Expert Systems with Applications 39 (2012) 4915–4926

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

An efficient color and texture based iris image retrieval technique

Umarani Jayaraman *, Surya Prakash, Phalguni Gupta

Department of Computer Science and Engineering, Indian Institute of Technology Kanpur, Kanpur 208016, India

ARTICLE INFO

Keywords:
Iris color indexing
Kd-tree
Iris texture retrieval
Color iris images

ABSTRACT

This paper proposes a hierarchical approach to retrieve an iris image efficiently from for a large iris database. This approach is a combination of both iris color and texture. Iris color is used for indexing and texture is used for retrieval of iris images from the indexed iris database. An index which is determined from the iris color is used to filter out the images that are not similar to the query image in color. Further, iris texture features of those filtered images, are used to determine the images which are similar to the query image. The iris color information helps to design an efficient indexing scheme based on color indices. The color indices are computed by averaging the intensity values of all red and blue color pixels. Kd-tree is used for real-time indexing based on color indices. The iris texture features are obtained through Speeded Up Robust Features (SURF) algorithm. These features are used to get the query's corresponding image at the top best match. The performance of the proposed indexing scheme is compared with two well known iris indexing schemes (Mehrotra, Majhi, & Gupta, 2010; Puhan & Sudha, 2008) on UPOL (Dobeš, Machala, Tichavský, & Pospíšil, 2004) and UBIRIS (Proencca & Alexandre, 2005) iris databases. It is observed that combination of iris color and texture improves the performance of iris recognition system.

1. Introduction

An iris based biometric system is used to authenticate an individual with the help of iris textural patterns. For a query iris image, the problem of iris identification system is to find the top t best matches from the iris database which consists of N iris images. For a large database, it becomes computationally expensive for making 1:N searches. In order to design a system which is powerful and fast, the matching engine needs to search in a reduced space in the database.

Recently, there exist indexing schemes of iris images which are based on iris texture. Yu, Zhang, Wang, and Yang (2005) have proposed a coarse iris classification technique using fractals that classifies iris images into four categories. The iris image is segmented into sixteen blocks, eight belonging to an upper group and remaining to a lower group. Fractal dimension value is calculated from the image blocks and takes the mean value of the fractal dimension as the upper and the lower group fractal dimensions. Finally, all the iris images are classified into four categories in accordance with the fractal dimensions of the upper and the lower groups.

Mukherjee and Ross (2008) have proposed two iris indexing schemes. The first scheme is based on an analysis of iris codes extracted from an individual iris. The iris code is a binary template whose dimensionality corresponds to the dimensions of the unwrapped iris structure. Iris codes are first projected onto a lower dimension (according to the number of significant principal components). Then k-means clustering approach is used to partition the reduced dimension iris code into multiple groups. The second scheme examines the textural content of the iris image using the Signed Pixel Level Difference Histogram (SPLDH) of the raw pixel intensities. These two indexing schemes have been tested on CASIA version 3.0 iris database.

Vatsa, Singh, and Noore (2008) have proposed an indexing scheme which consists of two phases. At first phase, it uses Euler code to generate a small subset of possible matches. In the next phase, it uses 2ν-SVM match score fusion algorithm to find the best matches from the list of possible matches obtained in first phase using textural and topological features of the iris images.

Mehrotra et al. (2010) have proposed geometric hashing based indexing scheme for iris images using SIFT (Lowe, 1999) feature points. The detected keypoints of an iris image are used for indexing. Each of the detected key feature points are hashed into its appropriate bin of the hash table using geometric hashing (Wolfson & Rigoutsos, 1997). When a query image comes for searching, extracted feature points are mapped into the hash table to its appropriate bin and matching is done only with the mapped bins.

There are some indexing schemes which are based on iris color. The first attempt to index iris images using color has been made by Fu, Caulfield, Yoo, and Atluri (2005). This technique is based on artificial color filtering and has used a set of nine artificial color filters to narrow down the search space. To make an artificial color

* Corresponding author. Tel.: +91 512 2597579; fax: +91 512 2597647.
E-mail addresses: umarani@cse.iitk.ac.in (U. Jayaraman), psurya@cse.iitk.ac.in (S. Prakash), pg@cse.iitk.ac.in (P. Gupta).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.10.025

filter, one must train a discriminator that assigns a pixel either to the class of interest or to some other class. However, the artificial color filtering is not an effective approach to index because the colors are generated artificially which are very much different from the natural iris color. Further, the performance of the indexing scheme has not been tested on any publicly available database.

Another iris color based indexing scheme has been made by Puhan and Sudha (2008). It also refers group based color indexing scheme which relies on the natural iris color. This technique converts the iris color images from *RGB* space to YC_bC_r color space and computes two types of color indices, namely blue and red indices using C_b and C_r components, respectively. The range of values of red color and blue color indices are partitioned individually into n groups. Depending upon the value of red and blue color indices, an image is assigned to one of these groups. During searching, for a query, a few groups from blue and red color indices are selected based on the blue and red color indices of query image. The nearest identities are declared based on the intersection between these groups. However, the problem with this group based indexing scheme is that a minor variation in red and blue color values may lead an image being assigned to a different group. Further, artificial color filtering and group based indexing scheme are solely dependent on color information for both indexing and recognising the iris images which may not be always helpful to distinguish the iris images of different subjects accurately.

Table 1 compares the existing iris indexing schemes with respect to indexing methodology, features used for indexing and recognition phase of iris system. Iris color, which identifies local features and largely dependent on view and resolution, can be efficiently used for indexing. It is used for indexing through which a large database is reduced to a small subset while iris texture is used to retrieve the corresponding iris image in the top best match from the small subset. In this paper, iris color information helps to design an efficient indexing scheme. The scheme is found to be robust in comparison with any other iris based indexing scheme because it is based on the combination of iris color and texture.

1.1. The role of iris color

Iris color is a highly complex phenomenon consisting of the combined effects of texture, pigmentation, fibrous tissue and blood vessels within the iris stroma, which together make up an individual's epigenetic constitution in this context. It is strongly pigmented with colors ranging from brown to green, blue, gray and hazel. There are many colors in different parts of the retina (Keating, 1988). Iris color is determined by the amount and the type of pigments present in the iris. It can provide a large amount of information about an individual and a classification of various colors may be useful in documenting pathological changes or determining the way that a person may respond to various ocular pharmaceuticals. A sample of 324 patients from 5 urban clinics in Massachusetts and Maryland, the investigators are reported on the database of iris color in Iris color as follows.

- 32% blue/grey irises.
- 15% blue/grey/green irises with brown/yellow specks.
- 12% green/light brown irises with minimal specks.
- 16% brown irises with specks.
- 25% dark brown irises.

This shows that use of iris color can improve the recognition of iris along with its texture. Also, color recognition of iris provides significant discrimination on its own, so it is worth to combine with spatial pattern recognition. Iris color can be an additional spatial pattern recognition to make the system even more reliable. There exist methods of textural based recognition of iris in Boles and Boashash (1998), Daugman (1993, 2003), Ganorkar et al. (2007), and Daugman et al. (2001).

1.2. Challenges

Human recognition based on iris is claimed to perform high in accuracy. However, all publicly available databases of iris images are captured in a controlled environment to ensure high quality. The iris images are captured with the help of a proper iris capture setup. Users are required to look into a camera from a fixed distance to capture the images. Images captured in an uncontrolled environment produce non-ideal iris images with varying image quality. If the eyes are not properly opened, certain regions of the iris cannot be captured due to occlusion and it affects the process of iris segmentation and consequently, the recognition performance. Images having partial occlusion, illumination, gaze direction, with different scale, presence of eyelids and eyelashes affect the performance of the system. Recently UBIRIS iris database has the images which are captured in an uncontrolled environment to produce non-ideal iris images with noise and varying image quality.

Fig. 1 shows various such noisy images from UBIRIS iris database. Fig. 1(a) is an example of iris texture occluded due to eyelids and eyelashes while Fig. 1(b) shows iris image of an individual having the motion blur and the presence of limited information Fig. 1(c) is an example of an iris image of an individual with a different gaze direction, Fig. 1(d) shows the images of an individual having the different illumination factor. The presence of noisy images increases the false rejection rate (FRR), thus decreasing the performance of the iris recognition system.

