

Information Fusion in Low-Resolution Iris Videos using Principal Components Transform*

Raghavender Jillela, Arun Ross
West Virginia University

{Raghavender.Jillela, Arun.Ross}@mail.wvu.edu

Patrick J. Flynn
University of Notre Dame

flynn@nd.edu

Abstract

The focus of this work is on improving the recognition performance of low-resolution iris video frames acquired under varying illumination. To facilitate this, an image-level fusion scheme with modest computational requirements is proposed. The proposed algorithm uses the evidence of multiple image frames of the same iris to extract discriminatory information via the Principal Components Transform (PCT). Experimental results on a subset of the MBGC NIR iris database demonstrate the utility of this scheme to achieve improved recognition accuracy when low-resolution probe images are compared against high-resolution gallery images.

Index Terms - Iris recognition, Low-resolution, Image-level fusion, Image averaging, Principal Components Transform (PCT).

1. Introduction

The iris is a thin membrane-like structure of the eye that regulates the size of the pupil and controls the amount of light impinging the retina. The iris exhibits a rich textural pattern due to its multi-layered structure consisting of fibrous tissues and blood vessels, and the presence of anatomical entities such as crypts, furrows, freckles and pigmentation. The texture of the iris is believed to be unique to each eye, and is thought to be stable throughout the lifespan of an individual [5]. Thus, human recognition using iris texture is considered to be highly reliable under ideal imaging con-

ditions. Large-scale experiments have confirmed this notion [2].

The key steps in a typical iris recognition system are (a) image acquisition, (b) iris segmentation, (c) iris normalization, (d) feature extraction, and (e) matching. In the image acquisition step, an image of the eye is captured using a sensor of adequate resolution. Depending on the field of view of the sensor, these images (or *frames* in the case of videos) can potentially include other regions of the eye such as the sclera, eyelashes, eyelid, or even the eyebrow. This additional information is excluded by the iris segmentation routine which isolates the iris from the rest of the image. During the normalization stage, the pixels within the nearly annular-shaped iris are geometrically re-mapped into a rectangular region. Textural features are then extracted from the normalized iris and encoded into a binary template. Finally, the match scores obtained by performing template matching are used for recognition.

Under ideal constrained acquisition conditions, iris recognition results in very good performance as has been borne out by several experiments [2]. However, such a performance can be guaranteed only when cooperative subjects are involved. Most iris recognition systems require the subjects to maintain a fixed gaze at a specified location for a short period of time, in order to acquire a high-quality image. In the case of moving subjects, the images may be of poor quality due to improper illumination, off-angle gaze, motion blur, occlusions, specular reflection, etc. Matey et al [9] discuss a system to successfully acquire iris images in less constrained environments. Such a system may require an elaborate setup, and can result in off-angle iris images if the facial profile of a subject is non-frontal.

The performance of an iris recognition system is negatively impacted when the spatial resolution of the acquired iris images is low. The resolution of the iris image depends on (a) the resolution of the sensor itself, and (b) the stand-off distance (distance of the subject from the sensor). While the effect of the first factor can be mitigated by using a high resolution sensor, it is difficult to handle the stand-off distance problem. An increase in the stand-off distance causes a de-

*Portions of this work were originally funded by the Center for Identification Technology Research (CITeR). Currently this work is sponsored under IARPA BAA 09-02 through the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-10-2-0013. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing official policies, either expressed or implied, of IARPA, the Army Research Laboratory, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

crease in the size (or pixel resolution) of the eye recorded in the image (at a fixed zoom-factor). Such a reduction in size can lower the textural quality of the iris in the image, which can in turn affect the performance of the recognition system. Figure 1 illustrates this effect.

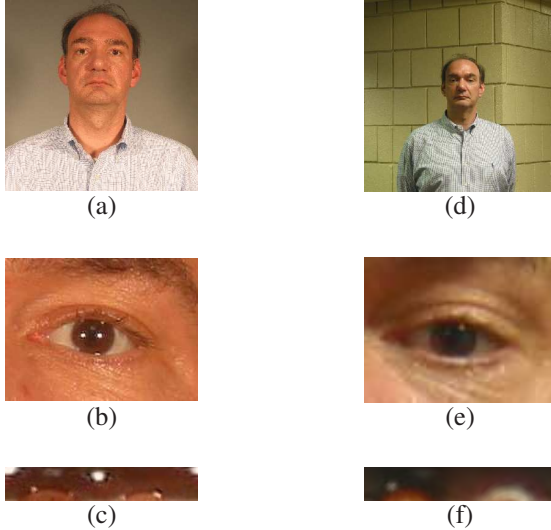


Figure 1. Panels (a), (b), and (c), respectively, show the face, the eye, and the normalized iris regions (in the visible spectrum) of a subject standing close to the sensor. The corresponding regions for a subject with larger stand-off distance are shown in panels (d), (e), and (f), respectively.

