Rathgeb *et al. EURASIP Journal on Information Security* (2016) 2016:26
DOI 10.1186/s13635-016-0049-9

EURASIP Journal on Information Security

RESEARCH Open Access

Unlinkable improved multi-biometric iris fuzzy vault

Christian Rathgeb[1*], Benjamin Tams[2], Johannes Wagner[1] and Christoph Busch[1]

Abstract

Iris recognition technologies are deployed in numerous large-scale nation-wide projects in order to provide robust and reliable biometric recognition of individuals. Moreover, the iris has been found to be rather stable over time, i.e. iris biometric reference data provides a strong and permanent link between individuals and their biometric traits. Hence, unprotected storage of (iris) biometric data provokes serious privacy threats, e.g. identity theft, limited re-newability, or cross-matching. Biometric cryptosystems grant a significant improvement in data privacy and increase the likelihood that individuals will effectively consent in the biometric system usage. However, the vast majority of proposed biometric cryptosystems do not guarantee desired properties of irreversibility, unlinkability, and re-newability without significantly degrading the biometric performance.

In this work, we propose an unlinkable multi-instance iris biometric cryptosystem based on the improved fuzzy vault scheme. The proposed system locks biometric feature sets extracted from binary iris biometric reference data, i.e. iris-codes, of the left and right irises in a single fuzzy vault. In order to retain the size of the protected template and authentication speed, the proposed fusion step combines the most discriminative parts of two iris-codes at feature level. It is shown that the proposed key-binding process enables the generation of irreversible protected templates which prevents from previously proposed cross-matching attacks. Further, we investigate the optimal choice among potential decoding strategies with respect to biometric performance and time of key retrieval. The fully reproducible system is integrated to two different publicly available iris recognition systems and evaluated on the CASIAv3-Interval and the IITDv1 iris databases. Compared to the corresponding unprotected recognition schemes, genuine match rates of approximately 95 and 97 % at which no false accepts are observed and maintained in a single- and multi-instance scenario, respectively. Moreover, the multi-iris system is shown to significantly improve privacy protection achieving security levels of approximately 70 bits at practical biometric performance.

Keywords: Biometrics, Iris recognition, Template protection, Multi-biometrics, Biometric cryptosystem, Fuzzy vault scheme

1 Introduction

Biometrics refers to the automated recognition of individuals based on their behavioral and biological characteristics, e.g. fingerprint or face [1, 2]. It is generally conceded that a substitute to biometrics for positive identification in integrated security applications is non-existent. In past decades, iris recognition [3, 4] emerged as a rapidly growing field of research. Due to its intricate structure, the iris constitutes one of the most powerful biometric characteristics. Existing approaches to iris recognition achieve auspicious biometric performance and deployments in diverse application scenarios, such as border control, underline its tremendous impact [5]. On national-sized biometric systems, the use of multiple biometric characteristics has been found to significantly improve the accuracy and reliability especially in challenging identification scenarios [6], while recognition systems based on a single biometric indicator often have to contend with unacceptable error rates [7]. However, improvement in biometric performance as a result of biometric fusion should be weighed against the associated overhead involved, such as additional sensing cost, i.e. it is preferred to combine biometric characteristics that can be acquired in a single presentation [8]. For instance, in the Unique

*Correspondence: christian.rathgeb@h-da.de
[1]da/sec – Biometrics and Internet Security Research Group, Hochschule Darmstadt, Darmstadt, Germany
Full list of author information is available at the end of the article

IDentification Authority of India (UIDAI) [9], which aims at registering all 1.2 billion Indian citizens, both irises and all fingers of a subject are acquired at enrollment. Similar national projects are also under way in Indonesia and in several smaller countries [10].

Generic iris recognition systems, see Fig. 1a, consist of four major modules: (1) image acquisition; (2) pre-processing; (3) feature extraction; and (4) comparison. At image acquisition, good-quality NIR images are required to provide a robust recognition, where most current deployments require subjects to fully cooperate with the system. At pre-processing, the pupil and the outer boundary of the iris are detected. Subsequently, general iris recognition algorithms transform the iris ring to a normalized rectangular texture on which image enhancement methods, e.g. histogram stretching, are applied. To complete the pre-processing, the parts of the iris texture which are occluded by eyelids, eyelashes, or reflections are detected and stored in an according noise mask. The vast majority of feature extraction algorithms follow the approach of Daugman [3], in which a binary feature vector, i.e. iris-code, is extracted by applying adequate filters to the iris texture in a row-wise manner. The data representation based on an iris-code offers compact storage and rapid Hamming distance-based (*HD*) comparison.

Recently, different researchers have shown that, in case an attacker has full knowledge of the employed feature extraction, iris-codes can be utilized in order to reconstruct images of subjects' iris textures [11, 12]. Such images can be presented to an iris recognition system in order to successfully launch presentation attacks [13]. In order to safeguard individuals' privacy, protection of biometric templates is of utmost importance, since unprotected storage of biometric templates poses serious privacy threats. Technologies of biometric template protection [14, 15], which are commonly categorized as biometric cryptosystems [16] and cancelable biometrics [17], offer solutions to privacy preserving biometric authentication. Biometric cryptosystems are designed to securely bind a digital key to a biometric or generate a digital key from a biometric signal [16] in order to meet the two major requirements of biometric template protection defined in the ISO/IEC IS 24745 [18]: (1) irreversibility, i.e. knowledge of a protected template can not be exploited to reconstruct a biometric sample which is equal or close (within a small margin of error) to an original captured sample of the same source; (2) unlinkability, i.e. different versions of protected biometric templates can be generated based on the same biometric data (renewability), while protected templates should not allow cross-matching.

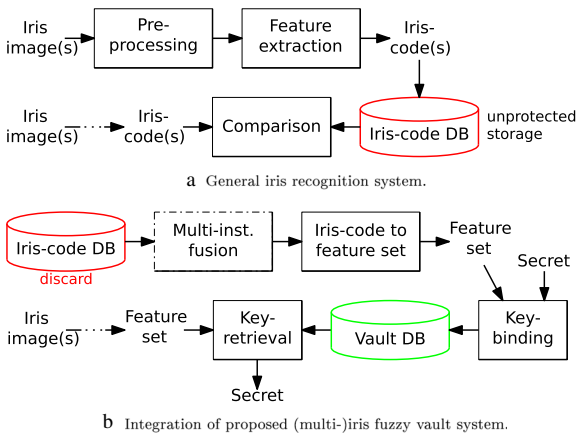


Fig. 1 Processing chains of **a** a conventional unprotected (multi-) iris recognition system and of **b** the proposed (multi-) iris fuzzy vault system

Apart from fulfilling the above properties, an ideal biometric cryptosystem shall not cause a decrease in biometric performance with respect to the corresponding unprotected system [19]. The vast majority of existing techniques do not satisfy desired template protection requirements in practice, mostly resulting in a trade-off between privacy protection and biometric performance [20]. The incorporation of multiple biometric characteristics to biometric cryptosystems has been found to improve biometric performance [7], while the protection of multi-biometric templates is especially crucial as they contain information regarding multiple characteristics of the same subject [21]. In contrast to conventional biometric systems, where fusion may take place at score or decision level [7], with respect to template protection schemes, feature-level fusion has been identified as most suitable. A separate storage of two or more protected biometric templates would enable parallelized attacks. In contrast, a single protected template, which has been obtained by means of feature-level fusion, is expected to improve privacy protection, since the fused template comprises more biometric information [21]. This is analogous to an access control system which requires multiple low strength (few bits) keys, where each key can be attacked individually. Such a system is less secure than one which uses a single key with a larger number of bits. Obviously, the development of multi-biometric cryptosystems is accompanied by further issues such as common data representation, storage requirement, or feature alignment [22].

1.1 Contribution of work

In this work, we present an unlinkable multi-biometric iris-cryptosystem based on the fuzzy vault scheme [23]. As shown in Fig. 1b, one property of the proposed scheme is that it can be seamlessly integrated to an existing iris recognition system. Since it is designed to protect iris-codes, a re-enrolment of already registered subjects is

not required, in contrast to proposed iris-based fuzzy vault schemes (see Section 2), which build upon specific feature extractors. The presented scheme is extended to a multi-iris template protection scheme where biometric fusion takes place at feature level, in order to maximize privacy protection. In order to minimize processing time as well as storage requirement, the fusion process is designed to retain the size of a single protected template. To maximize the entropy, and hence discriminativity, of the fused template, the most reliable feature parts of two iris-codes are combined, and based on a detailed investigation of decoding strategies, we choose an ideal trade-off between computational and biometric performance. Thereby, authentication times, which are of equal magnitude for single- and multi-iris fuzzy vaults, are kept low while accuracy is significantly improved.

The proposed system, which is evaluated on two publicly available iris databases for different iris biometric feature extractors, maintains biometric performance of corresponding unprotected baseline systems. In addition, security analysis show that the system is resistant to existing cross-matching attacks and provides cryptographically acceptable security levels confirming the soundness of the presented approach.

1.2 Organisation of article

The remainder of this article is organized as follows: Section 2 briefly summarizes related works regarding iris biometric cryptosystems and the fuzzy vault scheme. Section 3 provides a detailed description of key components of the proposed unlinkable improved (multi-) biometric iris fuzzy vault. Experiments are presented in Section 4. Finally, conclusions are drawn in Section 5.

2 Related work

2.1 Iris biometric cryptosystems

Among conceptual proposals [24, 25], the first implementation of an iris biometric cryptosystem, in which the fuzzy commitment scheme is employed [26], was presented in [27]. The fuzzy commitment scheme by Juels and Wattenberg [26] represents a cryptographic primitive, which combines techniques from the area of error correcting codes and cryptography. At key-binding, a prechosen key is prepared with error correcting codes and bound to a binary biometric feature vector of the same length by XORing both resulting in a difference vector. In addition, a hash of the key is stored together with the difference vector forming the commitment. During key retrieval, another binary feature vector is XORed with the stored difference vector and error correction decoding is applied. In case the presented feature vector is sufficiently "close" to the stored one, the correct key is returned, which can be tested using the stored hash. In [27], 2048-bit iris-codes are applied to bind and retrieve 140-bit keys prepared with Hadamard and Reed-Solomon error correction codes. On an in-house dataset comprising 700 iris images of 70 subjects, a FNMR of 0.47 % at a zero FMR is reported. To provide a more efficient error correction decoding in an iris-based fuzzy commitment scheme, two-dimensional iterative min-sum decoding was introduced in [28] achieving a FNMR of 5.62 % at a zero FMR on the ICE 2005 iris database [29], where bound keys consist of 40 bits. A context-based reliable component selection used to extract keys from iris-codes which are then bound to Hadamard codewords is presented in [30]. Further, diverse techniques to improve the performance and security of (iris) fuzzy commitment schemes have been proposed, e.g. [31–34]. In [35, 36], different attacks have been suggested, which utilize the fact that error correction codes underlie distinct structures. Statistical attacks based on so-called error correction code histograms have been successfully conducted against iris-based fuzzy commitment schemes in [37]. In [38], it was found that fuzzy commitment schemes leak information in bound keys and non-uniform templates. Suggestions to prevent from information leakage in fuzzy commitment schemes have been proposed in [39]. In addition, attacks via record multiplicity could be applied to decode stored commitments [40, 41]. In [33], a bit-permutation process was introduced to prevent from cross-matching attacks. It is, however, important to note that even if one uses record-specific bit-permutation processes, cross-matching may still be possible for records protecting very similar feature vectors and then even enables reversibility attacks from record multiplicity [42]. Considering proposed attacks, the security provided by iris-based fuzzy commitment is rather doubtable.

