Graduate Theses, Dissertations, and Problem Reports

2005

Image quality assessment for iris biometric

Nathan D. Kalka
West Virginia University

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Image Quality Assessment for Iris Biometric

Nathan D. Kalka

A Thesis submitted to the
College of Engineering and Mineral Resources
at West Virginia University
in partial fulfillment of the requirements
for the degree of

Master of Science
in
Computer Science

Natalia A. Schmid, D.Sc., Chair
Lawrence A. Hornak, Ph.D.
Bojan Cukic, Ph.D.

Lane Department of Computer Science
and Electrical Engineering

Morgantown, West Virginia
2005

Keywords: Iris image quality, iris recognition, quality factors, Dempster-Shafer
theory, belief function, defocus blur, motion blur

ABSTRACT

Image Quality Assessment for Iris Biometric

Nathan D. Kalka

Iris recognition, the ability to recognize and distinguish individuals by their iris pattern, is the most reliable biometric in terms of recognition and identification performance. However, performance of these systems is affected by poor quality imaging. In this work, we extend previous research efforts on iris quality assessment by analyzing the effect of seven quality factors: defocus blur, motion blur, off-angle, occlusion, specular reflection, lighting, and pixel-counts on the performance of traditional iris recognition system. We have concluded that defocus blur, motion blur, and off-angle are the factors that affect recognition performance the most. We further designed a fully automated iris image quality evaluation block that operates in two steps. First each factor is estimated individually, then the second step involves fusing the estimated factors by using Dempster-Shafer theory approach to evidential reasoning. The designed block is tested on two datasets, CASIA 1.0 and a dataset collected at WVU. Considerable improvement in recognition performance is demonstrated when removing poor quality images evaluated by our quality metric. The upper bound on processing complexity required to evaluate quality of a single image is $O(n^2 \log n)$, that of a 2D-Fast Fourier Transform.

Acknowledgements

This work wouldn't have been possible without the help, encouragement, tutelage, and guidance bestowed upon me over the past year of my Master's program. I would like to take this opportunity to thank those individuals for their time and patience.

Firstly, I would like to express a deep heartfelt thanks to my advisor and committee chair, *Dr. Natalia Schmid*. She has been very patient with me and has provided me with invaluable guidance and encouragement, not only towards the completion of this work, but as a researcher.

My graduate Committee members, *Dr. Bojan Cukic* and *Dr. Larry Hornak* deserve a heartfelt thanks as well, for their constructive criticisms and invaluable feedback.

To my family, *Mom*, *GQ*, *Corey*, *Jayme*, and *Jeremy*.

To *Vivek*, *Jinyu*, *Xiaohan*, and *Francesco*.

To my roommates, *Nick* and *Chris*, without your witty comments, life at 502 wouldn't be the same.

Finally, Lockheed Martin and the DOD Biometric Fusion Center are also deserving thanks, without their sponsorship this research wouldn't have been possible.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

The iris biometric has received much attention lately from both academia and industry mainly due to its viability as a reliable biometric in terms of verification and identification performance. Daugman has demonstrated that iris recognition can achieve some of the lowest error rates with respect to False Accept Rate & False Reject Rate. However, similar to other biometrics, iris recognition has its problems. These problems stem from the fact that non-ideal imaging results in bad verification and identification performance. For example, low quality iris data such as in Fig. 1.1 result in poor recognition performance provided that traditional processing [1] of iris images is applied. These images have variable lighting, defocus blur, off-angle, and heavy occlusion, which have a negative impact on even the best available segmentation algorithms such as those developed by Daugman & Wildes, which is also shown in Fig. 1.1.

(a) Daugman's Segmentation Results   (b) Wildes's Segmentation Results
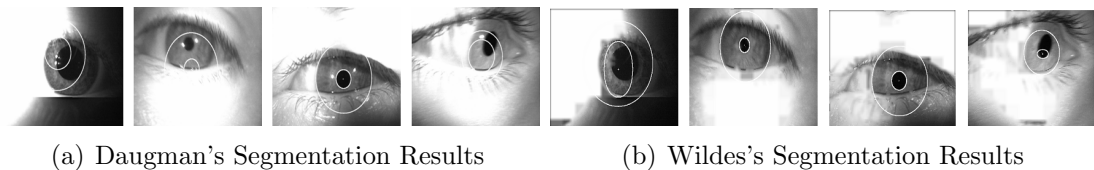
Figure 1.1: Segmentation Results

That being said, the primary motivation behind this research is to assess image quality for the purpose of successful recognition performance of an iris biometric system and demonstrate that quality must be considered when collecting biometric data.