1.3. Outline

In Section 2 the hierarchical framework of the proposed technique is discussed. Section 3 introduces the multidimensional color indices indexing scheme using Kd-tree (Bentley, 1975). Kd-tree is a

Table 1
Comparison of existing iris indexing schemes.

Research paper	Indexing methodology	Indexing features	Recognition features	Additional comments
Texture based approaches				
Yu et al. (2005)	Classification technique	Fractal dimensions	Fractals	Texture based
Mukherjee and Ross (2008)	*k*-Means clustering	Iris code Iris texture	Iris code and texture	Texture based
Vatsa et al. (2008)	Feature based	Euler code	2ν-SVM texture and topological	Texture based
Mehrotra et al. (2010)	Geometric hashing	SIFT Key feature points	SIFT Key feature points	Texture based
Color based approaches				
Fu et al. (2005)	Pattern recognition method	Artificial color filter	Artificial color	Color based
Puhan and Sudha (2008)	Group based method	Color indices	Color indices	Color based
Color and texture based approaches				
Proposed technique	Kd-tree based method	Color indices	Iris texture	Color and texture based

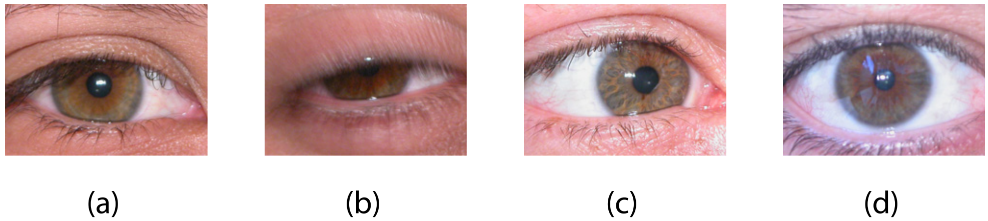

Fig. 1. Images representing the challenges of iris recognition from UBIRIS database.

data structure which can be used to index two dimensional data efficiently. Through this indexing scheme, the search space of the large iris database has been reduced effectively. In order to obtain the correct identity in the top best match from the reduced search space, an iris retrieval technique which uses iris texture has been proposed in Section 4. Section 5 shows the performance of the hierarchical approach to retrieve iris images tested on two publicly available databases, namely UPOL (Dobeš et al., 2004) and UBIRIS (Proencca & Alexandre, 2005). Also, performance of the proposed indexing scheme has been analysed against the group based color indexing scheme (Puhan & Sudha, 2008) where iris color is used for indexing. The experimental results show that the hierarchical approach works well even with the changes in the iris images due to gaze, illumination, partial occlusion and different scale. Conclusion is given in the last section. Table 2 shows the summary of the symbols and their definition used in this paper.

2. Hierarchical framework

The hierarchical framework consists of two phases, known as off line phase and online phase. Fig. 2 illustrates the hierarchical query framework. In the off line phase, color indices from each iris image in the database are extracted and stored in Kd-tree (Bentley, 1975). In the online phase, for a query image, the framework accesses the Kd-tree to filter out images that are not similar to the query image in color. Iris texture based features of those filtered images are used to identify the images similar to the query image by performing one to one match. It also helps to get the query's corresponding identity in the top best match.

2.1. Offline phase

During off line phase, iris color is extracted for each iris image to store in Kd-tree (Bentley, 1975). Iris color is the multidimensional (2-tuple) color indices. Kd-tree is formed to index two dimensional data, so the use of Kd-tree data structure is the best choice to index 2-tuple color indices efficiently. Given a discrete color space, color indices are the average of intensity values of all red and blue color pixels. These indices are invariant to translation, rotation about an axis perpendicular to the image and change slowly with rotation about other axes and occlusion (Shapiro & Stockman, 2003). Further, translation of an RGB image into the illumination invariant rg-chromaticity space allows the color indices to operate well in varying light levels. On the other hand, color indices for different iris images can differ significantly from each other. Therefore, the color indices can play an important role to index iris images.

Table 2
Summary of symbols and definitions.

Symbols	Definitions
DB	Database
N	Number of images in a database
q	Query iris image
I	Segmented iris region
k	Number of nearest neighbors to find
K	Subset of size k
H	Histogram

2.2. Online phase

In the online phase, the hierarchical framework consists of two steps process. In the first step, for a given query iris image 2-tuple color indices are extracted. Then the framework accesses the Kd-tree to filter out the database. As a result, a small subset is left (filtered images) whose images are very similar in color. The iris images obtained after the filtering step which is based on only the color indices, may not get the query's corresponding identity in the top best match because it ignores iris shape and texture. So in order to get the query's corresponding identity from the filtered images as the top best match, iris texture is used.

In the second step, iris texture features of those filtered images are used to perform one to one match to determine the top best match. To extract the iris texture, the proposed technique uses local feature descriptor algorithm, Speeded Up Robust Features (SURF) (Bay, Ess, Tuytelaars, & Van Gool, 2008). Local features are extracted by finding the key points in an iris image and forming descriptor vector around each detected key point.

3. Color based iris indexing scheme

In iris based biometric identification, features are extracted from each iris image. These features are mapped into a feature space and is defined to be a point $(f_1,\ldots,f_n)$ in the space. A metric is defined on the space and identification is done by finding the k nearest points in feature space to the given query point. All these feature points are indexed in such a way that it eliminates the comparison with all feature points in the space while searching.

In order to identify a subject using iris color, one can use a color histogram based method. A histogram of a color image can be represented by a vector $H = [h_1,\ldots,h_n]$ in which each element h_j contains the number of pixels having the color j in the image and can be considered as the probability density function (pdf) of the color values. Each of the different bins of the histogram can be considered as a different feature for recognition. The histogram H for each image is stored in the database. During identification for a query iris image, its nearest neighbors are found by comparing the histogram stored in the database using the sum of the squared differences. However, there are several issues with color histogram based method. One of the major issues is the high dimensionality of the color histogram. Even if one uses feature reduction such as pre-filtering and quantization methods, the feature space still occupies more than 100 dimensions. As a result, it is difficult to index such a high dimensional features. This larger dimension also increases the computational complexity. All well known color histogram techniques such as color histogram intersection (Swain & Ballard, 1991), color constant indexing (Funt & Finlayson, 1995), cumulative color histogram (Stricker & Orengo, 1995) and color

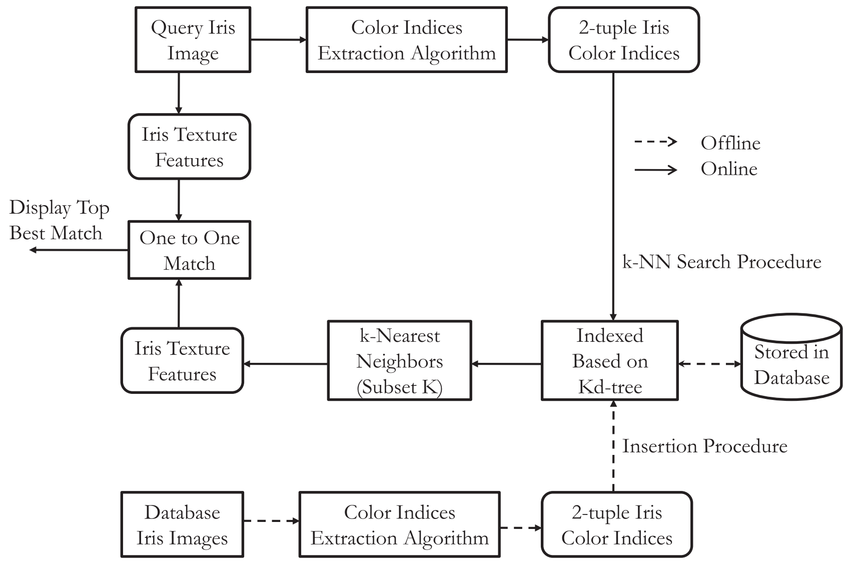

Fig. 2. Hierarchical framework.

correlograms (Huang, Kumar, Mitra, Zhu, & Zabih, 1997) suffer from the curse of dimensionality (Bengtsson, Bickel, & Li, 2008).

