)$ bits, respectively. The irreversible transform $h$ maps each binary column to its equivalent decimal value, setting to 1 the bit indexed by this value in the corresponding Bloom filter, as shown for two different codewords (i.e. columns stemming from face and iris) as part of Fig. 1. For each column $x \in \{0,1\}^w$, the mapping is defined as,

$$b[h(x)] = 1, \text{ with } h(x) = \sum_{p=0}^{w-1} x_p \cdot 2^p.$$  \hspace{1cm} (1)

The entire sequence of columns of each block is successively transformed to corresponding locations within Bloom filters; that is, protected templates are defined as sets of Bloom filters, $T_F = \{b_{F1}, b_{F2}, \ldots, b_{FK_F}\}$ and $T_I = \{b_{I1}, b_{I2}, \ldots, b_{IK_I}\}$, where $K_F$ and $K_I$ define the number of blocks, and according Bloom filters are of size $2^{w_F}$ and $2^{w_I}$. Within the original concept of Bloom filters [18] multiple hash functions are applied to large chunks of bytes. Since the size of codewords is small, $|x| \leq 10$, the proposed mapping is already collision-free and hash function would not add additional security. Focusing on the privacy protection of multiple biometrics, i.e. face and iris, the two protected templates denoted by $T_F$ and $T_I$ are required to have the same length: $|T_F| = |T_I|$. With respect to iris biometric features, this transform generates, to a certain extent, an alignment-free representation [13]. For face the alignment is done in the pre-processing step exploiting the horizontal baseline between the center of the eyes.

As the final step, both protected templates are fused by bit-wise ORing them, i.e. the final protected template $T$ is estimated as $T = T_F \lor T_I$. The fusion process conceals the origin of bits with respect to biometric characteristics. It will be shown that this simple fusion process highly improves the privacy protection provided by the Bloom filter-based transform. Since optimal choices of $w_F (w_I)$ depend on biometric characteristics and feature extraction algorithms (see [13, 14]), $|b_F| = |b_I|$ may not hold. However, in case $|b_F| \neq |b_I|$, $n \times 2^{\min\{w_F,w_I\}} = 2^{\max\{w_F,w_I\}}$ with $n \in \mathbb{N}$, holds. Note that according Bloom filters consist of $2^w$ bits, where $w = \max\{w_F,w_I\}$. The fused template $T_I$, which potentially results from a fusion of many small Bloom filters with a few large Bloom filters, is represented as, $T = \{b_{11}, b_{22}, \ldots, b_{\min\{K_F,K_I\}}\}$, where $\min\{K_F,K_I\}$ Bloom filters consist of a total number of $2^w$ bits $w = \max\{w_F,w_I\}$ bits.

2.2. Comparison of Protected Templates

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 [19]. Let $|b|$ denote the amount of bits within a Bloom filter $b$ set to 1. Then, the dissimilarity $\text{DS}$ between two Bloom filters $b_i$ and $b_j$ of same size is defined as [13],

$$\text{DS}(b_i, b_j) = \frac{\text{HD}(b_i, b_j)}{|b_i| + |b_j|} \quad |b_i|, |b_j| \neq 0.$$  \hspace{1cm} (2)

The final comparison score $S$ between two fused templates $T_i$ and $T_j$ is then defined as the average of all pair-wise
Bloom filter dissimilarity scores,

\[
S(T_i, T_j) = \frac{1}{\min\{K_F, K_I\}} \sum_{p=1}^{\min\{K_F, K_I\}} DS(b_{2p}, b_{2p+1}).
\]  

(3)

