Jurnal Teknologi

Full Paper

OPTIMIZED TRANSFORMATION AND MATCHING FOR IRISCODE BASED CANCELABLE BIOMETRICS SYSTEMS

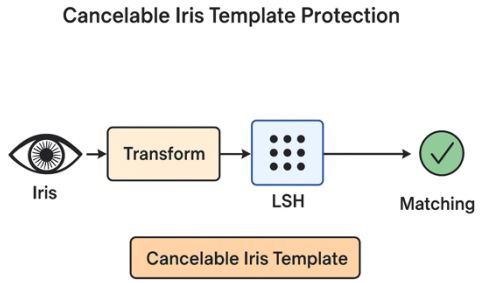
Tong-Yuen Chai[a], Muhammad Jahangir[a], Lim Ean Heng[a*], Cheang Kah Wai[b], Teik-Boon Tan[a]

[a]Department of Computer Science, Faculty of Information and Communication Technology, UTAR, 31900, Kampar, Perak, Malaysia
[b]Department of Information Systems, Faculty of Information and Communication Technology, UTAR, 31900, Kampar, Perak, Malaysia

Article history
Received
7th March 2025
Received in revised form
22th September 2025
Accepted
26th October 2025
Published
1st November 2026

*Corresponding author
ehlim@utar.edu.my

Graphical abstract



Abstract

Iris recognition is a reliable biometric model for identity identification, but reliability issues and cross-matching concerns can affect its accuracy. This research proposes a secure and cancelable iris template protection scheme to address these issues while maintaining system performance, privacy, and security. The scheme includes a matching technique for speed optimization and a transformation technique using locality sensitive hashing (LSH). The proposed framework improves the variance of original hamming distance in Iriscodes, discriminates between real and fake matches, and reduces vertical dependency, achieving an Equal Error Rate (EER) of 0.46% in empirical tests on CASIA v3-internal iris databases.

***Keywords*:** Cancelable biometrics, iris, template protection, security, locality sensitive hashing.

Abstrak

Pengumpulan maklumat mengenai aliran trafik adalah penting bagi Pengenalan iris adalah model biometrik yang boleh dipercayai untuk pengenalan identiti, tetapi isu kebolehpercayaan dan kebimbangan padanan silang boleh menjejaskan ketepatannya. Penyelidikan ini mencadangkan skema perlindungan templat iris yang selamat dan boleh dibatalkan untuk menangani isu-isu ini sambil mengekalkan prestasi sistem, privasi, dan keselamatan. Skema ini termasuk teknik pemadanan untuk pengoptimuman kelajuan dan teknik transformasi menggunakan penghashan sensitif lokasi. (LSH). Rangka kerja yang dicadangkan meningkatkan varians jarak hamming asal dalam Iriscodes, membezakan antara padanan sebenar dan palsu, serta mengurangkan kebergantungan menegak, mencapai EER sebanyak 0.46% dalam ujian empirik pada pangkalan data iris dalaman CASIA v3.

***Kata kunci*:** Pembatalan biometric, iris, perlindungan templat, keselamatan, hashing sensitive lokaliti.

© 2026 Penerbit UTM Press. All rights reserved

DOI: https://doi.org/10.11113/jurnalteknologi.v89.24239

1.0 INTRODUCTION

In biometrics, the physiological or behavioural characteristics of people are used for automated identification and verification. Throughout a person's lifetime, the iris does not change and is unaffected by genetic or environmental influence [1]. Daugman is recognised as the creator of the widely used iris model referred as Iriscode and as a pioneer in the practical use of automatic recognition of iris systems [2]. Furthermore, compared to other biometric features, iris patterns exhibit much higher entropy, which lowers the likelihood of erroneous matches between Iriscodes [3]. Because of its distinctive qualities, the iris is useful for identification and verification. However, the possibility of the Iriscode being made available to adversaries raises privacy issues. These problems must be resolved in order to maintain the reliability of the system [4-6]. The subject of biometric template protection (BTP) technology was developed to reduce such hazards [7,8] has been developed. Four key criteria should be addressed by an efficient biometric template protection system which is non-invertibility or irreversibility, renewability or revocability, non-linkability or unlinkability and performance preservation.

Two primary BTP (Biometric Template Protection) techniques have been proposed: feature transformation (cancelable biometrics) and biometric cryptosystems [9]. Feature transformation involves applying a function to the biometric template, generating a transformed version that is stored in the database. The parameters of the transformation function are commonly taken from a random key. The query features are subjected to the same transformation function during matching, and the converted query is contrasted with the altered template in the database. However, a biometric cryptosystem creates the safe biometric template by securely merging a secret key with biometric information. Alternately, it creates a cryptographic key solely from biometric data in a manner that makes it impossible to recover the biometric data or original key from the guarded sample. When the true biometric inquiry is submitted, the key can be obtained. Along with securing biometric templates, biometric cryptosystems also function as secret key managers.

The design of a robust cancelable transformation function continues to pose significant challenges, particularly in relation to alignment. This issue is especially pronounced in the context of Iriscodes, where the matching process necessitates iterative operations such as left and right bit shifts on the query Iriscode to accommodate intra-class variations. Such adjustments are indispensable for identifying the optimal correspondence among possible distortions induced by head tilt or cyclovergence [10]. Despite advances in acquisition techniques aimed at mitigating external constraints, residual alignment errors caused by minor rotational deviations remain unavoidable. In order to preserve non-invertibility within biometric template protection (BTP) schemes, additional mechanisms such as the incorporation of random matrices in bio-hashing (RMH) are frequently employed. However, in the absence of trustworthy pre-alignment, adding cancelable transformations with additional randomization can reduce intra-class consistency and inter-class separability, which will negatively impact recognition accuracy. This limitation has been observed across a range of contemporary iris-based BTP methodologies [11–13].

In order to avoid any potential compromise of the probing Iriscode during the matching process, it is important to deal with the requirement of pre-alignment in iris template protection. An ideal iris template protection strategy should not need any pre-alignment steps and should only use the protected iris template during the matching stage. Rathgeb et al. [14] presented an alignment-free cancelable transformation approach that demonstrates encouraging recognition performance. In a related direction, Chai et al. [27] proposed an alignment-free chaffed cancelable iris key binding scheme that addresses pre-alignment concerns while preserving security and matchability; their approach is particularly relevant to alignment-free BTP designs.

This strategy, however, represents security issues since it depends on a key space that is just 10 bits. Therefore, one of the biggest challenges for the deployment of safe iris identification systems in the future is to provide an alignment-free cancelable transformation for the human iris. Our suggested approach addresses the original Iriscode's pre-alignment requirement during the matching phase. It is obvious that the recognition performance would be subpar if different cancelable modifications were applied straight to the original Iriscode without accounting for the vertical dependence [15]. Although the mechanics should simply accept the protected iris template throughout the comparison process, rather the need to fall back to the raw Iriscode or a sequence of shifted Iriscodes, to establish a safe Iriscode-based biometric template protection (BTP) method. We provide two transformation and matching techniques in this work that meet this criterion. Our method's primary novelty is the immediate application of alignment to the protected template, which avoids the need to go back to the original Iriscode during the matching step. We make use of the vertical interdependence and horizontal independence of Iriscode's structure.