Focusing on iris, the first implementation of a fuzzy vault scheme (see Section 2.2) was presented in [43, 44] in which salient feature points are extracted from iris textures. Based on several enrolment samples, independent component analysis and k-means clustering is employed to extract 16 coefficients from most stable parts of the iris textures. On the BERC [45] and the CASIAv3-Interval iris database [46], a GMR of approximately 99 and 80 %, respectively, was achieved at a zero FMR employing 128 bit keys. In order to prevent from attacks via record multiplicity, an iris fuzzy vault which is hardened with an additional password is presented in [47, 48]. Image enhancement techniques are applied to extract iris fibres from which minutiae-like coordinates are extracted. Experiments are reported for the CASIAv1 [49] and the MMU iris database [50] achieving a GMR of at most 90 % at a zero FMR. In [51], a multi-biometric fuzzy vault based on fingerprint and iris is proposed. Secret keys, which are used to generate an iris-based fuzzy commitment, are directly encoded, fused with fingerprint data at feature level, and locked in the fuzzy vault scheme, achieving

a GMR of approximately 98 % at a FMR of 0.01 % on the in-house MSU-DBI fingerprint [52] and the CASIAv1 iris database [49] for a key size of 208 bit. However, the scheme requires an additional storage of corresponding difference vectors which allows for previously mentioned attacks. Further proposal of iris fuzzy vaults, e.g. [53–55], omit a detailed description of employed iris feature encoding or experimental protocols.

The majority of proposed approaches to iris biometric cryptosystems lack a thorough security analysis. It is important to note that the length of employed keys in biometric cryptosystems represents only an upper bound for the provided security. For instance, the scheme in [27], which is designed to bind and retrieve 140 bit keys, provides a security of approximately 40 bits [51]. This reduction is mainly caused by dependencies among neighbouring iris-code bits resulting in a large entropy loss.

Apart from iris biometric cryptosystems, different approaches to cancelable iris biometrics have been proposed, e.g. [56–60]. For further details on cancelable iris biometrics, the reader is referred to [15, 61].

2.2 Fuzzy vault scheme

By design, the fuzzy vault scheme [23, 62] enables protection and error-tolerant verification with feature sets. It is due to this property that led researchers to consider the fuzzy vault scheme as a promising tool for protecting fingerprint minutiae sets [63]. This preliminary analysis was followed by a series of working implementations for fingerprints most of which are minutiae-based [64, 65]. However, the fuzzy vault scheme as proposed by Juels and Sudan is vulnerable to a certain kind of linkage attacks, the correlation attack [66, 67], that very clearly conflicts with the unlinkability requirement and (even worst) with the irreversibility requirement of effective biometric template protection. Moreover, the use of public unprotected auxiliary alignment data can ease an attacker in performing linkage attacks. As a countermeasure, an implementation for absolutely pre-aligned minutiae that avoids any correlation between related records of the fuzzy vault scheme has been proposed [68]. Another advantage from the construction in [68] is that the improved fuzzy vault scheme by Dodis et al. [69] can be used for template protection which results in significantly more compact record sizes. Compared to earlier works [63, 64], security levels have been revealed to be approximately 40 bits for single-finger fuzzy vaults. It is important to note that other linkage attacks can be applied to the improved fuzzy vault scheme but they can, however, be effectively avoided [70].

It is a widely accepted hypothesis that the fingerprint modality does not contain a sufficient amount of effective entropy to resist attacks exploiting the distribution of fingerprint features, specifically false-accept attacks. From this perspective, we should increase the amount of entropy by considering implementations for fused templates extracted from more than one finger of a person [71] or, as a promising alternative, consider other biometric modalities such as a person's iris(es).

When designing a fuzzy vault-based cryptosystem, a practical decoding strategy is needed. Though, in the original fuzzy vault [23, 62], a Reed-Solomon decoder [72] has been proposed, its resulting error-correcting capabilities are not sufficient to achieve a practical implementation for single finger. As a consequence, the use of a Lagrange-based decoder has been proposed [64] and adopted for other single-finger implementations [65, 68]. A multi-finger implementation has been proposed in [71]. If a Lagrange-based decoder was chosen for the implementation, then the decoding complexity would become infeasible. A reasonable trade-off between decoding time and verification performance can be achieved using a Guruswami-Sudan-based decoder [71, 73] of which a Reed-Solomon decoder can be viewed as a special case. In this paper, these strategies are considered for key retrieval in Section 3.4.

3 Proposed system

3.1 From iris-codes to feature sets

As aforementioned, conventional iris recognition schemes extract two-dimensional binary feature vectors, i.e. iris-codes, from iris images [3, 4], cf. Fig. 5. Since the fuzzy vault scheme operates on feature sets, the first processing step of the proposed system represents a feature type transformation [74], in particular a binary-to-integer ordered-to-unordered transformation. In our scheme, an extracted iris-code is vertically divided into B blocks of size $w \times h$ bits, where each block, and hence each column, is associated with a block index $b = 0, \ldots, B-1$. Subsequently, transform f is applied to each column $\mathbf{c}_i \in \{0,1\}^h, i = 0, \ldots, w-1$, of a block with index $b, f(\mathbf{c}_i, b) = \text{int}(\mathbf{c}_i) \cdot \log_2 B + b$, where int denotes the conversion of a binary vector to its integer representation. Each block yields a vector $\mathbf{V}_b$ of feature values, $\mathbf{V}_b = (v_0, \ldots, v_{w-1})$, with $v_i = f(\mathbf{c}_i, b)$. For each vector, $\mathbf{V}_b, \mathbf{P}_b$ denotes the sets of unique vector elements, $\mathbf{P}_b = \{v_0, \ldots, v_l\}_{\neq}, l \leq w$. The set-based representation conceals the ordering of columns within each block; in addition, multiple entries are discarded. Since iris-code bits are not mutually independent [3], correlation between (neighbouring) columns is expected. Hence, a certain amount of columns within each processed block will be identical, as will be shown in experiments; see Section 4. Let $|\mathbf{P}_b| = l$ be the number of unique feature values of a transformed iris-code block. For a potential attacker, the reconstruction of the original iris-code block involves an arranging of l codewords to w positions. The degree of irreversibility provided by

transform f could be quantified by the number of possible blocks resulting in the same set of features which is estimated as,

$$\ell = \sum_{j=1}^{l} (-1)^{l-j} \binom{l}{j} j^w, \tag{1}$$

which follows from a simple application of the inclusion-exclusion principle. Large values of ℓ prevent an attacker from reconstructing (guessing) corresponding original iris-code blocks, which further improves privacy protection. It is important to note that, in contrast to fingerprints, where approximations of original fingerprints can be reconstructed from minutiae coordinates [75], this estimation only applies to the use of iris-codes. The final set of feature values $\mathbf{P}$ is estimated as the union of all sets of unique vector elements of all blocks, $\mathbf{P} = \bigcup_{b=0}^{B-1} \mathbf{P}_b$, where $|\mathbf{P}| = \sum_{b=0}^{B-1} |\mathbf{P}_b|$, since each block index b is unique. The entire procedure of generating $\mathbf{P}$ from an iris-code is illustrated in Fig. 2.

In the comparison stage, generic iris recognition systems apply circular bit shifts and estimate *HD* scores at different shifting positions, i.e. relative tilt angles, in order to compensate for head tilts. The minimal obtained *HD*, which corresponds to an optimal alignment, represents the final score. It is important to note that the number of shifting positions employed to determine an appropriate alignment between pair of iris-codes may vary depending on the application scenario. The proposed set-based representation is alignment-free to a certain extent, since equal columns within certain blocks are mapped to identical feature values. Hence, self-propagating errors caused by an inappropriate alignment of iris-codes are eliminated (radial neighbourhoods persist). The rotation-compensating property of the proposed system comes at the cost of location information of iris-code columns. In case iris-codes are strongly misaligned, columns will end up in different blocks yielding different feature sets. However, as will be shown in experiments, in case of iris images captured under favourable conditions, the proposed representation maintains distinctiveness if w and h are chosen appropriately.

3.2 Multi-instance single-algorithm fusion

An overview of the multi-instance single-algorithm fusion, which is designed to fuse biometric information extracted from the left and right iris of a subject, is depicted in Fig. 3. The proposed fusion technique is motivated by two facts:

1. Average decoding times generally increase with the size of a fuzzy vault. Hence, a fusion of the most discriminative parts of two iris-codes to one of the size of a single iris-code is preferable, in order to maintain reasonable decoding times.
2. Upper and lower iris parts are frequently occluded by eyelids or eyelashes, cf. Fig. 4 (even under active participation of captured subjects). It has been found that the use of a general noise mask reveals performance gains comparable to that obtained by subject-specific masks [76].