2.3. Irreversibility Analysis

Focusing on the mono-modal protected templates, \(T_F, T_I\), original positions of bit columns within according blocks are concealed, i.e. given a Bloom filter \(b\) of the final protected template \(T\) it is not clear from which column a distinct 1-bit in the protected template originates. In addition, high correlation between columns, especially neighboring ones, is expected. Consequently, a significant amount of columns are mapped to identical positions in Bloom filters even for small values of \(l\) (i.e. number of columns per block). We assume that after inserting \(l\) columns from a block there are \(|b|\) bits set to 1 within a Bloom filter. 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 binary code part involves arranging \(|b|\) codewords to \(l\) positions. For \(|b| \leq l\) the theoretical amount of possible sequences is recursively defined as \(f(|b|, l)\), where each of the \(|b|\) codewords have to appear at least once within \(l\) columns [13],

\[
f(|b|, l) = \begin{cases} 
1, & \text{if } |b| = 1, \\
|b|^l - \sum_{i=1}^{l-1} (|b|)^i \times f(i, l) & \text{otherwise.} 
\end{cases}
\]  

(4)

In other words, all sequences where less than \(|b|\) codewords appear are subtracted from the number of all possible sequences, \(|b|^l\), yielding a rapid increase of possible sequences even for small values of \(|b|\).

If we assume that a potential attacker knows the employed feature extraction algorithms and the amount of different columns within corresponding blocks contributing to the \(|b|\) bits set in a single Bloom filter \(b\) (in reality an attacker may only exploit certain statistics) the amount of possible sequences resulting in this fused Bloom filter can be theoretically estimated as,

\[
\left(\frac{|b|}{|b_F|}\right) \times f(|b_F|, l_F) \times \left(\frac{|b_F|}{|b_I| - (|b| - |b_F|)}\right) \times f(|b_I|, l_I).
\]  

(5)

In order to reconstruct both original binary templates, first the \(|b_F|\) bits originating from face biometric features have to be chosen from the total amount of bits set to 1, \(|b|\). Subsequently, the remaining \(|b_I| - (|b| - |b_F|)\) bits have to be chosen out of \(|b_F|\) (note that \(|b| - |b_F|\) bits must be selected in the second step, and are out of the possible combinations count). Finally, all sequences generating both bit sets have to be tested. The proposed fusion technique significantly increases the number of possible combinations which result in a single Bloom filter and, thus, improves the privacy protection provided by the fused template. In Sect. 3.3 we present a practical irreversibility estimation for the specific face and iris template sizes used in the experimental analysis.

3. EXPERIMENTAL EVALUATION

Biometric performance is estimated in terms of False Non-Match Rate (FNMR) at a targeted False Match Rate (FMR) and Equal Error Rate (EER) in accordance to the ISO/IEC IS 19795-1 [20]. The security of the protected system is configured by the targeted FMR and estimated as FMR\(^{-1}\), i.e. the required number of protected templates need to obtain a false acceptance [1] and privacy protection is defined as the number of different bit sequences yielding a distinct representation.

3.1. Experimental Setup

Experiments are carried out on a fused dataset using the face subcorpus included in the Desktop Dataset of the Multimodal
Table 1. Parameter choice for both feature extraction algorithms for both biometric characteristics.

<table>
<thead>
<tr>
<th>System</th>
<th>$W \times H$</th>
<th>$l$</th>
<th>$w$</th>
<th>$K$</th>
<th>$|T|$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face</td>
<td>32 $\times$ (40 $\times$ 60)</td>
<td>15</td>
<td>5</td>
<td>32</td>
<td>32 $2^{15}$</td>
</tr>
<tr>
<td>Iris</td>
<td>512 $\times$ 20</td>
<td>32</td>
<td>10</td>
<td>32</td>
<td>2 $2^{15}$</td>
</tr>
</tbody>
</table>

*for 32 (out of the original 80) sub-regions in the face image.

Table 2. Performance rates of original and protected systems (FNMRs are obtained at FMR=0.01%).

<table>
<thead>
<tr>
<th>System</th>
<th>Original Unprotected</th>
<th>Protected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face</td>
<td>70.079</td>
<td>6.536</td>
</tr>
<tr>
<td>Iris</td>
<td>96.587</td>
<td>0.929</td>
</tr>
<tr>
<td>Fusion</td>
<td>97.301</td>
<td>0.489</td>
</tr>
</tbody>
</table>

Table 3. Number of average bits set to one ($|b|$) and number of possible sequences an attacker would have to try in order to recover the face, the iris or both unprotected templates.

