Iris Recognition Using Discrete Cosine Transform and Relational Measures

Aditya Nigam^{1,2}^(⊠), Balender Kumar¹, Jyoti Triyar¹, and Phalguni Gupta^{1,3}

¹ Department of Computer Science and Engineering, Indian Institute of Technology Kanpur, Kanpur 208016, UP, India aditya@iitmandi.ac.in, {naditya,balendk,jyotit,pg}@cse.iitk.ac.in ² School of Computing and Electrical Engineering, Indian Institute of Technology Mandi, Mandi 175001, HP, India ³ National Institute of Technical Teacher's & Research, Salt Lake, Kolkata 700106, India

Abstract. Iris is one of the most discriminative biometric trait because it has random discriminating texture which does not change much, over a long time period. They are unique for all individuals, even for twins and the left and right eyes of the same individuals. In this paper an iris recognition system is presented that does iris segmentation, normalization, segregating of unwanted parts like occlusion, specular reflection and noise. Later iris images are enhanced and feature extraction and matching is performed. Iris features are extracted using Discrete Cosine Transform (DCT) and Relational Measure (RM). Later fusion of the dissimilarity scores of two feature extraction techniques has been proposed to get better performance. The results have been shown on large publicly available databases like CASIA-4.0 Interval, Lamp and self-collected IITK. The proposed fusion have achieved encouraging results.

Keywords: DCT \cdot Sobel operator \cdot Score-level fusion

1 Introduction

Personal authentication is a prime social requirement. Biometric based solutions are found to be much better than the traditional system of using passwords or identity card for authentication because they are easier to use and harder to circumvent. Most commonly used biometric traits in such systems are finger-print, facial features, iris, gait, hand-writing, retina, palm-prints, ear etc. The human iris is an annular part lying between pupil and sclera, and has good amount of irregular characteristics like freckles, furrows, ridges, stripes, etc. These characteristics are unique to each individual, even to different eyes of the same person. Also these textures of iris remain stable during lifetime of an individual. Moreover, it is an internal organ and externally visible so non-invasive acquisition can be done. Also they are much safely protected from damage as compared to fingerprint and palmprint.

[©] Springer International Publishing Switzerland 2015

G. Azzopardi and N. Petkov (Eds.): CAIP 2015, Part II, LNCS 9257, pp. 506–517, 2015.

DOI: 10.1007/978-3-319-23117-4_44

However, there are some challenges while using iris as biometric trait, like occlusion(hiding of data) due to eyelashes, eyelids and specular reflection and noise, which makes iris recognition inaccurate. In the proposed recognition system, first the image is acquired which is then segmented, normalized, denoised and enhanced. The features are extracted using Discrete Cosine Transform (DCT) and Relational Measures (RM). The matching scores of both the approaches are fused using weighted average. Figure 1 shows the flow-chart of the entire proposed iris recognition system. This paper is organized as follows. Section 2 gives an overview of some of the previous approaches used in iris recognition. Section 3 describes the proposed approach for the recognition system. Experimental results on standard databases are shown in Section 4. Conclusions are given in the last section.



Fig. 1. Overview of the Proposed Iris Recognition System

2 Literature Review

Iris is one of the most efficient and accurate biometric trait. It can provides high accuracy and less error rate than other biometric traits like fingerprint, palmprint, ear and face. Possibly, Flom and Safir [6] are the first one to propose the concept of iris recognition system. Daugman has used the concept of multi-scale Gabor filters to extract iris features [4], [3]. Despite of many advantages of iris, the system should also work over noisy iris images. In [2], [5], Noisy Iris Recognition Integrated Scheme(N-IRIS) has been proposed. Two local feature extraction techniques, i.e., Linear Binary Pattern (LBP) and Binary Large Objects(BLOBs) have been combined to design the scheme. Blobs are extracted from iris image using different LoG (Laplacian of Gaussian) filter banks which removes noise from the image as well as helps in better detection of blob. This feature extraction approach is invariant to rotation, translation as well as scale.