The findings of our study show that the human Iriscode has unique properties, particularly in the

DOI: https://doi.org/10.11113/jurnalteknologi.v89.24239

horizontal direction where the bits have a higher information value [16]. The iris texture, on the other hand, has natural correlations along the radial direction, which results in more reliance and less discriminability in vertically aligned Iriscodes. We provide a column-wise transformation strategy to address this issue and get along the requirement for pre-alignment during matching. With this method, the negative consequences of vertical reliance on recognition performance are reduced. Furthermore, we add a locality-sensitive hashing (LSH) based mapping function to the matching process, allowing us to concatenate many columns in the translated Iriscode. Our tests show that the converted Iriscode's recognition performance is enhanced and the vertical reliance is successfully reduced by this mapping function. Additionally, our suggested change greatly expands Iriscodes' key space to at least 40 bits without considerably degrading their original speed.

Further, we experimented with the suggested matching technique. The outcomes of these tests present that improving the matching effectiveness of the initial Iriscode can result in improved decidability scores. The measure of decidability [17] indicates how much a change, such as a decrease in the incorrect acceptance error rate, is countered by a rise in the false rejection error rate. With the suggested matching method, it is also feasible to improve the matching performance of the original Iriscode while getting a better decidability score and a lower equal error rate (EER) of approximately 0.5%. This has been confirmed through empirical experimental outcomes. Moreover, Chai et al. [28] demonstrated improved performance for protected iris systems using a secure hash algorithm (SHA) based confidence matrix, highlighting practical ways to boost recognition while maintaining template protection. Beside this conventional hashing and cryptography approaches, deep-learning solutions such as multi-head attention (MHA) and augmented random mapping (ARM) are also increasingly being adopted to improve iris recognition systems. MHA is a mechanism which enables models to capture long-range dependencies in iris texture data while ARM provides stronger cancelable transformations by mapping feature into randomized domains.

The suggested matching mechanism performed better than expected according to the experiments' results, which showed greater decidability scores. The decidability measure [18] determined the balance between raising the false rejection error rate and lowering the false acceptance error rate. The original Iriscode's matching performance was improved by the suggested matching technique, which led to greater decidability ratings. Additionally, the empirical testing findings showed that the suggested technique improved the overall matching performance of the original Iriscode while enabling a false acceptance rate of less than 0.5%.

2.0 METHODOLOGY

The basic premise of Locality Sensitive Hashing (LSH) is that when two points are close together, they tend to remain that way even after a projection operation. It acts as a hashing method for dimensional reduction, where related high-dimensional data is frequently and almost always mapped into the same bucket. After the hashing operation, the size of the buckets shrinks in comparison to the input data. LSH and traditional cryptographic hashing have different goals: LSH wants to increase the likelihood that comparable pieces of data will come together, whereas traditional cryptographic hashing wants to reduce the likelihood of collisions. The official definition of LSH [19] can be defined as follows:

$$P_{h\in H}\left(\left(h_i(\omega) = h_i(\omega')\right) \le P_1, \;\; if\; S(\omega,\omega') < R_1\right)$$
$$P_{h\in H}\left(\left(h_i(\omega) = h_i(\omega')\right) \ge P_2, \;\; if\; S(\omega,\omega') > R_2\right) \quad (1)$$

This pair of inequalities formalizes the core idea of LSH. The first inequality states that if two data points, ω and ω ' , are considered dissimilar (their similarity S (ω, ω ') is less than a threshold R_1), the probability of a hash collision is low, bounded by a value of P_1. Conversely, the second inequality asserts that if the two points are similar (their similarity is greater than a threshold R_2 , where $R_2 > R_1$), the probability that a random hash function from the family H maps them to the same bucket is high, bounded by a value of P_2 , where $P_2 > P_1$. This relationship ensures that the algorithm prioritizes collisions for truly similar items.

Consider the scenario $P_2 > P_1$, which combines input data M with two random inputs $\omega, \omega' \in M$, and H h family of hash functions. The locality sensitive hashing structure uses a similarity function S to determine which local hash functions to apply to the input data $H = h_i.M \to U$, and which output U mentions the hashed metric space. The scheme can be viewed as a probability distribution for a set of hash functions with the property that $P_{h\in H}[h(\omega) = h(\omega')] = S(\omega,\omega')$. The likelihood of a collision between the two hashed input data ω and ω^' was specifically mentioned by the similarity function S. LSH not only offers local hashing functions but also makes sure that two similar inputs will collide more frequently and be routed to the same bucket. However, a low risk of collision will be used to turn various inputs into hashes, guaranteeing that they are mapped into unique buckets.

The random bit sampling approach is one of the most successful techniques for building an LSH family [20]. Particularly, given a binary string $x \in \{0,1\}^k$, one create LSH family $H = \{h: \{0,1\}^k \to \{0,1\}$subject to the sampling function $h(x) = x_i$ where $i \in \{1, \dots, n\}$ is the index (location) chosen randomly over n indices while x_i refers to the i-th symbol of x. Consequently, this method produces a random binary string $v \in \{0,1\}^n$ using an LSH family of sampling functions $H = \{h_1, h_2, \dots, h_n\}$ .The expectation is that each

DOI: https://doi.org/10.11113/jurnalteknologi.v89.24239

sampling function will provide a single symbol value of v which is equivalent to the x_i, for instance, $v = [h_1(x), h_2(x), \ldots, h_n(x)]$. The LSH family may then be created using the following method for hamming distance:

$$\Pr[h(x) = h(y)] = 1 - (\frac{\|x \oplus y\|}{k})$$

This formula states that the probability of a hash collision for inputs x and y is inversely proportional to their Hamming distance. As the Hamming distance increases, the likelihood of a collision decreases.

The specifics of our suggested transformation for producing cancelable iris templates are presented in this portion. LSH sampling functions family is represented as $H = \{h_1, h_2, \ldots, h_n\}$. Let $x \in \{0,1\}^{a \times b}$ be the input Iriscode of size $a \times b$ where a the number is of rows and b is the number of columns. Let $m > 0$ be an integer subject to $n|m$ (i.e., n is divisible by m). Our suggested conversion for cancellable iris templates can be explained by function $F(x, H, m, n)$ with the following steps:

Function: $\boldsymbol{F(x, H, m, n)}$:

Step 1: This step randomly samples information from each column to reduce dimensionality while preserving discriminative features. Denote c_i as the i-th column of x, for example $c_1 = \{1\ 0\ 0\ 0\}$. Perform bit sampling strategy with $H = \{h_1, h_2, \ldots, h_n\}$ to produce hashed outputs $v = (v_1, \ldots, v_b)$ of every column. At this point, $v_i \in \{0,1\}^n$ denotes the output vector corresponds to then sampling functions for i-th column of x, i.e., $H(c_i) = v_i$.