In order to extract discriminative and stable feature sets of reasonable size, iris-code parts which originate from

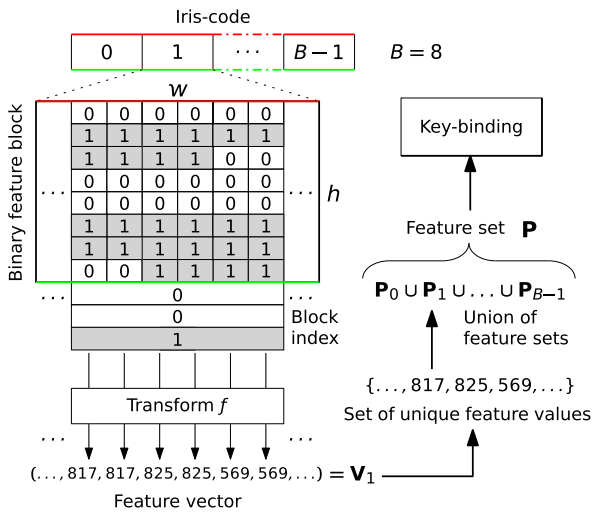


Fig. 2 Proposed iris-code to feature set transform: based on binary columns of a block and its corresponding block index, the transform f extracts feature elements from which a set of unique feature elements are generated. For example, for $B = 8$, column $01100110_2 = 102$, and index $001_2 = 1$, we get $f(102, 1) = 817$

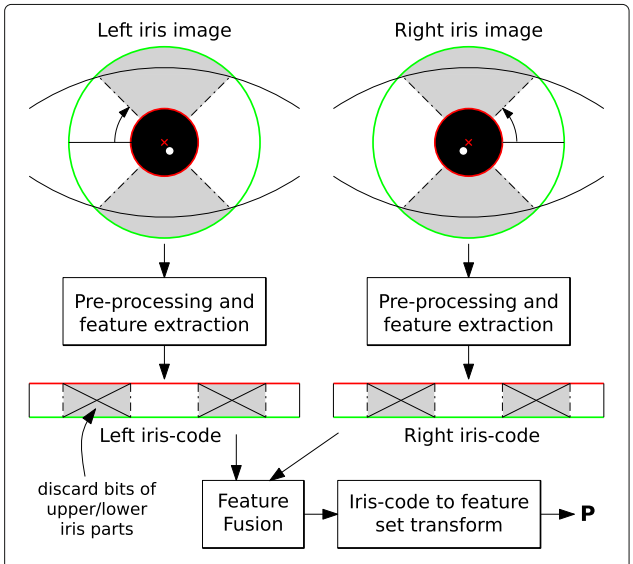
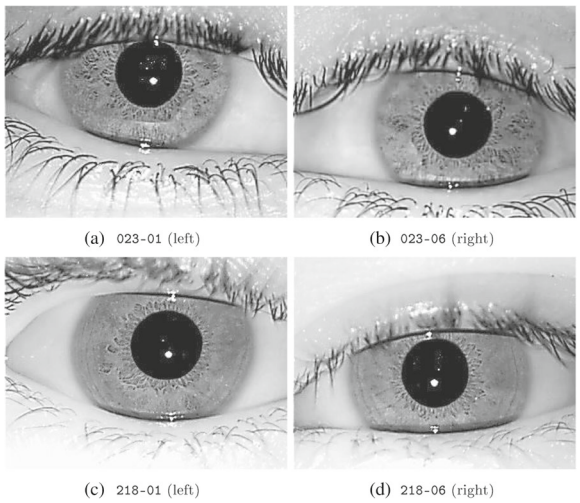


Fig. 3 Proposed multi-instance fusion: iris-code bits originating from upper and lower iris parts are discarded and remaining iris-code chunks are concatenated prior to extracting to extracting feature sets



(a) 023-01 (left) (b) 023-06 (right)

(c) 218-01 (left) (d) 218-06 (right)

Fig. 4 Sample pairs of left and right 320 × 240 pixel NIR iris images of the IITDv1 Iris Database *(numbers in captions* refer to identifiers). It can be observed that *upper* and *lower* iris parts are frequently occluded by eyelids and eyelashes

upper and lower iris parts (marked grey in Fig. 3) are discarded. Discarding iris-code blocks originating from those parts represents a static way of extracting half of the bits expected to be rather stable, which does not require the storage of any auxiliary data. Note that additional subject-specific data pointing at most reliable iris-code parts might enable linkage attacks. The resulting iris-code parts of the left and right iris of width $w \times K/2$ are concatenated, resulting in a single iris-code of dimension equal to that of an iris-code extracted from a single image (without deleting upper and lower iris-code parts). The stability of selected iris-code parts will be empirically analysed in experiments; see Section 4. Finally, the previously described method is employed to extract a single feature set. The presented feature-level fusion does not affect the size of the vault compared to single-instance scenario and, at the same time, retains biometric information expected to be most discriminative [4].

The proposed multi-instance fusion technique requires irises to be aligned properly. However, nowadays, iris sensors are capable of acquiring the left and right irises of a subject simultaneously (using two sensors), which results in an implicit alignment of both irises. Moreover, it is generally conceded that favorable acquisition conditions are considered a fundamental premise for biometric template protection [14]. Hence, the suggested trivial concatenation of iris-code parts is expected to preserve the alignment-free representation of subsequently processed iris-code blocks. Other fusion techniques, e.g. a random shuffling or interleaving of iris-code columns, would obscure neighbourhoods of columns within iris-code blocks and, thus, are expected to decrease the biometric performance of the system.

3.3 Key binding

In the first step of the binding process, a secret polynomial $\kappa \in \mathbf{F}[X]$ of degree smaller than k is chosen, and the hash $\mathrm{SHA}(\kappa)$ is stored. Given a feature set $\mathbf{P}$, extracted from both or a single iris image(s) of a subject, $\mathrm{SHA}(\kappa)$ is used as seed for a record-specific but public bijection $\sigma : \mathbf{F} \rightarrow \mathbf{F}$, which is applied to re-map the elements of $\mathbf{P}, \hat{\mathbf{P}} = \sigma(\mathbf{P}) = \{\sigma(v) | v \in \mathbf{P}\}$. This step is performed in order to prevent from the attack proposed in [70], which is based on the extended Euclidean algorithm. Let $V(X)$ and $W(X)$ be two related vault records protecting the feature sets $\mathbf{P}$ and $\mathbf{P}'$, respectively; unlinkability can be attacked efficiently and effectively, provided that

$$|\mathbf{P} \cap \mathbf{P}'| \geq (\max(|\mathbf{P}|, |\mathbf{P}'|) + k)/2. \tag{2}$$

In [70], it is also shown that the probability of Eq. 2 can be destroyed by applying the abovementioned re-mapping of feature elements. Note that the employment of these public maps does not affect the operational performance of the system. Moreover, as σ is generated based on $\mathrm{SHA}(\kappa)$, no additional data storage is required. Due to the assumed randomness of two maps σ and σ', the corresponding sets $\sigma(\mathbf{P})$ and $\sigma'(\mathbf{P}')$ are random and, based on the definition of the *hyper-geometric distribution*, the probability that with these sets Eq. 2 is fulfilled is equal to

$$1 - \binom{\rho}{|\mathbf{P}|}^{-1} \sum_{j=0}^{\omega_0 - 1} \binom{|\mathbf{P}'|}{j} \binom{\rho - |\mathbf{P}'|}{|\mathbf{P}| - j}, \tag{3}$$

where $\omega_0 = \lceil (|\mathbf{P}| + k)/2 \rceil$, $\rho = 2^{h + \log_2 B}$, and w.l.o.g. $|\mathbf{P}| \geq |\mathbf{P}'|$. In experiments, it will be shown that the probability of Eq. 3 is sufficiently low in a cryptographic sense; see Section 4.

The next step is performed based on the improved fuzzy vault scheme [69], which improves the original construction by means that it generates significantly more compact records. In the improved fuzzy vault, scheme genuine and chaff feature elements are encoded by a monic polynomial of degree $t = |\hat{\mathbf{P}}|$. The features in $\hat{\mathbf{P}}$, interpreted as elements of a finite field $\mathbf{F}$ where $|\mathbf{F}| = \rho$, are bound to the secret polynomial κ by computing $V(X) = \kappa(X) + \prod_{v \in \hat{\mathbf{P}}} (X - v)$, such that the pair $(V(X), \mathrm{SHA}(\kappa))$ builds our final record.

All elements of $\hat{\mathbf{P}}$ represent columns consisting of exactly $h + \log_2 B$ bits. The size of the vault increases with t and is upper bounded by the width of processed iris-codes, i.e. $|V(X)| \leq 512(h + \log_2 B)$ bits. In order to obscure the size of the vault, which might leak information about the protected iris-code, an additional zero-bit is appended to each element in $\hat{\mathbf{P}}$. Subsequently, random values, for which this additional bit is set to 1, are added to $\hat{\mathbf{P}}$ until a desired maximum vault size is reached. The maximum vault size might be defined as $512(h + 1 + \log_2 B)$

bits or by a maximum observed vault size $\max_{n\in N} t_n(h + 1 + \log_2 B)$, where N is the number of registered subjects.

It is important to note that the correlation attack [66, 67], which is a special linkage attack, cannot be applied due to the fact that the improved fuzzy vault scheme encodes a maximal number of chaff points. In such a way, no correlation can be exploited between feature sets protected by two (or more) related instances of the applied fuzzy vault records.

3.4 Polynomial reconstruction strategy and key retrieval

On authentication, an unlocking set $\mathbf{U} \subset \mathbf{F} \times \mathbf{F}$ of size u is built containing ω pairs being interpolated by the polynomial κ; we call these pairs *genuine*. If $\omega \geq k$, then it is possible to reconstruct the polynomial κ from $\mathbf{U}$ if we assume that we can verify the correctness of κ, e.g. by using the hash value $\mathrm{SHA}(\kappa)$. It may, however, not be feasible to reconstruct κ from $\mathbf{U}$ in case $\omega \geq k$. For example, we found from our training (see Section 4.2) that we can expect the unlocking set to be of size $u \approx 300$; furthermore, if we want to achieve a practical genuine acceptance rate, we need to be able to successfully recover κ from $\mathbf{U}$ when it contains $\omega = 100$ genuine pairs. In the following, we discuss different recovery strategies for this representative example to find a reasonable decoder.

3.4.1 Iterated Lagrange strategy

One way to reconstruct κ from $\mathbf{U}$ is to select k pairs from $\mathbf{U}$ and compute the unique polynomial κ^* of degree smaller than k that interpolates them. If all k selected pairs are genuine, then $\kappa* = \kappa$ which can be verified with overwhelming reliability by observing $\mathrm{SHA}(\kappa^*) = \mathrm{SHA}(\kappa)$. If not all k selected pairs are genuine, then most likely $\mathrm{SHA}(\kappa^*) \neq \mathrm{SHA}(\kappa)$ and we may repeat the procedure until $\mathrm{SHA}(\kappa^*) = \mathrm{SHA}(\kappa)$. This procedure is guaranteed to, eventually, reveal the secret polynomial κ if $\omega \geq k$ while for a single step the probability of success is equal to

$$p_L(u, \omega, k) = \binom{\omega}{k} \cdot \binom{u}{k}^{-1}. \tag{4}$$

However, the procedure can become very impractical to be performed. For instance, if $u = 300, \omega = 100$, and $k = 20$, then we have to expect to run 2^{34} iterations before κ is recovered. This is too costly for a practical polynomial reconstruction strategy.