1.2   Goals

The primary goal of this research is to design a fully automated image quality block that is capable of discriminating between good and poor quality images. Moreover, the quality metric should be able to predict performance. Finally, this research should provide insight on which factors negatively impact performance when using traditional iris recognition systems.

1.3   Requirements

Given the goals at hand, quality assessment should not require intensive preprocessing steps. This is extremely viable when collecting biometric data out in the field. The quality assessment should be able to provide feedback regarding the quality factors being measured if required. This allows for the characterization of datasets in terms of quality factors (i.e. CASIA primarily consists of occluded data). Assessment should only require one image rather than a sequence to evaluate quality.

1.4   Contribution

This work results in many contributions to the field of biometrics, namely the iris biometric in terms of quality. With respect to iris recognition, it is the first comprehensive work to evaluate which factors effect recognition performance on traditional iris recognition systems and Principle Component Analysis/Independant Component Analysis based systems. This is also the first work to incorporate seven quality factors in its quality evaluation as well as incorporating quality bounds on iris image quality. Finally, because the majority of the work done is related to the image processing

field, the quality metric can easily be modified to evaluate image quality for other biometrics such as face recognition.

1.5 Organization

The remainder of this work is organized as follows. Chapter 2 contains a literature review of biometrics and iris recognition. This review is necessary to help familiarize the user with nomenclature used in later chapters. Chapter 3 introduces quality assessment as well as providing a review of current iris quality literature. This chapter also gives a brief introduction to our approach to iris image quality. Chapter 4 discusses the synthetic studies for the quality factors used in this work and their effect on recognition performance. Chapter 5 outlines estimation procedures for the quality factors and the assumptions required to estimate those factors. Chapter 5 also introduces Dempster Shafer theory as an information fusion technique applied to quality assessment using Murphy's rule. Chapter 6 contains quality results for CASIA, and WVU datasets, as well as demonstrates the reliability of this quality metric by showing performance prediction in terms of verification performance. Chapter 7 concludes this work and proposes new ideas for future work related to iris quality assessment.

Chapter 2

Literature Review

Many identification systems authenticate individuals by associating them with a password and/or some form of physical key. These systems allow access based on "what you know" or "what you possess". The main problem with knowledge based systems is the difficulty arising from remembering passwords. This leads to individuals choosing simple passwords which could easily be guessed by malicious users. The main issue with possession based identification systems is that keys could be lost, stolen, forgotten, or misplaced. One approach that helps in alleviating the need to remember passwords or carry keys is based on "what you are" which is a combination of physiological and behavioral characteristics. This approach is known as biometrics.

2.1 Biometrics

Biometrics are automated methods of using physiological or behavioral characteristics to uniquely define individuals. Physiological biometrics can be described as those requiring the variability of the physical body, such as fingerprint, face,hand geometry, retina, ear and iris. Behavioral characteristics pertain to those biometrics which are affected by your behavior such as gait, voice, signature, and keystroke dynamics.

4

2.1.1 Biometric Characteristics

Both physiological and behavioral biometrics have desirable properties which determine their strengths and weaknesses such as: universality, uniqueness, permanence, collectability, performance, acceptability, and circumvention.

1. Universality - Corresponds to every individual having the characteristic.

2. Uniqueness - Refers to the fact that no two individuals should be the same with respect to the characteristic.

3. Permanence - Implies invariance to time.

4. Collectability - Implies that the characteristic can be quantitatively measured as well as how difficult the biometric is to collect.

5. Performance - Implies achievable and acceptable identification accuracy.

6. Acceptability - Refers to the extent to which individuals are willing to accept the biometric.

7. Circumvention - Refers to the difficulty in fooling the system with respect of spoofing and counter-navigation.

These seven characteristics determine the viability of a biometric. Ultimately the optimal biometric would score well with all characteristics, but such biometric exists to data.

