Multi Biometric Template Protection using Random Projection and Adaptive Bloom Filter

Sushma H. R.[1], Sandeep R.[2]
Cambridge Institute of Technology, K. R Puram, Bengaluru, India

Abstract **-** In multi biometric system, multiple biometric informations of same individual are consolidated. In this work, two methods have been used to extract the features from the image, namely: 1) Gabor filter and 2) Random projection (RP) techniques. In Gabor filter feature extraction method, the log-Gabor wavelet filters are convolved with normalized biometric image to extract feature code vectors. Where as in RP method, the random matrices are multiplied with normalized biometric image in order to extract feature code vectors. The multiple features are fused using non-invertible adaptive Bloom filter to protect the biometric template. The RP technique reduces the template dimension to 32 times, compared to the dimension of the template extracted using Gabor filter. Another advantage of using RP is that, it provides the templates protection at feature level itself. The performance of the proposed algorithm using the RP technique is comparable to that of the Gabor filtering technique. The experiments are conducted on CASIA-iris-interval database. The Hamming distances are calculated for matching the biometric templates. The performance of the biometric verification system developed using the proposed algorithm and Gabor filtering technique, is evaluated using the precision recall curve.

Keywords *- Multi instance biometric; iris recognition; Gabor filter; random projection; Bloom filter; biometric template protection;*

I. INTRODUCTION

The biometric field is concerned with recognizing the individual based on unique biological and behavioral characteristics. Based on various types of application, several biometric modalities such as fingerprint, iris, face, hand, voice and gait have been exploited by applying adequate sensors. In the enrolment process, the biometric templates are generated using distinctive feature extractors. During verification (authentication) or identification, the system processes another biometric trait to generate a template. This template is compared with the stored template in the database resulting in match/no match or acceptance/rejection, respectively. Most biometric systems used in real world applications are unimodal. The drawbacks of these systems includes noise in sensed data, non-universality or restricted degrees-of-freedom and spoof attack. In order to overcome these problems, the multi biometric recognition systems are used. Based on the nature of the sources of biometric information, a multi biometric system can be classified into five categories.

1. Multi sensor systems: Multiple sensors are used to capture single biometric trait of an individual.
2. Multi algorithm system: Multiple algorithms are used for feature extraction of same biometric trait.
3. Multi sample system: Multiple samples of the same biometric trait are acquired using single sensor.
4. Multi instance system: Multiple instances of the same biometric trait are extracted.
5. Multi modal systems: Multiple biometric traits of the same individual are used.

However, the Multi biometric systems also have several drawbacks compared to single biometric systems. One major issue regarding multi biometric recognition system is its hardware complexity. These systems are more expensive compared to unibiometric systems, as it involves multiple sensors. Furthermore, the use of multi biometric system increases the central storage to preserve multiple biometric templates of a single subject. The major drawback of multi biometric system is template protection. To be more precise, for each subject registered with multi biometric system, multiple biometric reference data has to be stored. The leakage of biometric template information to unauthorized individuals results in serious security and privacy threats. Therefore, the biometric templates needs to be protected. This article concentrates on multi biometric template protection techniques. Accordingly, a review of works related to multi biometric template protection methods are provided in the literature survey.

A. Motivation

Biometric systems are becoming popular because it provides more reliable identity management to several applications that render services to only legitimately enrolled users. The examples of such applications include performing remote financial transactions, boarding a commercial flight, granting access to nuclear facilities [3]. Traditional methods of establishing identity including knowledge based (eg. Passwords), token based (eg. ID cards) can be easily lost, stolen or manipulated by the impostors, thereby undermining the intended security. Biometric systems offers a natural and reliable solution to identity management of an individual based on their inherent biological and/or behavioral characteristics. Automatic recognition systems based on a single biometric modality often have to deal with unacceptable errors. multi biometric systems have improved the accuracy and reliability of unibiometric systems . However, the security of multi biometric templates is especially crucial as they contain information regarding multiple biometric traits of the same subject. The leakage of any kind of template information to unauthorized person results a serious security and privacy risks [4]. Therefore, biometric template protection technologies have been developed in order to protect privacy and security of the stored biometric data. The presented work is motivated by the requirements in the fields of multi

biometric recognition systems and biometric template protection schemes.

B. Problem Statement

The main problem of multi biometric systems is its hardware complexity, huge template dimensionality and protection of multiple templates. This article proposes a non-invertible method for protecting the multi biometric templates and at the same times addresses the problem of high dimensionality. The detailed description of the proposed method is given in section 7.

C. Contribution

The major contributions of the work are as follows.

1. The multi biometric (iris biometric trait) template protection is based on random projection (RP) technique. Here, the left and right iris features are extracted using random matrices.
- The templates generated using RP method gives 32 times lesser number of bits compared to Gabor filer method and also gives comparable performance.
2. The generated feature vectors are fused using adaptive Bloom filter.
3. Biometric verification application is developed using the proposed algorithm.
4. The performance of the application is evaluated using precision recall curve.

D. Organization

The remaining paper is organized as below:

- Section 2 presents biometric system's operation, application and limitations.
- Section 3 discusses about the multi biometric system and its design issues.
- Section 4 provides various multi biometric template protection schemes and literature survey of the existing work related to the field of the presented work.
- Section 5 briefs the information about the iris recognition system.
- Section 6 provides the required mathematical framework for the proposed algorithm.
- Section 7 presents the proposed method.
- Section 8 provides the experimental set-up and result analysis.
- Section 9 concludes the work carried out in this dissertation. Scope of the future work is also provided.

II. BIOMETRIC SYSTEM

The biometrics or the biometric authentication is a natural and reliable identity management system of an individual based on "who he is?" rather than "what he carries?" or "what he knows" [18]. Biometric systems automatically recognizes or verifies a person's identity based on his biological and behavioral characteristics such as fingerprint, iris, face, voice, gait and signature. These characteristic are referred as traits, modalities, identifiers or indicators in biometric literature.

The biometric system that acquires biometric information from an individual and extracts salient features from it, and then, compares this feature set against the stored reference feature set in the database, and executes an action based on the result of the comparison. Therefore, a generic biometric system operation can be divided into four main modules as below,

A. Sensor module

A good quality biometric scanner or reader is required to capture the raw biometric data of an individual.

