

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/284819268>

# Cancelable iris template generation using look-up table mapping

Article · April 2015

DOI: 10.1109/SPIN.2015.7095296

---

CITATIONS

17

READS

287

---

2 authors, including:



Rudresh Dwivedi

Netaji Subhas Institute of Technology

15 PUBLICATIONS 162 CITATIONS

SEE PROFILE

# Cancelable Iris Template Generation using Look-up Table Mapping

Rudresh Dwivedi, Somnath Dey  
Discipline of Computing Science  
Indian Institute of Technology Indore

Indore, India  
Email: {phd1301201006, somnathd}@iiti.ac.in

**Abstract**—One of the potential passiveness in a biometric system is the invasion of stored biometric template, which may lead to serious security and privacy thefts. The emerging need of biometric approaches is evolved from privacy invasion and irrevocable issues of identity thefts as passwords and tokens can be easily compromised. To address these issues, the notion of cancelable biometrics is introduced to signify biometric templates that can be canceled and replaced with the inclusion of another independent authentication factor. Cancelable biometric generation technique based on randomized look-up table mapping has been proposed in this paper. The technique uses a decimal vector, which is evaluated from a row vector. Row vector is generated by applying 1-D Log Gabor filter on the raw iris template. The decimal vector is then mapped to look-up table based on the position of bits. The proposed approach enables an operation of cancelable iris biometric systems at a high security level. Experiments which are carried out on the CASIA V3 Interval iris database confirm the effectiveness of the proposed approach. Although, the accuracy of the proposed approach is confirmed as 94.26% but the approach is efficient if we consider the security and non-revocable perspectives.

**Keywords**-cancelable, Look-up table, iris biometric, biometric

## I. INTRODUCTION

A biometric template derived from a user's biometric trait contains the user's concealed information. Thus, it may compromise the sensitive information (e.g., gender, possible disease) of a user. Intensive research [1], [2], [4] have been conducted to address the security, privacy and revocability issues of biometrics in recent years. One of the most promising approach [2] is proposed to address the above mentioned issues is the “cancelable biometrics” approach. The concept of cancelable biometrics depends on non-invertible transforms that can be applied to true biometric templates in order to generate multiple protected templates, which can be canceled and replaced in case of compromise. In order to satisfy the requirements of the cancelable biometrics construct, cancelable biometric techniques depend on other validation factors such as password keys and/or user specific tokens in the transformation process. Unlike passwords and tokens, biometric templates cannot be canceled and replaced easily if they are compromised (stolen by an imposter). In addition, many privacy concerns have been raised regarding use of biometrics in human authentication or identification [3]. Another way of generating a cancelable biometric traits is by mixing an artificial pattern by generating a synthetic biometric pattern

[4]. Davida et al. [5] proposed a majority decoding scheme for iris biometrics. In non-invertible transformed-based approach, instead of storing the original biometric, the biometric is transformed using a one-way function and stored into the database. The transformation occurs in the same feature space as the original biometric [5], [6]. However, non-invertible transforms may cause a loss in accuracy. Registration-based cancelable template approaches [7] depend on the accurate image alignment, which is very hard to achieve. In the error-correcting code approaches [8], [2], codeword and decoding functions are established from the biometric templates during the enrollment phase. The authentication phase utilizes input biometric data to compute or recover the codeword. Error-correcting approaches have suffered from the privacy theft issues as input data can be easily compromised. The existing approaches address two important observations: (1) A number of approaches failed to yield a perfect, non-revocable and highly secure cancelable template (2) Cryptography methods are not suitable for cancelable templates as they reform the template and generate a poor matching rate.

The contribution of this work is to propose an approach to generate cancelable iris pattern using iris code evaluated by 1-D log Gabor filter. After the necessary iris preprocessing, such as segmentation and feature extraction, the iris code is intentionally transformed into a row vector. In the proposed approach, row vector is partitioned into a set of fixed length words and these words are mapped into a decimal vector. Finally a look-up table is used to generate cancelable iris template. The rotation invariance is also taken into account to retain the accuracy of the proposed system. Experimental evaluations confirm the efficiency of the proposed work.