In [7], [12], the Gabor filter alongwith its response to the image has been discussed. Depending on different spatial frequencies and orientations, Gabor filter can be used effectively for extracting features. Also as the number of Gabor filters for extracting features increases, the more effective discriminative feature vector is extracted. In [9], a new approach of extracting iris features using local frequency variations between adjacent patches of enhanced normalized iris image has been proposed. Overlapping rectangular blocks with some orientation are considered as patches. In this, the patches are averaged widthwise to reduce the noise that gives 1-D signal on which window is applied to reduce spectral leakage and then Fast Fourier Transform (FFT) is applied to obtain spectral coefficients. The differences between the frequency magnitudes of adjacent patches are binarized using zero crossings. This approach gives better performance parameters than other existing state-of-the-art approaches of [3] and [8].

3 Proposed Technique

The iris segmentation is done using the technique proposed in [1]. Iris region is normalized to a fixed size strip in order to deal with iris dilations. One of the major hurdles in iris recognition is occlusion (hiding of iris) due to eyelids, eyelashes, specular reflection and shadows. It hides most of the useful iris texture and introduces irrelevant parts like eyelids and eyelashes which are not even an integral part of every iris image.

3.1 Occlusion Mask Creation

Occlusion detection is done in three steps: eyelid detection followed by eyelash and reflection detection.

[A] Eyelid Detection: Statistically, upper eyelid can be found at the center of left half while lower eyelid at the center of the right half of the normalized image. Eyelids have almost uniform texture and a boundary flooded with eyelashes and shadows. These challenges have motivated us to use region-growing approach to determine the eyelids. It uses only the texture information to separate eyelid region from the rest. Region growing is a morphological flooding operation which helps to find objects of uniform texture. A set S of seed points is selected from the image lying in the region S_d , that is required to be detected. All pixels in 8-neighborhood of any pixel in the set S. Pixels which are having this difference less than a certain threshold are added to S. This process is iterated until no pixel



Fig. 2. Application of Region-Growing

can be added further. Finally, S covers the desired region S_d . Figure 2 shows the results of the region-growing algorithm on an image after a few iterations.

Eyelid detection from the normalized iris strip of size $r \times c$ requires two seed points for region-growing, one for each lower and upper eyelid. They are selected as $(r, \frac{c}{4})$ and $(r, \frac{3c}{4})$ for upper and lower eyelid respectively as shown in Fig. 2. These two seed points are chosen because after normalization, upper and lower eyelids are centered mostly at $(\frac{\pi}{2})^{\circ}$ and $(\frac{3\pi}{2})^{\circ}$ angles w.r.t. x-axis. Regiongrowing begins with these seeds using a low threshold and expands the region until a dissimilar region is encountered. This gives the expected lower and upper eyelid regions. Detected eyelids are shown in Figure 3. If region grows beyond a limit, it indicates that there is no eyelid. Finally, a binary mask is generated in which all eyelid pixels are set to 1, as shown in Figure 3.



Fig. 3. Eyelid Regions: Arrows Denote the Direction for Region-Growing

[B] Eyelash Detection: There are two types of eyelashes: separable and multiple. Separable eyelashes are like thin threads whereas multiple eyelashes constitute a shadow like region. Eyelashes have high contrast with their surrounding pixels, but having low intensity. As a result, standard deviation of gray values within a small region around separable eyelashes is high. The standard deviation for every pixel in a normalized image is computed using its 8-neighborhood. It is high in areas where there are separable eyelashes. Multiple eyelashes may have high standard deviation, but they also have dark intensity value. Hence, the low gray value intensity is also given some weight. The computed standard deviation for each pixel is normalized using max - min normalization method and is saved in a 2D-array SD. If SD is used alone for segregating eyelash regions, then multiple eyelashes may not be detected and iris texture which has large standard deviation at some points gets wrongly classified as eyelashes. Hence for

each pixel, a fused value F(i, j) is computed which considers both the computed standard deviation as well as the gray value intensity of that pixel defined as :

$$F(i,j) = 0.5 \times SD(i,j) + 0.5 \times (1 - N(i,j))$$
(1)

where N(i, j) is the normalized gray intensity values (0 - 1) and SD(i, j) is the standard deviation computed using 8 neighborhood pixel intensities for the pixel (i, j). This fused value F(i, j) boosts up the gap between eyelash and non-eyelash part. The image histogram F_H of F has two distinct clusters: a cluster of low values of F consisting of the iris pixels and the second cluster with high values of F representing eyelash pixels. To identify the two clusters, Otsu thresholding is applied on the histogram F_H of F to obtain binary eyelash mask. It determines two clusters in a histogram by considering all possible pairs of clusters and chooses that clustering threshold that minimizes the intra-cluster variance. It thus separates the eyelash portion from the iris portion. The detected eyelash of an iris image is shown in Figure 4(c).