Step 2: The goal of this step is to group the sampled bits into manageable chunks and represent them in decimal form for compactness. For each v_i, set $r = n/m$ and divide v_i into r-tuple sub-strings (each sub-string will contain m bits). At that point, transform the sub-strings to their particular decimal values. This leads to a mapping of $v_i \rightarrow T_i \in [0, 2^m - 1]^r$.

Step 3: This final step ensures security by making the template non-invertible. Construct the template as $T = [T_1, \ldots, T_b] \in [0, 2^m - 1]^{r \times b}$. Perform element-wise one-way hashing on T with a randomly generated matrix (salting). Here, a mixture of random salt and a decimal number constitutes the final output for each element in T.

Each stage of our suggested transformation model for cancellable iris recognition is shown in Figure 1.

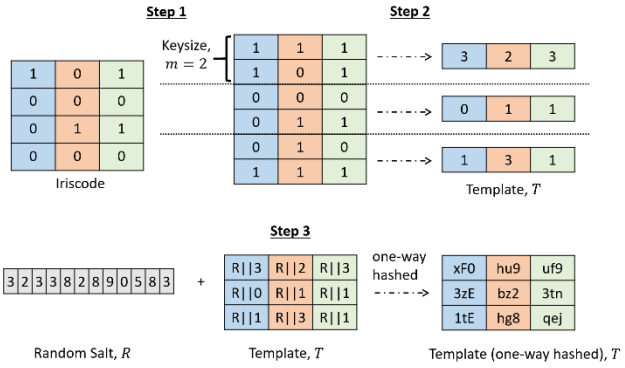


Figure 1 A case study of the suggested alteration is m=2 and n=6

Convert the two provided iris templates $(T, T') \in [0, 2^m - 1]^{r \times b}$, the similarity between T and T' is possible by calculating their hamming distance, i.e., the number of symbols where T and T' are unalike. Matching for the changed templates may be carried out iteratively by using several left and right shifting's, much like the conventional Iriscode. These adjustments are made to accommodate the various Iriscodes' varying orientations. Therefore, choosing the match having the lowest error rate will result in the best match for T'.

Based on the research in [21-23], Iriscodes have built-in correlations in the radial direction since the furrow or ciliary pattern tends to affect the vertically adjacent bits. The discriminative data of the iris texture is commonly dispersed in the horizontal plane. According to the columns of the unwrapped human iris in its 2D-polar equivalent space, as observed using Daugman's rubber sheet model, it is hypothesized that the vertically adjacent bits in an iriscode are dependant [24]. In light of these findings, additional limitations are added to the matching approach between two distinct altered templates (T, T'). With the help of these restrictions, the altered iris template's vertical reliance will be reduced, enhancing security and system performance.

Using the saved iris template following the suggested transformation $T \in [0, 2^m - 1]^{r \times b}$and a query Iriscode$x' \in \{0,1\}^{a \times b}$. The recommended matching technique can be addressed by using a function$F'(T, x', H, m, n, n_c, q, \tau)$with T, x', H as the inputs and integers m, n, n_c, q, τ as the parameters where $b|n_c$.

Function$\boldsymbol{F'(T, x', H, m, n, n_c, q, \tau)}$:

Step 1 Concatenate Tuples: Divide the stored iris template T into b/n_c–tuples, respectively composed of n_c columns of T. Concatenate all n_c fields in each tuple to create a single data column. This operation causes the converted template's dimension to

DOI: https://doi.org/10.11113/jurnalteknologi.v89.24239

change from $T \in [0, 2^m - 1]^{r \times b}$ to $T \in [0, 2^m - 1]^{(rn_c) \times (b/n_c)}$.

Step 2 Compute Hamming Distances: Make conversion of $F'(x', H, m, n) \rightarrow T'$ onto the query Iriscode to receive the modified iris template $T' \in [0, 2^m - 1]^{r \times b}$.

Step 3 Probability of Match: Concatenation should be repeated in step 1 for T' to create a particular information column. This alters the dimension of the transformed query template to $T' \in [0, 2^m - 1]^{(rn_c) \times (b/n_c)}$.

Step 4 Iterative Scoring: Initialize a score count, $z = 0$. Let T_i and T_i'relate to the same bit position at i-th column of T and T'correspondingly the$T_i \in T$ and $T_i' \in T'$.Calculate the hamming distance $\|T_i \oplus T_i'\|$, refers as integers of various signs among T_i and T_i'. If $\|T_i \oplus T_i'\| \leq \tau$, set $z = z + 1$. Otherwise, do nothing.

Step 5 Final Similarity Score: Set T' as the bit-shifted transformed template. Subsequently placing the bit-shifting technique (see [33], Section 4.4), repeat steps 3 & 4 for q number of times (q is equivalent to the number of bit-shifting applied), result in a set of score counts $(z_1, \ldots, z_q)$. Stabilize the last comparison score $s \in [0,1]$ as $s = max(z_1, \ldots, z_q)/(b/n_c)$.

The phases of our suggested matching technique for cancelable iris recognition are highlighted in Figure 2.

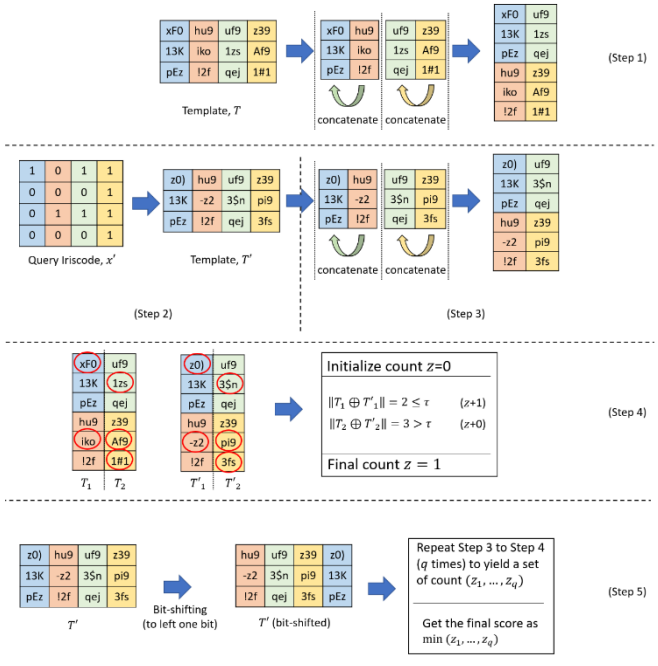


Figure 2 An illustration of the suggested matching method with $m = 2, \tau = 2, n = 6, n_c = 2$ and q= 1