This paper is organized as follows. Section II summarizes related work. In Section III, the proposed approach is described in detail. Experimental setup is presented in Section IV. Section V shows the results and discuss about the parameters involved. Section VI concludes the paper.

## II. RELATED WORK

The different existing approaches to cancelable biometric have been discussed here. Davida et al. [5] investigated the feasibility of biometrics as an enabling technology for secure system and application development with the use of cryptographic mechanisms. The scheme does not require the

co-operation of an on-line database for the security infrastructure. However, the approach is incapable to provide two major requirements: reusability and cancelability. Juels and Wattenberg [6] improved the Davida et. al. approach [5]. Their scheme involves the extension of a biometric template into an error-correcting codeword through the addition of check bits. In contrast, the fuzzy commitment method [6], as applied to biometric templates, treats the template itself without any modification as a corrupted codeword. This difference in perspective yields several advantages. Jules and Sudan [10] proposed the fuzzy vault scheme with the improvisation on Juels and Wattenberg's approach [6] using Reed-Solomon error correcting code. Linnartz and Tuyls [11] developed  $\delta$ -contracting and  $\epsilon$ -preserving functions to preprocess the measurement data for authentication. The approach does not allow learning any parameter information about the user unless the user willingly releases these parameters. Clancy et al. [12] applied Juels and Sudan's [10] fuzzy vault for building and analyzing a secure authentication scheme using a private key stored on a smart card. The polynomial parameters of the vault are deployed in a manner that the attacker's vault unlocking complexity is maximized, having zero unlocking complexity with a matched fingerprint and a considerable amount of error. BioHash approach of cancelable biometric mixes a set of user-specific random vectors with biometric features. BioHash fuses the biometric template with user-specific Tokenised Random Numbers (TRN) [8] to conjure a set of non-invertible binary bit strings. Tohari et.al. [9] proposed a method in which a projection line is constructed with minutiae points for cancelable fingerprint template design. Jin et al. [13] explored a minutiae descriptor, termed as Minutiae Vicinity Decomposition (MVD) to attain a set of randomized geometrical invariant features fused with random projection. The dissimilarity of randomized MVD is then developed by user-specific minutia propinquity scheme and implanted into a Hamming space by means of Graph-based Hamming Embedding method [14]. A problem with biometric authentication systems arises when the data associated with a biometric feature has been compromised. Another way of generating a cancelable biometric template [2], [5], [8], [15], [18] is to supply a different set of distortion parameters using any non-invertible transform.

### III. PROPOSED METHOD

In the proposed method, the iris image is preprocessed to localize and normalize using traditional algorithms [23], [24], which localize the iris pattern. Iris code features are extracted using 1-D Log-Gabor filter [22] with phase quantization and rotation invariant codes are generated. The bit-wise biometric template is converted into 1-D row vector which is divided into number of word of size  $M$ . Here,  $M$  is evaluated by the statistics of the different histograms of a template. A decimal vector is generated by partitioning the row vector using the definition of  $M$ . A look-up table is maintained to map the decimal vector and certain digits are selected from the look-up table to generate the final template. Figure 1 shows the block diagram of the proposed method.