When the input images are of poor quality, fusion methods can be used to enhance the recognition performance. Biometric fusion refers to the process of aggregating the information needed for reliable recognition from multiple sources of evidence [10]. Depending on the type of information available, fusion can be performed at various levels in a biometric system (e.g., image-level, feature-level, score-level, etc.). In this work, an image-level fusion scheme is presented that uses the information contained in the multiple frames of an iris video. The use of multi-frame iris fusion has several benefits:

1. In many image-level fusion techniques, registration (or alignment) of the input images into a single coordinate system of reference is very important. Registration of images obtained at different time instances or from different sensors, is a challenging task. However, the frames extracted from a given iris video are likely to be aligned. As a result, the errors caused by improper image registration can be greatly reduced.
2. The frames in an iris video contain information related to the spatio-temporal activity of the iris and its surrounding region over a short period of time. As this information is continuous, good quality frames can be

selectively chosen for fusion while avoiding poor quality frames.

In this work, image-level fusion is performed in two stages: (i) by first applying Principal Components Transform (PCT) to the individual frames, and (ii) then averaging the resulting images. The performance of image-level fusion is compared against that of score-level fusion. Experimental results, in both cases, indicate that the fused outputs provide better recognition performance than their corresponding low-resolution source images.

2. Image-level Fusion

Super-resolution techniques may be used to perform image-level fusion. Super-resolution is the process of generating an image with a higher resolution than the corresponding source images. The information from individual frames can be fused into a single composite image with higher resolution, resulting in better recognition performance. Although much work has been done in the face recognition domain, super-resolution of iris images has not been widely discussed in the literature. This is due to the stochastic nature of the iris texture which does not lend itself to traditional super-resolution schemes. Fahmy [4] describes an interleaving process to generate a high resolution iris image from a low resolution face video. Iris regions of equal sizes are segmented from the low-resolution frames of a face video. These iris regions are registered using a cross correlation model, and interleaved with each other to form an image of higher resolution. This process is iterated multiple times to generate a high resolution iris image.

Huang et al. [7] propose a learning based algorithm to improve the resolution of normalized iris images. The algorithm is trained using a large number of image pairs consisting of low-resolution normalized iris images and their corresponding high-resolution versions. In the training stage, each low-resolution normalized iris image is tessellated into multiple blocks, and the relation of each block with its corresponding high-resolution pair is modeled using Markov networks. In the testing stage, a high-resolution output is generated by upsampling the input low-resolution image and restoring the lost frequency information based on the best matching training blocks from the database.

While the input in the above mentioned approaches is a static set of individual images, Hollingsworth et al. [6] use a set of frames extracted from an iris video. A set of 10 good quality frames are chosen automatically and the iris is segmented and normalized in each of them. These normalized irides are then fused on a pixel-by-pixel basis, by using an operator (e.g., *mean*, *median*, etc.). Consider a set of n images $\{I_1, I_2, \dots, I_n\}$ each of size $M \times N$. The intensity of the final fused image I_{fused} at a location (i, j)

can be obtained via the mean operator as

$$I_{fused}(i, j) = \frac{1}{n} \sum_{p=1}^n I_p(i, j), \quad (1)$$

where $1 \leq i \leq M$ and $1 \leq j \leq N$. This technique can be viewed as a *pixel-level* fusion scheme, where the pixel intensity at a given location of the output is dependent only on the corresponding pixel intensities of the input images. Although the technique is simple, the recognition performance of the resulting output is greatly improved [6]. However, such an output strongly depends on the following factors:

1. Number of observations: The output is typically more reliable if the number of the input samples, n , is large. If n is small, the output can be a weak estimate.
2. Accuracy of observations: If a majority of the input images contain noise, the quality of the output image cannot be expected to improve over the input images. For example, if a large number of input images are blurred at a specific region, it cannot be rectified in the resulting output.

Furthermore, the input images should be perfectly registered. If the registration is inaccurate, the output would be an approximate or a smoothed representation of the actual scene. Perfect registration of iris images obtained in non-ideal environments is a challenging problem. In iris images, imperfect registration can perturb the texture of the output and reduce the matching performance of the system. Thus, it has to be ensured that the fusion scheme does not alter the textural richness of the iris.

The proposed approach performs image-level fusion in two stages. In the first stage, an image reconstruction scheme based on the Principal Components Transform is used to re-estimate the input images. In the second stage, the reconstructed input images are fused by an image averaging scheme using the *mean* operator. The recognition performance obtained using the output generated by the two-stage approach is observed to be higher than the performances obtained after applying either of the stages.