B. Quality evaluation and feature extraction module

The quality of the biometric information captured by the sensor is first evaluated in order to determine its suitability for further processing.

C. Matching and decision making module

During enrolment, the extracted features are stored in the database and matched against query data to generate match scores. The decision module is also encapsulated in a matcher module, in which the match scores are used to either validate a claimed identity or provide the identity of an individual.

D. Database module

The feature set extracted during enrolment process from the raw biometric sample (template) is stored in the database.

The applications of the biometric systems can be categorized into three main groups:

1. Commercial applications such as electronic data security, e-commerce, computer network login, internet access, medical records management and distance learning.
2. Government applications such as driver's license, border control, welfare disbursement, social security and passport control.
3. Forensic applications such as criminal investigation, corpse identification and parenthood determination.

Depending up on the application, the functionality of the biometric system can be classified as identification and verification. For detailed explanation about identification vs verification refer [1,3].

Though biometric systems have several advantages to both government and civilian authentication applications over password and token based approaches, it is essential to consider vulnerabilities and limitations of these systems when implementing them in real world applications involving a large number of users (say in the order of millions) . Some of the commonly encountered challenges of the biometric systems are listed below.

1. Noise in sensed data

Due to the defective or improperly maintained sensors, noise may be present in the captured biometric data.

2. Intra class variations

Biometric instances of an individual usually show large intra user variations. These variations in biometric systems are mainly due to improper interaction of an individual with the sensor, changes in the biometric characteristics of a person over a period of time, use of different sensors during enrolment and verification.

3. Inter class similarities

Inter class or inter user similarity refers to overlap of feature samples corresponding to multiple individuals. The lack of distinctiveness in the biometric feature set restricts the discriminative ability of the biometric system and leads to an increase in the false acceptance rate.

4. Non universality

A biometric trait is said to be universal if every individual in the target population is able to present the biometric sample for recognition. However, most of the biometric traits are not truly universal. People with hand-related disabilities, manual workers with many cuts and bruises on their fingertips, and people with very oily or dry fingers [5] cannot be enrolled in a fingerprint recognition system. Similarly, persons who are from eye abnormalities or diseases like glaucoma, cataract, and those having long eye lashes cannot provide good quality iris images for automatic recognition [6].

III. MULTI BIOMETRIC SYSTEM

Multi biometric systems are designed to recognize an individual based on multiple information captured from multiple biometric sources. Due to the presence of multiple evidences, the multi biometric systems overcome many of the limitations of unibiometric system and also tends to be more accurate. multi biometric systems offers the following advantages over traditional (uni biometric) systems.

1. Combining the evidence obtained from different sources using effective fusion methodology, significantly improves the overall accuracy of the biometric system.
2. Multi biometric systems addresses the issue of non-universality or inadequate population coverage.
3. The availability of multiple sources of information effectively addresses the effect of noisy data.

The multi biometric system is designed based on the application requirements. The major issues that need to be considered while designing a multi biometric systems are discussed below.

1. Different sources of biometric information

Sources of biometric information includes multiple sensors, multiple feature extraction algorithms, multiple samples of the same biometric trait, multiple instances of a biometric trait and multiple biometric traits. Based on the application, the system designer has to decide which of these sources should be used.

2. Acquisition and processing sequence of biometric information

The multiple sources of information can be acquired and processed in serial (cascade or sequential), parallel or hierarchical (tree-like) sequence. For a given application scenario, an appropriate acquisition and processing architecture must be selected.

3. Fusion technology

The process of combing information provided by different biometric sources is known as biometric fusion. Depending upon the type of information that is fused, the fusion scheme can be grouped as sensor level, feature level, score level and decision level fusion. For more detailed explanation about fusion technologies refer [1].

4. Type of information for fusion

The designer has to decide what type of information or trait (i.e., features, match scores, decisions) should be fused. A number of techniques are available for fusion of biometric information provided by the multiple sources.

Though multi biometric systems offer several advantages over unibiometric systems such as better recognition accuracy, increased population coverage, better security and flexibility, the design of a multi biometric system is not an easy task. multi biometric system designing has several challenging issues because it is very difficult to predict the ideal sources of biometric information and the ideal fusion strategy for a particular application.

IV. LITERATURE SURVEY

One of the most possible harmful attacks on a biometric system is against the biometric templates. Attacks on the template can head to the following four susceptibilities:

1. An impostor can gain an unauthorized access by replacing the template.
2. By creating physical spoof from an unprotected template, a deceiver can gain unauthorized access to the system.
3. The stolen template can be reproduced to the matcher to gain illegal access.
4. One can covertly track a person without his/her permission by cross checking the template across different database.

Because of these reasons, the biometric templates (or the raw biometric images) should not be stored in clear text form and spoof proof techniques are required to securely store the templates, so that both the security of the application and the users privacy can be preserved. This section provides a brief review of template protection schemes and existing works in the field of multi biometric template protection.

A. Review of Template Protection methodologies

An optimal biometric template protection scheme should contain the following four properties [7].

1. Heterogeneity

The secure template must not allow cross matching across databases, thereby protecting the user's privacy.

2. Revocability

It should be straightforward to revoke a conceded template and reissue a new one based on the same biometric information.

3. Security

It must be computationally difficult to obtain the original biometric template from the protected template. This property avoids an impostor from creating a physical spoof of the biometric trait from a stolen template.

4. Performance

The biometric template protection scheme should not weaken the performance of the biometric system.