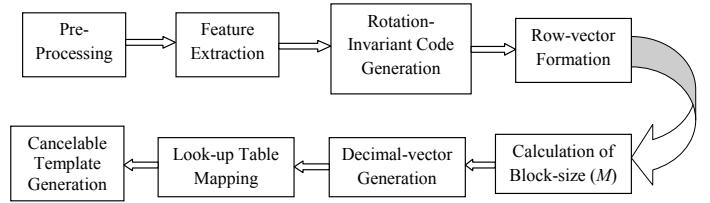


Fig. 1: Block diagram of the proposed method

#### A. Pre-processing

The pre-processing involves iris segmentation followed by image enhancement. Iris segmentation is required to separate the iris region from the eye image and remove the eyelids, eyelashes and other noises to enhance the matching performance. In our work, circular Hough transformation [20] is applied for finding the circles using the parameters: radius and center co-ordinates. Eyelids are detected from the image using parabolic curve parameter instead of the circle parameters afore-discussed [19]. Iris may be captured in different size with varying imaging distance. Due to illumination variations, the radial size of the pupil may change accordingly. The resulting deformation of the iris texture will affect the performance of subsequent feature extraction and matching stages. Therefore, the iris region needs to be normalized to compensate for these variations. Daugman's rubber sheet model [21] is used for normalization. The homogenous rubber sheet model remaps each point within the iris region to polar coordinates. The normalized iris image has low contrast and non-uniform illumination caused by the light source position. Local histogram analysis [24] is applied to the normalized iris image to reduce the effect of non-uniform illumination and well-distributed texture image is obtained. Reflection regions are characterized by high intensity values close to 255. A simple thresholding operation is performed to remove the reflection noise. Figure 2 shows the enhanced iris image.



Fig. 2: Enhanced image after normalization

#### B. Feature Extraction

Normalized iris image is transformed into a 0-1 form of matrix after convolving with quadrature 1-D Gabor filter [22]. The Gabor wavelet is a complex valued function which has real and imaginary parts. The Gabor coefficient values are coded with either 1 or 0 depending the sign of the coefficient. The Gabor function is represented in Eq. (1).

$$F(x, \omega, \sigma) = \frac{1}{\sqrt{2\Pi}\sigma} \exp\left(\frac{-x^2}{2\sigma^2} + j\omega x\right) \quad (1)$$

Here,  $\sigma$  is the spatial spread with  $\theta$  orientation and  $\omega$  frequency component [22]. The function produces real and imaginary components which are phase quantized to get iris code in the form of 0-1.

### C. Rotation Invariant Code Generation

An iris pattern can be affected by rotation with a maximum rotation of 8 columns producing the maximum 5.625 degree rotation to an iris image. It is quite difficult to match a rotated image as this degrades the performance of the system [23]. Even for genuine, it may result in poor matching. Therefore, rotation invariance mechanism must be employed. To deploy rotation invariance, we have taken 8 images per subject. One image from these images is selected as a reference image. We have computed Hamming distances between the reference image and other images by shifting one column. The whole circular iris pattern is considered to have 512 columns, so shifting of one column of image is equivalent to 0.703125 degree to a maximum of 8 columns generating 5.625 degree rotation. One right shift in the iris code account for .703125 degree rotation as shown in Fig. 3. It has been analyzed that Hamming distance between images without rotation invariant is more than the images with rotation invariant. Considering this shifting, iris code having minimum Hamming distance is taken into account for further processing.

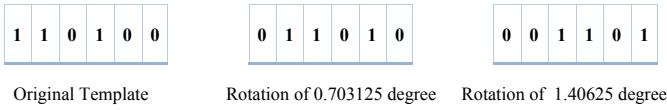


Fig. 3: Rotation on templates with different degrees

#### D. Row Vector Formation

The rotation free templates are shift invariant in comparison to iris code. Rotation invariant templates are transformed into a row vector to make ease in computation. Furthermore, row vector can easily be deployed to any transformation. The row vector is created through merging the next row to previous one. A row vector of  $1 \times 24$  is obtained from the iris code of  $4 \times 6$  as shown in Fig. 4.