[C] Reflection Detection: Pixels which exceed a threshold value in grayscale image are declared as reflections because reflections are very bright in every acquisition setting. Also, since occlusion due to reflection is not a major component, it is chosen not to do complex computation to remove reflection. Detected reflection from a sample image is shown in Figure 4(d). A binary mask $Mask_{reflection}$ (reflection mask) is generated in which pixels affected by reflection are set to 1. Final occlusion mask is generated by addition (logical OR) of the binary masks of eyelid, eyelash and reflection. Detected occlusion of a sample image is shown in Figure 4.



Fig. 4. Overall Occlusion Mask

3.2 Iris Enhancement

The iris texture is enhanced in such a way that it increases its richness as well as its discriminative power. The iris ROI is divided into blocks and the mean



Fig. 5. Iris Texture Enhancement

of each block is considered as the coarse illumination of that block. This mean is expanded to the original block size as shown in Fig. 5(b). Selection of block size plays an important role. It should be such that the mean of the block truly represents the illumination effect of the block. So, larger block may produce improper estimate. We have seen that a block size of 8×8 is the best choice for our experiment. The estimated illumination of each block is subtracted from the corresponding block of the original image to obtain the uniformly illuminated *ROI* as shown in Fig. 5(c). The contrast of the resultant *ROI* is enhanced using Contrast Limited Adaptive Histogram Equalization (*CLAHE*). It removes the artificially induced blocking effect using bilinear interpolation and enhances the contrast of image without introducing much external noise. Finally, Wiener filter is applied to reduce constant power additive noise and the enhanced iris texture is obtained shown in Figure 5(d).

3.3 Feature Extraction

The main aim of iris recognition system is to minimize intra-class differences and to maximize the inter-class differences. In this paper, two different types of feature extraction techniques *viz*. DCT and RM are discussed. The matching scores of both the techniques are fused by taking weighted average of their scores.

[A] Feature Extraction Using DCT: This paper proposes the feature extraction using DCT with some parameters optimized for best performance. A nonconventional technique of applying 1-D DCT on overlapping blocks of particular size for extracting feature variations has been proposed. It is observed that DCT coefficients are robust when applied over values lying between -128 to 127, so 128 is subtracted from each pixel value of the enhanced normalized image. This results in a matrix levelled off by 128 from each pixel entry such that all entries in the matrix lie between -128 to 127.

[A.1] Segmentation Into Rectangular Blocks: The levelled-off matrix is divided into rectangular blocks of size $M \times N$ with overlapping of half the width between vertically adjacent blocks and overlapping of half the length between

horizontally adjacent blocks as shown in Figure 6(a). These rectangular blocks form the basis of extracting features in our proposed approach. Parameters like length and width of rectangular blocks are tuned to achieve optimum performance.

[A.2] Coding of Rectangular Blocks: The rectangular block is first averaged across its width. This gives a one-dimensional intensity signal of size $1 \times N$. Formally, a rectangular block of width M and length N, averaged across width gives a 1-D intensity signal R' of size $1 \times N$ which can be represented by

$$R'_{j} = \sum_{k=1}^{M} R_{j,k} \quad where \ j = 1, 2, ..., N$$
(2)

Averaging smoothens the image and reduces the effect of noise and other image artifacts. The obtained intensity signal R' is windowed using Hanning window of size N to reduce spectral leakage during the transformation. Application of averaging and windowing also results in reduction of resolution of the image in the horizontal direction. Also image registration becomes easier for broad patches, thereby making iris recognition rotation invariant [10]. The generated 1-D DCT coefficient matrix CM of each rectangular block is binarized using zero crossing to give a binary sub-feature vector B as given below :

$$B_j = \begin{cases} 1 & \text{if } CM_j > 0 & \text{where } j=1, 2, ..., N \\ 0 & \text{otherwise} \end{cases}$$
(3)