2.1.2 Biometric System Outline

Biometric systems typically operate in two modes: verification and identification. Verification requires the user to assert an identity. The claimed users biometric is retrieved from a database and then compared to the input biometric. During identification no identity is claimed; the user is compared against all templates in the database.
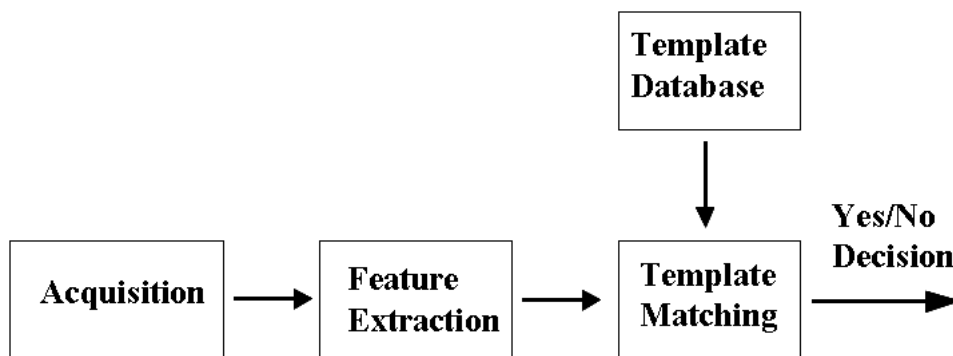
Figure 2.1: Biometric Block Diagram

Fig. 2.1 illustrates a generalized block diagram of a biometric system. The first step in any biometric system is biometric acquisition. This is typically done by some form of biometric reader. The medium (image, video, latent fingerprint) to which biometric data is collected depends on the biometric. The next stage, feature extraction, extracts features from the input medium gathered in the acquisition stage. The extracted features compromise a template.During verification the extracted template is only compared to the claimed user's template. The final output of the system is a yes/no decision.

2.1.3   Performance Analysis

The result of typical decisions made by biometric systems are of two types: genuine and imposter. These decisions are usually represented statistically by two distributions: genuine distribution and imposter distribution. Using these distributions, performance of the system can be represented by the following error rates:

1. FAR - False Accept Rate is characterized by imposter users being accepted as genuine users.

2. FRR - False Reject Rate is characterized by a genuine users being falsely rejected as an imposter.

3. EER - Equal Error Rate is characterized as the point in which FAR is equal to FRR.

(a) Generalized Matching Distributions

(b) Generalized Receiver Operating Curve
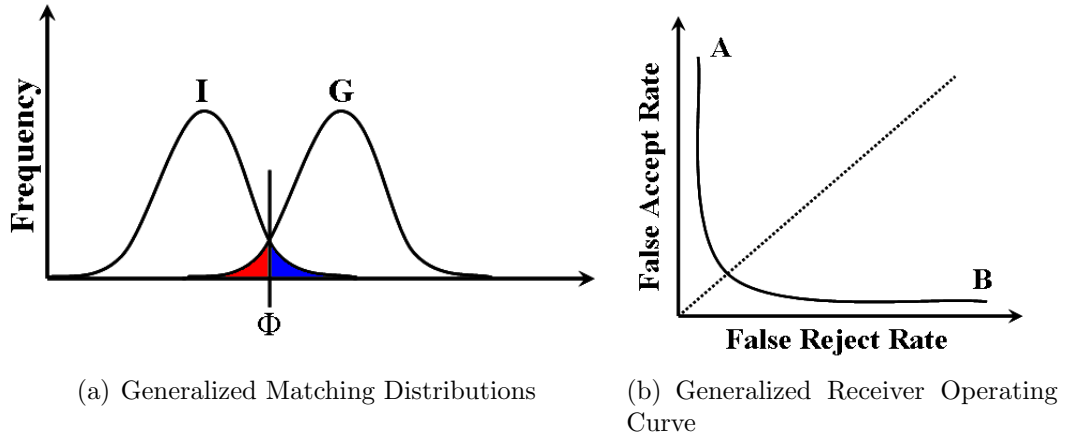
Figure 2.2: Generalized Matching Distributions and Generalized ROC