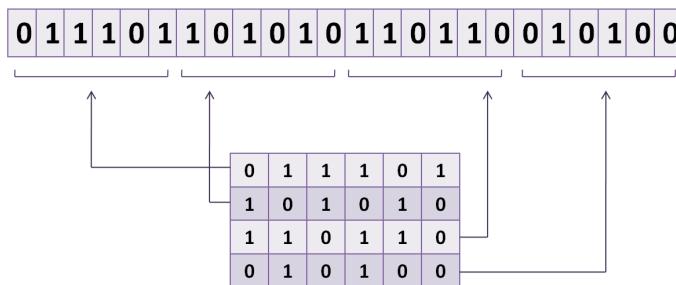


Fig. 4: Example of creating row vector

### *E. Calculation of Block Size M*

If row vector has large number of bits, it is difficult to apply any transformation function into all bits. Therefore, row vector is partitioned into a number of blocks of size  $M$ . The value of  $M$  may be chosen statically or dynamically and different for every image. But the dynamic approach is more secure to use as it is difficult to hack the value of  $M$ , which is the key to be used. We determine  $M$  for every image based on the histograms generated from the iris code. In our method, we take histograms of 0's, 1's, 00's, 11's, 000's, 111's and so on. The value of  $M$  is chosen by generating histogram of these bit's combinations. For the different images of the same person, histograms will differ with less value, so we take the range of these values. The block size is chosen in such a way that it should have uniform histogram distribution. We can generate any number of cancelable templates by changing the value of  $M$ . Large value of  $M$  may lead to large decimal vector and look-up table, which can influence the performance of method. Therefore,  $M$  should be chosen in such a way it produces high recognition rates.

### *F. Decimal Vector Generation*

The decimal vector is constructed by partitioning row vector using a fixed length word of size  $M$ . The decimal vector has the same number of positive integers as number of words in row vector. The conversion of a word from binary to positive integer seize the left most bit as the most significant bit. These integer numbers are represented into this vector. Using this vector, we can easily map each word to an unique row of look-up table. Let the value of  $M$  for given row vector be 4. Therefore, the row vector is divided into 6 words, each having 4 bits as shown in Fig. 5. The decimal vector will have size  $2^m$  with positive integers in range of 0 to  $2^m - 1$ .



Fig. 5: Partitioned vector

If the word length is large, then the decimal vector has large positive integers. Therefore, a suitable normalization function can be used to convert decimal vector into suitable range. The value of  $M$  should be the multiple of 2 to the power to get the consistent words of  $M$  bits. If it is not then the partition will not be a perfect and some bits will be left over. There is no need to store the  $M$  for images, as we calculate it dynamically. This will increase the security and non-revocability of the method.

#### G. Look-up Table Mapping

To distinguish between different words, we map the decimal vector to a corresponding word utilizing a look-up table. More than one word can be mapped to the same positive integer, so reverse mapping is very difficult. The row vector words are mapped to the corresponding decimal value. There can be many words having decimal value 0. If decimal value of

a vector is 0, then the word can be mapped to next decimal value. The mapping for given row vector is illustrated in Fig. 6.

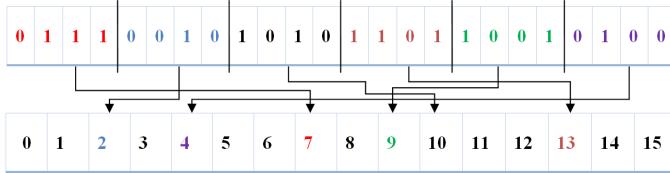


Fig. 6: Mapping of word to decimal vector

A Look-up table is generated of random values 0 and 1. The minimum size (rows) of the table depends on the value of  $M$ . For example if we have word length 4, then the maximum decimal integer is 15. So the table must have entries greater than or equal to 15. There is a possibility that all entries of a particular row or more than one row are 0. In this situation, the use of these entries are vulnerable to privacy invasion, as this makes imposters task easy. Therefore, look-up table should be such that the number of 0's and 1's are approximately same in a randomized manner. For the given row vector, the size of look-up table will be  $R \times C$ , here  $R$  represents the size of the decimal vector which is to be mapped and  $C$  represents the size of each word. In our case  $R$  is 13, and  $C$  is assumed to be 4. Each row of the look-up table is filled by the 0 or 1 which is generated randomly.