The template protection schemes can be broadly divided into two categories [7],[1] .

a. Feature Transformation

In the feature transform scheme, a transformation function (F) is applied on the biometric template (T) and only the transformed template (F(T ;K)) is stored in the database The parameters of the transformation function are typically drawn from a random key (K) or a secret password. The same transformation function is applied on the query template (Q) and the transformed query (F(Q;K)) is directly matched against the transformed template (F(T ;K)). Depending on the nature of the transformation function F, the feature transform approach can be further divided as *salting* and *non-invertible* transforms. In salting, the transformation function F is invertible, i.e., if an adversary gains access to the random key and the transformed template, he can reclaim the original biometric template (or a close approximation of it). Hence, the security of the salting scheme depends on the secrecy of the key or password. In case of non-invertible approach, the transformation is a one-way function on the template and it is computationally difficult to invert a transformed template even if the secret key is known [8][9][10].

b. Biometric Cryptosystems

In biometric cryptosystem approach, some universal information about the biometric feature is stored. This universal information is also referred as helper data and hence, biometric cryptosystems are also known as helper data-based approach. While the helper data does not (is not supposed to) report any significant information about the original biometric template, it is needed during matching, to extract a cryptographic key from the query biometric template. Matching is established indirectly by verifying the validness of the extracted key.

Biometric cryptosystems are further categorised as key binding and key generation, depending on how the helper data is gathered. When the helper data is gathered by binding a key (that is independent of the biometric features) with the biometric template, it is referred as key-binding biometric cryptosystem. In key binding system, the matching is performed by recovering a key from the helper data of the query biometric feature. If the helper data is drawn only from the biometric template and the cryptographic key is directly obtained from the helper data and the query biometric features, then it is called key generation biometric cryptosystem. The majority of existing approaches implement cryptographic schemes such as fuzzy commitment and fuzzy vault for biometric protection [11],[12] and [13].

In [14], the authors introduces a new non invertible approach to protect the biometric data called the Coverage-Effort (CE) curve. The curve measures the amount of effort (or number of guesses) required by an impostor to recover a certain fraction of original biometric data. In [15], the authors used adaptive Bloom filters on binary iris feature vectors. The authors utilized Bloom filters to establish rotation invariant transformation to the iris codes. Bloom filter based transform enables (1) Biometric template protection (2) Biometric data compression and (3) Improved biometric identification. The detailed description of this algorithm is provided in subsection C of section 6. A new biometric template protection method is proposed in [16]. The authors used RP technique to build a non invertible transformation of the biometric template, that meets the template protection scheme requirements [7] of revocability , security, diversity and performance. More information about the RP and random matrix generation is provided in subsection A of section 6. In [1], an iris cryptosystem is proposed to secure the iris code templates. These templates are obtained by demodulating the iris pattern using quadrature 2D Gabor wavelets. Since the iris code is a fixed length binary vector in which the relative order information between the bits is needed for matching therefore iris codes cannot be directly secured using fuzzy fault. In order to overcome this problem, the author in [1] construct the iris cryptosystem in two steps. In the first step, a invertible transformation to the iris code based on a randomly generated transformation key is applied. Since the transformation is invertible, the security of the template depends on the security of the transformation key. In the second step, the transformation key is represented as an unordered set and fuzzy vault is constructed to secure the iris code.

V. IRIS RECOGNITION SYSTEM

The iris is a thin circular structure, which lies between the cornea and the lens of the human eye [2]. A front view of the human eye is shown in Fig. 5.1. The iris consists of a number of zones, the lowest is the epithelium zone, which contains thick pigmentation cells. The stromal zone lies above the epithelium zone, and contains blood vessels, pigment cells and the two iris muscles. The color of the iris is determined by the density of the stromal pigmentation. The externally visible surface of the multi-zoned iris contains two layers, which often differ in color . An outer ciliary layer and an inner pupillary layer, and these two layers are divided by the collarette which appears as a zigzag pattern.

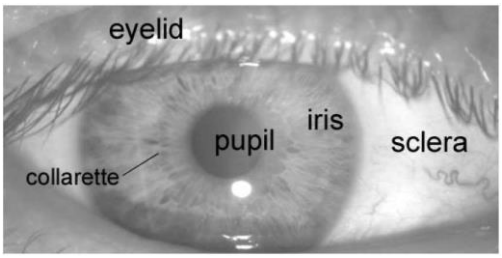


Fig.5.1: Front view of human eye

The epigenetic nature of iris patterns makes two eyes of an individual completely independent of each other, and identical twins also possess uncorrelated iris patterns. Compared to other biometrics such as face, palm and finger prints, irises have enormous pattern variability. A single iris scan can analyze more than 200 different spots of the iris, such as furrows, corona, rings and freckles. Another notable advantage of iris compared to other biometrics, such as voice

and facial features, is that the iris biometric is stable and remains unchanged for a person's lifetime . The basic steps involved in iris recognition system are briefly described as follows:

1. Image acquisition: The eye images are captured using high quality cameras and sensors.

2. Iris segmentation: In the segmentation stage, the actual iris is isolated from the rest of the eye image.

3. Normalization: The normalization fixes the dimensions of an eye images in order to allow comparisons.

4. Feature extraction: In feature extraction stage, the most significant information present in an iris pattern is extracted in order to provide accurate recognition of an individual.

5. Matching: During matching, a metric is calculated for the generated feature template which gives a measure of similarity between two iris templates.

VI. MATHEMATICAL FRAMEWORK

This section provides the complete mathematical framework of the techniques utilized in the proposed method which is discussed in the next section.

A. Random Projection

The biometric recognition systems usually deal with large amount of data. This data has to be archived or exchanged between numerous users and systems, consuming expensive resources such as storage space and transmission bandwidth [18]. In order to handle the available data adequately, the dimensionality needs to be reduced. The RP method is an efficient dimensionality reduction tool. In RP, the original high dimensional data is projected onto a lower dimensional subspace using a random matrix whose columns have unit lengths. The concept of RP can be mathematically understood as follows: The original d-dimensional data is projected on to a k-dimensional (k << d) subspace, using a k × d random matrix R, whose columns have unit lengths [19]. Using matrix notation, the concept can be represented as

$$\mathbf{A}_{[\mathbf{k}\times\mathbf{N}]} = \mathbf{R}_{[\mathbf{k}\times\mathbf{d}]} \times \mathbf{X}_{[\mathbf{d}\times\mathbf{N}]} \qquad (6.1)$$

where $\mathbf{X}_{[\mathbf{d}\times\mathbf{N}]}$ is the original d-dimensional data matrix, N is the total number of points and k is the desired dimension.