Fig. 2.2 illustrates generalized matching distributions (a) as well as a generalized ROC (b). FAR and FRR can be derived from the overlap region in fig. 2.2. Equation's (2.1) & (2.2) illustrate these error rates:

$$FAR = \int_{\Phi}^{\infty} I_{imposter\ distribution} \tag{2.1}$$

$$FRR = \int_{-\infty}^{\Phi} G_{genuine\ distribution} \tag{2.2}$$

The equal error rate (represented as Φ in fig. 2.2a) is the point at which FAR = FRR. Besides these metrics, there are some uncommon measures that are more often than not overlooked. Such metrics include:

1. FTE - Failure to Enroll is characterized as the acquired input metric not matching the enrollment criteria.

2. FTA - Failure to Acquire is characterized as the sensor being unable to detect the input biometric signal.

Fig. 2.2 (b) illustrates a generalized Receiver Operating Curve. The dashed line represents the equal error rate, the point at which FRR = FAR. This curve provides more information than just FAR or FRR alone because it is an characterization of the system performance at different operating points. For example, applications requiring high security should operate around the B region because the FAR is low. The disadvantage of operating at this region is the FRR is high. Operating at region A yields high FAR but low FRR. This region accommodates those forensic applications in which identifying a individual is of utmost importance. The drawback is having to examine a large number of false accepts.

One other performance measure known as d' characterizes performance by measuring the separation in the genuine and imposter distributions. The following expression:

$$d' = \frac{|\mu_{genuine} - \mu_{imposter}|}{\sqrt{\frac{\sigma^2_{genuine} + \sigma^2_{imposter}}{2}}} \tag{2.3}$$

uses first and second order statistics to determine the separation of the genuine and imposter distributions. $\mu_{genuine}$ & $\mu_{imposter}$ represent the means for both genuine and imposter distributions respectively while $\sigma^2_{genuine}$ & $\sigma^2_{imposter}$ represent the variances of the genuine and imposter distributions respectively. High d' values imply more separation of the genuine and imposter distributions while low d' values imply the opposite. The underlying assumption with d' is that both imposter and genuine distributions are gaussian. If the distributions are not Gaussian then equation (2.3) must be modified to accommodate the underlying distribution.

2.2   Iris Recognition

This section gives only a brief introduction to iris recognition systems. More rigorous explanations can be found in [1],[2],[3]. Throughout the rest of this work, the word "Traditional" is used to describe iris recognition systems based on Daugman's algorithms.

Traditional iris recognition systems are composed of five processing blocks: Acquisition, Localization, Normalization, Encoding, and Matching which are described in the following sections.

2.2.1 Acquisition

The acquisition stage, captures the iris image in near infrared light ranging from 700-900nm. The typical distance from camera to user is about one meter, with user cooperation (i.e. Fixed position with user looking into camera).

2.2.2 Localization

Localization represents the process of segmenting the pupil, sclera, and eyelid regions. Pupil and iris detection/segmentation in a traditional system is carried out by using an integro-differential operator that acts as a circular edge detector. Other segmentation methods employ various forms of edge detectors, active contours/snakes and Hough transforms which are explained in [4],[2],[5].

2.2.3 Normalization

This block normalizes the segmented iris region. Normalization is carried out to represent the segmented iris region with regard to invariance of size, position and orientation. Daugman transforms the coordinate system from cartesian coordinates to a doubly dimensionless nonconcentric polar coordinate system.
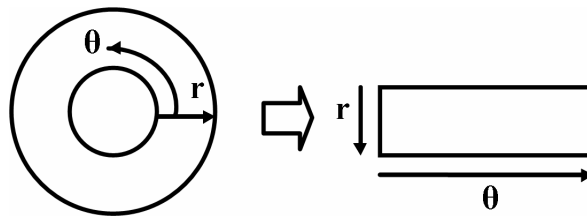
Figure 2.3: Daugman's Rubber Sheet Model

Fig. 2.3 illustrates the normalization process. Θ represents the angle (between 0 and 360) and **r** represents radial resolution (between 0 and 1).