This paper proposes a Kd-tree (Bentley, 1975) based indexing scheme for iris database. Kd-tree has been used to index two dimensional data efficiently. The problem of features having higher dimension obtained from the color histogram can be handled efficiently by a simple technique known as color indices which can be computed by averaging the intensity values of all blue and red color pixels. In the proposed indexing scheme, two dimensional data are blue and red indices of iris images computed under YC_bC_r color space. Let I be the segmented iris region for any given iris image. The iris region I which is originally in RGB color space is converted to YC_bC_r color space. It uses the technique presented in Gonzalez and Eddins (2009) as follows

$$Y(x,y) = 16 + 1/256 \cdot (65.7 \cdot R(x,y) + 129.0 \cdot G(x,y) + 25.0 \cdot B(x,y))$$

$$C_b(x,y) = 128 + 1/256 \cdot (-37.9 \cdot R(x,y) - 74.4 \cdot G(x,y) + 112.4 \cdot B(x,y))$$

$$C_r(x,y) = 128 + 1/256 \cdot (112.4 \cdot R(x,y) - 94.1 \cdot G(x,y) - 18.2 \cdot B(x,y))$$

where $R(x,y)$, $G(x,y)$ and $B(x,y)$ are red, green and blue pixel values of (x,y) over RGB color space, respectively. Then these blue and red indices of I under YC_bC_r color space can be obtained by

$$b_I = \frac{1}{|I|} \sum_{\forall (x,y) \in I} C_b(x,y)$$

$$r_I = \frac{1}{|I|} \sum_{\forall (x,y) \in I} C_r(x,y)$$

where $C_b(x,y)$ and $C_r(x,y)$ are the chrominance of blue and red color, respectively for pixel (x,y) of I. Importantly, since YC_bC_r color space separates illumination from color, conversion from RGB to YC_bC_r color space allows the color indices to operate well in varying light levels. As a result, high indexing performance can be achieved. It can be noted that, computing color indices by converting RGB color space into rg-chromaticity color space or in the RGB color space itself may not achieve high indexing performance because of varying illumination in the images. An example of color indices computed on the segmented iris region I is shown in Fig. 3.

Let $DB = \{D_1, D_2, \ldots, D_N\}$ be the database of N iris images where D_L, $L = 1, 2 \ldots, N$, is a 2-tuple containing blue and red indices of the iris region I over YC_bC_r color space, *i.e.* $D_L = (b_L, r_L)$. Given the database DB, one can create a Kd-tree by inserting blue and red indices (b_L, r_L) of N images by invoking *Insert* procedure proposed in Bentley (1975). For a query image q, blue and red indices are computed over YC_bC_r color space and *k-NN Search* (Bentley, 1975) is used to obtain a subset K of size k images which are nearest to the query image q. The subset K contains all iris images satisfying $(\forall i \in K)$,

$$\|q - i\| \leqslant \|q - n\|, \quad \forall n \in (DB - K) \tag{1}$$

where $\|\cdot\|$ is a distance measure.

4. Texture based iris retrieval technique

Iris images in the subset K are arranged based on the Euclidean distance of color image from the query image. However, the color indices which are the average of red and blue color values of an iris image may not get the query's corresponding identity in the top best match. So in order to get the corresponding identity from the subset K in the top best match, texture is used. This section proposes a technique to retrieve the top best match from a subset K based on the iris texture patterns. The pattern of iris region I which has richer information can help to improve the matching performance. Such an improvement of iris recognition system has been achieved due to integration of a local feature descriptor, Speeded-Up Robust Features (SURF) (Bay et al., 2008). Local features are extracted by finding the key points in an image and forming descriptor vector around each detected key point.

SURF has more discriminative power than any other local feature descriptor such as SIFT (Lowe, 1999). Also it can be computed more efficiently and results lower dimensional features and hence the matching is much faster. In contrast to SIFT which approximates Laplacian of Gaussian (LOG) with difference of Gaussians (DOG), SURF approximates second order Gaussian derivatives with box filters. For each pixel (x,y) in an iris region I, Hessian matrix at scale σ is obtained. The determinant of Hessian matrix is used to select location and scale. The local maxima found using approximated Hessian matrix determinant are interpolated in scale and image space. Fig. 4 shows detected SURF key points of an iris region I of UPOL database. SURF descriptors are obtained by taking a rectangular window around each detected key point. The window is splitted into 4×4 sub-regions. For each sub-region, Haar wavelet

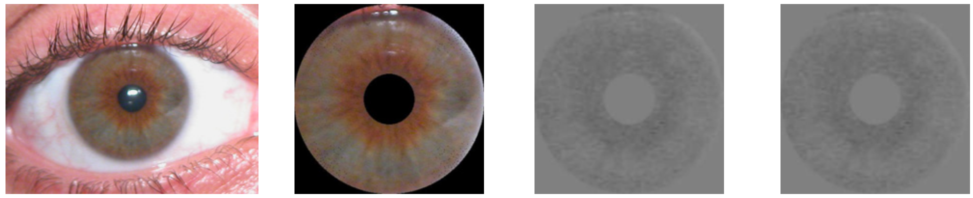

Fig. 3. Computing the color indices: (a) input iris image, (b) segmented iris, (c) blue index, (d) red index.

responses are extracted. The wavelet response in horizontal (d_x) and vertical (d_y) directions and the absolute values of wavelet responses $|d_x|$ and $|d_y|$ for each sub region are summed up to find the polarity of image intensity changes. Hence, feature vector for each sub-region is given by

$$f = \left(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|\right) \tag{2}$$

All descriptor vectors from all 4×4 sub regions are concatenated to have feature descriptor of length 64. The descriptor vector of all key points forms the feature vector F_J of Jth iris image in the subset K. Finally, query image q is matched with all the images in the subset K using these descriptor vectors. The image which has the maximum matching points is displayed as the top best match for the given query image.

5. Performance evaluation

To determine the performance of the proposed indexing scheme, four measures, namely, *hitrate*, *bin-miss rate*, *penetration rate*, and *Cumulative Match Characteristic curve* are used.

- The *hitrate* (H_r) is defined by $H_r = \left(\frac{X}{L}\right) \times 100\%$ where X is the number of times that a query's corresponding identity is occurred at top best match and L is the total number of attempts.
- The *bin-miss rate* (B_r) is defined as $B_r = 100 - H_r$
- The *penetration rate* (P_r) is defined $P_r = \left(\frac{1}{X}\sum_{i=1}^{X}\frac{k_i}{N}\right) \times 100\%$ where X is the total number of query images correctly identified, k_i is the number of images retrieved per query and N is the database size.
- *Cumulative Match Characteristic (CMC) curve* represents the relationship between probability of identification accuracy against rank (top position).

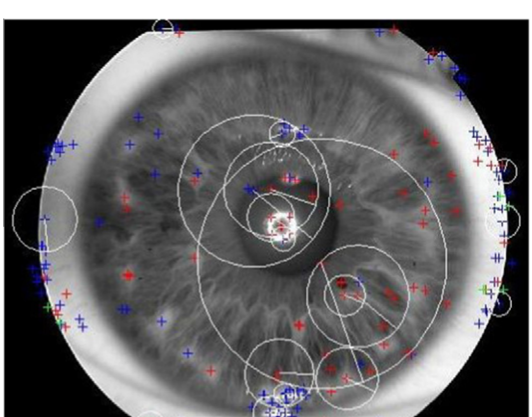

Fig. 4. Detected SURF keypoints of an iris image.

5.1. Experiment 1: UPOL database

There are 384 iris images collected from 64 subjects in UPOL database (Dobeš et al., 2004). It contains 3 left and 3 right iris images of each subject. Few samples of iris images from this database are shown in Fig. 5. For each subject among 3 left/right iris images 2 images are considered for training and the remaining one image is used for testing.