The key idea of RP is based on the Johnson-Lindenstrauss lemma(JL lemma): For any $0 < \varepsilon < 1$ and any integer n, let k be a positive integer such that,

$$\mathbf{k}^3 \geq \mathbf{4}\left(\frac{\boldsymbol{\varepsilon}^2}{\mathbf{2}} - \frac{\boldsymbol{\varepsilon}^3}{\mathbf{3}}\right)^{-1} \ln n \qquad (6.2)$$

Then, for any set V of n points in $\mathbf{R}^{\mathbf{d}}$, there is a map $\mathbf{f}: \mathbf{R}^{\mathbf{d}} \rightarrow \mathbf{R}^{\mathbf{k}}$ such that for all u, v, ∈ V [20].

$$(1-\varepsilon)\,\|u-v\|^2 \leq \|f(u)-f(v)\|^2 \leq (1+\varepsilon)\,\|u-v\|^2 \qquad (6.3)$$

where f(u) and f(v) are the projections of u and v respectively.

The above lemma states that if a set of n points in a vector space $\mathbf{R}^{\mathbf{d}}$ is projected orthogonally onto a selected lower dimensional subspace $\mathbf{R}^{\mathbf{k}}$, then the distance between the points are approximately preserved. For complete proof of lemma refer [20]. The formation of random matrix R is one of the crucial points of interest. Gram-Schmidth orthogonalization is time consuming and it returns set of orthogonal vectors if and only if the input vectors are linearly independent. There are many choices for the random matrix whose elements $\mathbf{r}_{j,k} \sim N(0,1)$ have normal distribution but the problem of this type of RP is its computational complexity due to the dense nature of the projection matrix. A new approach suggested by Achlioptas [21] uses two simpler distributions that generate sparse projection matrices with elements drawn independent and identically distributed as:

$$\mathbf{r}_{\mathbf{i,j}} = \begin{cases} +1; \text{ with probability } \frac{1}{2} \\ -1; \text{ with probability } \frac{1}{2} \end{cases} \qquad (6.4)$$

OR

$$\mathbf{r}_{\mathbf{i,j}} = \sqrt{3}\begin{cases} +1; \text{ with probability } \frac{1}{6} \\ 0; \text{ with probability } \frac{2}{3} \\ -1; \text{ with probability } \frac{1}{6} \end{cases} \qquad (6.5)$$

The distributions shown in Eq. 6.4 and Eq. 6.5 reduces computational time for the calculation of R·X. For the second distribution speedup is threefold because it is sparse and only one third of the operations are required. In this article, the sparse projection matrix presented in Eq.6.4 is used.

B. Gabor Filter

The Gabor wavelet transform extract both spatial and frequency information of a signal. Several mathematical and biological properties of the Gabor wavelet motivates the researchers to use it for various image processing applications [22].

- Mathematical motivation

The Gabor wavelet transform provides optimal resolution in both time (spatial) and frequency domain. Besides, it has been found to yield distortion tolerance for pattern recognition tasks.

- Biological motivation

The simple cells of the visual cortex of mammalian brains are best modeled as a family of self-similar 2D Gabor wavelets.

A Gabor wavelet filter is constructed by modulating a sine/cosine wave with a Gaussian (see Fig: 6.1). This provides the optimum conjoint localization in both space and frequency. The decomposition of a signal is accomplished using a quadrature pair of Gabor filters, with cosine modulated by Gaussian specifies a real part and sine modulated by a Gaussian specifies a imaginary part. The real and imaginary filters are also called as the even and odd symmetric components respectively [23].

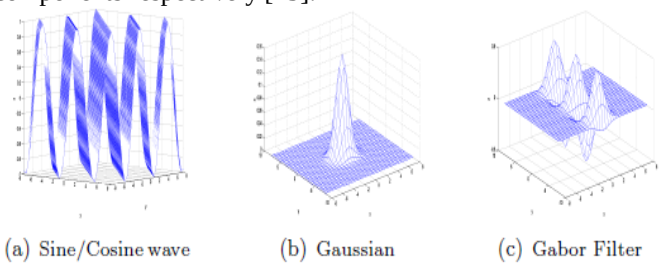


Fig.6.1: Gabor filter generation

The center frequency of the filter is specified by the frequency of the sine /cosine wave, and the bandwidth of the filter is specified by the width of the Gaussian function. A 2D Gabor filter [17] over an image domain (x, y) is represented as

$$G(x,y) = e^{-\pi\left[\frac{(x-x_0)^2}{a^2}+\frac{(y-y_0)^2}{\beta^2}\right]} e^{-2\pi i\left[u_0(x-x_0)+v_0(y-y_0)\right]} \quad (6.6)$$

where (x0, y0) specify position in the image, (α, β) specify the effective width and length, and (u0, v0) specify modulation, which has spatial frequency $\omega_0 = \sqrt{u_0^2 + v_0^2}$.

A disadvantage of the Gabor filter is that whenever the filter bandwidth is larger than one octave the even symmetric filter will have a DC component. However, zero DC component can be obtained for any bandwidth by using a Gabor filter which is Gaussian on a logarithmic scale, this is known as the log-Gabor filter [23]. The frequency response of a log-Gabor filter is given as

$$G(f) = \exp\left(\frac{-(\log(f/f_0))^2}{2(\log(\sigma/f_0))^2}\right) \quad (6.7)$$

where f_0 represents the center frequency, and σ gives the bandwidth of the filter.

C. *Adaptive Bloom Filters*

Bloom filter (b) is basically a probabilistic data structure representing a set S of length n. Initially, all bits are set to 0. Basic Bloom filter supports two operations ADD and TEST. The ADD operation simply insert an element to the set. The Test is use to check whether a given element is in the set or not. The general block diagram of the basic Bloom filter is given in Fig. 6.2(a). To represent a set S Bloom filter utilizes K independent hash functions h1, h2...hk with range [0.n − 1]. For each element x ∈ S , bit at positions hi(x) of Bloom filter b are set to 1 for 1 ≤ i ≤ k as shown in Fig. 6.2(b). In order to test whether an element y is in S, it has to be checked whether all positions of hi(y) in b are set to 1.