2.2.4 Encoding

Encoding is characterized by projecting the segmented iris region onto complex-valued 2-D Gabor wavelets at varying scales, orientations and frequencies as demonstrated by Daugman [1]. The end result is quantized to form a binary template termed an iris code. This code is then used in the matching block. Other methods encode segmented iris regions by employing Gaussian pyramids such as in the Wildes system [4] or PCA/ICA encoding implemented by Dorairaj in [6],[7]. The PCA/ICA encoding techniques are of particular interest because they will be used along with the Gabor encoding technique in chapters 4 and 6 for comparison analysis.

2.2.5 Matching

Matching is done on two iris templates by use of the Hamming distance metric, which is essentially a standard measure for comparison of binary strings. Equation (2.4) expresses the hamming distance between two iris codes x and y. n is the length of the vectorized templates and $\bigotimes$ represents the exclusive-or operator.

$$Hamming\ Distance\ =\ \frac{1}{n}\sum_{i=1}^{n}x_i\bigotimes y_i \qquad (2.4)$$

PCA and ICA encoding employ a Euclidean distance metric which is expressed in equation (2.5). Here x and y represent the projected coefficients. The final output from this block is a yes/no decision determining if the comparison resulted in a match.

$$Euclidean\ Distance\ =\ \sqrt{\sum_{i=1}^{n}(x_i-y_i)^2} \qquad (2.5)$$

Chapter 3

Quality Assessment

In general, quality assessment refers to the evaluation, grading and measurement process to assess design and performance. Wang et al. in [8] discusses the importance and application of image quality metrics. They describe their various uses in acquisition systems to monitor and adjust themselves to obtain the best quality data. They also explain their use in benchmarking image processing systems and algorithms. They finally describe their use in optimizing parameter setting such as design and configuration of visual communication systems. All of these applications can easily be generalized with regard to biometric systems.

These applications play an important role in all biometric systems namely because of their impact on system performance which has been demonstrated for numerous biometrics such as face, fingerprint [9] [10], and iris [11] [12] [13]. Generalizing from above, biometric quality assessment metrics can help tune capturing systems and monitor collections (such as selecting images from a video sequence). Quality assessment can also be used as a primary discriminator when fusing biometric data, along with the ability to predict performance.

This work describes a quality assessment methodology for an iris biometric based on the input to the system. Before explaining the intricacies of this algorithm, a literature review of current iris quality metrics is presented, followed by an introduction to our approach.

3.1   Previous Works

Previous work on iris image quality can be placed into two categories: local and global analysis. Zhu in [14] evaluates quality by analyzing the coefficients of particular areas of iris texture by employing discrete wavelet decomposition. Chen et al. in [15] classify iris quality by measuring the energy of concentric iris bands obtained using 2-D wavelets. The Hamming distance metric is then modified, giving more weight to bands containing higher energy. They demonstrate a performance improvement of about 20% and 10% in equal error rate, respectively, for CASIA & WVU datasets. Ma et al. in [12] analyze the Fourier spectra of local iris regions to characterize defocus, motion and occlusion. Zhang [13] examines the sharpness of the region between the pupil and the iris. Daugman [1] and Kang [11] characterize quality by quantifying the energy of high spatial frequencies over the entire image region.

The major drawback of most existing approaches is that evaluation of iris image quality is reduced to estimation of a single [1], [11], [13], [15] or a pair of factors [12], such as defocus blur, motion blur, and occlusion. In addition, previous literature on evaluation of iris quality involves some form of segmentation with the intent of local analysis on the iris texture [15], [14], [12].

3.2   Our Approach

This research introduces a comprehensive approach to assess image quality for an iris biometric. We identify a broad range of factors including defocus blur, motion blur, occlusion, specular reflection, lighting, off-angle, and pixel-counts. We then analyze their effects on traditional iris recognition systems (our interpretations) as well as a PCA and an ICA encoding based systems. The intent of this analysis is to evaluate the importance of these factors in terms of performance degradation as well as gain insight on how to reasonably quantify each factor.