3.4.2 Iterated Reed-Solomon strategy

Alternatively, we may apply a *Reed-Solomon decoder* (e.g. see [72]). This class of algorithms is capable of recovering κ from $\mathbf{U}$ efficiently (by means of deterministic polynomial time) if $\omega \geq (u+k)/2$. On the other hand, these class of algorithms will fail to recover κ from $\mathbf{U}$ for $\omega < (u + k)/2$. In order to establish a decoding mechanism reasonably dealing with these cases, it is conceivable to choose a c-sized subset $\mathbf{U}_0 \subset \mathbf{U}$ randomly where $|\mathbf{U}| \geq c \geq k$, apply the Reed-Solomon decoder to $\mathbf{U}_0$, and if successfully revealing κ^* with $\mathrm{SHA}(\kappa^*) = \mathrm{SHA}(\kappa)$, output the recovered polynomial; otherwise, the procedure is repeated up to a predefined number of steps. This mechanism will succeed eventually if $\omega \geq (c + k)/2$ which improves upon the bound $\omega \geq (u + k)/2$ since $c \geq u$. The success probability for a single step of the procedure is equal to

$$p_{\mathrm{RS}}(u, \omega, k, c) := \binom{u}{c}^{-1} \sum_{j=\lfloor (c+k)/2 \rfloor}^{\min(u,\omega,c)} \binom{\omega}{j} \cdot \binom{u-\omega}{c-j}. \tag{5}$$

If in addition, the time $z_{RS}(c, k)$ for a single iteration is known, then the expected time for the iterated Reed-Solomon strategy to successfully recover κ can be calculated as

$$\frac{\log(0.5)}{\log(1 - p_{RS}(c, \omega, k, c))} \cdot z_{\mathrm{RS}}(c, k). \tag{6}$$

In order to assess feasibility of the iterated Reed-Solomon decoding strategy, we consider our case example where $(u, \omega, k) = (300, 100, 20)$. For each $c = k, ..., 2\omega - k$ computed with Eq. (5), we applied an experimentally supported analysis on the estimated effort for the iterated Reed-Solomon decoding strategy to successfully decode the vault. For each c, we measured the time $z_{RS}(c, k)$ a Reed-Solomon decoder applied to an unlocking set of size c would consume and estimated the expected time for a successful recovery using Eq. (6). We found that for $c = 36$, the effort becomes minimal with $z_{\mathrm{RS}}(c, k)$ being very close to zero and $p_{\mathrm{RS}}(u, \omega, k, c) \approx 2^{-27}$. Yet, even though significantly improving upon the iterated Lagrange decoding strategy, an effort of 27 bits is still too costly for a practical implementation.

3.4.3 Iterated Guruswami-Sudan strategy

The bound $\omega \geq (u + k)/2$ can be significantly improved by applying a Guruswami-Sudan algorithm [73]. This class of algorithms can potentially recover κ from $\mathbf{U}$ provided that $\omega > \sqrt{u \cdot (k-1)}$. However, in practice, implementations of the Guruswami-Sudan decoder may be too time-consuming if one aims at recovering exactly up to $u - \sqrt{u \cdot (k-1)}$ errors. The number of errors that one can tolerate depends on an additional parameter controlling input to a Guruswami-Sudan algorithm called *multiplicity*. The higher the multiplicity, the more errors the algorithm can tolerate while also being more inefficient. The lower the multiplicity, the more efficient is the algorithm but being able to tolerate fewer errors, e.g. $\approx u - \sqrt{2 \cdot u \cdot (k-1)}$ for a single multiplicity. Yet, this can still be a significant improvement as compared to a Reed-Solomon decoder.

In the same way as for the iterated Reed-Solomon strategy, we could apply an iterated Guruswami-Sudan strategy which, given u, ω, k, and c, will recover κ from $\mathbf{U}$ in a single step with probability

$$p_{\mathrm{GS}}(u,\omega,k,c) := \binom{u}{c}^{-1} \sum_{j=\lfloor\sqrt{c\cdot(k-1)}+1\rfloor}^{\min(u,\omega,c)} \binom{\omega}{j} \cdot \binom{u-\omega}{c-j}. \quad (7)$$

For $(u, \omega, k) = (300, 100, 20)$ we found that $c = 297$ yields the best iterated Guruswami-Sudan strategy (for a single multiplicity) where each step consumes $z_{GS}(c, k)$ time (a fraction of a second on a standard workstation, see Section 4.3) with success probability $p_{GS}(u, \omega, k, c) \approx 0.96$ %. It has furthermore been measured that a single step is approximately 2^9 times harder to perform than a single Lagrange interpolation. This results in an overall difficulty to recover κ from $\mathbf{U}$ with more than 95 % probability of only 9 bits and clearly shows that this strategy outperforms the iterated Lagrange and Reed-Solomon strategy.

3.4.4 Proposed strategy

The different decoding strategies that we discussed above have been tested for $(u, \omega, k) = (300, 100, 20)$ which is a typical property of an unlocking set being built on a genuine verification. Since a Guruswami-Sudan strategy clearly is the best choice from the above experiments, we choose a Guruswami-Sudan strategy. Furthermore, since $c \approx u$ has been found as the optimal choice, we will only apply a single step of the Guruswami-Sudan algorithm to the entire unlocking set $\mathbf{U}$ and no subsets thereof. Moreover, we use a multiplicity of three which yields a decoder with a reasonable trade-off of error tolerance and decoding time, i.e. less than a second [71].

Our discussion emphasizes that it is hard to choose an optimal polynomial recovery strategy, mainly due to the fact that one can either optimize the decoding time or the success probability but not both (unless a predefined trade-off between the two criteria can be selected). From our discussion, we have chosen the final decoding strategy from experimental observations and a final educated guess. However, it must be stressed that a Guruswami-Sudan strategy is not necessarily the best strategy. For example, whenever one wants to successfully decode even if $\omega = k$, then the iterated Lagrange strategy is optimal, and if $\omega \geq (u + k)/2$, then the Reed-Solomon decoding strategy yields to the best decoder. It seems therefore reasonable to develop a dynamic decoder switching between different strategies depending on the expected values for ω given u and k. This is an interesting topic; the analysis of which we leave open for future research. The reader should be aware of the fact that from an attacker's point of view, other strategies can be much more efficient—e.g. strategies that employ the statistics of the protected features as, for example, a false-accept attack.

4 Experimental evaluation

4.1 Database and experimental method

Experiments are carried out on the CASIAv3-Interval [46] and the IITDv1 iris database [77], where all left eye images of the CASIAv3 database are used for training purposes (see Section 4.2) and all five images of left and right eyes of the IITDv1 dataset are used for performance evaluations. Both datasets comprise good-quality NIR iris images of size 320 × 240 pixel. Number of subjects, images, and resulting amounts of genuine and impostor authentication attempts are summarized in Table 1. Sample pairs of left and right iris images of the IITDv1 database are depicted in Fig. 4.

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. 5a–c applying the weighted adaptive Hough algorithm proposed in [78]. In the feature extraction stage, two different iris recognition algorithms are employed where normalized enhanced 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 analysed). The first feature extraction method follows the Daugman-like 1D-LogGabor feature extraction algorithm of Masek [79] (LG) and the second follows the algorithm proposed by Ma et al. [80] (QSW) based on a quadratic spline wavelet. Both feature extraction techniques generate iris-codes of 512 × 20 = 10,240 bit. Sample iris-codes generated by both feature extraction methods are shown in Fig. 5d, e. Custom implementations of employed segmentation and feature extractors are freely available in the University of Salzburg Iris Toolkit (USIT) [81]. For further details on the employed feature extraction algorithms, the reader is referred to [10].

In accordance with IS ISO/IEC 19795-1:2006 [82], biometric performance is evaluated in terms of genuine match rate (GMR), false non-match rate (FNMR), false match rate (FMR), and equal error rate (EER). In addition, for the proposed fuzzy vault schemes, we estimate average

Table 1 Overview of employed databases, corresponding number of subjects, iris images, and the resulting number of genuine and impostor comparisons

Set	Database	Subjects	Images	Gen.	Imp.
Training	CASIAv3[a]	198	1334	4295	19,503
Evaluation	IITDv1[b]	224	2240	2240	24,976

Note that the genuine and impostor scores have been derived following the recommendations of Mansfield and Wayman [86]

[a] Left eyes

[b] Left and right eyes

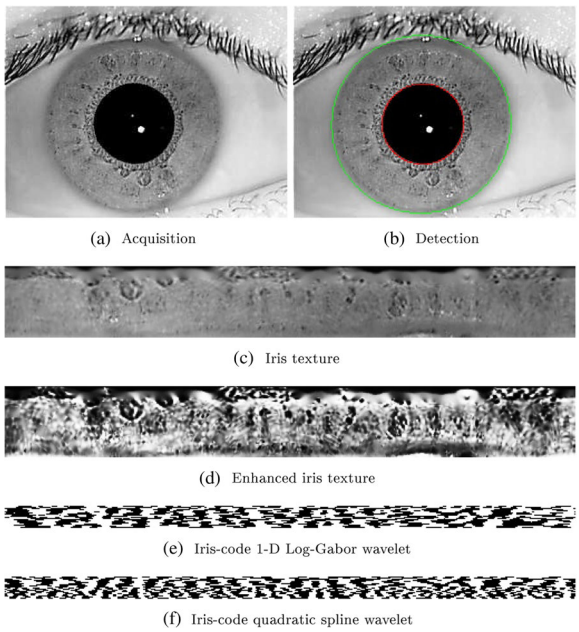


Fig. 5 Iris recognition processing chain. **a** Iris image. **b** Iris detection. **c** Normalized iris texture. **d** Pre-processed iris texture. **e**, **f** Iris-codes of applied feature extractor for sample image 001-02 of the IITDv1 database

genuine decoding time (GDT) and average impostor decoding time (IDT) and provided brute-force security (BFS) in bits. All experiments were performed on a single core of an Intel Core i7-3610QM CPU with 3.2 GHz on a standard workstation with sufficient RAM.