3. Principal Components Transform

Principal Components Analysis (PCA) has been widely used in the field of automatic face recognition. Turk and Pentland [11] view each face image as a point in a high-dimensional sub-space whose coordinates correspond to the pixel intensities. Based on a large set of registered training face images, PCA is used to determine a set of orthogonal basis vectors, referred to as *eigenfaces* or *principal components*, that correspond to directions of maximum variance in the original sub-space. Subsequently, any given face image can be represented as a weighted sum of such eigenfaces.

In the field of iris recognition, Dorairaj et. al. [3] use PCA to determine a set of basis vectors for iris images. However, in the absence of a common morphology in normalized iris (unlike face which has some common landmarks across images), the basis vectors do not have a trivial physical interpretation.

In this work, PCA is used in a different manner than what has been typically used in the biometrics literature. To avoid confusion, the PCA technique used in [11] (and other publications) is referred to as the *conventional PCA*, while the technique used in this work is referred to as *Principal Components Transform (PCT)*. The major differences between the conventional PCA, and the PCT approach used in this work are listed below:

1. Given a set of n images, each having a spatial resolution of $M \times N$, conventional PCA represents every image as a point in the MN dimensional space. In this work, an n dimensional space is considered in which each *pixel intensity vector* is a point. The pixel intensity vector, \vec{V}_{pq} , that contains pixel intensity values across all given images at a location (p, q) , is defined as:

$$\vec{V}_{pq} = [I_1(p, q), I_2(p, q), \dots, I_n(p, q)]^T, \quad (2)$$

where $I_j(p, q)$ denotes the pixel intensity value of an image I_j at a location (p, q) , and $j = \{1, \dots, n\}$, $p = \{1, \dots, M\}$, $q = \{1, \dots, N\}$. This variation in representation can be easily obtained by considering a different scheme for stacking (or arranging) the pixel values of the images, as shown in Figure 2(a).

2. Conventional PCA is typically applied on multiple images of different subjects, to highlight the variance information among the images. On the other hand, PCT is applied on multiple images of the same subject, to highlight the variance information among the pixel intensity vectors. Consequently, PCT seeks to extract *discriminatory* pixels from the iris frames.

Consider a set of n images $\{I_1, I_2, \dots, I_n\}$, each having a spatial resolution of $M \times N$ pixels. Every image I_j , is transformed to a row vector, \vec{I}_j , of size $1 \times MN$ where $j = \{1, 2, \dots, n\}$. An image data matrix \mathbf{X} is obtained by stacking¹ the n row vectors, one per row, as shown below:

$$\mathbf{X} = \begin{bmatrix} \vec{I}_1 \\ \vec{I}_2 \\ \vdots \\ \vec{I}_n \end{bmatrix}. \quad (3)$$

¹The major difference between the conventional PCA and the PCT approach lies in the stacking process used to generate the image data matrix. Turk and Pentland [11] stack the images as column vectors into \mathbf{X} , whereas PCT considers the images as row vectors.

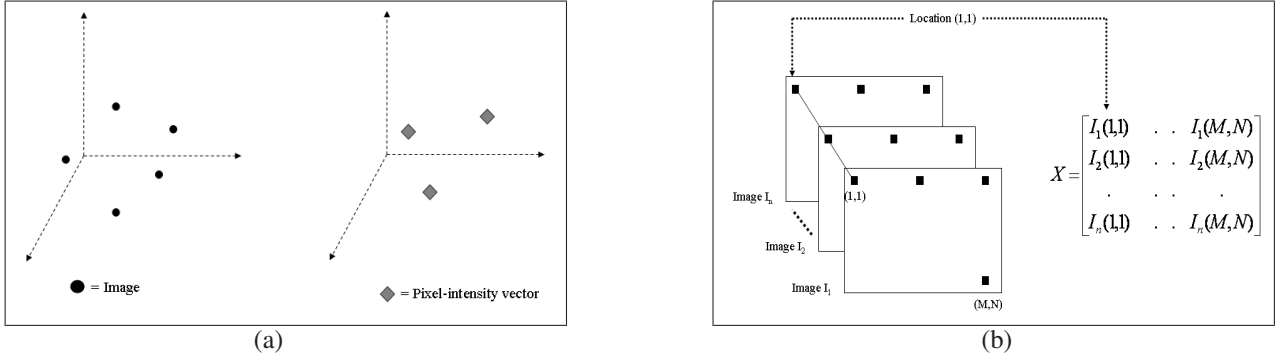


Figure 2. (a) Left: In an MN dimensional space, the conventional PCA considers each image as a point (denoted by a circle). Right: In an n dimensional space, the PCT approach considers every pixel intensity vector as a point (denoted by a diamond). (b) The process of stacking image pixels used to generate image data matrix \mathbf{X} for the PCT approach.