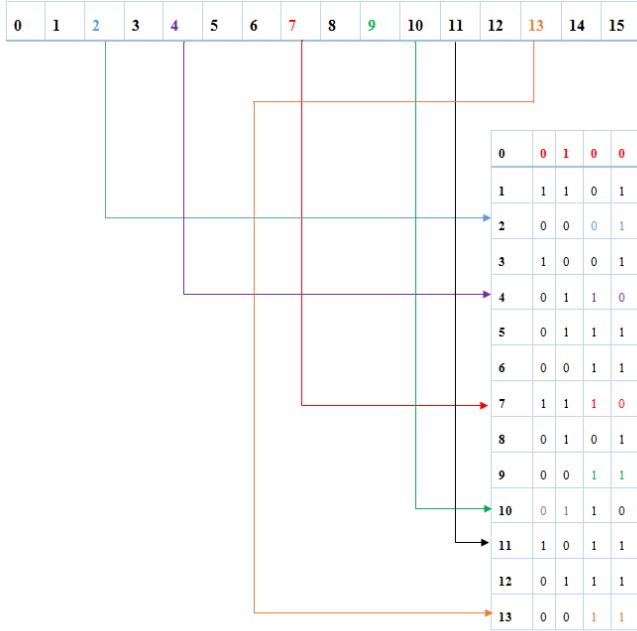


Fig. 7: Mapping from decimal vector to look-up table

#### H. Generate Cancelable Templates

The mapping of decimal vector is performed to the corresponding row of the randomly generated look-up table. Figure

7 illustrates the mapping procedure. Then  $d$  bits are selected from each row of the look-up table. Here,  $d$  represents the number of bits deployed into the final template. For example if  $d = 2$ , then 2 bits from any position in each row will be taken from look-up table for the final template generation. At the utmost the  $d$  bits from every row of look-up table is stored in database. Therefore, the final template consists of 12 bits if we choose 2 bits from each word as depicted in Fig. 8.

0	1	1	0	1	0	1	1	0	1	1
---	---	---	---	---	---	---	---	---	---	---

Fig. 8: Final template

Numerous templates can be generated for the different values of  $M$ , but for every  $M$  look-up table will be fixed. To perform matching, the proposed method has been deployed. We calculate the Hamming distance between the stored templates and the evaluated template. The accept or reject decision is confirmed using the Hamming distance and a threshold.

#### IV. EXPERIMENTAL SETUP

In order to test the effectiveness of our proposed method, we have used publicly available CASIA-iris V3-Interval [16]. The database includes 2639 images captured from 249 different subjects. The proposed work is implemented using MATLAB 7.8.0 (R2010b) of MathWorks, Inc, USA with machine specification illustrated in Table 1. We evaluate the efficiency of the proposed method by means of False Accept Rate (FAR), False Reject Rate (FRR) and Equal Error Rate (EER).

Manufacturer	Dell Computers
Model	XPS L501 X
RAM	4 GB
Hard Drive capacity	512 GB
Operating System	Windows 7 Home Premium
System type	64 bit
Processor	Intel core i5 CPU
Clock	2.53 GHz

TABLE I: Machine specifications

The proposed method uses two different parameters, which are required to generate the different cancelable templates. These are also responsible for the efficiency of proposed method.

(i) The word length  $M$  is used to divide the row vector into fix length segments. The different values of  $M$  produce the different Hamming distances for the same subject; therefore the parameter  $M$  has an impact on the efficiency of the proposed method. Moreover, a change in  $M$  may result in the generation of the different biometric templates. The results for the different values of  $M$  is given in section V.

(ii) Look-up table can also be considered as a parameter that affects the efficiency and revocability of the proposed method. The different look-up tables can be designed using randomly generated values. Sometime it may happen that table is biased for either 0 or 1 that makes the proposed approach non-revocable. The results for the different values of look-up table is given in section V.

## V. RESULTS & DISCUSSION

Implementation and testing results of proposed approach are analyzed with respect to two parameters; value of  $M$  and look-up table. In this section, first we present the result for the different values of  $M$ . Then, we present the result for the different look-up tables.