Its resistance to recovery attacks, i.e. the effort for an imposter given a vault record to recover the original feature sets or the secret polynomial, represents a fundamental aspect of a fuzzy vault implementation. Generally, the fuzzy vault can be attacked by a brute-force attack, where the attacker repeatedly samples k points from the vault and tries to interpolate the secret polynomial from these. The expected number of attempts of this attack can be estimated by combinatorial means [83]. In contrast, the false-accept attack exploits the specific distribution of the biometric features, by repeatedly simulating verifications employing features of randomly chosen (real) iris-codes. The success probability of the false-accept attack is equal to the FMR provided by the employed parameter setting. In order to optimize the success rate, the attacker can deviate from the parameters used in actual operation; however, in the proposed scheme, the number of decoding iterations is the only parameter that is not already fixed at the time of enrolment. It has been proven in [68] that the expected number of decoding attempts of the false-accept attack is minimized for using one iteration. Moreover, it has been shown that an attacker's success probability is maximized by using simple polynomial interpolation as opposed to using a Reed-Solomon decoder. Hence, we report the security in terms of false-accept security (FA security) using this optimal strategy.

Estimating high security levels assumes sharp estimations of FMRs when they are close to zero. In biometric systems with deterministic verification algorithm, the FMR can only be estimated down to the magnitude of $1/I$, where I is the number of impostor verifications performed in the evaluation. However, the verification of our implementation is probabilistic as soon as the unlocking set contains more than k points. This property allows us to give heuristic estimates of FMRs that are much smaller than $1/I$: For each single impostor verification, we compute the success probability based on the size of the unlocking set and the number of correct points contained, and finally, we estimate the FMR as the mean over all verifications. For details, we refer to [68].

4.2 Training stage

In the training stage ω, the number of genuine elements in the unlocking set is employed as comparison score between genuine and impostor pairs of feature sets. Since biometric cryptosystems should be operated at reasonable low FMRs, configurations which achieve the highest GMRs at a zero FMR are preferred. Biometric performance in terms of GMR at a zero FMR for different values of h and w is shown in Table 2. Feature blocks are obtained from the uppermost h adjacent rows of iris-codes bits originating from inner bands of the iris, i.e. the number of blocks is estimated as $K = 512/w$ for all configurations. Two configurations are identified as optimal, $h = 10$ and $w = 16$ as well as $h = 10$ and $w = 32$. Note that the training step is performed on a disjoint dataset where a single instance, i.e. iris image, is used per subject to calculate ω between pairs feature sets.

4.3 Performance estimation

Sample pairs of left and right iris images of the evaluation dataset are shown in Fig. 4. In the baseline systems,

Table 2 Biometric performance in terms of GMR (%) at zero FMR for both feature extractors for different parameter settings obtained in the training for left eyes of the CASIAv3 database (settings considered as optimal are marked bold)

h	w	LG	QSW	h	w	LG	QSW
8	64	74.7381	82.9336	8	16	92.4796	94.0861
10	64	85.4016	93.0617	**10**	**16**	**95.7159**	**95.9721**
12	64	90.3609	93.7602	12	16	94.4820	95.6228
14	64	89.9651	92.7590	14	16	93.5274	94.5518
8	32	88.7311	94.1793	8	8	89.4063	88.9872
10	**32**	**94.2724**	**96.6007**	10	8	90.3143	89.9418
12	32	94.9476	95.8789	12	8	89.9185	89.7555
14	32	93.0617	95.2503	14	8	87.8463	88.4284

HD-based dissimilarity score are estimated, where ±8 circular bit shifts are performed to obtain a minimum *HD*. For both feature extraction methods, Fig. 7 visualizes the probability of mis-matching bits between genuine iris-code comparisons. Similar typical two-dome pattern of unstable bits are clearly visible for both feature extractors. Obtained biometric performance is summarized in Table 3 and Fig. 6. As expected, a sum-rule-based score-level fusion of comparison scores between iris-code obtained from both iris images significantly improves biometric performance, yielding EERs below 0.5 %. Note that, in the context of score-level fusion, using the sum-rule with proper normalization, which is implicitly provided due to the use of a single feature extraction per fusion, has been observed to result in competitive performance [8]. Table 4 shows obtained biometric performance after applying the proposed transform for different parameter settings. Again, a sum-rule-based fusion of ω scores across all iris-code blocks is performed for using both iris images. On the one hand, we observe that the set-based representation does not affect (or might even improve) the biometric performance in case a single iris is used. On the other hand, if both irises are employed, a sum-rule-based fusion of *HD*-based scores reveals better biometric performance, compared to the proposed representation.

For the single-instance scenario (using iris images of the left eye), the obtained GMRs and corresponding FMRs, provided security, GDTs, and IDTs for different polynomial degrees k are summarized in Tables 5 and 6 for the LG and QSW feature extraction, respectively. According relations between FNMR and FMR are plotted in Fig. 8. Note that scales for FNMR and FMR have been adjusted to improve visibility. As can be observed, the biometric performance of both unprotected baseline systems is either maintained or slightly improved. Moreover, for best performing configurations, with respect to the polynomial degree k, GDTs and IDTs are significantly smaller than 1 s.

Focusing on the proposed multi-biometric fuzzy vault, the probabilities of mismatching bits shown in Fig. 7 confirm the soundness of the proposed selection of iris-code parts. Obtained performance for both feature extractors using different configurations is summarized in Tables 7

Table 3 Baseline performance in terms of GMR at zero FMR and EER of the original unprotected systems estimating *HD*-scores of left eyes (top) and score-level fusion of both eyes (bottom) on IITDv1 database

Algorithm	Iris	GMR@FMR=0 (%)	EER (%)
LG	Left	91.3392	1.310
	Both	94.2857	0.403
QSW	Left	87.2767	1.368
	Both	92.6785	0.466

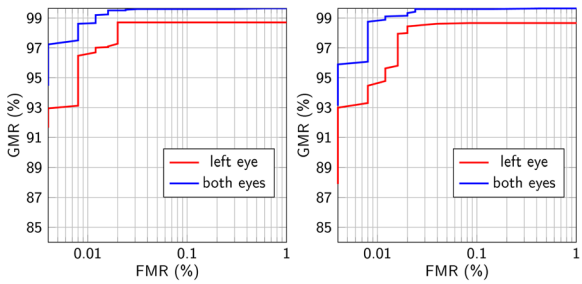


Fig. 6 ROC curves of the original unprotected systems estimating *HD* scores of the left eyes and score-level fusion of both eyes on IITDv1 database

and 8, respectively. Compared to obtained performance rates for a sum-rule-based fusion of ω scores across all iris-code blocks shown in Table 4, the presented fusion strategy clearly improves recognition accuracy. According FNMRs and FMRs are plotted in Fig. 9. As can be observed, biometric performance is significantly improved for multi-biometric fuzzy vaults achieving better rates than according baseline systems, which confirms that the proposed fusion retains highly discriminative feature sets. In addition, decoding times increase only sightly due to the compact representation of fused feature sets. By comparing relations between FNMRs and FMRs of single-iris and multi-iris fuzzy vaults, cf. Figures 8 and 9, it can be observed that, for reasonable values of k, both rates significantly decrease across all configurations.

For chosen parameters $h = 10$ and $w = 16, 32$, the vault size is estimated as $512(h + 1 + \log_2(512/w))$. For single- and multi-iris vaults, this results in $512 \times 16 = 8192$ and $512 \times 15 = 7680$ bits for $w = 16$ and $w = 32$, respectively. Hence, compared to the original 512×20 bit iris-codes, resulting vault sizes are even smaller, since only the upper (more discriminative) half of the original iris-code

Table 4 Biometric performance in terms of GMR (%) at zero FMR and EER (%) and privacy protection in terms of l and ℓ provided by the employed transform for both feature extractors for different parameter settings obtained for left eyes and a score-level fusion of both eyes on the IITDv1 database ($h = 10$)

Algorithm	Iris	w	GMR@FMR=0 (%)	EER (%)	l	ℓ
LG	Left	16	93.6294	1.705	8.68	$\sim 2^{48}$
		32	93.5972	1.691	17.57	$\sim 2^{127}$
	Both	16	94.3694	1.161	8.68	$\sim 2^{48}$
		32	94.2793	1.425	17.56	$\sim 2^{127}$
QSW	Left	16	86.2613	1.835	10.71	$\sim 2^{50}$
		32	85.7786	2.255	21.44	$\sim 2^{130}$
	Both	16	87.0270	1.350	10.72	$\sim 2^{50}$
		32	86.5315	1.485	21.45	$\sim 2^{130}$

Table 5 Performance of the iris fuzzy vault in terms of GMR, FMR, FA-security, GDT, and IDT in relation to the polynomial degree k, applying the LG feature extraction with $w = 16$ (top) and $w = 32$ (bottom) for left eyes of the IITDv1 database

k	GMR (%)	FMR (%)	FA-sec. (bits)	GDT (sec)	IDT (sec)
4	98.4375	0.02001920	20.8021	49.539	0.978
6	97.7232	0.00800769	27.3634	9.247	0.617
8	97.0089	0.00800769	32.1446	3.330	0.460
10	96.0714	0.00400384	36.8344	1.648	0.366
12	94.8661	0.00400384	41.5671	0.973	0.313
16	92.0982	0	51.2525	0.488	0.246
24	83.8393	0	71.7633	0.234	0.181
32	71.7857	0	94.1555	0.169	0.152
4	98.6607	0.0240231	17.6631	65.286	1.267
6	98.3929	0.0200192	25.0975	9.683	0.577
8	97.5446	0.0160154	30.41	3.440	0.418
10	97.0089	0.0120115	35.0868	1.692	0.345
12	96.1161	0.00800769	39.7419	1.006	0.294
16	93.9286	0	49.2163	0.477	0.218
24	86.3393	0	69.0849	0.229	0.163
32	75.9375	0	90.5816	0.158	0.135

Table 6 Performance of the iris fuzzy vault in terms of GMR, FMR, FA-security, GDT, and IDT in relation to the polynomial degree k, applying the QSW feature extraction with $w = 16$ (top) and $w = 32$ (bottom) for left eyes of the IITDv1 database

k	GMR (%)	FMR (%)	FA-sec. (bits)	GDT (sec)	IDT (sec)
4	97.9911	0.0120115	20.5745	72.571	1.741
6	97.0536	0.00400384	27.5663	10.906	1.016
8	95.8929	0.00400384	32.2375	4.034	0.743
10	94.4643	0.00400384	36.8065	2.103	0.606
12	92.5893	0.00400384	41.4282	1.329	0.530
16	88.1696	0	50.8645	0.701	0.423
24	75.2679	0	70.5884	0.372	0.333
32	59.8214	0	91.6024	0.270	0.252
4	98.2143	0.0600577	17.111	71.130	1.641
6	97.7679	0.0240231	24.9017	12.898	1.010
8	96.875	0.0120115	30.5012	4.568	0.724
10	95.8482	0.0120115	34.9417	2.285	0.598
12	94.8214	0.00400384	39.2536	1.413	0.489
16	90.8929	0.00400384	47.9585	0.707	0.396
24	80.8036	0	66.0467	0.361	0.322
32	66.5179	0	85.2412	0.268	0.250

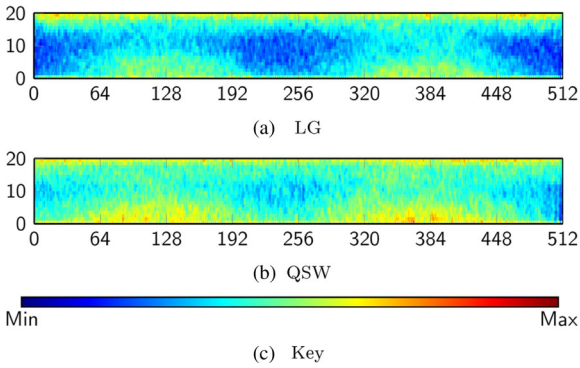


Fig. 7 Heatmaps illustrating the probability of mis-matching bits between genuine iris-code comparisons for both feature extraction methods on the IITDv1 database

is employed. As aforementioned, vault sizes may be further reduced based on an maximum observed number of feature elements in extracted point sets.