<table>
<thead>
<tr>
<th>System</th>
<th>Mono-modal BF</th>
<th>Multi-modal BF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>b</td>
<td>$</td>
</tr>
<tr>
<td>Face</td>
<td>6.33 $2^{10}$</td>
<td>202.56  $2^{11}$</td>
</tr>
<tr>
<td>Iris</td>
<td>16.73 $2^{10}$</td>
<td>16.73 $2^{10}$</td>
</tr>
<tr>
<td>Fusion</td>
<td>-</td>
<td>215.24 $2^{15}$</td>
</tr>
</tbody>
</table>

BioSecure Database$^1$ [21], which comprises four frontal face images of 210 subjects, captured in two time-spaced acquisition sessions, and the IITD Iris Database version 1.0$^2$ which comprises 1120 NIR images from 224 different subjects, where only the first four left eye images of the first 210 subjects are considered. For each subject, four pairs of face and eye images are formed for conducting genuine comparisons, and a single pair of face and eye images is applied for imposter comparisons, leading to a total number of $4 \times 3 \times 210 / 2 = 1,260$ genuine comparisons and $210 \times 210 / 2 = 22,050$ imposter comparisons.

For feature extraction as well as parameter selection, we refer to the configurations identified as most adequate in [13, 14], where $\|T_F = T_I\|$ holds (see last column in Table 1). In facial images eyes were automatically annotated using VeriLook SDK 4.0, developed by Neurotechnology$^3$. The FaceRecLib of the free signal and image processing toolbox Bob$^4$ [22] is utilized in order to extract facial features based on local Gabor pattern histogram sequences [23]. In this particular implementation, each image is divided into $8 \times 10$ sub-regions where only the central $4 \times 8$ sub-regions are used. The feature extraction generates and concatenates $32 \times 40 = 1,280$ 59-bit histograms, which are padded with a 0 prior to the binarization step, in order to obtain 60 bins per histograms, a non-prime number that allows a further division of each sub-region into $40/w_F \times 60/l_F$ sub-blocks for each of the 32 sub-regions. Histograms are simply binarized by setting bins greater than zero to one.

Focusing on eye images, at pre-processing the iris of a given sample image is detected, un-wrapped to an enhanced rectangular texture of $512 \times 64$ pixels applying the weighted adaptive Hough algorithm proposed in [24]. In the feature extraction stage, iris-codes are extracted based on a dyadic wavelet transform provided in a custom implementation within the University of Salzburg Iris Toolkit v1.0$^5$. This feature extraction algorithm, which was proposed by Ma et al. [25], a dyadic wavelet transform is performed on 10 signals obtained from the according texture stripes. The final code is $512 \times 20 = 10,240$ bits. For both characteristics, template comparisons are performed by calculating the fractional Hamming distance ($\pm 8$ circular bit-shifts are applied in each direction in order to compensate for rotations in the unprotected system). In the unprotected system the fusion is performed at score level (without normalisation) employing the sum-rule with equal weights of scores.

The parameter configurations are summarized in Table 1, where it can be observed that both protected templates are formed of $K_F \times 2^{w_F} = K_I \times 2^{w_I} = 2^{15}$ bits. For face and iris, a number of 15 5-bit and 32 10-bit columns are transformed to Bloom filters of size $2^5$ bit and $2^{10}$ bits, respectively. It is important to note that a total number of $15 \times 1,024 / 32 = 480$ columns of the binary face template and 32 columns of the iris-code are transformed to a single Bloom filter. It is important to note that the distribution of bits originating from face and iris, which contribute to the final protected template, is rather unbalanced.

3.2. Performance and Security Evaluation

The performance rates of original (unprotected) and protected systems are summarized in Table 2. As expected, the score level fusion of face and iris significantly improves the biometric performance in the original system yielding a decrease of the EER by 47% compared to the iris biometric system. Furthermore, we observe that the Bloom filter protection scheme maintains (and even slightly improves) biometric performance for face and iris as reported in [13, 14], however, we do not consider this performance gain as significant. More importantly, biometric performance is maintained in the fusion scenario, i.e. the proposed feature level fusion obtains comparable performance with respect to the score level fusion based on the sum-rule.