Defocus blur, motion blur, occlusion, specular reflection, lighting variation, off-angle, and pixel-counts are then quantified. Although the individual factors provide useful insight to quality, our primary goal was to develop a single metric that took

in the weight of all factors. To accomplish this we adopted a framework based on Dempster-Shafer theory. This framework allows us to aggregate quality bounds from the combination of the individual factors using Murphy's rule. Finally performance of our quality metric is tested on CASIA and WVU datasets. Chapters 4 and 5 explain in detail procedures carried out for factor performance degradation and factor estimation.

3.3 Datasets

The datasets evaluated in this study include: CASIA 1.0 [16], and WVU.

3.3.1 CASIA Details

The CASIA 1.0 dataset [16], contains 756 images from 108 different eyes with 7 images per eye. Images were taken using a self-developed capturing device with a resolution of 320 x 280. The pupils are synthetically masked to remove specular reflections.

3.3.2 WVU Details

The WVU dataset, consisted of 2495 images from 356 different eyes. The number of acquisitions per eye ranges from 2 to 17. Images were captured using an OKI IrisPass-H hand-held device with a resolution of 640 x 480.

Chapter 4

Synthetic Studies

There are currently no iris image databases publicly available that provide enough data with the factors to be studied. In regard to that, we selected a subset of iris images from CASIA and WVU datasets (10 users per dataset, 2 images per user) that yielded high quality images (determined by visual inspection). We then synthetically degraded image quality and evaluated recognition performance. To evaluate the influence of individual quality factors we invoked three algorithms (i) a traditional Gabor filter based iris encoding algorithm (our interpretation of Daugman's algorithm) [1], (ii) global Principle Components Analysis (PCA) encoding method and, (iii) the global Independent Component Analysis (ICA)-based encoding method introduced in [6],[7]. We intentionally use three distinct iris encoding techniques to simultaneously analyze the robustness of recognition system response on encoding techniques. The corresponding metrics that we used as measures of performance are Hamming and Euclidean distances which were introduced in chapter 2. Each figure demonstrating the degradation of performance contains two plots: an error-bar plot of mean genuine scores and an error-bar plot of mean imposter scores displayed as functions of the parameter characterizing the strength of a quality factor under study.

4.1 Procedures

As mentioned previously each user has two templates of fairly good quality (from visual evaluation). One template is never degraded while the other is synthetically degraded at varying strengths. For Gabor encoding, the good template is compared to the degraded template. During PCA & ICA encoding, training is done on the good template and testing is done on the synthetically degraded template. The quality factors under consideration are:

- Defocus Blur

- Motion Blur

- Off-Angle

- Lighting Variation

- Occlusion

- Specular Reflection

- Pixel Counts

The following sections describe each factor as well as show their effect on performance for three encoding techniques.

4.2 Defocus

Defocus blur can result from many sources, but in general, defocus occurs when the focal point is outside the "depth of field" of the object to be captured. The further an object is from this depth of field the higher the degree of defocus. Depth of field is affected by aperture size, the smaller the aperture size the greater the depth of field. To simulate this factor we convolve a sequence of low-pass Gaussian filters with our iris images. The images in Fig. 4.1 show the effect of the filtering.

(a) Kernel Size = 3   (b) Kernel Size = 7   (c) Kernel Size = 13
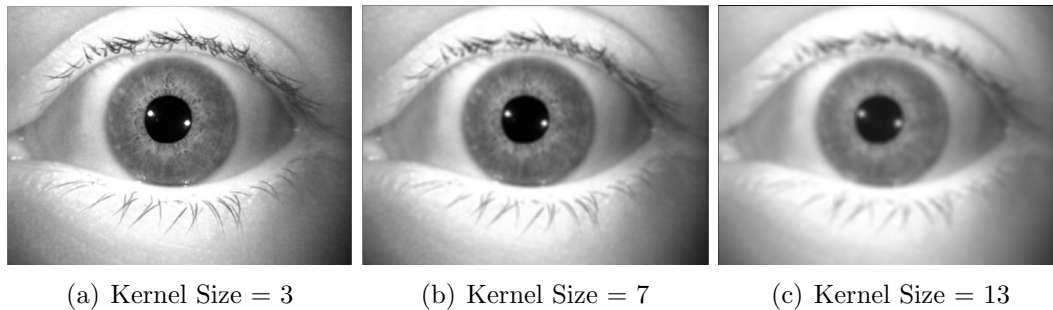
Figure 4.1: Synthetically Defocused Images

Fig. 4.2 shows the effect of defocus blur on (a) Gabor,(b) pca , and (c) ica encoding techniques. The graphs consist of genuine and imposter scores with error bars represented by the colors blue and red respectively. Synthetic blur level corresponds to the size of filter ($\sigma = 5$ for all sizes). It is interesting to note that for all techniques, small blur levels decrease Hamming and Euclidean distances. This is attributed to the slight smoothing effect of the Gaussian filter which denoises image content.