Given a bit sampling function $h \in H$. The i-th column of the Iriscodes x and y are denoted as $c_i \in x$ and $c_i' \in y$ respectively. If x and y are two different iriscodes, a dissimilarity of $\varepsilon_d = \|c_i \oplus c_i'\|/k$ will be obtained. Referring to Eq. 2, The likelihood of the bits sampling function aligning may be calculated as:

$$h(c_i) \neq h(c_i') = \varepsilon_d \tag{1}$$

Following our suggested transformation, the stored iris template and the query iris template are displayed by $T$and T'. Any symbol at the i-th column of the transformed templates T and T' shall consist of m bits as shown in Figure 1. Considering that a symbol's likelihood of being different across two modified templates (in the same column) relies on m independent bit sampling functions $(h_1, \ldots, h_m)$ can be further described as:

$$(h(c_i) \neq h(c_i'))^m = (\varepsilon_d)^m \tag{2}$$

It might be argued that various pairs of columns inside the Iriscodes x and y could lead to various distances, i.e., $\|c_i \oplus c_i'\| \neq \|c_j \oplus c_j'\|$ for $(i \neq j)$ mapping the transformed template from $(T, T') \in [0, 2^m - 1]^{r \times b}$ to $(T, T') \in [0, 2^m - 1]^{(rn_c) \times (b/n_c)}$ (refer to *Step 1* and *Step 3* under the function $F'(x', H, m, n)$).To analyze this issue, determine the shortest area amongst the i-th columns of (T, T')

$$\|T_i \oplus T_i'\| \geq \min((\|c_1 \oplus c_1'\| /k)^m, \ldots, (\|c_{n_c} \oplus c_{n_c}'\|/k)^m) rn_c \tag{3}$$

$$\tag{4}$$

However, the greatest difference between the i-th columns of (T, T') follows:

$$\|T_i \oplus T_i'\| \geq \max((\|c_1 \oplus c_1'\|/ /k)^m, \ldots, (\|c_{n_c} \oplus c_{n_c}'\|/k)^m) rn_c \tag{5}$$

Equations 4 and 5's modest computed value for $\|T_i \oplus T_i'\|$ imply a greater value of z (see *Step 4*'s discussion of the function $F'(T, x', H, m, n, n_c, q, \tau)$) with the knowledge that Iriscodes *(x,y)* demonstrate a short pair-wise hamming distance over their columns. Whereas, given that the input Iriscodes demonstrate significant pair-wise hamming distance over their columns, a high value for $\|T_i \oplus T_i'\|$ predicts a smaller score count of z. This supports our claim that the suggested matching mechanism's distance-preserving characteristic is true. This characteristic assures high similarity scores for Iriscodes with low pairwise distances over their columns, a measure of

DOI: https://doi.org/10.11113/jurnalteknologi.v89.24239

similarity. On the other hand, as previously stated in the prior section, Iriscodes with tiny pairwise Hamming distances over their columns, suggesting dissimilarity, would produce lower similarity scores. Let:

$$\varepsilon_{min} = \min\left(\left(\|c_1 \oplus c_1'\|/k\right)^m, \ldots, \left(\|c_{n_c} \oplus c_{n_c}'\|/k\right)^m\right)$$

and

$$\varepsilon_{max} = \max\left(\left(\|c_1 \oplus c_1'\|/k\right)^m, \ldots, \left(\|c_{n_c} \oplus c_{n_c}'\|/k\right)^m\right)$$

respectively. In terms of any random $\varepsilon^* > 0$, n_c, and r follows that $\tau = rn_c\varepsilon^*$, the parameter can be stated. It means $\tau \in [rn_c\varepsilon_{min}, rn_c\varepsilon_{max}]$ for any value of $\varepsilon^* \in [\varepsilon_{min}, \varepsilon_{max}]$. The probability result of the matching may be represented by a binomial distribution, as detailed below, with a mean equal to and variance equal to $\tau(1-\varepsilon^*)$, as the random bit sampling for the individual columns of the Iriscode is independent, leading to:

$$\Pr[\|T_i \oplus T_i'\| \leq \tau] = \sum_{i=1}^{rn_c} \binom{rn_c}{i} (\varepsilon^*)^i (1-\varepsilon^*)^{rn_c - i}. \quad (6)$$

This formula is binomial cumulative probability of observing at most τ differing bits out of r n_c independent comparisons when each bit differs with probability ϵ^*.

The solution for $\Pr[\|T_i \oplus T_i'\| \leq \tau]$may be determined with a suitable value of $\varepsilon^* \in [\varepsilon_{min}, \varepsilon_{max}]$, this suggests a suitable value being used of $\tau \in [rn_c\varepsilon_{min}, rn_c\varepsilon_{max}]$. More specifically, the relationship between $\Pr[\|T_i \oplus T_i'\| \leq \tau]$ versus ε^* is illustrated in Figure 3.

The analysis and its formulation have shown that selecting an appropriate value for τ ensures that Iriscodes with small pairwise Hamming distances over their columns (at most τ) have a high probability of yielding an expected similarity score (s) that is extremely close to one.

Only Eq. 6 has a solution that gives the numerical value of$\tau \in [rn_c\varepsilon_{min}, rn_c\varepsilon_{max}]$. In terms of the pair-wise input column distance distribution of Iriscodes (x, y), for instance, $\varepsilon_{min}$and ε_{max} must be known. Ideally, we wish that $|\varepsilon_{max} - \varepsilon_{\min}|$may be as vast as feasible so that a larger variety ofτ can be chosen for the solution of Eq. 6.

Nevertheless, it is impractical to assume that $|\varepsilon_{max} - \varepsilon_{\min}|$ may be accurately calculated without having perfect knowledge of how the input Iriscodes' columns' pair-wise distances are distributed. A naïve approach to maximize the range of $|\varepsilon_{max} - \varepsilon_{\min}|$ through increasing the value of $n_c$are more applicable which will lead to additional columns being concatenated? As a result, the variability of the pair-wise distance distribution of the columns will rise as $|\varepsilon_{max} - \varepsilon_{\min}|$.