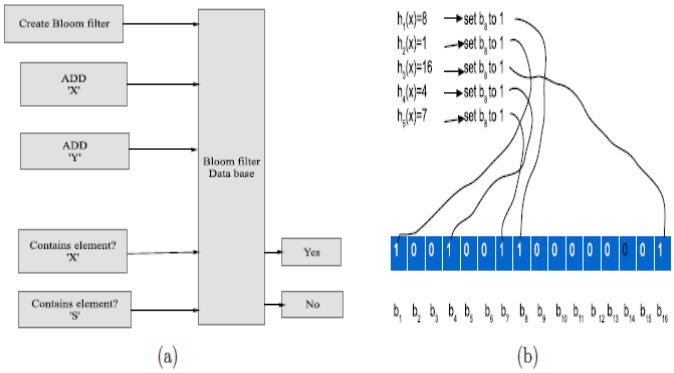


Fig.6.2: Basic bloom filter functionality

In this work, the Bloom filter is adapted in two ways,

1. One transform function h is applied to each element x ∈ S instead of multiple hash functions.
2. For a given Bloom filter b of length n, the set S is restricted to insert only *l* elements where $l \leq n$.

In this article adaptive Bloom filter is used for mixing an iris code generated from left and right eye of an individual. The operation of Bloom filter based mixing transform is as shown in fig. 6.3. The binary two dimensional feature vectors of width W and height H are extracted from both eyes of the subject. The iris codes are then divided into K blocks of equal size, where each column consists of w ≤ H bits. The entire column of each block is transformed to corresponding locations within the Bloom filters of length n = 2w. The transformation function h is implemented by mapping columns within 2D iris codes to the indexes of their decimal value, which is shown for two different code vectors as part in fig. 6.3 for each column x ∈ {0, 1}w, the mapping is defined as

$$b\left[h(x) \oplus AD\right] = 1, \text{with } h(x) = \sum_{i=0}^{w-1} x_i \cdot 2^i \quad (6.8)$$

where AD represents an application specific secret key which is incorporated to ensure unlinkability. Bloom filter based mixing transform fulfils two major requirement of biometric information protection namely, 1) Unlinkability 2) Irreversability.

1. *Unlinkability*

Unlinkability is provided by incorporating subject and application specific secret key AD, which XORed with a processed code word x before mapping it to a Bloom filter b.

2. *Irreversability*

Original positions of code words within iris codes are concealed, i.e for a given bloom filter b it is very difficult to tell which column of code vector is set 1 in protected template. By mixing code words of different feature vectors, it is even not clear which feature vector is set 1 in protected template.

Inevitably, a significant amount of code words can be mapped to identical position in Bloom filter even for small values of ***l.***

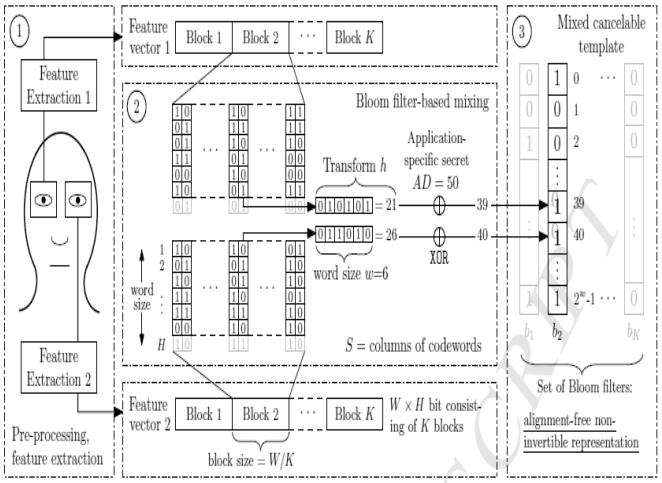


Fig.6.3: Operation of the Bloom filter based transform which mixes feature vectors of two different biometric instances [2]

VII. PROPOSED METHOD

In the proposed work two methods are used to extract feature vectors from the normalized image namely, 1) Gabor filters and 2) Random projection. As mentioned in problem statement section, the huge dimensionality of multi biometric template is addressed by using RP technique. In the proposed method, the multi instance (left and right iris) type multi biometric system is used and thereby addressing the hardware complexity and user inconvenience in using multiple sensors. Fusion of multiple information and template protection is achieved using Adaptive Bloom filters. The dimensions of the features extracted using RP technique is less compared to the features extracted using Gabor filtering technique. The ideology of the proposed method in the article can be expressed using the block diagram as shown in Fig. 7.1 and Fig.7.2.