(a) Gabor   (b) PCA   (c) ICA
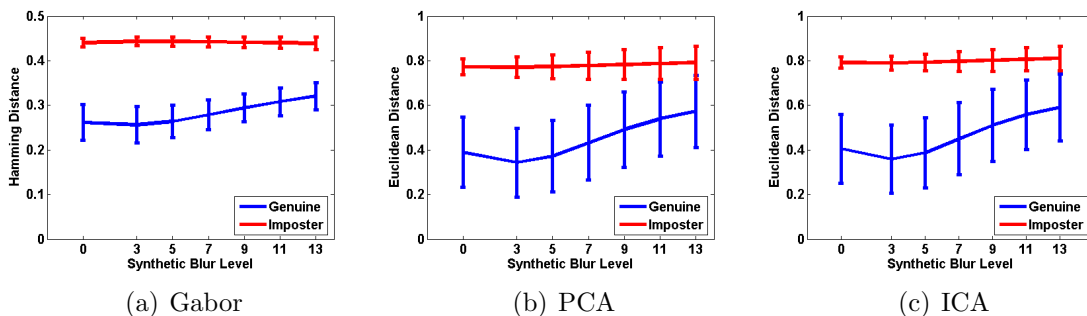
Figure 4.2: Defocus Blur Results

Strong blur levels on the other hand decrease the separation between the genuine and imposter scores for both distance metrics. The reason for this can be explained as follows: informative iris texture is primarily composed of high spatial frequencies, which correspond to iris features such as crypts, furrows and contour fibers. The resulting smoothing effect from the Gaussian filter at high blur levels suppresses the high frequency information and hence the resulting degradation in performance.

4.3 Motion

Motion blur can result either from the relative motion of an object or relative motion of the camera during exposure time. In general, there are two types of motion blur, linear and non-linear. Linear motion blur can be thought of as smearing in only one direction while non-linear involves smearing in multiple directions at different strengths.

(a) Smear Length = 5   (b) Smear Length = 15   (c) Smear Length = 25
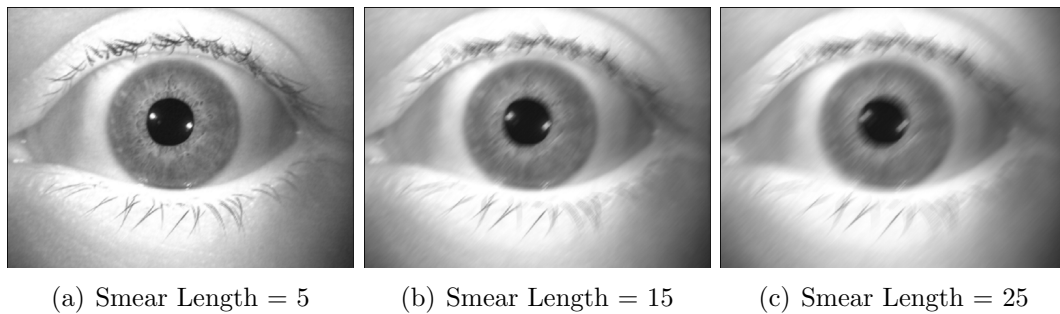
Figure 4.3: Synthetically Motion Blurred Images ($\Theta = 45$)

We consider only linear motion blur. With that in mind, we synthetically create linear motion blur by modeling two parameters: direction and strength of pixel-smear.