The size of \mathbf{X} will be $n \times MN$. In other words, every column of the image data matrix \mathbf{X} is a *pixel intensity vector* at a particular location. For this data, the empirical mean vector $\vec{m}_{\mathbf{X}}$ is computed along each dimension. The resulting row vector of size $1 \times MN$ is given by the following equation:

$$\vec{m}_{\mathbf{X}} = \frac{\sum_{p=1}^n \vec{I}_p}{n}. \quad (4)$$

The covariance matrix $\mathbf{C}_{\mathbf{X}}$ for the image data can be computed by the equation:

$$\mathbf{C}_{\mathbf{X}} = \frac{1}{n}(\mathbf{X} - I\vec{m}_{\mathbf{X}})(\mathbf{X} - I\vec{m}_{\mathbf{X}})^T \quad (5)$$

where I is an identity matrix of size $n \times 1$. The size of the covariance matrix $\mathbf{C}_{\mathbf{X}}$ will be $n \times n$.

The eigenvectors of the covariance matrix $\mathbf{C}_{\mathbf{X}}$ are obtained by decomposing it into its canonical form. Using this information, the input data can be transformed into a new feature space by the equation:

$$\mathbf{Y} = \mathbf{A}(\mathbf{X} - I\vec{m}_{\mathbf{X}}), \quad (6)$$

where \mathbf{A} is an $n \times n$ matrix, whose rows contain the normalized eigenvectors of $\mathbf{C}_{\mathbf{X}}$. At any point of time, the original data \mathbf{X} can be recovered by performing a simple inverse transformation given by:

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{Y} + I\vec{m}_{\mathbf{X}}. \quad (7)$$

As the rows of \mathbf{A} are ortho-normal vectors, $\mathbf{A}^{-1} = \mathbf{A}^T$. Hence the above equation becomes:

$$\mathbf{X} = \mathbf{A}^T\mathbf{Y} + I\vec{m}_{\mathbf{X}}. \quad (8)$$

If only the most significant q principal components of the data are retained, \mathbf{A} becomes a $q \times n$ matrix, denoted as \mathbf{A}_q . The transformed data obtained by using only the selected q principal components is given by the following equation:

$$\tilde{\mathbf{Y}} = \mathbf{A}_q(\mathbf{X} - I\vec{m}_{\mathbf{X}}), \quad (9)$$

where the size of $\tilde{\mathbf{Y}}$ is $q \times MN$. Even though some of the eigenvectors are discarded, it is still possible to recover the initial input data by considering the following approximation:

$$\tilde{\mathbf{X}} = \mathbf{A}_q^T \tilde{\mathbf{Y}} + I\vec{m}_{\mathbf{X}}, \quad (10)$$

Consider an input of 6 normalized iris images, as shown in the left panel of Figure 3. The principal components of this data, arranged in the descending order of magnitude, are shown in the central panel of Figure 3. It can be noticed that the principal components of higher magnitude account for the maximum variability in the input data. If all the principal components are used, the original data can be reconstructed. The images reconstructed using only the top q (in this case, $q = 2$) principal components are shown in the right panel of Figure 3. The above process results in an approximation of the initial data, since all the eigenvectors are not used during reconstruction. However, the impact of dropping the lowest-valued eigenvectors is less significant on the reconstructed data.

4. Information Fusion

The following steps describe the process by which the proposed image-level information fusion scheme is applied to low resolution iris video frames:

1. Let V be a low-resolution iris video containing n frames denoted by $F = \{f_1, f_2, \dots, f_n\}$.
2. A set of k good quality frames are manually selected from the available frame set. It is not necessary for the selected frames to be successive in the video stream.
3. The selected frames are processed to segment and normalize (un-wrap) the iris to equal sized rectangular entities.