5.1.1. Color based indexing

To improve the time required for achieving similar type of accuracy as in exhaustive search, the databases of iris images are indexed using its color. The blue and the red indices over YC_bC_r color space are computed for all images. It has been observed that variation of each such color indices for different images of the same subject is very small while that between two distinct subjects is large for blue and red indices. This shows that the 2-tuple consisting of blue and red indices over the YC_bC_r color space can play an important role to determine a good index for each iris image.

The efficiency of identification system is also measured in terms of *bin-miss rate* and *penetration rate*. The *bin-miss rate* is obtained for various subset size k lying between 1 and 128 which is total database size (N). Fig. 6(a) shows relationship between *bin-miss rate* and subset size k while Fig. 6(b) shows the relationship between *penetration rate* and subset size k for both left and right iris databases. From these two graphs, it can be observed that *bin-miss rate* decreases with the increase in the value of subset size while *penetration rate* increases with the subset size. This is obvious because increase in the k-nearest neighbors for a given query increases the *penetration rate*. Thus there exists a trade off between these two performance measures. Increasing the value of subset size reduces the *bin-miss rate* while increases the *penetration rate*. Low *bin-miss rate* results in a less secure identification system while high *penetration rate* results in more response time. Hence, we have to choose the subset size (optimum) in such a way that both are at acceptable level.

Fig. 7(a) and (b) show the relationship between two performance measures for left and right iris database, respectively. The value of the subset size is considered as an optimum where two curves intersect *i.e.* $BMR = PR$ which is shown by arrow in Fig. 7. For the optimum subset size, equal *bin-miss rate* and *penetration rate* are achieved. In our experiment, we are choosing the subset size 15 and 11 such that both *bin-miss rate* and *penetration rate* are equal for left and right database, respectively. Further increase in the value of subset size to get zero *bin-miss rate* increases the *penetration rate*. In this case, for the subset size of 20 and 15 the *bin-miss rate* is zero for left and right iris database, respectively. However, after 20 and 15 increases in the values of subset size increases only the *penetration rate*. The *penetration rate* becomes 1 when k is 128 which is database size (N) and is shown in Fig. 7. Hence optimum subset size is chosen as it gets the required results with low *penetration rate*. After indexing using color, a small subset

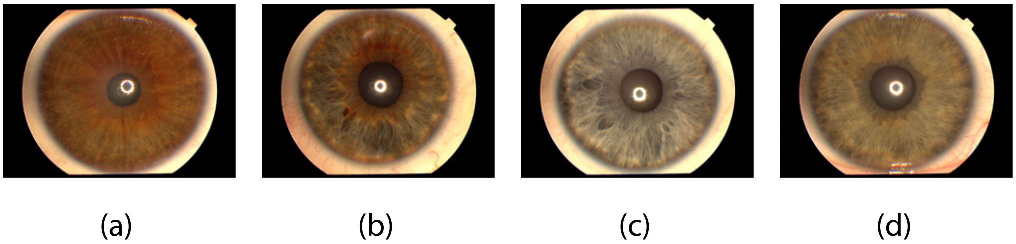

Fig. 5. Sample iris images from UPOL database.

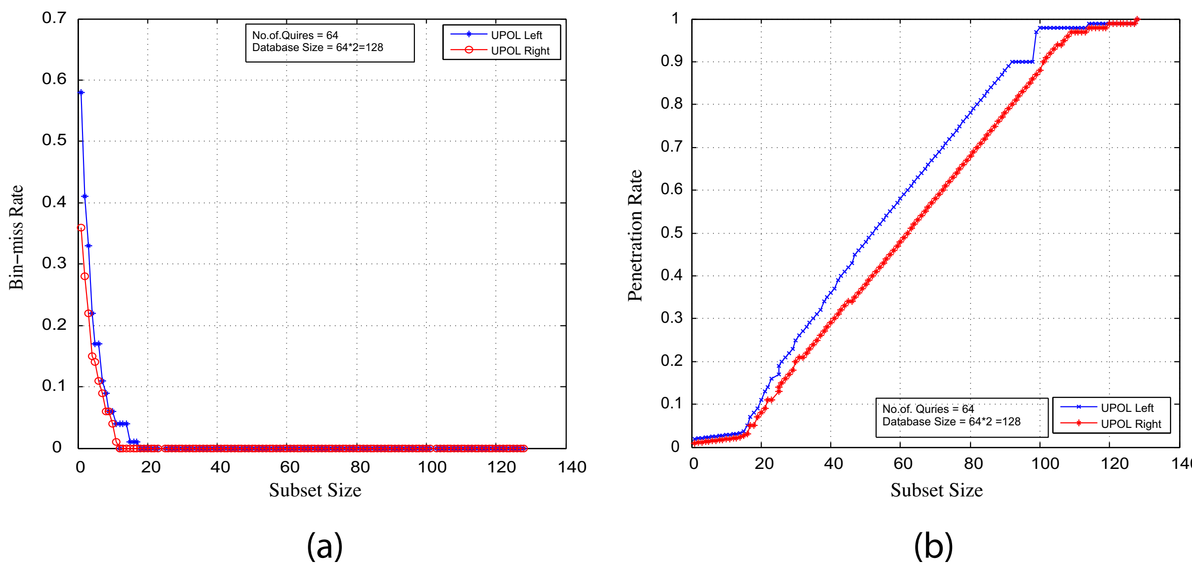

Fig. 6. UPOL database, (a) relationship between bin-miss rate and subset, (b) relationship between penetration rate and subset.

K consisting of *k* nearest images in color from the database is retained for query *q*.

5.1.2. Texture based retrieval

At this stage, the proposed texture based iris retrieval technique has been used to get the identities of a query as the top best match. The Cumulative Match Characteristic (CMC) curve obtained for an optimum subset size *k* after indexing with the help of color is shown in Fig. 8(a). It can be observed that queries are correctly identified with the probability of 98.9% and 98.0% for left and right iris databases. Also, the correct identities are found in top 15th and 11th positions for left and right iris databases, respectively. Similarly, the CMC curve obtained with the help of color followed by texture is shown in Fig. 8(b). It can be noted that for all the query images, corresponding identities are found in 1st position with the probability of 98.9% and 98.0% for left and right iris databases. Table 3 shows *hitrate* and *penetration rate* achieved by the proposed indexing scheme for the UPOL database.

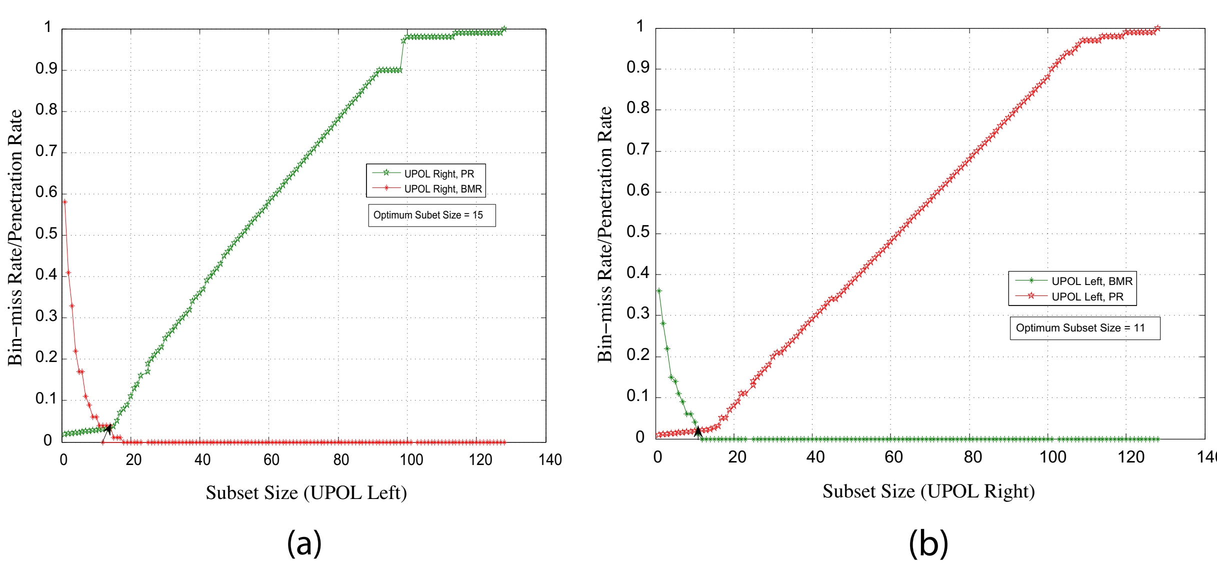

Fig. 7. UPOL database, (a) optimum subset size for UPOL-left, (b) optimum subset SIZe for UPOL-right.