4.4 Security analysis

For chosen parameter settings, the average number of unique feature elements per iris-code block, l, and the corresponding amount of different sequences of iris-code columns which result in a single set of feature values, ℓ, according to Eq. 1, is summarized in Table 4. The large values of ℓ indicate a certain degree of irreversibility,

Table 7 Performance of the multi-iris fuzzy vault in terms of GMR, FMR, FA-security, GDT, and IDT in relation to the polynomial degree k, applying the LG feature extraction with $w = 16$ (top) and $w = 32$ (bottom) using left and right eyes of the IITDv1 database

k	GMR (%)	FMR (%)	FA-sec. (bits)	GDT (sec)	IDT (sec)
4	99.2411	0.0120155	20.2336	111.180	1.445
6	98.6161	0.00800769	28.6599	20.131	0.926
8	97.6786	0.00400384	34.6343	6.935	0.702
10	96.9643	0	40.0863	3.301	0.562
12	96.25	0	45.5598	1.944	0.476
16	93.7946	0	56.7434	0.887	0.385
24	86.2054	0	80.3626	0.376	0.270
32	73.6161	0	106.131	0.268	0.240
4	99.4196	0.0840807	16.7213		
6	99.0179	0.0160154	24.6816	21.535	0.869
8	98.5268	0.00800769	31.6944	7.284	0.654
10	97.9464	0.00800769	37.3526	3.479	0.536
12	97.1875	0.00400384	42.4997	1.970	0.449
16	95.4911	0	52.7801	0.895	0.352
24	88.7946	0	74.3242	0.375	0.247
32	78.4821	0	97.5726	0.254	0.224

Table 8 Performance of the multi-iris fuzzy vault in terms of GMR, FMR, FA-security, GDT, and IDT in relation to the polynomial degree k, applying the QSW feature extraction with $w = 16$ (top) and $w = 32$ (bottom) using left and right eyes of the IITDv1 database

k	GMR (%)	FMR (%)	FA-sec. (bits)	GDT (sec)	IDT (sec)
4	98.9286	0.00800769	20.0627	96.125	2.319
6	97.7232	0.00400384	29.1706	17.313	1.459
8	96.6071	0.00400384	35.8048	6.238	1.077
10	95.0446	0	41.486	3.056	0.871
12	93.4375	0	47.1429	1.813	0.746
16	86.6071	0	58.6874	0.910	0.532
24	71.6518	0	82.9584	0.464	0.399
32	51.4732	0	109.121	0.330	0.314
4	99.4196	0.212204	16.378	106.957	2.156
6	98.9286	0.0160154	24.3893	21.143	1.501
8	97.6339	0.00800769	31.8845	7.280	1.142
10	97.1429	0.00400384	38.1659	3.423	0.784
12	95.5357	0.00400384	43.5714	1.964	0.683
16	91.6964	0	54.0004	0.998	0.533
24	79.0625	0	75.5451	0.476	0.402
32	60.4018	0	98.4993	0.340	0.312

which will hamper a systematic reconstruction of iris-code blocks from the proposed set-based representation, even if an attacker succeeds in unlocking the vault. A more thorough analysis of potential reconstruction attacks is beyond the scope of this work.

For all employed parameter settings $h = 10, w = 16, 32, k = 4, ..., 32$, and $256 \leq |\mathbf{P}'| \leq |\mathbf{P}| \leq 512$, one can verify experimentally that the probability of Eq. 2 never becomes larger than 2^{-155}. This clearly shows that the cross-matching attack from [70] can be effectively prevented.

We estimate the security against recovery attacks by the expected number of decoding attempts that an attacker has to perform with a Lagrange-based decoder after having built the unlocking set using a query template [68]. In this way, our notion of security is a measure of resistance against a certain false-accept attack by also accounting for a decoding complexity. However, it is important to note that there may be more efficient strategies for an attacker to perform. For example, the attacker may rely on a Guruswami-Sudan-based decoder as in the verification protocol or on a completely different (currently unknown) strategy. However, we stress that the security achieved with a Guruswami-Sudan-based decoder cannot

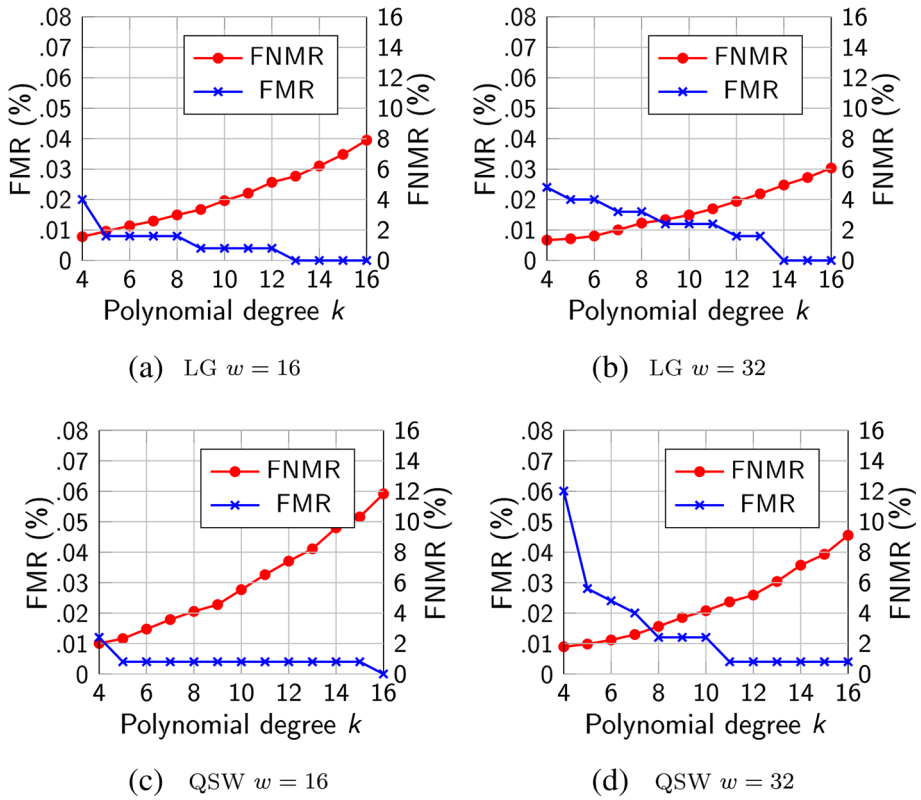


Fig. 8 Biometric performance in terms of FNMR and FMR in relation to reasonable values of the polynomial degree k for optimal parameter settings for both feature extractors using left eyes of the IITDv1 database

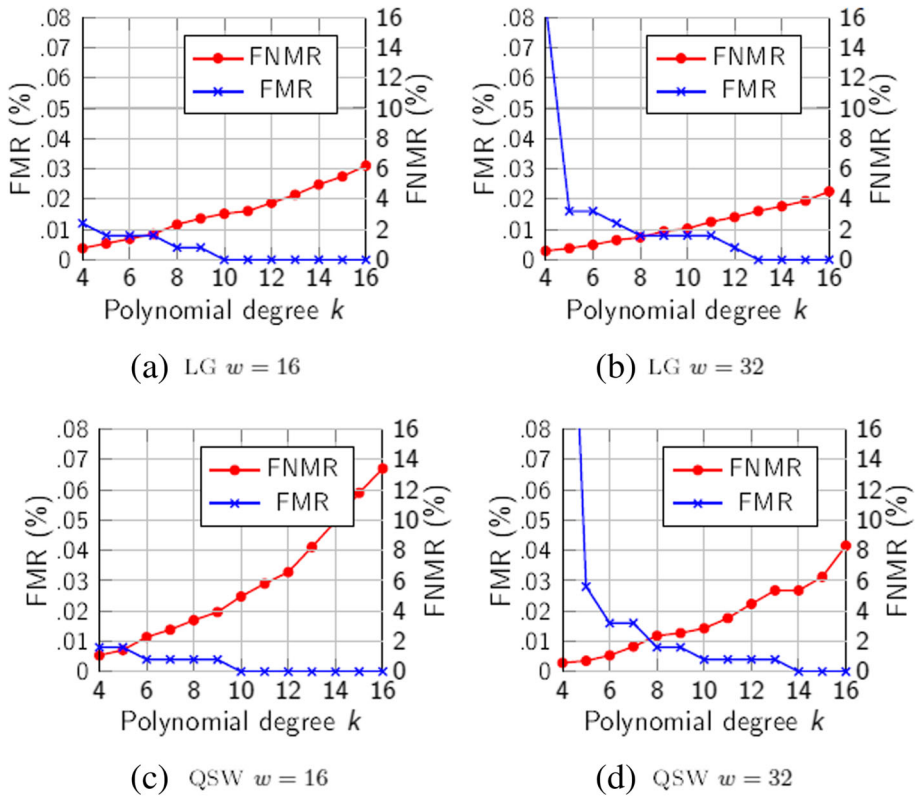


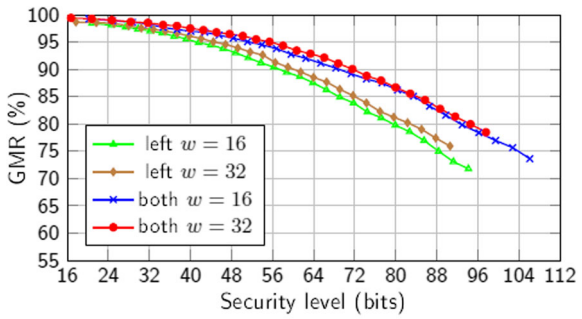
Fig. 9 Biometric performance in terms of FNMR and FMR in relation to reasonable values of the polynomial degree k for optimal parameter settings for both feature extractors using left and right eyes of the IITDv1 database

be estimated beyond the bound of 2^{15} times the decoding complexity—even if the true security is significantly higher. The method from [68] allows to estimate a higher false-accept security; but, the reader should be aware that the resulting security estimates are heuristic false-accept security estimates. Yet, these estimates are still much more realistic and plausible than security estimates against brute-force attacks. Figure 10 compares obtained GMRs versus the security levels, for different parameter settings of the proposed (multi-)biometric fuzzy vault schemes for both employed feature extraction algorithms. Firstly, we observe that security levels provided by iris biometric fuzzy vaults using a single iris appear satisfactory. For GMRs above 90 %, security levels of approximately 50 bits are obtained. Secondly, the use of both irises is shown to be beneficial. On the one hand, it improves biometric performance: at a fixed security level, higher GMRs are obtained when using both eyes; on the other hand, privacy protection is increased: for a fixed GMR, higher security levels are achieved; Hence, depending on the chosen size of k, recognition accuracy and/or privacy protection is improved. At operation points which achieve reasonable GMRs of approximately 90 and 95 %, the security level is increased from 50 to almost 70 bits and 40 to 50 bits, respectively.