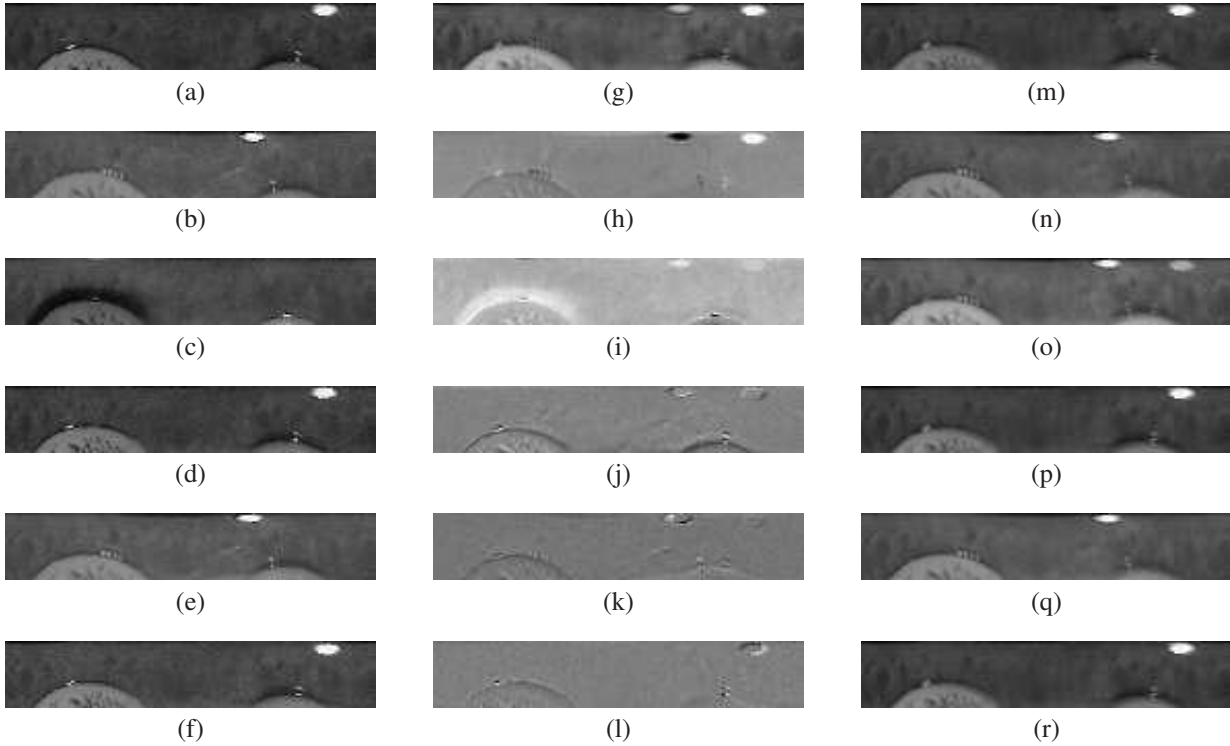


Figure 3. (a)-(f): Normalized input iris images, (g)-(l): principal components of the input data arranged in the decreasing order of magnitude, (m)-(r): reconstructed output obtained by using the PCT approach with $q = 2$. Images have been scaled to fit the document.

4. PCT is applied on these normalized frames to obtain the reconstructed frames ².

5. The reconstructed frames, are further fused by image averaging process, yielding a single output image. This output, instead of the original input, is used during the recognition process.

By using the evidence of multiple frames, the PCT based on pixel intensity vectors, projects the iris image onto a lower manifold where its discriminatory information is optimized. This optimization is accomplished using the eigenvectors of the covariance matrix of pixel intensities. While other types of manifold analysis techniques (e.g., NMU) can be used, in the interest of computational complexity, the PCT scheme is adopted in this work. By reconstructing the frames based on the eigenvectors, the salient information is extracted. A schematic of the technique is shown in Figure 4.

²The transformation alters the content of the input frames but does not change their spatial resolution.

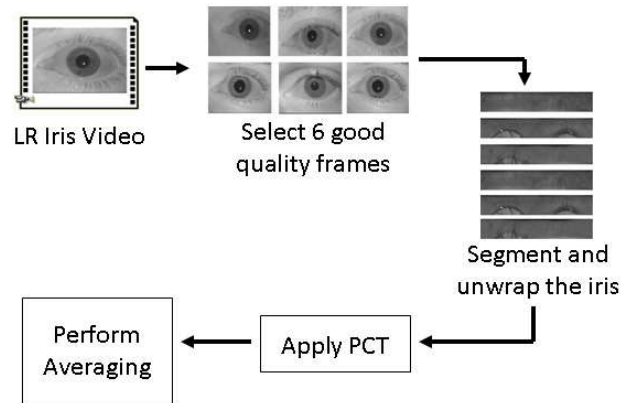


Figure 4. Proposed image-level fusion scheme.

5. Experiments

5.1. Database

A subset of the Multi-Biometric Grand Challenge (MBGC) database ³ containing Near Infrared (NIR) iris videos was used for the experiments. The iris video streams of multiple subjects are recorded in MPEG-4 format under varying illumination. A set of 110 right iris videos were

³www.face.nist.gov/mbgc/

selected for the study by considering 1 video each of 110 unique subjects. Frames are extracted from the videos and saved in BMP format without any compression. Each video contained 300 frames on an average, with every frame having a spatial resolution of 640×480 pixels.