### A. Effect of parameter $M$

Intra-class Hamming distance represents a match of one template of a subject with other templates of the same subject. The FAR, FRR and EER have been determined for 20 subjects using the different values of  $M$ . In our experiment we consider the values of  $M$  as 2,4,8,16. The efficiency of the proposed method is evaluated using these parameters.

Inter-class Hamming distance is calculated by matching one template of a subject with all the templates in database. Ideally, Hamming distances for inter-class should be greater than intra-class as bit difference will be more in case of inter-class. Results for different values of  $M$  are shown in Fig. 10. Initially, when  $M=2$  there are 16384 different words of 2 bit each, so the variation in bits is less. This leads to very less difference in intra-class comparisons. For  $M=4$ , the bit difference increases. Higher the  $M$  value, more will be the bit difference. For less  $M$  values, the efficiency is more. This increases the intra-class variation resulting ease in matching of template. Threshold,

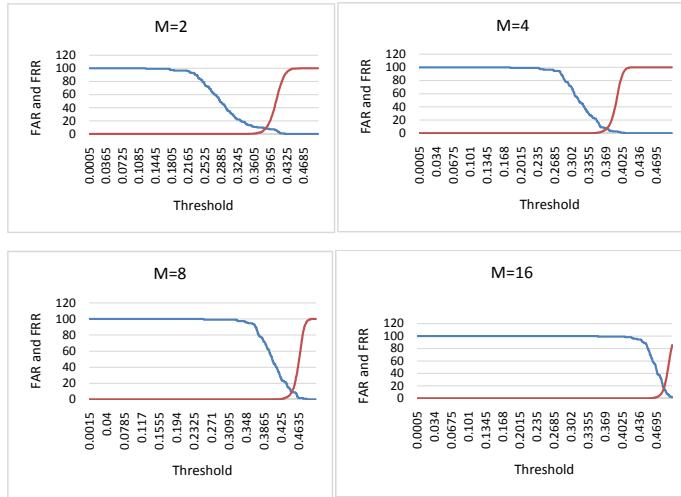


Fig. 9: FAR/FRR and Threshold graph for different values of  $M$

EER and efficiency for respective  $M$  values are depicted in Table 2. For less  $M$  values, the efficiency is more. For the proposed approach, maximum efficiency of 94.26% is obtained for  $M=4$ .

### B. Effect of Look-up Table

The final template is generated by considering the number of bits from look-up table, therefore the bit sequence in look-up table have an impact on the matching performance. The same look-up table is maintained for the final template generation. A modification in look-up table leads to alter the random bits

Serial no.	$M$	Threshold	EER	Efficiency (100-EER)
1	2	0.3845	7.84	92.16
2	4	0.3714	5.74	94.26
3	8	0.4075	9.14	90.86
4	16	0.3940	7.35	92.65

TABLE II: Analysis with the different values of  $M$

which results a change in the template generation. For the different look-up tables, we may have discrete thresholds and performances due to change in sequence of bits. Though, this difference is minute, still it has the effect on performance of proposed approach. We have taken  $M$  as 16 for 20 different people and generated the final templates by taking different look-up tables. The results for the different lookup tables are given in Fig. 11. The approach does not allow to save any parameter except look-up table, so hacker has to know all about the method as well as the different values of  $M$  compromise the security of iris template.

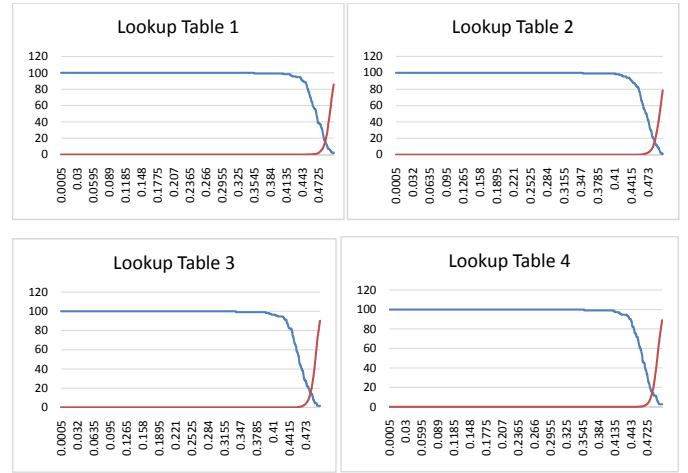


Fig. 10: FAR/FRR and Threshold for different Look-up tables

Threshold, EER and efficiency for respective look-up tables is illustrated in Table 3.