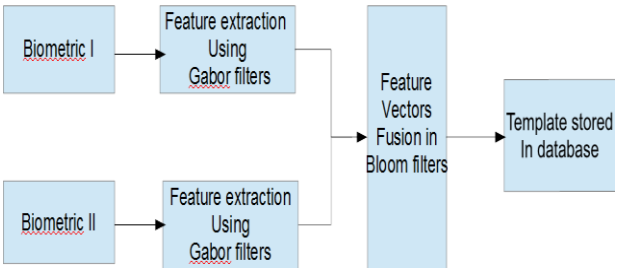


Fig.7.1: Gabor filter based multi biometric template protection

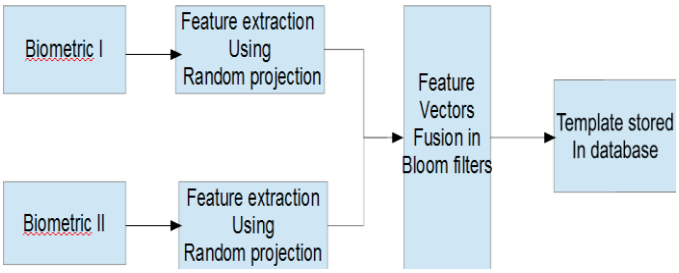
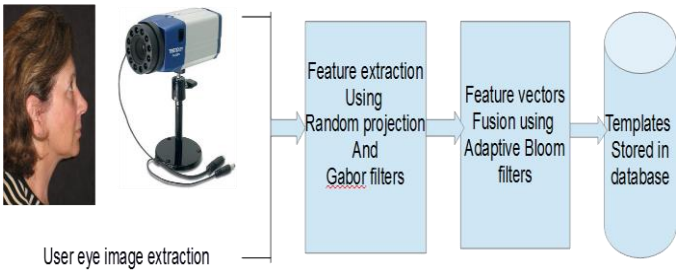
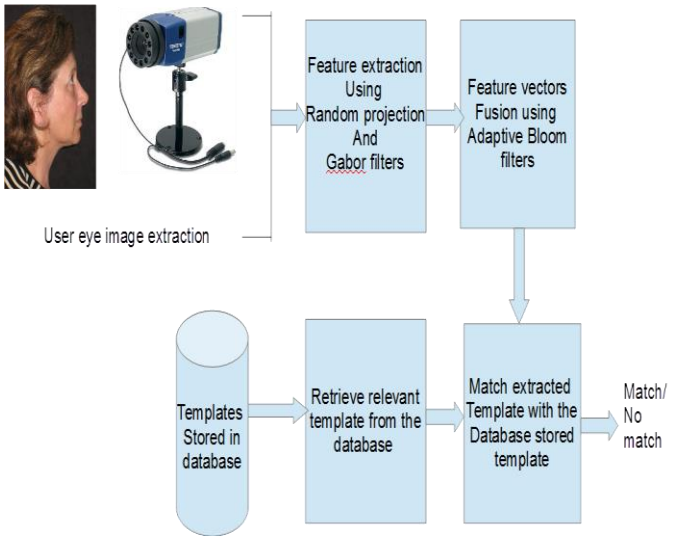


Fig.7.2: Random projection based multi biometric template protection

The verification application is developed using the proposed method. The block diagram depiction of the application is shown in Fig. 7.3 (a) and Fig. 7.3(b). The precision-recall curve is used to evaluate the performance of the proposed methods.



(a) Iris biometric enrollment



(b) Iris verification

Figure 7.3: verification application developed using proposed methods. During enrollment RP and Gabor filter methods are used to extract iris features and adaptive Bloom filter is used for feature vectors fusion. During verification the template generated using similar function is matched against stored database.

VIII. EXPERIMENTAL RESULTS AND EVALUATION

The experiments are carried out using CASIA-Iris-Interval database. It contains a total of 2,639 images from both left and right eye of 249 subjects. Each eye has a resolution of 320 × 380. Images in the CASIA iris database do not contain light reflections because, the images are captured in indoor environment using CASIA close up infra-red light illumination cameras. Sample images of left and right eyes of six different subjects are shown in Fig. 8.1.

This section provides brief description of basic steps involved in iris recognition system such as segmentation, normalization, feature extraction and matching. The method of performance evaluation is provided at the end of the section. For iris segmentation and feature extraction using the Gabor filter, an open source iris recognition system developed by Libor Masek [23] is used.

At segmentation stage, the circular Hough transform is applied to detect iris and pupil boundaries. This includes, first applying the Canny edge detection to generate an edge map. Gradients in vertical direction are weighted for the outer iris/sclera boundary. Vertical and horizontal gradients are weighted equally for the inner iris/pupil boundary. In order to make the circle detection process more effective and accurate, the two stage Hough transform algorithm is employed. Hough transform for the iris/sclera boundary is applied first. Later, the Hough transform for the iris/pupil boundary within the iris region is applied, instead of the complete eye region, because the pupil always lies within the iris region.

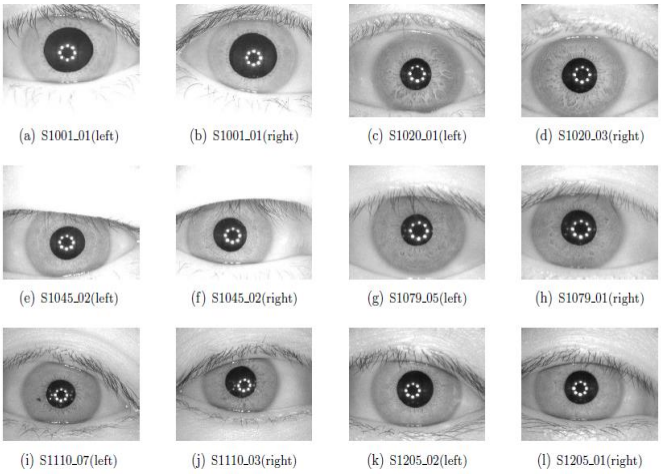


Fig.8.1: Sample pairs of left and right 320×380 resolution eye image of CASIA database. Numbers in captions refer to identifiers.

After the completion of the segmentation process, six parameters are stored, the radius, and x and y center coordinates for both circles. The iris isolated eye image is shown in Fig. 8.2(b). The top and bottom black regions denote the detected eyelash.

For normalization of iris region the Daugman's rubber sheet model is used.

This model remaps each point within the iris region to a pair of polar coordinates (r, θ) where r is on the interval [0, 1] and θ is an angle in the range [0, 2π]. The center of the pupil is considered as the reference point. Number of data points selected along each radial line is referred as the radial resolution. The number of radial lines going around the iris region is referred as the angular resolution.

In the feature encoding stage, two different feature extraction algorithms namely, 1) Gabor filter method, 2)random projection method are employed on the normalized iris texture. The first feature extraction method is implemented by convolving the normalized iris pattern with 1D logGabor wavelets. The 2D normalized iris pattern is broken up into number of 1D signals and then, these 1D signals are convolved with 1D logGabor wavelets. The rows of the 2D normalized iris pattern are taken as 1D signal. Each row of 2D iris pattern represents the angular direction and each column represents the radial direction of the iris region. For feature extraction, the angular direction is taken rather than radial one, since maximum independence occurs in angular direction. The total number of bits in the template will be *radial resolution*×*angular resolution*×*2*×*the number of filters* used.

A template size of 10 × 240 pixels is obtained in Gabor filter method by choosing radial resolution of 10 pixels, angular resolution of 40 pixels and number of filters used are 3. Secondly, that is, RP feature extraction method, the sparse projection matrix is multiplied with the normalized iris texture to obtain the feature code vector. In this, the dimension of the template is reduced to 5 × 40 which is almost 12 times lesser compared to the former feature extraction method.