5.2. Experiment 2: noisy UBIRIS database

The performance of the proposed indexing scheme is also evaluated on noisy UBIRIS (Proencca & Alexandre, 2005) database and is analysed against the known group based color indexing scheme (Puhan & Sudha, 2008). UBIRIS database consists of 1860 iris images collected from 372 subjects (5 samples each) in two Sessions. For Session-I, images have been captured from 241 persons under controlled environments in order to minimize noise factors, specially those relative to reflections, luminosity and contrast. For Session-II, images are captured from 131 persons by changing the location to introduce natural luminosity factor. This has enabled the appearance of heterogeneous images with respect to reflections, contrast, luminosity and focus problems. All the images from both Sessions have been classified with respect to three parameters. The classification statistics are given in Table 4. Sample images for both good and bad visible iris along with their segmented iris are shown in Figs. 9 and 10, respectively. For noisy UBIRIS database, iris region (I) from the eyelid has been segmented using the method proposed in Kimme, Ballard, and Sklansky (1975). It has been observed that there are some iris images in the database which are occluded more than 75%. This type of images has not been considered for testing the proposed technique. If one considers only those subjects having at least two good segmented iris regions (I), one can get a total of 237 out of 241 different subjects from Session-I. Similarly from Session-II, a total of 130 out of 131 different subjects are found. For each subject, one image or more images are taken for training and remaining images are considered for testing.

5.2.1. Color based indexing

Fig. 11(a) shows relationship between *bin-miss rate* and subset size while Fig. 11(b) provides the relationship between *penetration rate* and subset size. The *bin-miss rate* is obtained for various subset size k lying between 1 and 308 for Session-I and between 1 and 156 for Session-II, respectively which are the database size (N). Similar to UPOL database, it could be observed from these graphs that *bin-miss rate* decreases with the increase in the value of subset while *penetration rate* increases with the subset size. Fig. 12(a) and (b) shows the relationship between two performance measures for Session-I and Session-II, respectively. In our experiment, we are choosing the optimum subset size as 38 and 31 for Session-I and Session-II iris databases, respectively. Further increase in the value of subset size to get zero *bin-miss rate* increases the *penetration rate*. In this case, for the subset size of 117 and 50 the *bin-miss rate* is zero for Session-I and Session-II iris database, respectively. However, after 117 and 50 increases in the values of subset size increases only the *penetration rate*. The *penetration rate* becomes 1 when k is 308 and 156 for Session-I and Session-II, respectively which are the database size (N) and is shown in Fig. 12. Hence optimum subset size gets the required results with low *penetration rate*. After indexing using color, a small subset K consisting of k nearest images in color from the database is retained for query q.

5.2.2. Texture based retrieval

At this stage, the proposed texture based iris retrieval technique has been used to get the identities of a query as the top best match. The CMC curve obtained for an optimum subset size k after indexing with the help of color is shown in Fig. 13(a). It could be observed that queries are correctly identified with the probability of 95.8% and 96.2% for Session-I and Session-II iris databases. Also, the correct identities are found in top 38th and 31st positions for Session-I and Session-II iris databases, respectively. Similarly, the CMC curve obtained with the help of color followed by texture is shown in Fig. 13(b). It could be noted that for all the query images, corresponding identities are found in 1st position with the

Table 3
Identification result by proposed indexing scheme for UPOL database.

Indexing scheme	Left		Right	
	Hitrate (%)	PR (%)	Hitrate (%)	PR (%)
Proposed technique	98.91	1.09	98.00	2.0

Table 4
Classification of images with respect to three parameters.

Parameter	Good (%)	Average (%)	Bad (%)
Focus	73.83	17.53	8.63
Reflections	58.87	36.78	4.34
Visible iris	36.73	47.83	15.44

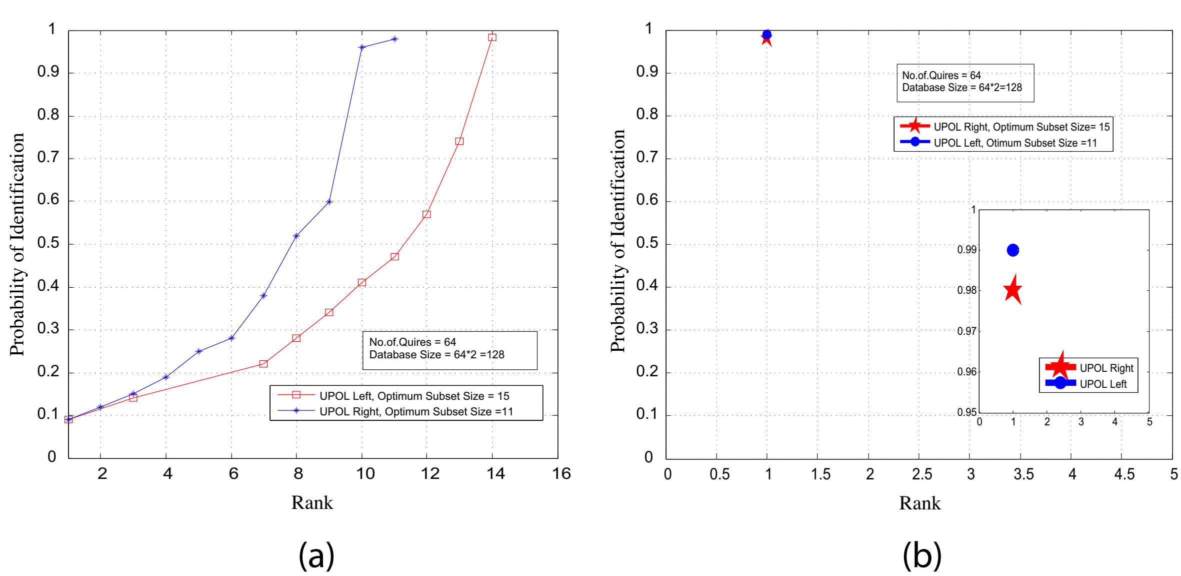

Fig. 8. CMC curve for optimum subset (BMR = PR), (a) indexed using color, (b) indexed using color followed by texture.

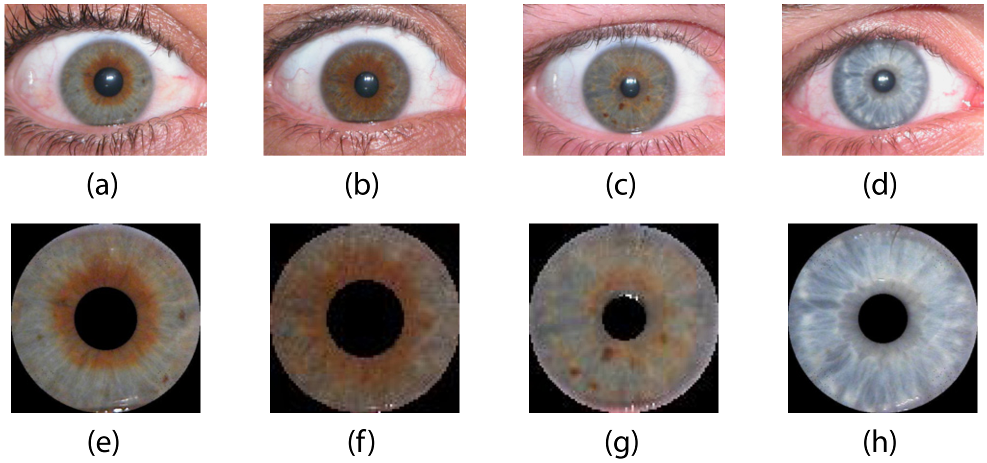

Fig. 9. Good visible iris samples from UBIRIS database (segmentation success cases).