A gallery set comprising a total of 440 images was formed by selecting 4 frames per subject. The value of k , which represents the number of low-resolution probe frames, was empirically chosen as 6. Thus, a set of 660 frames (6 frames each, for 110 videos) were chosen as the probe set. The frame selection process was performed manually, based on factors that impact the amount of information useful for recognition (e.g., percentage of iris visible in the frame, specular reflection, blur, occlusion due to eyelashes and eyelids, etc.). As the selected frames need not be successive, it is possible to observe contraction or dilation of the pupil due to varying illumination. However, the proposed approach can handle such variations by a small margin (as was verified empirically).

5.2. Pre-processing

To generate the low-resolution data, the original probe set of 640×480 pixel resolution (referred to as *ProbeSet L1*) is sub-sampled. Sub-sampling was performed by an averaging operator to reduce the size of a frame by a factor of 4: $1/2$ the length and $1/2$ the width. The sub-sampling operation was performed on every frame of the *ProbeSet L1* to obtain a lower-resolution frameset. This process was used iteratively to generate multiple framesets of the following resolutions: 320×240 , 160×120 , and 80×60 (referred to as *ProbeSet L2*, *ProbeSet L3*, and *ProbeSet L4*, respectively). Any resolution below 80×60 pixels (*ProbeSet L4*) was considered too low to work with. The average diameter of the iris in *ProbeSet 1*, *ProbeSet 2*, *ProbeSet 3*, and *ProbeSet 4* was approximately 220, 110, 50, and 20 pixels, respectively. Figure 5 shows a sample frame at various resolutions.

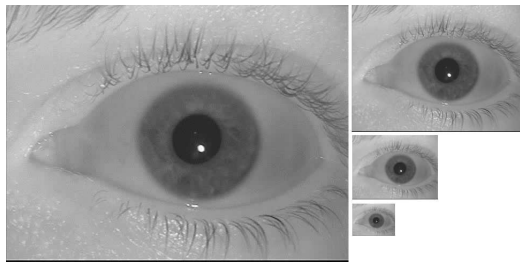


Figure 5. A sample right iris frame of resolution 640×480 down-sampled to resolutions of 320×240 , 160×120 , and 80×60 . Images have been scaled to fit the document.

Iris segmentation was performed independently on each frame for the various probe sets. Segmenting the iris be-

comes difficult with the decreasing resolution of the image. To avoid the effect of incorrect iris segmentation on the recognition performance, a semi-automated segmentation scheme is adopted. This is performed by a human observer by marking the boundaries of the iris in a frame. Simultaneously, a noise mask is created, which records the locations of eyelids and eyelashes that occlude the segmented iris. Both the segmented iris region and the noise masks are normalized using Daugman's rubber sheet model [1]. In this work, the segmented irides in all probe sets (varying resolution) are normalized to a fixed size: 32×180 pixels. Usually, the most suitable size for normalizing an iris image is based on the radii of the pupil and iris. However, matching normalized irides of different sizes is not possible using the existing Daugman's approach [1]. In general, matching iris images of different resolutions is still an open problem. The loss of textural information caused by normalizing frames of varying resolutions to a fixed size can be noticed in Figure 6.

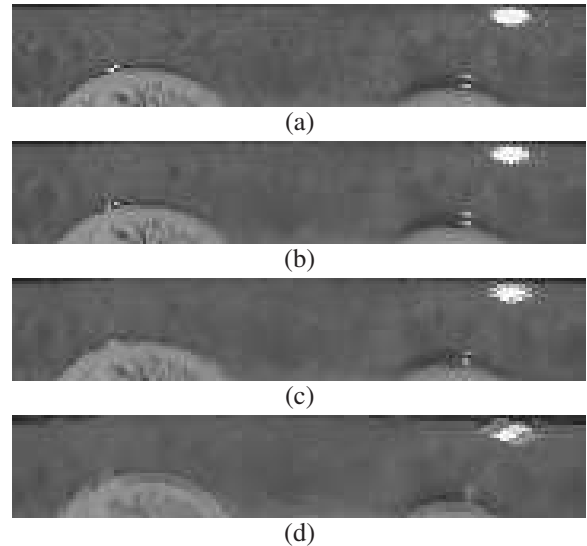


Figure 6. Normalized probe images corresponding to an iris frame at multiple resolutions: (a) 640×480 , (b) 320×240 , (c) 160×120 , and (d) 80×60 . Images have been scaled to fit the document.

To extract the textural features of the iris, a two dimensional Gabor filter is convolved with the unwrapped iris image. The output of this convolution operation contains both the real and imaginary responses. A phase demodulation process is used to encode these responses to a binary biometric template, often called as an *IrisCode*. Hamming distance is used to measure the dissimilarity between two *IrisCodes* while masking the corresponding noisy regions. An open source MATLAB implementation [8] was used with minor modifications to perform above mentioned operations.