A second level occlusion mask based on the feature vector calculation is generated as follows: The corresponding occlusion mask *occmask* is also divided into blocks in the same way as the levelled-off matrix. The summation across its width gives a block or patch of size $1 \times N$. If the summation along its width is more than 80% of the width of the block (M), then the bit is masked (set to 1) in the second level occlusion mask; otherwise it is left unmasked. The next bit is then added to the next row and so on. It gives a block of size $N \times 1$. This is done for each overlapping $M \times N$ block in the occlusion mask. Second level mask is required as the feature vector is block-based and not pixel-based. Figure 6(b) illustrates the steps involved in generating feature vector of enhanced normalized iris image. The steps to calculate the feature vector and second level occlusion mask is summarized in Algorithm 1.

[B] Feature Extraction Using Relational Measures. The relational measures approach has been used as second feature extraction technique. Relational Measures are features which are based on relational operators like \langle , \rangle and =. Unlike giving the exact difference between any two quantities, the concept of relational measures is based on finding the relative difference between the two. This encoding into bits is fast and also takes less memory. Also the iris texture has lot of variations in texture; so relational measures concept can be used to encode iris. Vertically and horizontally overlapping regions are chosen from the enhanced normalized image. A central region of size $b \times b$ is chosen. Its four



(a) Overlapping Rectangular (b) Feature vector generation from enhanced nor-Blocks and parameters malized iris image (width M=4, length N=6)

Fig. 6. Feature Extraction using DCT

neighboring regions of same size is taken but at a particular distance d, where d is large as compared to b. A symmetric 2-D Gaussian filter centrally clipped to size $b \times b$ is put and convoluted over each of these five regions. The response of central region is greater than its neighbor, then the bit is encoded as 1, otherwise it is set to 0. In this way, four bits of code are obtained for each central region named as RM bits. This is then iterated for other vertical and horizontal overlapping regions over the entire image. All such RM bits concatenated together gives the 2-D binary template. Second level mask is also generated from the raw occlusion mask of iris based on the feature vector calculation. If the central block has more than 80% of the occluded pixels, then the RM bits for that block are encoded as [1 1 1 1], i.e., masked; otherwise it is encoded as [0 0 0 0], i.e., unmasked. This is repeated for overlapping central blocks according to parameters chosen. This second-level mask is required because the feature vector is block-based and not pixel-based.

3.4 Feature Matching

The feature vector templates and corresponding second level occlusion masks are used in matching. Matching between two iris images using their respective feature vector templates and second level occlusion masks can be done by computing a dissimilarity score between them. The dissimilarity score is calculated using hamming distance metric. Consider two templates t_1 and t_2 of same size, say

Algorithm 1 FeatExtract(I, occmask)

- **Require:** Enhanced Normalized image I of size $Rows \times Cols$, Occlusion Mask occmask of same size as I, M is the width of the rectangular block, N is the length of the rectangular block
- Ensure: Feature vector template Feat and Second level occlusion mask Mask
- 1: $I_{new} \leftarrow$ Level off image by subtracting 128 from each pixel value
- 2: Initialize a variable HannFil to a Hanning window of size N
- $3: \ Feat \leftarrow AllocateZero$
- $4:\ Mask \leftarrow AllocateZero$
- 5: Divide I_{new} and occmask into rectangular blocks $B_{i,j}$ and $occBlk_{i,j}$ of size $M \times N$ with overlapping of M/2 between vertically adjacent blocks and overlapping of N/2between horizontally adjacent blocks
- 6: for each rectangular block $B_{i,j}$ and $occBlk_{i,j}$ do
- 7: $MeanBlk_i \leftarrow Compute the mean of B_{i,j} across width$
- 8: $HannBlk_i \leftarrow ElementWiseMultiplication (HannFil, Transpose of MeanBlk_i)$
- 9: $dctBlk_i \leftarrow DCT (HannBlk)$ //Extract the 1-D DCT coefficients
- 10: Binarize the 1-D DCT coefficients using zero-crossing to give sub-feature vector f_i
- 11: //Calculating Second level occlusion mask
- 12: $sumarr_i \leftarrow Compute the sum of occBlk_{i,j} across width$
- 13: for each sum in $sumarr_i$ do
- 14: **if** sum > 0.8 * M **then**
- 15: $resMask \leftarrow 1$ //bit masked
- 16: else
- 17: $resMask \leftarrow 0$ //bit unmasked
- 18: end if
- 19: Concatenate the bits resMask vertically to give a sub-mask $maskBlk_i$
- 20: end for
- 21: Concatenate the sub-features $\{f_i\}$ to give the final feature vector Feat
- 22: Concatenate the sub-masks $\{maskBlk_i\}$ to give the final second-level occlusion mask Mask
- 23: end for
- 24: Return (*Feat*, *Mask*)