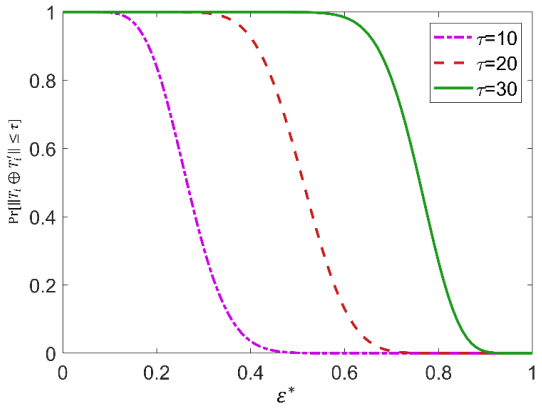
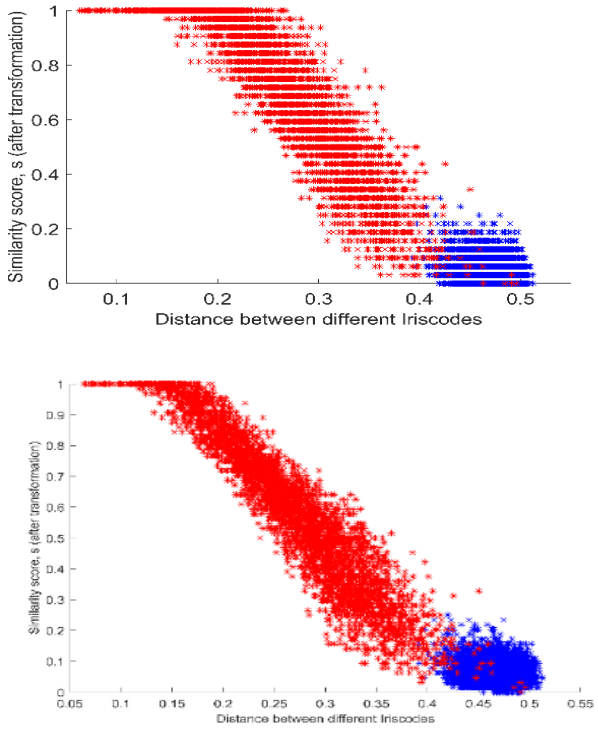


Figure 3 The relation of $\Pr[\|T_i \oplus T_i'\| \leq \tau]$ vs ε^*

We presented the relationship between the output similarity score (s) and the normalised original Hamming distance [0,1] of several Iriscodes in Figure 4 to further illustrate the effect of n_c on the variance of the pair-wise distance distribution of the input Iriscodes. The findings demonstrate that a higher n_c number results in a wider dispersion of points along the x-axis. This increasing spread suggests that the input Iriscodes' pair-wise distance variation for the columns has also risen. As a consequence, different Iriscodes may be identified based on the pair-wise distances between their columns, following the probability distribution stated in Eq and seen in Figure 3.



DOI: https://doi.org/10.11113/jurnalteknologi.v89.24239

Figure 4 Comparison of the score for similarity, s compared to the original normalised hamming distance of several Iriscodes with the same parameter configuration of $m = 10$, and $n = 50$

3.0 RESULTS AND DISCUSSION

The research makes the use of 2639 iris photographs from 396 distinct people in the CASIA v3-interval iris collection. Matching each template with templates from other iris samples of the same individual produces 4406 genuine comparisons for intra-class comparison. Inter-class comparisons create impostor comparisons by matching each template with the initial template of several subjects, producing 199110 imposter comparisons. Metrics used to evaluate recognition performance include the FAR (false acceptance rate) and FRR (false rejection rate). FAR stands for imposter samples that are mistakenly identified as genuine users, whereas FRR is for genuine samples that are mistakenly identified as imposters. Evaluation also uses the Equal Error Rate (EER), which signifies the error rate when FAR equals FRR.

Table 1 Performance of our proposed scheme with varying τ and m

τ		99	95	90	80	60	40	20
$m = 5$	FRR,%	12.4	6.50	2.33	0.65	0.61	1.15	3.54
	FAR,%	33.4	6.27	2.00	0.78	0.56	1.00	1.85
$m = 10$	FRR,%	3.08	3.08	0.68	0.65	0.86	2.33	4.26
	FAR,%	3.04	0.74	0.79	0.93	2.76	7.04	3.97
$m = 15$	FRR,%	0.93	0.83	1.27	2.65	7.73	8.55	24.5
	FAR,%	1.05	0.86	1.59	2.64	3.73	7.91	0.97
$m = 20$	FRR,%	0.70	1.18	2.22	3.42	7.12	14.5	37.0
	FAR,%	0.74	1.30	1.70	4.10	8.57	4.46	0.52

Experiments are conducted with 20×512 Iriscodes as the input, various values of, expanding key sizes, m=5,10,15,20, and constant r=50. In order to account for the left (-) and right (+) shifting of bits from positions $-4, -3, -2, -1, 0, 1, 2, 3, 4$, we set n_c=2 and q=9. The results are shown in Table 1, which show that choosing the right parameters is necessary to give the recommended system a promising recognition performance with low FRR and FAR.

The lowest EER (achieved) for each parameter value examined is shown in Table 2. As "m" increases, performance is seen to somewhat decrease. The necessity for greater "m" values to improve non-invertibility and subsequently, stronger security, is considered to be the cause of this trade-off between security and recognition. The non-invertibility of our proposed method is guaranteed by the one-way hashing of the modified template (T, T'). We propose a novel approach based on cryptographic one-way hashing. During the matching procedure between modified templates, the number of non-collided entries among the pair-wise (columns) of the updated template is counted. To calculate the overall similarity matching score, only the column of the query templates that shows the majority of "non-collided items," i.e., $\|T_i \oplus T_i'\| \leq \tau$, will be counted (z).

Table 2 Performance of our proposed scheme with different parameter settings

n_c	r	Performance EER,%				
		$m = 5$	$m = 10$	$m = 15$	$m = 20$	$m = 40$
4	50	**0.46 (τ = 125)**	0.69 (τ = 180)	0.76 (τ = 190)	0.79 (τ = 199)	1.97 (τ =199)
	100	0.49 (τ = 250)	0.69 (τ = 380)	0.75 (τ = 385)	0.84 (τ = 390)	1.67 (τ =399)
	200	0.48 (τ = 500)	0.71 (τ = 780)	**0.71 (τ = 790)**	**0.79 (τ = 790)**	1.36 (τ =799)
	800	0.70 (τ = 3120)	0.88 (τ = 3140	0.78 (τ = 3160)	0.81 (τ = 3180)	**0.92 (τ =3199)**
8	50	0.58 (τ = 250)	0.71 (τ = 360)	0.81 (τ = 385)	1.08 (τ = 396)	2.32 (τ =199)
	100	0.56 (τ = 650)	0.81 (τ = 760)	0.91 (τ = 770)	1.13 (τ = 785)	1.62 (τ =399)
	200	0.56 (τ = 1000)	**0.66 (τ = 1480)**	0.85 (τ = 1540)	1.00 (τ = 1580)	1.32 (τ =799)
	800	0.60 (τ = 3120)	0.70 (τ = 3140)	0.88 (τ = 3160)	1.11 (τ = 3180)	1.36 (τ =3199)
16	50	0.55 (τ = 500)	0.84 (τ = 720)	0.98 (τ = 780)	1.18 (τ = 790)	2.50 (τ = 3299)
	100	0.58 (τ = 1000)	0.96 (τ = 1520)	1.01 (τ = 1555)	1.24 (τ = 1575)	2.55 (τ = 3299)
	200	0.58 (τ = 2000)	0.67 (τ = 2960)	1.05 (τ = 3130)	1.24 (τ = 3160)	2.36 (τ = 3299
	800	0.74 (τ = 3120)	0.70 (τ = 3140)	1.18 (τ = 3160)	1.31 (τ = 3180)	2.70 (τ = 3299

A crucial prerequisite must be established in order for a hacker to successfully reverse the original Iriscode. The attacker needs to locate the hashed entry's m-bit pre-image in order to recover the original Iriscode. For instance, the binary equivalent of the initial entry of the transform template x=7 (where m=3) may be 111, which represents various bits in the original Iriscode. Since x is one-way hashed, or hash(x), the attacker must look for a pre-image x' such that hash(x')=hash(x). Given x', reverse permutation' using the available bit-sampling techniques can swiftly reverse m=3 bits of the original Iriscode information.