5.3. Recognition Accuracy

Two iris recognition software packages, IrisBEE⁴ and VeriEye⁵, were initially used to observe the matching performance at various image resolutions. However, neither package could generate the iris templates or perform matching on probe images below a resolution of 320×240 . Thus, the performance evaluation in this work is conducted using an open source MATLAB implementation [8] for iris encoding and matching. Receiver Operating Characteristic (ROC) curves are used to evaluate and compare recognition performance. Every frame in the probe set is matched against all the gallery frames. A total of 2,640 genuine scores and 287,760 impostor scores were obtained by the matching process for each probe set. The performances obtained by matching the probe sets with the gallery are shown in Figure 7.

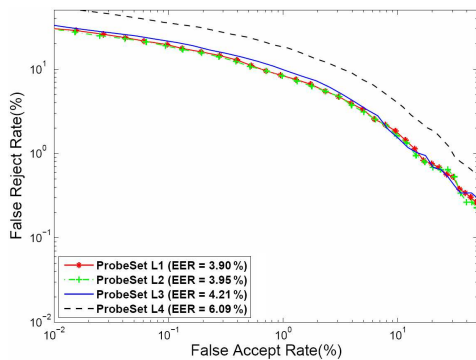


Figure 7. ROC curves obtained by matching probe sets with the gallery.

From Figure 7, it is observed that the recognition performance drops significantly for *ProbeSet L4*. Therefore, the proposed algorithm is used on this set to improve its performance. As interpolation is a commonly used technique to upsample low-resolution images, the matching performance obtained after interpolation is used as the baseline. Hence, the recognition performance obtained by interpolating *ProbeSet L4* to a resolution of 160×120 is used as a baseline. Figure 8 shows the recognition performances of the new framesets obtained by applying (a) only PCT, (b) only image averaging, and (c) the proposed approach, on normalized frames of *ProbeSet L4*. From the figure, it can be noticed that the Equal Error Rate of *ProbeSet L4* is reduced from 6.09% to 1.76%. This suggests that the recognition performance of the proposed approach is better than performances obtained by the individual stages involved (PCT and averaging).

The genuine and impostor match score distributions of *ProbeSet L4* before and after applying the proposed technique are shown in Figure 9. From the results, it is observed

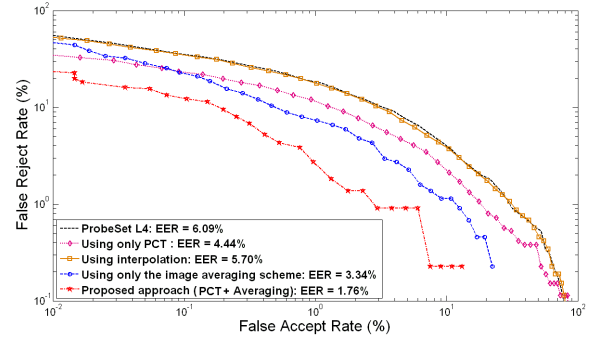


Figure 8. ROC curves obtained before and after applying the proposed technique. Note that by applying the proposed approach, the Equal Error Rate of *ProbeSet L4* is reduced from 6.09% to 1.76%.

that the genuine match score distribution shifts toward zero, indicating a reduction in false reject rate (FRR).

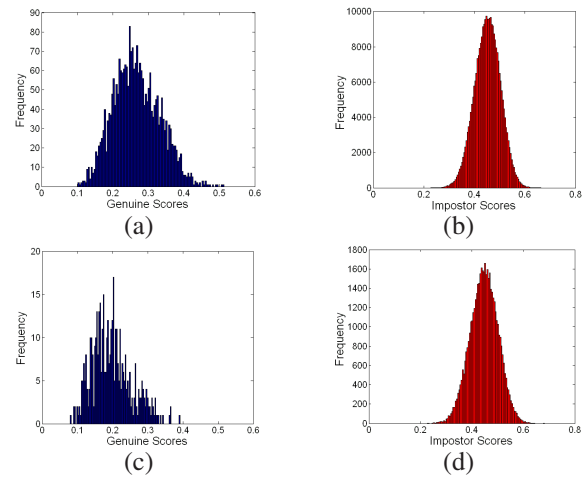


Figure 9. Genuine and impostor match score distributions: (a), (b) before, and (c), (d) after applying the proposed technique.

To further evaluate the performance of the proposed technique, the above experiment was repeated with the same setup using the left iris videos of 100 subjects. The performances obtained using this data are summarized in Table 1. From the results obtained using both left and right iris videos, it can be stated that the recognition performance of low resolution frames can be significantly improved using the proposed technique.