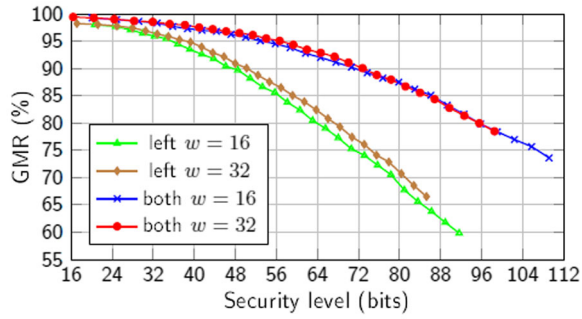
Another very important security aspect concerns the risk of correlation attacks on two or more vault records of the same user. Since we use the improved fuzzy vault scheme, which effectively uses all finite field elements as vault points [69], the correlation attack from [66] cannot be applied. On the other hand, there are specific correlation attacks against the improved fuzzy vault scheme based on solving systems of polynomial equations [84] or deploying the extended Euclidean algorithm [70]. However, these attacks only work if in both vault records the features are represented by the same finite field elements, and, hence, are prevented by our use of a record-specific permutation σ of the field elements [68].

5 Conclusions

The link between individuals and their biometric (physiological) characteristics is considered strong and permanent, particularly for iris [3]. Therefore, iris-codes must be protected in order to safeguard individuals' privacy and biometric systems' security. Albeit this is rarely the case in operational systems, favourable capture conditions (resulting in low intra-class variance) are considered a fundamental premise for biometric cryptosystems. Consequentially, compared to conventional recognition



(a) LG



(b) QSW

Fig. 10 GMRs versus FA-security of proposed (multi) iris biometric fuzzy vaults for both feature extraction algorithms and different parameter settings on the IITDv1 database

systems, biometric cryptosystems suffer from a significant decrease in biometric performance. Multi-biometric cryptosystems, which should be designed to combine multiple biometric characteristics at feature level [21], might bridge the gap between biometric performance and privacy protection [19].

The key advantages of the proposed (multi-)iris fuzzy vault scheme can be summarized in three terms: (1) in contrast to the vast majority of biometric cryptosystems, the proposed system maintains biometric performance obtained by the corresponding unprotected iris recognition system in a single- or multi-instance scenario; (2) the suggested fusion technique combines most discriminative information of two iris-codes to a single fuzzy vault and key retrieval is performed based on an analysis of suitable decoding strategies providing fast decoding times; (3) in contrast to existing proposals the presented (multi-) iris fuzzy vault scheme is designed to protect iris-codes, i.e. the protection of an iris biometric database does not require re-enrolment registered subjects.

Furthermore, it is shown that the proposed (multi-)iris fuzzy vault schemes obtain unrivalled security levels in comparison to fuzzy vaults based on other biometric characteristics. Even minutiae-based multi-finger fuzzy vaults using four fingers do not achieve higher security levels [71]. Like in the vast majority of biometric cryptosystems, obtained decoding times may limit the system to be operated in verification mode while protected biometric templates still enable a compact storage. A security analysis for the scenario where an adversary might be in possession of one iris-code prior to attacking a multi-iris fuzzy vault might be subject to future research. Finally, we note that the presented work is based on open-source software and evaluated on public biometric databases, i.e. represents fully reproducible research [85].

Acknowledgements
This work has been partially supported by the German Federal Ministry of Education and Research (BMBF) within the Center for Research in Security and Privacy (CRISP).

Competing interests
The authors declare that they have no competing interests.

Author details
[1] da/sec – Biometrics and Internet Security Research Group, Hochschule Darmstadt, Darmstadt, Germany. [2] secunet Security Networks AG, Essen, Germany.

Received: 13 January 2016 Accepted: 13 October 2016
Published online: 02 November 2016

References

1. AK Jain, A Ross, S Prabhakar, An introduction to biometric recognition. IEEE Trans. Circ. Syst. Video Technol. **14**, 4–20 (2004)
2. AK Jain, P Flynn, AA Ross, *Handbook of Biometrics*. (Springer, New York, 2008)
3. J Daugman, How iris recognition works. IEEE Trans. Circ. Syst. Video Technol. **14**(1), 21–30 (2004)
4. KW Bowyer, K Hollingsworth, PJ Flynn, Image understanding for iris biometrics: a survey. Elsevier Comput. Vis. Image Underst. **110**(2), 281–307 (2007)
5. A Ross, Iris recognition: the path forward. Computer. **43**, 30–35 (2010)
6. A Ross, K Nandakumar, AK Jain, *Handbook of Multibiometrics*. (Springer, Secaucus, NJ, USA, 2006)
7. A Ross, AK Jain, Information fusion in biometrics. Pattern Recogn. Lett. **24**, 2115–2125 (2003)
8. AK Jain, B Klare, AA Ross, in *In Proc. of the 8th Int. Conf. on Biometrics 2015 (ICB'15)*. Guidelines for best practices in biometrics research (IEEE, 2015), pp. 1–5
9. Unique Identification Authority of India, Aadhaar: http://uidai.gov.in/. retrieved July, 2016. http://uidai.gov.in/
10. C Rathgeb, A Uhl, P Wild, *Iris Recognition: From Segmentation to Template Security. Advances in Information Security*, vol. 59. (Springer, New York, 2013)
11. S Venugopalan, M Savvides, How to generate spoofed irises from an iris code template. IEEE Trans. Inf. Forensics Secur. **6**, 385–395 (2011)
12. J Galbally, A Ross, M Gomez-Barrero, J Fiérrez, J Ortega-Garcia, Iris image reconstruction from binary templates: an efficient probabilistic approach based on genetic algorithms. Comp. Vision Image Underst. **117**(10), 1512–1525 (2013)
13. S Marcel, M Nixon, SZ Li, *Handbook of Biometric Anti-Spoofing*. (Spring, New York, 2014)
14. AK Jain, K Nandakumar, A Nagar, Biometric template security. EURASIP J. Adv. Sig. Process. **2008**, 1–17 (2008)
15. C Rathgeb, A Uhl, A survey on biometric cryptosystems and cancelable biometrics. EURASIP J. Inf. Secur. **2011**(3) (2011)
16. U Uludag, S Pankanti, S Prabhakar, AK Jain, Biometric cryptosystems: issues and challenges. Proc. IEEE. **92**(6), 948–960 (2004)
17. NK Ratha, JH Connell, RM Bolle, Enhancing security and privacy in biometrics-based authentication systems. IBM Syst. J. **40**(3), 614–634 (2001)
18. ISO/IEC JTC1 SC27 Security Techniques:ISO/IEC 24745:2011. Information Technology - Security Techniques - Biometric Information Protection