 $X \times Y$ and their second level occlusion masks o_1 and o_2 of same size as that of their feature vector templates, then Hamming distance hd between the templates is calculated using the formula

$$hd(t_1, t_2, o_1, o_2) = \frac{\sum_{i=1}^X \sum_{j=1}^Y [t_1(i, j) \oplus t_2(i, j)] | [o_1(i, j) + o_2(i, j)]}{X \times Y - \sum_{i=1}^X \sum_{j=1}^Y [o_1(i, j) + o_2(i, j)]}$$
(4)

where the operators \oplus , | and + represent binary XOR, NAND and OR operations respectively. Second level occlusion masks are considered while calculating the dissimilarity score between the two iris images. It enables us to perform matching only in valid bits and ignore the occluded parts of iris image. The value of hd is zero if both feature templates are similar, i.e., have all bits of same value. Hence for genuine matching hd should be low. [3.4.1] Robustness Against Rotation: While acquiring image, there can be some amount of rotation in the image. Rotation of the eye in Cartesian coordinate-space corresponds to horizontal translation in the normalized image. When a probe template is matched with a gallery template, the gallery template is circularly shifted in horizontal direction to get the minimum hamming distance which is taken as the final dissimilarity score. When gallery template is rotated, its corresponding second level mask is also rotated.

4 Experimental Results

The proposed iris recognition system has been tested on two publicly available CASIA-4.0 Interval and CASIA-4.0 Lamp databases and also over our own IITK database. The iris database is divided into two sets - gallery set and probe set. All images of the probe set are matched against the images of the gallery set. System performance is tested in terms of CRR and EER [11].

[A] Databases: CASIA-4.0 Interval consists of 2639 images of 249 subjects of size 320×280 pixels taken in two sessions. First three images are taken in the gallery set and rest in the probe set. So total there are 1047 gallery images and 1508 probe images in this database. CASIA-4.0 Lamp Database consists of 16,212 images of 411 people of size 640×480 pixels collected in one session with variable illumination conditions with lamp being switched on/off. Each subject has 20 images. First 10 images per subject have been taken in the gallery set and rest 10 images in the probe set. So total there are 7830 images in both the gallery and the probe sets in this database. The IITK Database consists of 20,420 images of 1021 subjects of size 640×480 pixels collected in two sessions. In each session, 10 images per subject have been collected, with 5 images for each eye. Images in first session are taken in the gallery set and images of second session are considered in the probe set. So finally there are 10,210 images in both the gallery and the probe sets in this database.

[B] Performance Analysis of the Proposed System: The fusion of matching scores of DCT approach and RM approach is done on the basis of weights determined empirically which gives the best system performance. All parametric evaluation is done over a small validation set consisting of only first 1000 images of that dataset optimized w.r.t performance. Higher weight is given to matching scores of DCT approach as compared RM because the DCT approach performance is better than RM. Table 1 shows the individual performance parameters of the both approaches as well as fusion performance parameters.

In both DCT and RM approaches, some matchings have been discarded in which the individual mask or the combined mask is more than 85% of the image size. This has been done to avoid inaccuracies caused due to heavy occlusion. So in all, around 1.5% of the overall matchings are discarded. The Rank - 10 accuracy of the proposed system over all databases got saturated to 100%.