6. Score-level Fusion

The match scores generated by comparing a gallery image against a multi-frame probe set can be fused by employing a score-level fusion scheme [10]. In this work, score-level fusion is applied to (a) *ProbeSet L4* and (b) the cor-

⁴www.iris.nist.gov/ice/

⁵www.neurotechnology.com/verieye.html

Table 1. Equal Error Rates obtained using the left iris video data.

| Probe Set | EER |
|--|-------|
| ProbeSet L1 | 3.16% |
| ProbeSet L2 | 2.96% |
| ProbeSet L3 | 3.10% |
| ProbeSet L4 | 5.45% |
| On interpolating ProbeSet L4 | 3.38% |
| On applying PCT on ProbeSet L4 | 2.03% |
| On applying averaging on ProbeSet L4 | 2.58% |
| On applying the proposed approach (PCT + averaging) on ProbeSet L4 | 1.48% |

responding frame set obtained after applying PCT. Given a set of match scores $\{S_1, S_2, \dots, S_n\}$ obtained by matching n probe frames $\{f_1, f_2, \dots, f_n\}$ against a gallery image, a new score is generated by the sum rule that merely takes the average of these scores. The EERs obtained by invoking score-level fusion on the right and left iris video data are summarized in Table 2. The score-level fusion of the PCT output reduces the Equal Error Rate of *ProbeSet L4* from 6.09% to 1.45% for the right iris videos, and from 5.45% to 1.46% for the left iris videos, respectively. The results indicate that the recognition performance can be further improved by using score-level fusion.

Table 2. Equal Error Rates obtained before and after applying score-level fusion on the right and left iris video data.

| | Right Iris (110 subjects) | | Left Iris (100 subjects) | |
|--------------|---------------------------|-------|--------------------------|-------|
| | Before | After | Before | After |
| ProbeSet L4 | 6.09% | 2.59% | 5.45% | 2.25% |
| PCT frameset | 4.44% | 1.45% | 2.03% | 1.46% |

7. Conclusions and Future Work

An image-level fusion scheme is proposed which improves the recognition performance of low-resolution iris images. By using the proposed approach on a low-resolution iris database with an average iris diameter of 20 pixels, the equal error rate is reduced from 6.09% to 1.76% for the right iris video data and from 5.45% to 1.48% for the left iris video data, respectively. The use of complex routines to enhance low-resolution iris videos is avoided. Future work would include improving the recognition performance further by considering other sub-space analysis techniques for fusion (such as Linear Discriminant Analysis, Independent Component Analysis, etc.). The performance of the proposed technique when the input images are not properly registered should be investigated. A large scale evaluation of the proposed approach on larger and more challenging databases is to be performed.

References

- [1] J. Daugman. How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):21–30, January 2004.
- [2] J. Daugman. Probing the uniqueness and randomness of iriscodes: Results from 200 billion iris pair comparisons. *Proceedings of the IEEE*, 94(11):1927–1935, November 2006.
- [3] V. Dorairaj, N. Schmid, and G. Fahmy. Performance evaluation of iris based recognition system implementing PCA and ICA encoding techniques. In *Proceedings of SPIE Conference on Biometric Technology for Human Identification II*, volume 5779, pages 51–58, 2005.
- [4] G. Fahmy. Super-resolution construction of iris images from a visual low resolution face video. *National Radio Science Conference*, pages 1–6, March 2007.
- [5] L. Flom and A. Safir. Iris Recognition System. US Patent 4641349, February 1987.
- [6] K. Hollingsworth, T. Peters, K. Bowyer, and P. Flynn. Iris recognition using signal-level fusion of frames from video. *IEEE Transactions on Information Forensics and Security*, 4(4):837–848, December 2009.
- [7] J. Huang, L. Ma, T. Tan, and Y. Wang. Learning based resolution enhancement of iris images. *British Machine Vision Conference*, pages 153–162, 2003.
- [8] L. Masek and P. Kovesi. Matlab source code for a biometric identification system based on iris patterns, The School of Computer Science and Software Engineering, The University of Western Australia, 2003.
- [9] J. Matey, O. Naroditsky, K. Hanna, R. Kolczynski, D. Lolaccono, S. Mangru, M. Tinker, T. Zappia, and W. Zhao. Iris on the move: Acquisition of images for iris recognition in less constrained environments. *Proceedings of the IEEE*, 94(11):1936–1947, November 2006.
- [10] A. Ross, K. Nandakumar, and A. Jain. *Handbook of Multi-biometrics*. Springer Publishers, 2006.
- [11] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of cognitive neuroscience*, 3(1):71–86, 1991.