International Organization for Standardization (ISO, 2011). International Organization for Standardization
19. K Nandakumar, AK Jain, Biometric template protection: bridging the performance gap between theory and practice. IEEE Sig. Process. Mag. - Spec. Issue Biom. Secur. Priv. **32**(5), 1–12 (2015)
20. Y Wang, S Rane, SC Draper, P Ishwar, A theoretical analysis of authentication, privacy, and reusability across secure biometric systems. Trans. Inf. Forensics Secur. **7**(6), 1825–1840 (2012)
21. A Nagar, K Nandakumar, AK Jain, Multibiometric cryptosystems based on feature-level fusion. Trans. Inf. Forensics Secur. **7**(1), 255–268 (2012)
22. C Rathgeb, C Busch, in *New Trends and Developments in Biometrics*. Multibiometric template protection: issues and challenges (InTech, Rijeka, 2012), pp. 173–190
23. A Juels, M Sudan, in *Proc. IEEE Int'l Symp. on Information Theory*. A fuzzy vault scheme (IEEE, 2002), p. 408
24. G Davida, Y Frankel, B Matt, On enabling secure applications through off-line biometric identification. Proc. IEEE, Symp. Secur. Priv, 148–157 (1998)
25. G Davida, Y Frankel, B Matt, On the relation of error correction and cryptography to an off line biometric based identication scheme. Proc. of WCC99, Work. Coding Cryptogr, 129–138 (1999)
26. A Juels, M Wattenberg, in *Proc. 6th ACM Conf. on Computer and Communications Security*. A fuzzy commitment scheme (ACM, 1999), pp. 28–36
27. F Hao, R Anderson, J Daugman, Combining cryptography with biometrics effectively. IEEE Trans. Comput. **55**(9), 1081–1088 (2006)
28. J Bringer, H Chabanne, G Cohen, B Kindarji, G Zemor, Theoretical and practical boundaries of binary secure sketches. IEEE Trans. Inf. Forensics. Secur. **3**, 673–683 (2008)
29. National Institute of Standards and Technology (NIST): Iris Challenge Evaluation (ICE) 2005. http://www.nist.gov/itl/iad/ig/ice.cfm. (2005). Accessed Oct 2016
30. C Rathgeb, A Uhl, Context-based biometric key-generation for iris. IET Computer Vision (Spec. Issue Futur. Trends Biom. Process.), IET. **5**(6) (2012)
31. ABJ Teoh, J Kim, Secure biometric template protection in fuzzy commitment scheme. IEICE Electron. Express. **4**(23), 724–730 (2007)
32. L Zhang, Z Sun, T Tan, S Hu, in *In Proc. of the 3rd Int. Conf. on Biometrics 2009 (ICB'09) LNCS: 5558*. Robust biometric key extraction based on iris cryptosystem (Springer, 2009), pp. 1060–1070
33. ERC Kelkboom, J Breebaart, TAM Kevenaar, I Buhan, RNJ Veldhuis, Preventing the decodability attack based cross-matching in a fuzzy commitment scheme. Trans. Inf Forensics Secur. **6**(1), 107–121 (2011)
34. E Maiorana, P Campisi, A Neri, in *Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'14)*. Iris template protection using a digital modulation paradigm (IEEE, 2014), pp. 3759–3763
35. A Cavoukian, A Stoianov, in *Biometrics: Fundamentals, Theory, and Systems*. Biometric encryption: the new breed of untraceable biometrics (Wiley, New York, 2009)
36. A Stoianov, T Kevenaar, M van der Veen, in *2009 IEEE Toronto International Conference on Science and Technology for Humanity (TIC-STH)*. Security issues of biometric encryption (IEEE, 2009), pp. 34–39
37. C Rathgeb, A Uhl, Statistical attack against fuzzy commitment scheme. IET Biom. **1**(2), 94–104 (2012)
38. T Ignatenko, FMJ Willems, Information leakage in fuzzy commitment schemes. IEEE Trans. Inf. Forensics Secur. **5**(2), 337–348 (2010)
39. P Failla, Y Sutcu, M Barni, in *Proc. of the 12th ACM Workshop on Multimedia and Security. MM&Sec '10*. esketch: a privacy-preserving fuzzy commitment scheme for authentication using encrypted biometrics (ACM, 2010), pp. 241–246
40. K Simoens, P Tuyls, B Preneel, in *Proc. of the 30th IEEE Symposium on Security and Privacy*. Privacy weaknesses in biometric sketches (IEEE, 2009), pp. 188–203
41. I Buhan-Dulman, JG Merchan, E Kelkboom, in *Proc. of IEEE Workshop on Information Forensics and Security (WIFS)*. Efficient strategies for playing the indistinguishability game for fuzzy sketches (IEEE, 2010)
42. B Tams, Decodability attack against the fuzzy commitment scheme with public feature transforms. CoRR. **abs/1406.1154** (2014)
43. YJ Lee, K Bae, SJ Lee, KR Park, J Kim, in *Proc. of Second Int. Conf. on Biometrics*. Biometric key binding: fuzzy vault based on iris images (Springer, 2007), pp. 800–808
44. YJ Lee, KR Park, SJ Lee, K Bae, J Kim, A new method for generating an invariant iris private key based on the fuzzy vault system. Trans. Syst. Man, Cybern. Part B: Cybern. **38**(5), 1302–1313 (2008)
45. Biometrics Engineering Research Center (BERC): Iris database (version 1) (2008)
46. Chinese Academy of Sciences' Institute of Automation: CASIA Iris Image Database V3.0 — Interval. http://biometrics.idealtest.org (2002). Accessed Oct 2016
47. ES Reddy, I Ramesh Babu, in *8th Int. Conf. on Computer and Information Technology Workshops, 2008*. Performance of iris based hard fuzzy vault (IEEE, 2008), pp. 248–253
48. ES Reddy, I Ramesh Babu, in *Second Asia Int. Conf. on Modeling Simulation AICMS'08*. Authentication using fuzzy vault based on iris textures (IEEE, 2008), pp. 361–368
49. Chinese Academy of Sciences' Institute of Automation: CASIA Iris Image Database V1.0. http://biometrics.idealtest.org (2002). Accessed Oct 2016
50. Multimedia University: MMU Iris Image Database. http://pesona.mmu.edu.my/ccteo (2004). Accessed Oct 2016
51. K Nandakumar, AK Jain, in *2nd Int. Conf. on Biometrics: Theory, Applications and Systems (BTAS'08)*. Multibiometric template security using fuzzy vault (IEEE, 2008), pp. 1–6
52. AK Jain, S Prabhakar, A Ross, Fingerprint matching: data acquisition and performance evaluation. Michigan State Univ., Tech. Rep., TR99-14 (1999)
53. M Fouad, A El Saddik, J Zhao, E Petriu, in *Int. Systems Conf. (SysCon'11)*. A fuzzy vault implementation for securing revocable iris templates (IEEE, 2011), pp. 491–494
54. R Álvarez Mariño, FH Álvarez, LH Encinas, A crypto-biometric scheme based on iris-templates with fuzzy extractors. Inf. Sci. **195**, 91–102 (2012)
55. S Sowkarthika, N Radha, in *7th Int. Conf. on Intelligent Systems and Control (ISCO'13)*. Securing iris and fingerprint templates using fuzzy vault and symmetric algorithm (IEEE, 2013), pp. 189–193
56. J Zuo, NK Ratha, JH Connel, in *Proc. of the 19th Int. Conf. on Pattern Recognition 2008 (ICPR'08)*. Cancelable iris biometric (IEEE, 2008), pp. 1–4
57. S Kanade, D Petrovska-Delacretaz, B Dorizzi, in *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, CVPR '09*. Cancelable iris biometrics and using error correcting codes to reduce variability in biometric data (IEEE, 2009), pp. 120–127
58. J Hämmerle-Uhl, E Pschernig, A Uhl, in *Proc. of the 12th Int. Information Security Conf. (ISC'09). LNCS*, ed. by P Samarati, M Yung, F Martinelli, and CA Ardagna. Cancelable iris biometrics using block re-mapping and image warping, vol. 5735 (Springer, New York, 2009), pp. 135–142
59. JK Pillai, VM Patel, R Chellappa, NK Ratha, Secure and robust iris recognition using random projections and sparse representations. Trans. Pattern Anal. Mach. Intell. **33**(9), 1877–1893 (2011)
60. C Rathgeb, C Busch, Cancelable multi-biometrics: mixing iris-codes based on adaptive bloom filters. Comput. Secur. **42**, 1–12 (2014)
61. C Rathgeb, A Uhl, P Wild, in *Iris Biometrics. Advances in Information Security*. Cancelable iris biometrics, vol. 59 (Springer, New York, 2013), pp. 223–231
62. A Juels, M Sudan, A fuzzy vault scheme. Des. Codes Cryptogr. **38**(2), 237–257 (2006)
63. TC Clancy, N Kiyavash, DJ Lin, in *Proc. ACM SIGMM Workshop on Biometrics Methods and Applications. WBMA '03*. Secure smartcard-based fingerprint authentication (ACM, New York, NY, USA, 2003), pp. 45–52
64. K Nandakumar, AK Jain, S Pankanti, Fingerprint-based fuzzy vault: implementation and performance. IEEE Trans. Inf. Forensics Secur. **2**(4), 744–757 (2007)
65. A Nagar, K Nandakumar, AK Jain, A hybrid biometric cryptosystem for securing fingerprint minutiae templates. Pattern Recogn. Lett. **31**, 733–741 (2010)
66. WJ Scheirer, TE Boult, in *Proc. of Biometrics Symp*. Cracking fuzzy vaults and biometric encryption (IEEE, 2007), pp. 1–6
67. A Kholmatov, B Yanikoglu, in *Proc. SPIE*. Realization of correlation attack against the fuzzy vault scheme, vol. 6819 (SPIE, 2008)
68. B Tams, P Mihăilescu, A Munk, Security considerations in minutiae-based fuzzy vaults. IEEE Trans. Inf. Forensics Secur. **10**(5), 985–998 (2015)
69. Y Dodis, R Ostrovsky, L Reyzin, A Smith, Fuzzy extractors: how to generate strong keys from biometrics and other noisy data. SIAM J. Comput. **38**(1), 97–139 (2008)
70. J Merkle, B Tams, Security of the improved fuzzy vault scheme in the presence of record multiplicity (full version). CoRR (2013). https://arxiv.org/abs/1312.5225. Accessed Oct 2016

71. B Tams, Unlinkable minutiae-based fuzzy vault for multiple fingerprints. IET Biom. **5**(3), 170–180 (2016). IET Digital Library
72. S Gao, in *Communications, Information and Network Security, V.Bhargava, H.V.Poor, V.Tarokh, and S.Yoon*. A new algorithm for decoding reed-solomon codes (Dordrecht, 2002), pp. 55–68
73. V Guruswami, M Sudan, Improved decoding of reed-solomon and algebraic-geometric codes. IEEE Trans. Inf. Theory. **45**, 1757–1767 (1998)
74. M-H Lim, ABJ Teoh, J Kim, Biometric feature-type transformation: making templates compatible for template protection. IEEE Sig. Process. Mag. **32**(5), 77–87 (2015)
75. R Cappelli, A Lumini, D Maio, D Maltoni, Fingerprint image reconstruction from standard templates. IEEE Trans. Pattern Anal. Mach. Intell. **29**(9), 1489–1503 (2007)
76. W Dong, Z Sun, T Tan, Iris matching based on personalized weight map. IEEE Trans. Pattern Anal. Mach. Intell. **33**(9), 1744–1757 (2011)
77. Indian Institute of Technology Delhi (IITD): IITD Iris Database version 1.0 (2007). http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Iris.htm. Accessed Oct 2016
78. A Uhl, P Wild, in *Proc. 5th Int'l Conf. on Biometrics*. Weighted adaptive hough and ellipsopolar transforms for real-time iris segmentation, (2012), pp. 1–8
79. L Masek, *Recognition of human iris patterns for biometric identification. Master's thesis*. (University of Western Australia, 2003)
80. L Ma, T Tan, Y Wang, D Zhang, Efficient iris recognition by characterizing key local variations. IEEE Trans. Image Process. **13**(6), 739–750 (2004)
81. USIT – University of Salzburg Iris Toolkit. http://www.wavelab.at/sources/Rathgeb12e. version 1.0.x. Accessed Oct 2016
82. 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 (2006). International Organization for Standardization and International Electrotechnical Committee
83. P Mihăilescu, A Munk, B Tams, in *Proc. of BIOSIG*. The fuzzy vault for fingerprints is vulnerable to brute force attack (IEEE, 2009), pp. 43–54
84. M Blanton, M Aliasgari, Analysis of reusability of secure sketches and fuzzy extractors. IEEE Trans. Inf. Forensics Secur. **8**(9), 1433–1445 (2013)
85. P Vandewalle, J Kovacevic, M Vetterli, Reproducible research in signal processing - What, why, and how. IEEE Sig. Process. Mag. **26**(3), 37–47 (2009)
86. AJ Mansfield, JL Wayman, Best practices in testing and reporting performance of biometric devices. Technical report. U.K. Goverment Biometrics Working Group (August 2002)