$^1$Biosecure Multimodal Database, http://biosecure.it-sudparis.eu/AB


$^5$USIT – University of Salzburg Iris Toolkit v1.0, http://www.wavelab.at/sources/
leading to 57.82% and 47.73% of the columns occurring more than once, modality face and iris systems, an average of approximately
 sequences originating a given template are shown in Table 3.

As previously mentioned for a targeted FNMR, security is estimated as [26]. As can be seen in Fig. 2 (bottom) for the protected monomodal face and iris systems, an average of approximately 57.82% and 47.73% of the columns occur more than once, leading to $15 \cdot (1 - 0.5782) = 6.33$ bits set to one for face ($|b_F| = 32 \cdot 6.33 = 202.56$) and $|b_I| = 32 \cdot (1 - 0.4773) = 16.73$ for iris. Therefore, according to Fig. 2 (top), there are $2^{39}$ and $2^{126}$ sequences that can generate each of the $2^{10}$ and $2^5$ Bloom filters, respectively.

### 3.3. Irreversibility Analysis

The average number of words mapped into one fused Bloom filter is $15 \cdot 32 + 32 = 512$. Therefore, the number of bits set to one is $|b| = 512 \cdot (1 - 0.5796) = 215.24$. According to Eq. 5 and the values described above, the total of possible face and iris sequences is estimated as,

$$2^{215/202} \times 2^{44} \times \left(\frac{202}{17} - (215 - 202)\right) \times 2^{126} \sim 2^{263}$$

On the other hand, an eventual attacker who wants to recover only the face or the iris would have to try the following number of sequences,

$$\left(\frac{215}{202}\right) \times 2^{44} \sim 2^{111} \text{ for face, } \left(\frac{215}{17}\right) \times 2^{126} \sim 2^{208} \text{ for iris.}$$

Those numbers must be multiplied by $K$ Bloom filters of the according protected template in order to recover the whole unprotected template.

As may be observed, the number of possible sequences, and therefore the privacy, of the single traits considerably increases (from $2^{49}$ to $2^{116}$ for face, and from $2^{108}$ to $2^{213}$ for iris) between the mono-modal and multi-modal protected templates. Furthermore, the total amount of sequences an eventual attacker would have to try in order to recover both traits raises to $\sim 2^{208}$ (!). It should be noted that, following this procedure, the original binary unprotected templates would be recovered, but not the original image. An inverse engineering process would have to be applied to those templates in order to obtain the images [26].

### 4. CONCLUSION AND FUTURE WORK

In this work, we proposed a generic feature level fusion of protected templates, obtained from irreversible Bloom filter-based transforms, and applied it to face and iris samples. A detailed privacy analysis of the presented approach is given and the experimental evaluation shows that the protected multi-biometric system maintains the biometric performance (EER $\sim 0.4\%$) compared to the score level fusion of the original unprotected systems. The privacy of the user is also increased: an eventual attacker would have to try up to $\sim 2^{208}$ sequences in order to recover the unprotected binary templates. Furthermore, the presented fusion highly reduced the amount of required storage. For the used parameter configuration, original facial templates of 76,800 bits and original iris-codes of 10,240 bits are compressed to a single protected template of size 32,768 bits, i.e. a compression of 63% with respect to the original size is obtained.
To achieve full multi-biometric template protection, which is out of scope in this paper, some ideas have been proposed, e.g. in [16, 17]. Future work will be focused on incorporating application-specific non-linear transforms to columns of binary templates prior to the Bloom filter-based transform to provide unlinkability.

5. ACKNOWLEDGEMENTS

This work has been partially supported by projects Bio-Shield (TEC2012-34881) from Spanish MINECO, FIDELITY (FP7-SEC-284862) and BEAT (FP7-SEC-284989) from EU, the Center for Advanced Security Research Darmstadt (CASED) and Cátedra UAM-Telefónica. Marta Gomez-Barrero is supported by a FPU Fellowship from Spanish MECD.

6. REFERENCES