In light of the previously stated, the non-invertibility of our suggested approach depends on how difficult it is computationally to find a preimage of the one-way hash function for any given key size "m." To facilitate template matching, there is a dimensional compatibility condition that each modified template's key length must be the same. The one-way hashing function can only search for preimages after $O(2^m)$ trials, as a result of this limitation. Therefore, a bigger "m" is required to provide a convincing non-invertibility justification for our suggested system. We recommend using a random salt as an additional input to the one-way function to prevent attackers from constructing hash tables prior of time for all 2^m altered template elements. This approach strengthens defences against attacks using precomputed tables, such as

DOI: https://doi.org/10.11113/jurnalteknologi.v89.24239

rainbow tables [38] It is very difficult for attackers to build outsized hash tables for successful attacks without annoying users when utilising random salts in the one-way function. Due to their uniqueness, salts prevent the use of duplicate passwords by producing distinct salted hashes for the same converted entries [39].

For each modified template, a fresh random salt can be produced in a variety of systems or applications. It then enters a one-way cryptographic hash algorithm after being mixed with the template. It is not necessary to encrypt or store the salt without include the password hash. Even though an attacker uses salts and hash values to access the database, they still require a pre-image x' to learn anything about the original iriscode.

Correlation in between Iriscodes: The random distribution shown in biometric sources may not be uniform. There is a degree of resemblance through correlation while producing Iriscodes for various people. When n_c=1, it is difficult to determine whether the query Iriscode (column) in Figure 5(a) belongs to the genuine user or an impostor. This suggests that attackers can try to reverse the original Iriscode while intercepting the converted template by taking advantage of the vertical dependence in Iriscodes. Attackers might use the suggested transformation and matching approach while sampling a random Iriscode to accomplish this. Since hashing exposes information, they could examine any data that overlapped both recorded data and queries templates and use reverse engineering to reconstruct the original Iriscode. In light of these attacks and the correlation between Iriscodes, the proposed scheme's invertibility is dependent upon the attacker's ability to identify an Iriscode that has at least one element that matches the stored template. It is crucial to ensure the attacker cannot retrieve a single or numerous colliding entries by randomly sampling an Iriscode in order to show resilience against this kind of attack. Reduce the amount of hash elements in the stored template (thereby minimising information loss) to ensure higher non-invertibility against coupled Iriscodes. This step guarantees strong defence against prospective assaults and improves the proposed scheme's overall security against various but associated Iriscodes.

The parameters should be minimised in favour of lowering the amount of hash entries. The performance of the suggested strategy in Table 3 below as an (EER,%) was recorded with various parameter values for $r = 5, 10, 20, 30$, with constant $\tau = r - 1$ and $n_c = 1$. The results show that, when "r" is minimised (with fixed values for and n_c), performance degrades, which leads to a trade-off between non-invertibility and system performance. Figures 6 show the average number of colliding items for each of the altered templates' respective columns (1 to 512). The number of colliding entries for various columns in the modified templates clearly shows many peaks, showing a significant association amongst those columns in the Iriscodes. An average of less than one colliding entry over the one-way hashed template can be achieved by using a smaller value of "r" to lessen this correlation. This tactic provides increased system security and aids in resisting the aforementioned assault.

Table 3 Performance of our proposed scheme with different parameter settings

Performance EER,%	r = 10	r = 15	r = 20	r = 25	r = 30	r = 35	r = 40
	0.92	0.87	0.83	0.66	0.55	0.53	0.48

The performance effectiveness of the proposed LSH-based framework can be further contextualized by comparing its performance with other established cancelable biometric methods, particularly in terms of Equal Error Rate (EER). The proposed framework achieves an EER of 0.46% on the CASIA v3-internal iris database. As shown in the table below, this performance is highly competitive with other published works:

Table 4 Performance of our proposed scheme with respect to its counterpart

No	Method	Biometric Modality	EER %
1	Proposed Framework	Iris	0.46%
2	Soliman et al. [29]	Iris	0.58%
3	Enhanced Logistic Map [30]	Iris	0.117%
4	Unimodal CBS (MVD) [31]	Fingerprint	1.77%

The proposed framework demonstrates a strong performance, with its 0.46% EER being better than or comparable to several other cancelable iris and multi-biometric systems. The results indicate that the LSH-based approach is a viable and effective method for creating secure and accurate cancelable biometric templates.

The LSH-based approach offers practical advantages for operational biometric systems. LSH reduces search complexity compared to brute-force matching, enabling sub-linear retrieval time in large-scale databases and making it suitable where linear matching is a bottleneck. There is a direct trade-off between speed and accuracy: the number of bit-shifting iterations q improves robustness to misalignment (increasing true positives) but also raises computation. One-way hashing with random salt strengthens security by preventing precomputed table (rainbow table) attacks and ensuring template non-invertibility. Finally, because our method operates on the protected template and does not require reverting to the original IrisCode for alignment, it simplifies deployment and improves user experience while reducing exposure of raw biometric data.

DOI: https://doi.org/10.11113/jurnalteknologi.v89.24239

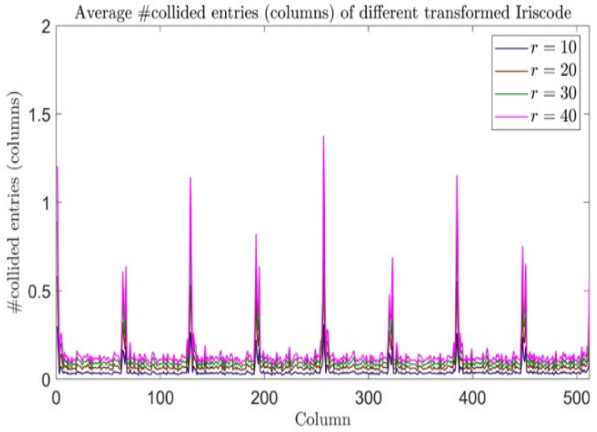
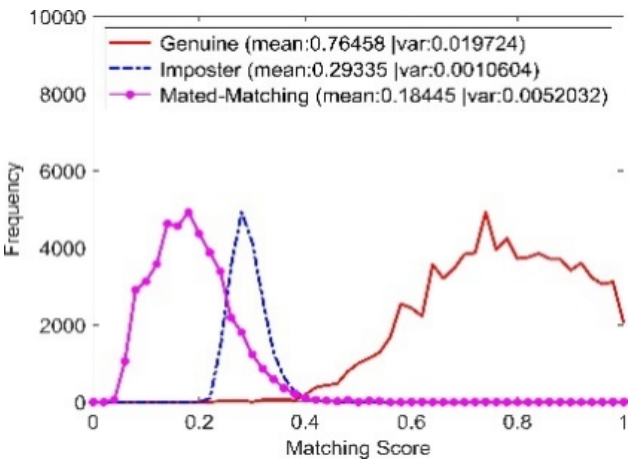