[C] Comparative Performance Analysis: The proposed approach has been compared with that of Daugman's recognition system [3]. All the pre-processing

Database	Proposed DCT		RM		Weightage		Fused	
	$\operatorname{CRR}(\%)$	EER(%)	CRR(%)	EER(%)	DCT	RM	$\operatorname{CRR}(\%)$	EER(%)
Interval	99.40	1.81	99.07	2.26	0.75	0.25	99.40	1.52
Lamp	98.69	3.89	98.69	4.21	0.72	0.28	98.91	2.91
IITK	98.46	2.07	98.66	2.12	0.60	0.40	98.92	1.52

Table 1. Fused Result with weightage given to matching scores of DCT and RM

Table 2. Comparison of Results on various Databases with different approaches

Databasa	CRR(%)				EER(%)			
Database	Gabor	RM	DCT	Fused	Gabor	RM	DCT	Fused
Interval	99.47	99.07	99.40	99.40	1.88	2.26	1.81	1.52
Lamp	98.90	98.69	98.69	98.91	5.59	4.21	3.89	2.91
IITK	98.85	98.66	98.46	98.92	2.49	2.12	2.07	1.52

stages including segmentation, normalization and occlusion masking have been kept common. They differ only in their feature extraction phase. The matching scores of both DCT and RM approaches have been fused using weighted average to get better performance results. All these approaches have been tested on all three databases. Table 2 shows the performance metrics of the four approaches (Gabor, DCT, RM and Fusion approaches). The ROC graphs of the system on all three databases comparing the four approaches are shown in Figure 7(a), Figure 7(b) and Figure 7(c) respectively. From these figures, it can be seen that DCT approach performers better than Gabor-filtering and RM approaches. The fusion approach of DCT and RM has the best performance because weak classifier fusion works better than individuals.



Fig. 7. ROC Graph based Performance Comparison of all the four approaches

5 Conclusions

This paper presents an iris recognition system which has been tested on three databases to claim its superior performance. It has presented the segmentation, normalization, occlusion mask detection, denoising and enhancement as preprocessing steps. A non-conventional technique based on 1-D DCT has been used to extract robust iris features. Another feature extraction technique of Relational Measures (RM) is used that is based on calculating intensity relationships between local regions and encoding them on the basis of relative difference of intensities. Matching of images is done by using Hamming distance metric which gives a dissimilarity score. Score-level fusion technique is used to compensate for some images which have been rejected by one classifier while accepted by other. Such a fusion has shown much improved accuracy with less error rates.

References

- Bendale, A., Nigam, A., Prakash, S., Gupta, P.: Iris segmentation using improved hough transform. In: Huang, D.-S., Gupta, P., Zhang, X., Premaratne, P. (eds.) ICIC 2012. CCIS, vol. 304, pp. 408–415. Springer, Heidelberg (2012)
- Chenhong, L., Zhaoyang, L.: Efficient iris recognition by computing discriminable textons, vol. 2, pp. 1164–1167 (2005)
- Daugman, J.: High confidence visual recognition of persons by a test of statistical independence. IEEE Transactions on Pattern Analysis and Machine Intelligence 15(11), 1148–1161 (1993)
- Daugman, J.: Statistical richness of visual phase information: update on recognizing persons by iris patterns. International Journal of Computer Vision 45(1), 25–38 (2001)
- De Marsico, M., Nappi, M., Riccio, D.: Noisy iris recognition integrated scheme. Pattern Recogn. Lett. 33(8), 1006–1011 (2012)
- 6. Flom, L., Safir, A.: Iris recognition system, February 3, 1987. US Patent 4,641,349
- Grigorescu, S.E., Petkov, N., Kruizinga, P.: Comparison of texture features based on gabor filters. IEEE Transactions on Image Processing 11(10), 1160–1167 (2002)
- Ma, L., Tan, T., Wang, Y., Zhang, D.: Efficient iris recognition by characterizing key local variations. IEEE Transactions on Image Processing 13(6), 739–750 (2004)
- Monro, D., Zhang, Z.: An effective human iris code with low complexity, vol. 3, p. III-277 (2005)
- Monro, D., Zhang, Z.: An effective human iris code with low complexity. In: IEEE International Conference on Image Processing, ICIP 2005, vol. 3, p. III-277. IEEE (2005)
- Nigam, A., Gupta, P.: Iris recognition using consistent corner optical flow. In: Lee, K.M., Matsushita, Y., Rehg, J.M., Hu, Z. (eds.) ACCV 2012, Part I. LNCS, vol. 7724, pp. 358–369. Springer, Heidelberg (2013)
- 12. Prasad, V.S.N., Domke, J.: Gabor filter visualization (2005)