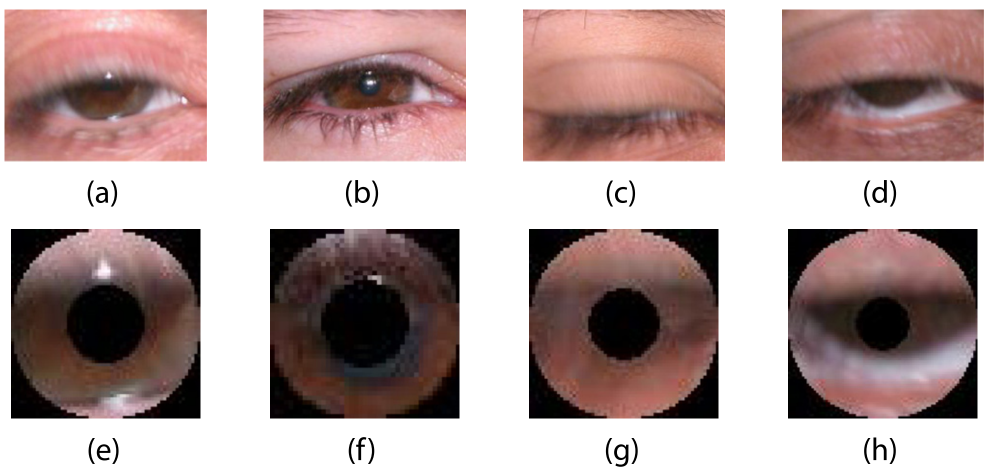

Fig. 10. Poorly visible iris samples from UBIRIS database (segmentation failed cases).

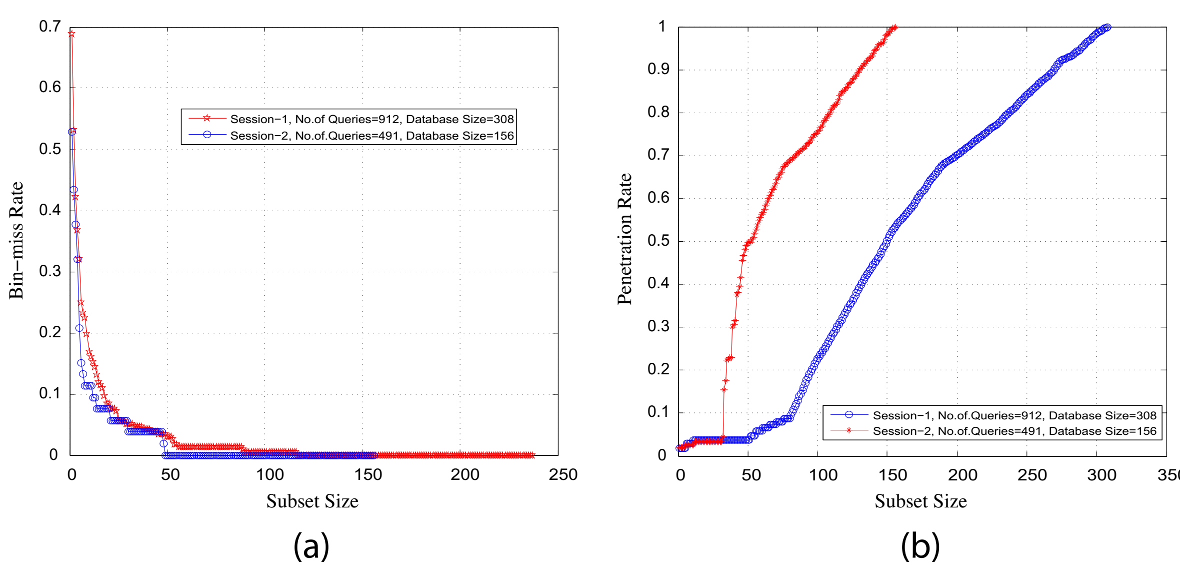

Fig. 11. UBIRIS database (a) relationship between *bin-miss rate* and subset size (b) relationship between *penetration rate* and subset size.

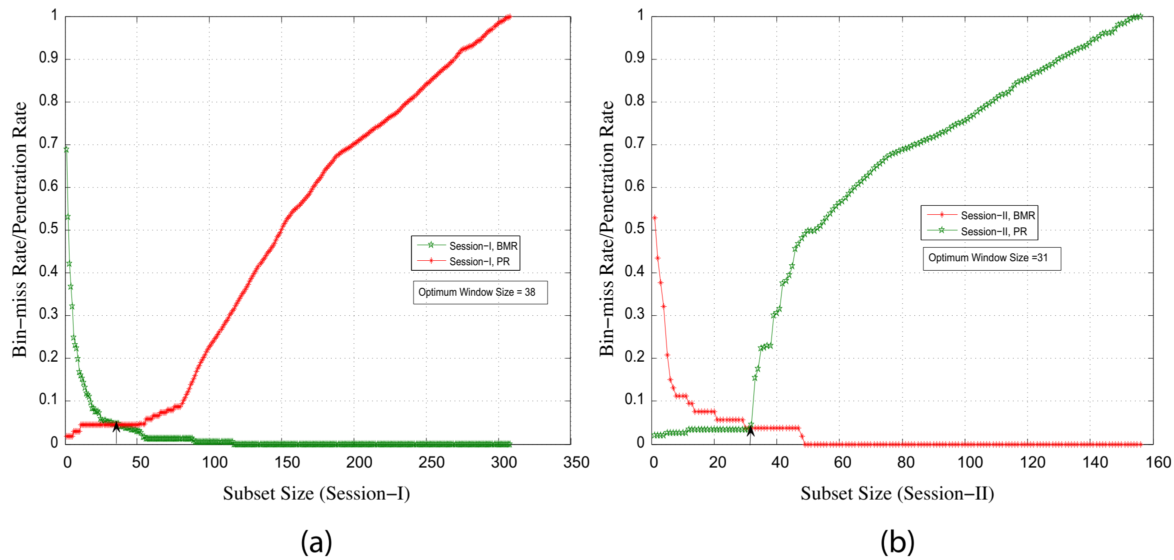

Fig. 12. UBIRIS database (a) optimum subset size for Session-I, (b) optimum subset size for session-II.

probability of 95.8% and 96.2% for Session-I and Session-II iris databases.

5.3. Time complexity

In the proposed indexing scheme a Kd-tree is constructed for N iris images, each having 2-tuple color indices. It uses linear median-finding algorithm to compute the median at each level to partition the space and it takes $O(N\log N)$ time complexity to construct the tree. And for a given iris query image, Kd-tree performs $k-NN$ search in order to produce subset size of k. It is clear that on an average $O(\log N)$ searches are necessary because any nearest neighbor search requires to traverse at least one leaf of the tree. During iris retrieval, one to one match between query and all the images in subset of size k is performed. Hence the total time taken to display the top best match is $O(\log N + k)$ where k is the subset size (*i.e.* the number of images compared during iris retrieval).

5.4. Robust to change in gaze, illumination and partial occlusion

Experiments have been performed to test the sensitivity of the proposed indexing scheme due to change in gaze, illumination and partial occlusion. Fig. 14 shows sample images from UBIRIS database with change in gaze, illumination and partial occlusion along with its corresponding segmented iris region considered for the experiment. The experiment shows that the proposed indexing scheme is capable of indexing a large iris database by eliminating most of the possible matches only with 2-tuple color indices. It has been observed that recognition accuracy is fairly insensitive to change in gaze, partial occlusion and illumination. Hence the proposed indexing scheme increases the robustness of iris recognition system. In general, use of color indices is the best choice for indexing as it is invariant to translation and rotation about the viewing axis and change only slowly due to change of angle of view. Fig. 15 shows the matching performance for UBIRIS database obtained for the optimum subset size of 38 and 31 for Session-I and