Figure 5 Correlation between different one-way hashed templates

We use the identical setup earlier to get 4406 mated-matching scores in order to assess the algorithm pair's revocability. These results are acquired by intra-class comparisons between the modified templates, each of which was created using a different sampling function from the LSH family. Different values of m = 10; 20; 30 are used to evaluate the revocability of the proposed system. The mated-scores distribution (included 4406 separate sets of H) and the genuine and imposter comparable score distribution via single sampling function set were shown together in a graph. Figure 7 shows three distinct graphs produced by various parameter values and constants $n_s = 2; r = 50; q = 9$. In this test, a considerable degree of overlap between the distributions of the imposter's similarity scores and the mated-matching scores is seen. This shows that despite being created from the same subject, the updated templates are sufficiently different. The newly altered samples, created with various combinations of H, perform more like imposter than genuine similarity scores. This result confirms that the suggested strategy is revocable since it successfully creates fresh altered templates to replace compromised ones.



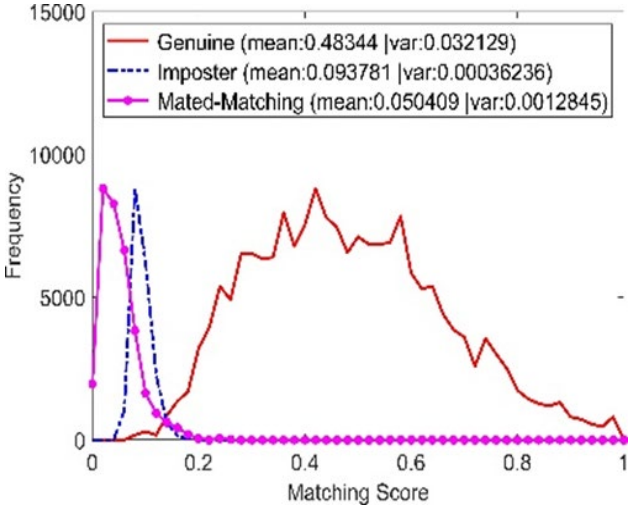


Figure 6 Evaluate the revocability of the proposed scheme

Numerous changed templates created using the identical method, Iriscode should be distinct from one another, as shown by unlinkability. To assess the unlinkability approach is suggested by Gomez et al. [25] is implemented. To assess the unlinkability, we use this method on the mated and non-mated score distributions. As opposed to non-mated scores, which require matching converted templates from many subjects with different sets of H, mated scores are achieved by matching converted templates from a single subject using different sets of H. When the score patterns of the mated and unmated groups coincide, the biometric system's unsink-ability characteristic is realized [26].

Let $P(s|M_s)$ be the known conditional probability of a similarity score $s \in [0,1]$ that is a member of the group of mated matches M_s and $P(s|M'_s)$ indicates the likelihood that a similarity score s under known conditions relate to the non-mated group M'_s. Two transformed templates are said to have linkage if it is more likely that two templates stem from mated sample (M_s) than from non-mated samples (M'_s) given a score s: $P(M_s|s) > P(M'_s|s)$. On the other hand, two transformed templates are said to be completely unlink-able (belong to different biometric sample) if $P(M_s|s) \leq P(M'_s|s)$. The unlink-ability character can be categorized by the local linkability: $D(s) = 2\frac{\omega LR(s)}{1+\omega LR(s)} - 1$, given that $\omega LR(s) = \frac{P(s|M_s)}{P(s|M'_s)} > 1$ where $LR(s)$ is the likelihood ratio between the known probabilities $P(s|M_s)/P(s|M'_s)$ and $\omega = P(M_s)/P(M'_s)$ denotes the proportion among the unidentified possibilities of the *mated samples* and *non-mated samples* distributions. We can assume that $P(M_s) = P(M'_s)$, thus set $\omega = 1$. The system's overall linkability can be further defined as:

DOI: https://doi.org/10.11113/jurnalteknologi.v89.24239

$$D_{sys} = \int D(s).P(s|M_s\ )ds \quad (7)$$

This measurement is within the range of $D_{sys} \in [0,1]$ represents zero, complete unlink-ability and singularity for systems that may be linked entirely. Consequently, it is essential to demonstrate that $D_{sys}$is negligibly tiny in order to achieve unlinkability of a BTP scheme. The four graphs in Figure 7 each contain 4406 mated-matching values and 199110 non-mated matching scores that were produced using various parameter values with constant$n_c = 2; r = 50; q = 9$ and (a) $\tau = 20,$ (b) $\tau = 1,$ (c) $\tau = 1$ respectively. Every score distribution, both mated and unmated, has extensive overlapping and a hardly tiny value of $D_{sys} = 0.01,0.03,0.01$ respectively. As a result, we claim that the suggested approach satisfies the unlinkability requirements.

4.0 CONCLUSION

Protected Iris recognition systems are becoming more and more necessary because of the increased need for reliable and safe identification verification. We provide an effective cancelable iris template protection system for increased security in this study. Through a new cancelable transformation and matching technique, our method overcomes pre-alignment and vertical dependence difficulties in Iriscodes. This makes it unnecessary to go back to the original Iriscode during the matching step and enables immediate bits-shifting onto the altered iris template. Our method maintains great performance with no degradation thanks to the robust collisions probability guaranteed by the core bit-sampling and Locality Sensitive Hashing (LSH) technology. This claim is supported by the good equal error rates (EERs) obtained on the CASIA v3-interval iris database for various parameter combinations. For Iriscode key sizes ranging from 5 to 40 bits, the best EER was 0.46. We analyse the security characteristics of our model, such as unlinkability, non-revocability, and invertibility and demonstrate how it can survive several attacks, such as SHA (Stolen Hash Attack), MHA (Mask Hill-Climbing Attack), and ARM (Attack via Record Multiplicity). In conclusion, the iris BTP translation and comparison method we propose provide strong theoretical security while retaining tolerable performance accuracy. We want to further develop and improve this model for future applications in multi-modal biometrics.