For the proposed system, the adaptive Bloom filter is used to combine the multi instance feature vectors in both the methods. In former method, the template of size 10×240 (here H=10, W=240) is divided into 4 blocks (i.e K=4) of length l = 60, which are mapped to 4 Bloom filters of size n = 2^w = 2^{10} = 1024bits. In the latter method, the dimension reduced template of size 5× 40 (here H=5,W=40) is divided into 4 blocks (K=4) of length l = 10 bits. And these bits are mapped to 4 Bloom filters of size n = 2^w = 2^5 = 32bits. After fusion, the template obtained from the Gabor filter method consists of $K \cdot 2^w = 4 \cdot 2^{10}$ = 4096 bits. And the template obtained from RP method consists of $K \cdot 2^w = 4 \cdot 2^5$ = 128 bits. The template size in RP method is reduced by 32 times compared to Gabor filter method. The templates generated are stored in the database and used for identity verification application. The Hamming distance(HD) is used as a metric for recognition. Using the HD between two templates, a decision can be made as whether the two templates are generated from same or different irises. The HD is defined as the sum of disagreeing bits over N (the total number of bits in the template).

$$HD = \frac{1}{N}\sum_{j=1}^{N} X_j \neq Y_j \qquad (8.1)$$

The HD between the stored template and query template is calculated. Obtained HD can be used to decide whether the claimed identity is genuine or impostor. The Table 8.1 and 8.2 shows the example of how the genuine and impostor comparison is performed in the article. The diagonal values in the table indicates the HD obtained during genuine

verification of each template, the values above and below the diagonal indicates the impostor comparison. Table 8.1 and 8.2 shows the result for only five comparison. Similarly, the HD of remaining templates in the database are calculated.

Precision recall curve is used to evaluate the performance of the proposed verification system. In pattern recognition system with binary classification, the precision (also called positive predictive value) is defined as

$$\text{Precision} = \frac{\text{Number of relavant instances retrieved}}{\text{Total number of retrieved instances}}$$

Recall(also called sensitivity) is defined as

$$\text{Recall} = \frac{\text{Number of relavant instances retrieved}}{\text{Total number of relevant instances}}$$

	Sub1	Sub2	Sub3	Sub4	Sub5
Sub1	0	0.0234	0.0078	0.0078	0.0156
Sub2	0.0234	0.0156	0	0.0078	0.0234
Sub3	0.0078	0	0.0156	0.0078	0.0313
Sub4	0.0078	0.0078	0.0078	0	0.0313
Sub5	0.0156	0.0234	0.0313	0.0313	0.0234

Table 8.1: Identity matching using Random projection method

	Sub1	Sub2	Sub3	Sub4	Sub5
Sub1	0.0032	0.0049	0.0049	0.0046	0.0042
Sub2	0.0049	0.0063	0.0054	0.0061	0.0051
Sub3	0.0049	0.0054	0.0066	0.0068	0.0059
Sub4	0.0046	0.0061	0.0068	0.0063	0.0049
Sub5	0.0042	0.0051	0.0059	0.0049	0.0054

Table 8.2: Identity matching using Gabor filter method

Table 8.3 shows the recall and precision values obtained for Gabor filter and random projection verification systems by setting different thresholds. Fig.8.3 shows the precision-recall curve for the biometric verification system using two aforementioned methods. The graph obtained indicates that both the systems has comparable performance.

Recall	Precision(Gabor)	Precision(RP)
0.2	0.2	0.1379
0.3	0.1739	0.1379
0.5	0.0867	0.0835
0.6	0.0757	0.0788
0.8	0.0632	0.0756
0.9	0.0674	0.0756
1	0.0674	0.0743

Table 8.3: Table containing normalized recall and precision

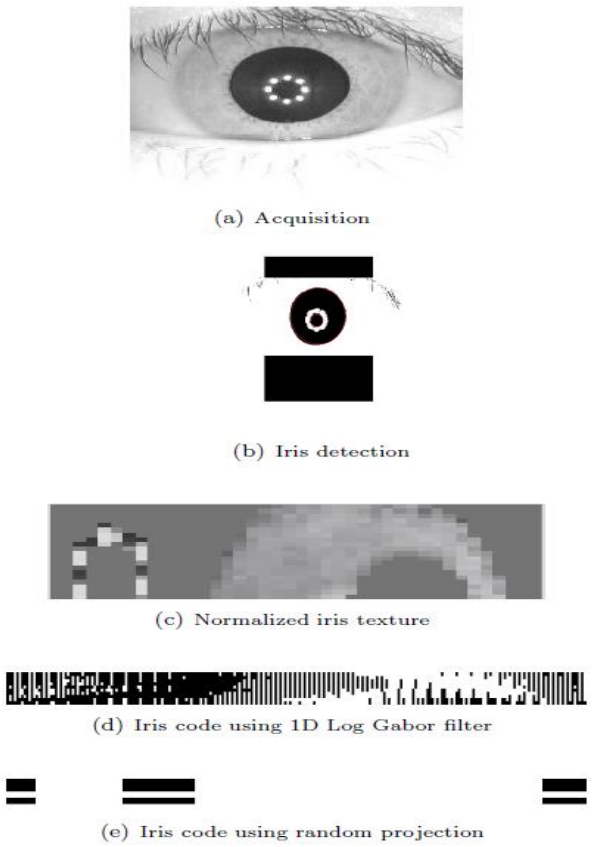


Fig.8.2: Iris detection, feature extraction methods applied to S1001L 01 of CASIA database

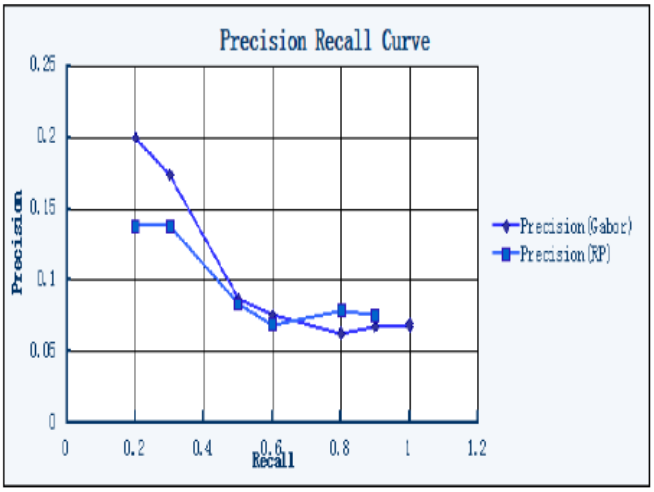


Fig.8.3: Precision recall curve for 1) Gabor filter based iris recognition system and 2) Random projection based iris recognition system.