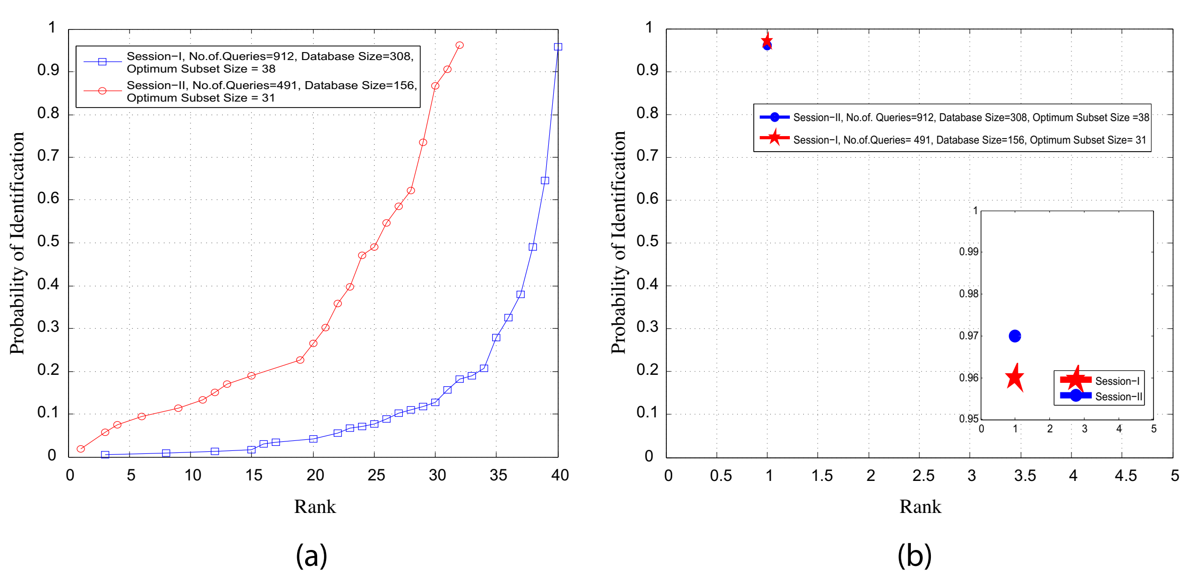

Fig. 13. CMC curve for optimum subset (BMR = PR), (a) indexed using color, (b) indexed using color followed by texture.

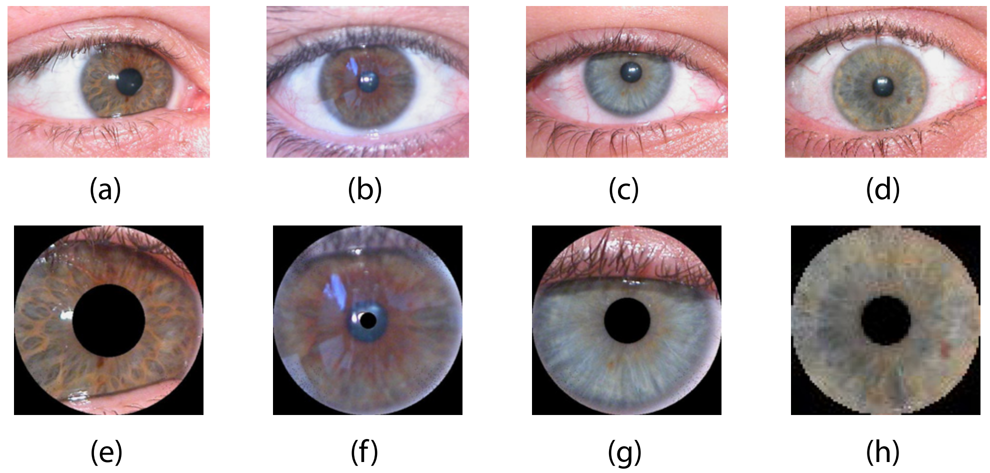

Fig. 14. UBIRIS iris images, (a) change in gaze, (b) illumination, (c) partial occlusion, (d) different scale (200 × 150), (a)–(c) 800 × 600, (e)–(h) its corresponding segmented region.

Session-II, respectively. First level in Fig. 15 considers the input query images while Second level shows its matched identities when only color information is used along with its position. Third level shows its matched identities when both color and texture are used. It can be observed that query's corresponding identities always lie in the first position when both color and texture are used.

5.5. Comparison with other methods

The proposed indexing scheme has been compared with the well known iris indexing schemes to analyse its performance. UBIRIS database consists of more noisy iris images which have been considered by many authors for experimental purpose. Hence, the proposed indexing scheme considers only this database for further comparison. From Table 1, it can be seen that some of the existing methods fall under clustering and classification techniques or that are not tested on publicly available databases and hence are not considered for comparison. Further, among three methods such as feature based (Euler code), group based color indices and geometric hashing methods the feature based method which is of 4-tuple Euler code, is generated for every iris image and is stored in the database. At the time of searching, the Euler codes of query and database are compared sequentially in order to filter out some images. The search time has been reduced because it has compared only with 4- tuples. But actually the method does not make use of any indexing scheme to reduce the search space (*i.e.*) it is still $O(N)$ time.

Hence the proposed indexing scheme has been compared with two indexing schemes such as group based color indices (Puhan & Sudha, 2008) and geometric hashing based indexing scheme (Mehrotra et al., 2010). For this, we have implemented geometric hashing based indexing scheme as described in Wolfson and Rigoutsos (1997). Fig. 16(a) and (b) shows the relationship between *bin-miss rate* and *penetration rate* for three indexing schemes for Session-I and Session-II, respectively. An efficient indexing system should have both *bin-miss rate* and *penetration rate* as low as possible. It can be observed from Fig. 16(a) and (b), the proposed indexing scheme has low *bin-miss rate* and *penetration rate* when compared to other two indexing schemes. This shows the effectiveness of the proposed indexing scheme.

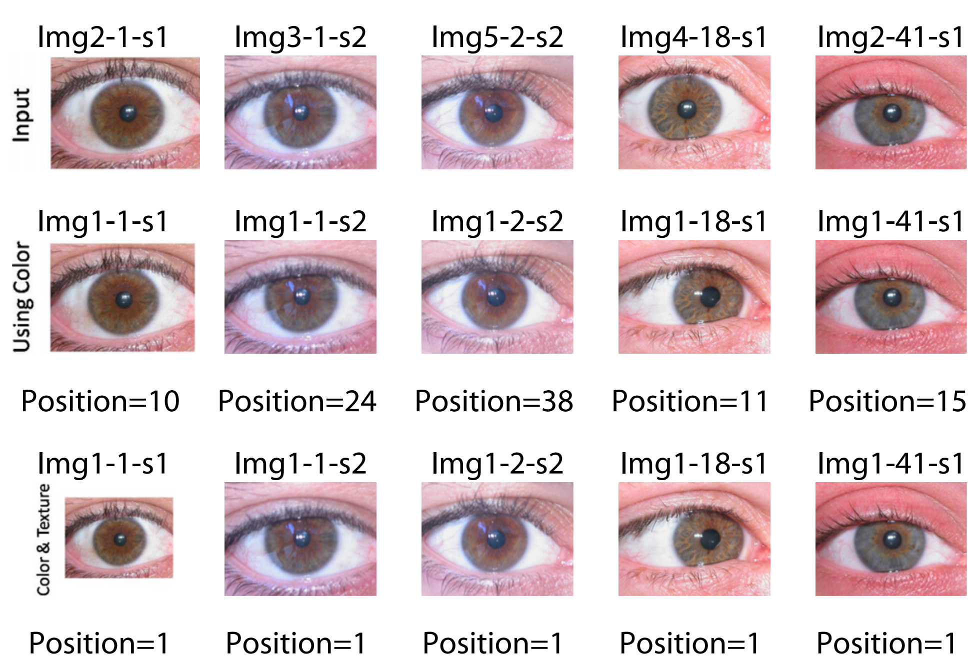

Fig. 15. Matched results obtained for the subset size 50, using only color, using color and texture.