Acknowledgement

We would like to express our sincere gratitude to Universiti Tunku Abdul Rahman for providing the necessary resources and facilities to conduct this research.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Data availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Ethics declaration

Formal ethical approval was not obtained for this study. However, the authors confirm that the research was conducted in accordance with established ethical principles and good research practices, with due consideration given to participant rights, informed consent, confidentiality, and responsible data handling.

References

[1] Daugman, J. How iris recognition works. IEEE Transactions on circuits and systems for video technology 2004, 14, p.21-30.
[2] Daugman, J. Probing the uniqueness and randomness of iriscodes: Results from 200 billion iris pair comparisons. Proceedings of the IEEE 2006, 94, p.1927-1935.
[3] Venugopalan, S.; Savvides, M. How to generate spoofed irises from an iris code template. IEEE Transactions on Information Forensics and Security 2011, 6, p.385-395.
[4] Galbally, J.; Ross, A.; Gomez-Barrero, M.; Fierrez, J.; Ortega-Garcia, J. Iris image reconstruction from binary templates: An efficient probabilistic approach based on genetic algorithms. Computer Vision and Image Understanding 2013, 117, p.1512-1525.
[5] Cappelli, R.; Maio, D.; Lumini, A.; Maltoni, D. Fingerprint image reconstruction from standard templates. IEEE transactions on pattern analysis and machine intelligence 2007, 29, p.1489-1503.
[6] Natgunanathan, I.; Mehmood, A.; Xiang, Y.; Beliakov, G.; Yearwood, J. Protection of privacy in biometric data. IEEE access 2016, 4, p.880-892.
[7] Jain, A.K.; Nandakumar, K.; Nagar, A. Biometric template security. EURASIP Journal on Advances in Signal Processing 2008, 2008, p.113.
[8] Daugman, J.G. High confidence visual recognition of persons by a test of statistical independence. IEEE transactions on pattern analysis and machine intelligence 1993, 15, p.1148-1161.
[9] Lai, Y.-L.; Jin, Z.; Goi, B.-M.; Chai, T.-Y.; Yap, W.-S. Iris Cancellable Template Generation Based on Indexing-First-One Hashing. In Proceedings of International Conference on Network and System Security; p.440-463.
[10] Sadhya, D.; Raman, B. Generation of cancelable iris templates via randomized bit sampling. IEEE Transactions on Information Forensics and Security 2019, 14, p.2972-2986.
[11] Zuo, J.; Ratha, N.K.; Connell, J.H. Cancelable iris biometric. In Proceedings of Pattern Recognition, 2008. ICPR 2008. 19th International Conference on; p. 1-4.
[12] Rathgeb, C.; Breitinger, F.; Busch, C. Alignment-free cancelable iris biometric templates based on adaptive

DOI: https://doi.org/10.11113/jurnalteknologi.v89.24239

bloom filters. In Proceedings of Biometrics (ICB), 2013 International Conference on; p. 1-8.
[13] Hu, Y.; Sirlantzis, K.; Howells, G. Optimal generation of iris codes for iris recognition. IEEE Transactions on Information Forensics and Security 2016, 12, p.157-171.
[14] Liu, J.; Sun, Z.; Tan, T. Code-level information fusion of low-resolution iris image sequences for personal identification at a distance. In Proceedings of 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS); p. 1-6.
[15] Ratha, N.K.; Chikkerur, S.; Connell, J.H.; Bolle, R.M. Generating cancelable fingerprint templates. IEEE Transactions on pattern analysis and machine intelligence 2007, 29, 561-572, doi:10.1109/TPAMI.2007.1004.
[16] Quan, F.; Fei, S.; Anni, C.; Feifei, Z. Cracking cancelable fingerprint template of Ratha. In Proceedings of Computer Science and Computational Technology, 2008. ISCSCT'08. International Symposium on; pp. 572-575.
[17] Hämmerle-Uhl, J.; Pschernig, E.; Uhl, A. Cancelable Iris Biometrics Using Block Re-mapping and Image Warping. In Proceedings of ISC; p. 135-142.
[18] Chin, C.S.; Jin, A.T.B.; Ling, D.N.C. High security iris verification system based on random secret integration. Computer Vision and Image Understanding 2006, 102, p.169-177.
[19] Pillai, J.K.; Patel, V.M.; Chellappa, R.; Ratha, N.K. Sectored random projections for cancelable iris biometrics. In Proceedings of Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on; p. 1838-1841.
[20] Masek, L. Recognition of human iris patterns for biometric identification. 2003. University of Western Australia 2006.
[21] Kong, A.; Cheung, K.-H.; Zhang, D.; Kamel, M.; You, J. An analysis of BioHashing and its variants. Pattern Recognition 2006, 39, p.1359-1368.
[22] Lacharme, P.; Cherrier, E.; Rosenberger, C. Preimage attack on biohashing. In Proceedings of Security and Cryptography (SECRYPT), 2013 International Conference on; p. 1-8.
[23] Maltoni, D.; Maio, D.; Jain, A.K.; Prabhakar, S. Handbook of fingerprint recognition; Springer Science & Business Media: 2009.
[24] Jenisch, S.; Uhl, A. Security analysis of a cancelable iris recognition system based on block remapping. In Proceedings of Image Processing (ICIP), 2011 18th IEEE International Conference on; p. 3213-3216.
[25] Ouda, O.; Tsumura, N.; Nakaguchi, T. On the security of bioencoding based cancelable biometrics. IEICE TRANSACTIONS on Information and Systems 2011, 94, p.1768-1777.
[26] Ouda, O.; Tsumura, N.; Nakaguchi, T. Tokenless cancelable biometrics scheme for protecting iris codes. In Proceedings of 2010 20th international conference on pattern recognition; p. 882-885.
[27] Chai, T.Y., Goi, B.M., Tay, Y.H. and Jin, Z., 2019. A new design for alignment-free chaffed cancelable iris key binding scheme. Symmetry, 11(2), p.164.
[28] Chai, T.Y., Goi, B.M. and Yap, W.S., 2021. Towards Better Performance for Protected Iris Biometric System with Confidence Matrix. Symmetry, 13(5), p. 910.
[29] Soliman, R. F., Amin, M. & Abd El-Samie, F. E., 2019. A modified cancelable biometrics scheme using random projection. Ann. Data Sci. 6, p. 223–236.
[30] Soliman, R.F., Ramadan, N., Amin, M., Ahmed, H.H., El-Khamy, S. and Abd El-Samie, F.E., 2020. Efficient cancelable Iris recognition scheme based on modified logistic map. Proceedings of the National Academy of Sciences, India section a: physical sciences, 90(1), p.101-107.
[31] Nassar, R.M., Khalaf, A.A., El-Banby, G.M., El-Samie, F.E.A., Hussein, A.I. and El-Shafai, W., 2023. Cancelable template generation based on quantization concepts. *Opto-Electronics Review*, *31*(3).