Serial no.	Look-up table	Threshold	EER	Efficiency (100-EER)
1	1	0.3845	5.77	94.23
2	2	0.4075	6.11	93.89
3	3	0.3940	5.89	94.11
4	4	0.3714	5.74	94.26

TABLE III: Analysis with different Look-up tables

### C. Results

The final testing is performed on 100 subjects with 9 to 5 images per subject. For intra-class, we acquire one reference image of each template which is matched with other images of same person resulting a total of 388 matching. In case of inter-class matching one template of a person is matched with other templates of other persons. A total of 48312 matching are performed with some overlapping area representing some

inter-class Hamming distances are in intra-class. The efficiency is achieved as 94.26 %. The final testing output is presented in Fig. 12.

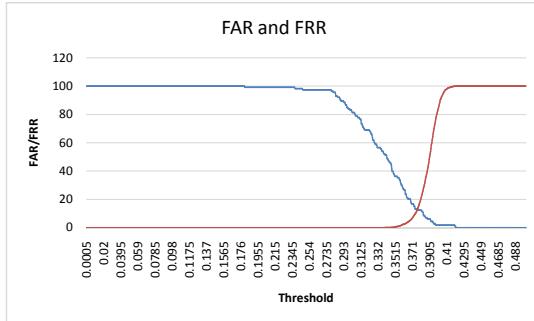


Fig. 11: final testing results

- The proposed method is analyzed by taking the different values of parameters  $M$  and look-up tables. The FAR, FRR, EER and efficiency of proposed method for the different values of  $M$  and look-up tables have been determined. The maximum efficiency has achieved when  $M$  is set to 4. The efficiency will not be same in each case due to different values of  $M$  and look-up tables. The maximum efficiency of 94.26 % has been achieved for look-up table 4.
- ## VI. CONCLUSION
- An approach for generation of cancelable iris biometric templates is proposed which is tested using CASIA-V3 interval database. The proposed method generates bit-strings or cancelable templates by mapping the decimal vector into look-up table. From the generated cancelable template, it is very difficult to regenerate the original iris code. In our approach, iris code is generated using 1-D Log-Gabor filter which is further partitioned into a number of words of size  $M$ . The value of  $M$  is calculated from the histograms of the bit patterns in the iris code. This makes difficult to guess the value of  $M$  for the hacker. It also ensures the irrevocability of the method. A look-up table is maintained to map these positive integers and certain digits are selected from each row of the look-up table to generate the final template. As the look-up table mapping is performed by many-to-one, it is very difficult to regenerate the original iris code. We have tested our approach with CASIA V3 Interval iris database and performed a detailed analysis of the proposed method with respect to different parameters. It is evident from the experimental results, that our approach is able to provide 94.26 % accuracy and also handles security and revocable issues. The accuracy is affected by the preprocessing task as it cannot properly segment the iris part for poor quality images. Hence, higher accuracy can be achieved by improving the segmentation process.
- ## REFERENCES
- [1] Y. Sui, X. Zou, E.Y. Du and F. Li “Design and analysis of a highly user-friendly, secure, privacy-preserving, and revocable authentication method”, IEEE Transactions on Computers, vol. 63, no.4, pp. 902-916, April 2014.
  - [2] R.M. Bolle, J.H. Connell and N.K. Ratha “Biometric perils and patches”, Pattern Recognition, vol. 35, no. 12, pp. 2727-2738, 2002.
  - [3] O. Ouda, N.