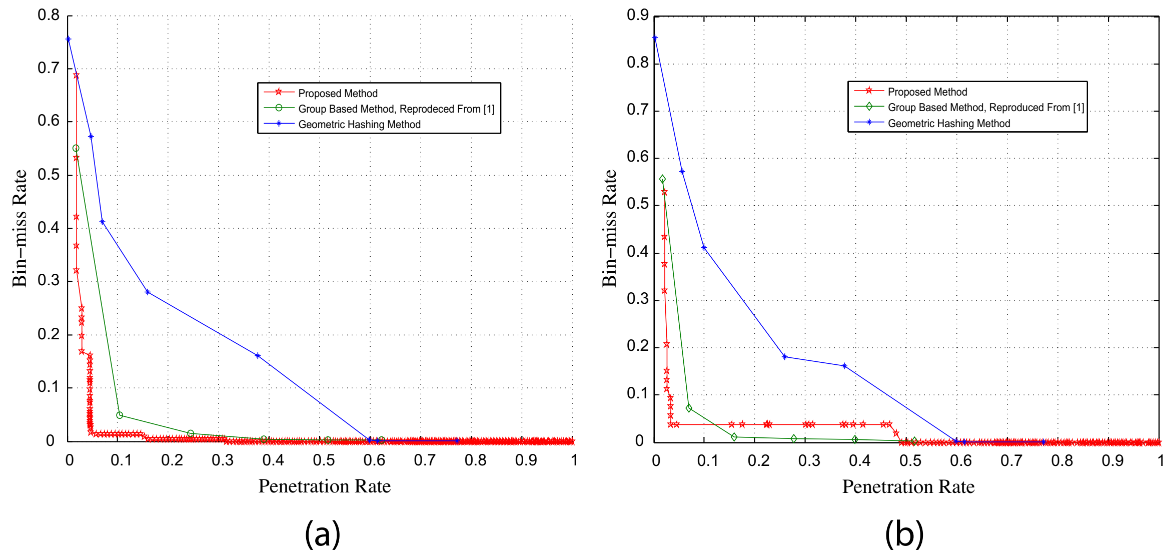

Fig. 16. Comparison graph for UBIRIS database, (a) Session-I, (b) Session-II.

Table 5
Comparison table for UBIRIS database.

Indexing schemes	Session-I		Session-II	
	Hitrate (%)	PR (%)	Hitrate (%)	PR (%)
Geometric hashing	97.67	39.21	98.01	36.81
Group based method reproduced from Puhan and Sudha (2008)	98.50	24.44	98.70	15.93
Proposed technique	98.70	8.29	98.71	7.14

Further, the proposed indexing scheme is achieved a *hitrate* of 95.8% for Session-I with equal *bin-miss rate* and *penetration rate* which is 4.2% for optimum subset size. Similarly, for Session-II a *hitrate* of 96.2% is achieved with equal *bin-miss rate* and *penetration rate* which is 3.8%. In order to compare the proposed indexing scheme with two other indexing schemes for a *hitrate* of 98.7%, *bin-miss rate* and *penetration rate* are calculated. It can be noted that *bin-miss rate* and *penetration rate* are found to be not equal for such a *hitrate*. Table 5 shows the *hitrate* and the *penetration rate* for all three indexing schemes. The proposed indexing scheme is found to be outperform among all the indexing schemes.

6. Conclusions

This paper has proposed a hierarchical approach to retrieve an iris image which is based on a new indexing scheme for a large iris database. It is a combination of both iris color and texture. Iris color is used for indexing and texture is used for retrieval of iris images from the indexed iris database. An index which is determined from the iris color is used to filter out the images that are not similar to the query image in color. Further, iris texture features of those filtered images are used to determine the images which are similar to the query image. It has been shown that during iris indexing, the value of color indices does not change even with the change in gaze, illumination, partial occlusion and scale which makes the proposed indexing scheme superior to other indexing schemes. Also SURF local descriptor which is used for iris retrieval to get the queries' identities in the top best match, enhances the performance of the iris recognition system. The proposed indexing scheme has been analysed against two indexing schemes such as group based color indexing scheme (Puhan & Sudha, 2008) and geometric hashing based indexing scheme (Mehrotra et al., 2010) and found the better performance with respect to *hitrate* and *penetration rate*.

References

Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer Vision and Image Understanding, 110*(3), 346–359.

Bengtsson, T., Bickel, P., & Li, B. (2008). Curse-of-dimensionality revisited: Collapse of the particle filter in very large scale systems. *IMS Collections, 2*(3034), 316–334.

Bentley, J. L. (1975). Multidimensional binary search trees used for associative searching. *Communications of the ACM, 18*(9), 509–517.

Boles, W., & Boashash, B. (1998). A human identification technique using images of the iris and wavelet transform. *IEEE Transactions on Signal Processing, 46*(4), 1185–1188.

Daugman, J. G. (1993). High confidence visual recognition of persons by a test of statistical independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 15*(11), 1148–1161.

Daugman, J. (2003). The importance of being random: Statistical principles of iris recognition. *Pattern Recognition, 2*(36), 279–291.

Daugman, J., & Downing, C. (2001). Epigenetic randomness, complexity and singularity of human iris patterns. *Proceedings of the Royal Society of London Series B-Biological Sciences, 268*(1477), 1737–1740.

Dobeš, M., Machala, L., Tichavský, P., & Pospíšil, J. (2004). Human eye iris recognition using the mutual information. *Optik, 115*(9), 399–405.

Fu, J., Caulfield, H. J., Yoo, S.-M., & Atluri, V. (2005). Use of artificial color filtering to improve iris recognition and searching. *Pattern Recognition Letter, 26*(14), 2244–2251.

Funt, B. V., & Finlayson, G. D. (1995). Color constant color indexing. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 17*(5), 522–529.

Ganorkar, S. R., & Ghatol, A. A. (2007). Iris pecognition: An emerging biometric technology. In *Proceedings of the 6th WSEAS international conference on signal processing, robotics and automation (ISPRA'07)* (pp. 91–96).

Gonzalez, R. W. R. C, & Eddins, S. (2009). *Digital Image Processing*, Pearson Education.

Huang, J., Kumar, R., Mitra, M., Zhu, W.-J., & Zabih, R. (1997). Image indexing using color correlograms. In *Proceedings of the Computer Vision and Pattern Recognition (CVPR'97)* (pp. 762–768).

Iris color, <http://www.eyecarecontacts.com/eyecolor.html>.

Keating, M. P. (1988). *Geometric, physical, and visual optics*. Butterworths.

Kimme, C., Ballard, D., & Sklansky, J. (1975). Finding circles by an array of accumulators. *Communications of the ACM, 18*(2), 120–122.

Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *Proceedings of the international conference on computer vision (ICCV'99)* (pp. 1150–1157).

Mehrotra, H., Majhi, B., & Gupta, P. (2010). Robust iris indexing scheme using geometric hashing of SIFT keypoints. *Journal of Network and Computer Applications, 33*(3), 300–313.

Mukherjee, R., & Ross, A. (2008). Indexing iris images. In *Proceedings of the international conference on pattern recognition (ICPR'08)* (pp. 1–4).
Proencca, H., & Alexandre, L. A. (2005). UBIRIS: A noisy iris image database. In *Proceedings of the international conference on image analysis and processing (ICIAP'05)* (pp. 970–977).
Puhan, N. B., & Sudha, N. (2008). A novel iris database indexing method using the iris color. In *Proceedings of the 3rd international conference on industrial electronics and applications (ICIEA'08)* (pp. 1886–1891).
Shapiro, L. G., & Stockman, G. C. (2003). *Computer vision*. Prentice Hall.
Stricker, M., & Orengo, M. (1995). Similarity of color images. In *Proceedings of the storage and retrieval for image and video databases (SPIE'95)* (pp. 381–392).
Swain, M. J., & Ballard, D. H. (1991). Color indexing. *International Journal of Computer Vision, 7*(1), 11–32.
Vatsa, M., Singh, R., & Noore, A. (2008). Improving iris recognition performance using segmentation, quality enhancement, match score fusion, and indexing. *IEEE Transactions on Systems, Man, and Cybernetics, Part B, 38*(4), 1021–1035.
Wolfson, H. J., & Rigoutsos, I. (1997). Geometric hashing: An overview. *IEEE Computational Science and Engineering, 4*(4), 10–21.
Yu, L., Zhang, D., Wang, K., & Yang, W. (2005). Coarse iris classification using box-counting to estimate fractal dimensions. *Pattern Recognition, 38*(11), 1791–1798.