IX. CONCLUSION

In this thesis, two techniques namely, the Gabor filters and the random projections (RP) are used for extraction of features from the iris. The CASIA-iris-interval database was used in both the methods. The adaptive Bloom filter was used for the fusion of features extracted from the multiple instances of the iris. In this work, the dimensionality of the iris template was

reduced using RP. The multi instance biometric was implemented in order to improve the accuracy of the system. The biometric verification system is developed using the proposed method and Gabor filtering method. The performance of the biometric verification system is evaluated using precision recall curves. In the biometric verification system, the proposed method using RP provides comparable performance to that of the Gabor filtering technique. The implementation of the multi modal biometrics for template generation and protection is the scope of the future work.

X. REFERENCES

[1] K. Nandakumar, "Multibiometric systems: Fusion strategies and template security," Tech. Rep.,2008.

[2] C. Rathgeb and C. Busch, "Cancelable multi-biometrics: Mixing iris-codes based on adaptive bloom filters," Computers & Security, vol. 42, pp. 1–12, 2014.

[3] A. A. Ross, K. Nandakumar, and A. K. Jain, Handbook of Multibiometrics (International Series on Biometrics). Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

[4] C. Rathgeb and C. Busch, Multi-Biometric Template Protection: Issues and Challenges. INTECH Open Access Publisher, 2012.

[5] [Online] Available: http://sequoyah.nist.gov/pub/nistinternalreports/NISTAPPNov02.pdf, November 2002

[6] [Online] Available: http ://news.bbc.co.uk/2/hi/uknews/politics/3693375.stm

[7] A. K. Jain, K. Nandakumar, and A. Nagar, "Biometric template security," EURASIP J. Adv. Signal Process, vol. 2008, pp. 113:1–113:17, Jan. 2008.

[8] A. B. J. Teoh, Y. W. Kuan, and S. Lee, "Cancellable biometrics and annotations on biohash," Pattern Recogn., vol. 41, no. 6, pp. 2034–2044, Jun. 2008.

[9] N. K. Ratha, S. Chikkerur, J. H. Connell, and R. M. Bolle, "Generating cancelable fingerprint templates." IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 4, pp. 561–572, 2007.

[10] Y. Sutcu, H. T. Sencar, and N. Memon, "A secure biometric authentication scheme based on robust hashing," in Proceedings of the 7th Workshop on Multimedia and Security, ser. MM&Sec '05. New York, NY, USA: ACM, 2005, pp. 111–116.

[11] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," in ACM CCS'99. ACM Press, 1999, pp. 28–36.

[12] L. Wu, P. Xiao, S. Jiang, and X. Yang, "A fuzzy vault scheme for feature fusion." in CCBR, ser. Lecture Notes in Computer Science, Z. Sun, J.-H. Lai, X. Chen, and T. Tan, Eds., vol. 7098. Springer, 2011, pp. 237–243.

[13] K. Nandakumar, A. K. Jain, and S. Pankanti, "Fingerprint-based fuzzy vault: Implementation and performance," Trans. Info. For. Sec., vol. 2, no. 4, pp. 744–757,Dec.2007.

[14] A. K. J. Abhishek Nagar, "On the security of non-invertible fingerprint template transforms," 2008.

[15] C. Rathgeb, F. Breitinger, C. Busch, and H. Baier, "On the application of bloom filters to iris biometrics," IET Biometrics, Dec. 2013, to appear.

[16] C. Moujahdi, S. Ghouzali, M. Mikram, M. Rziza, and G. Bebis, "Spiral cube for biometric template protection." in ICISP, ser. Lecture Notes in Computer Science, A. Elmoataz, D. Mammass, O. Lezoray, F. Nouboud, and D. Aboutajdine, Eds., vol. 7340. Springer, 2012, pp. 235–244.

[17] J. Daugman, "How iris recognition works." in ICIP (1), 2002,pp.3336.

[18] E. Bingham and H. Mannila, "Random projection in dimensionality reduction: Applications to image and text data," in Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

[19] M. K. Vildana Sulic, "Efficient dimensionality reduction using random projection," 2010.

[20] S. Dasgupta and A. Gupta, "An elementary proof of a theorem of johnson and lindenstrauss," Random Struct. Algorithms, vol. 22, no. 1, pp. 60–65, Jan. 2003.

[21] D. Achlioptas, "Database-friendly random projections: Johnson-lindenstrauss with binarycoins," J. Comput. Syst. Sci., vol. 66, no. 4, pp. 671–687, 2003.

[22] W. lun Chao, "Gabor wavelet transform and its application."

[23] L. Masek, "Recognition of human iris patterns for biometric identification," Tech. Rep., 2003.

Sushma H.R. has completed B.E. degree in Electronics and Instrumentation Engineering from Kuvempu University in the year 2008. Currently, she is pursuing the M.Tech. degree in Signal processing from Visvesvaraya Technological University, Belagavi,India. Her current research interests include signal processing and biomedical image processing.

Sandeep R. received the B.A. degree in Hindi from Mysore Hindi Prachar Parishad, the B.E. degree in Electronics and Communication Engineering and the M.Tech. Degree in Electronics Engineering from Visvesvaraya Technological University in the year 2001, 2006 and 2009 respectively. Currently, he is pursuing Ph.D. in the department of Electronics and Electrical Engineering, Indian Institute of Technology Guwahati, India and working as Associate Professor at Cambridge Institute of Technology (Affliated to Visvesvaraya Technological University), Bengaluru, India. He has both academic and industry experience. His current research interests include perceptual image hashing, perceptual video hashing, biometric hashing and biomedical image processing.