(a) Gabor   (b) PCA   (c) ICA
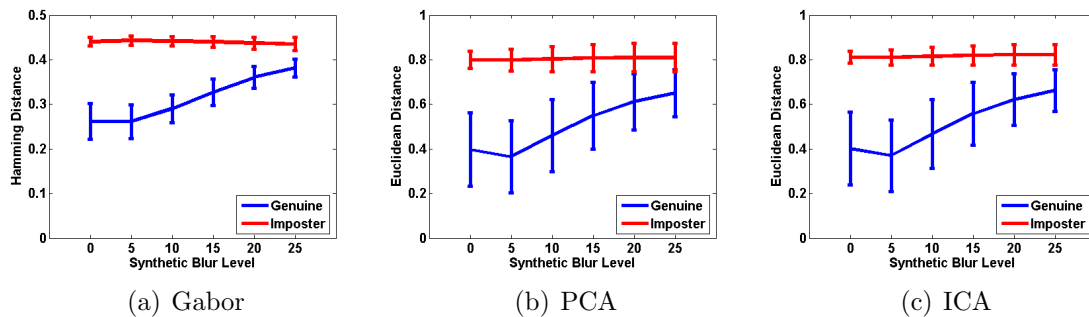
Figure 4.4: Motion Blur Results

Direction, denoted by Θ, corresponds to the linear direction of the blur, ranging from 0-180. Pixel-smear corresponds to the "strength" of the linear motion blur, ranging from 0-25. The iris images in Fig. 4.3 illustrate linear motion blur along the same direction but varying smear strengths. Notice the displacement of the

specular reflections in 4.3 (a) and 4.3 (c) which are displaced as a result of the linear motion blur by approximately 20 pixels. This displacement is somewhat of a "feature" in Fourier space for detecting linear motion blur in iris images. Fig. 4.4 shows the performance results for linear motion blur. All encoding techniques suffer from performance degradation with respect to genuine & imposter separation as the pixel smear increases. The main difference between defocus and motion with respect to Gabor encoding is that degradation is faster with motion blur, while PCA & ICA are equally influenced.

4.4   Off-Angle

Iris images which are not frontal view images are of special interest. Off-angle degradation can result from non-cooperative users or when capturing iris's at a distance. For evaluating the effect of off-angle on performance, initial testing is done on 36 iris classes from the WVU's off-angle iris image database. The database has 208 iris classes, four images per each class including two from frontal views, one from 15 degree view, and one 30 degree view. The initial angle values are those assigned during the data collection. Fig. 4.5 illustrates some of the images from this dataset. The first image 4.5 (a) is of frontal view while 4.5 (b) and 4.5 (c) are of 15 and 30 degrees respectively.


(a) $\Theta = 0$ (b) $\Theta = 15$ (c) $\Theta = 30$
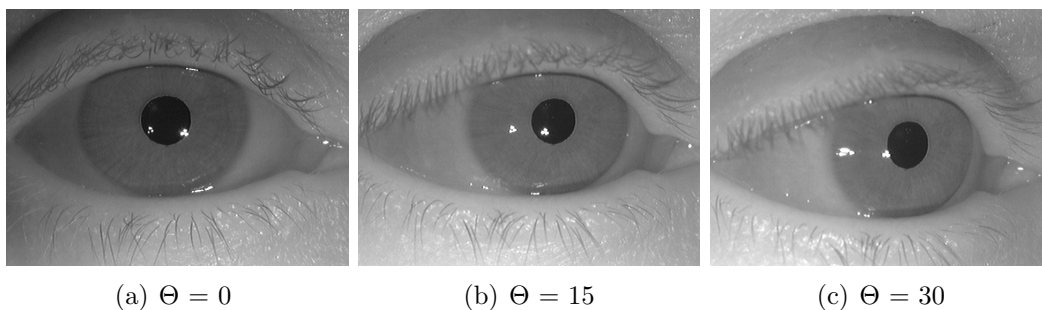
Figure 4.5: Sample Images from WVU Off-Angle Database

To evaluate the effect of off-angle, training is done on frontal view images and testing is done on off-angle images. The dependence of matching score values on the angle for

three encoding techniques are displayed in Fig. 4.6. For all encoding techniques one may summarize that the most influential performance degradation is at 15 degrees. After 15 degrees degradation starts to plateau.

(a) Gabor   (b) PCA   (c) ICA
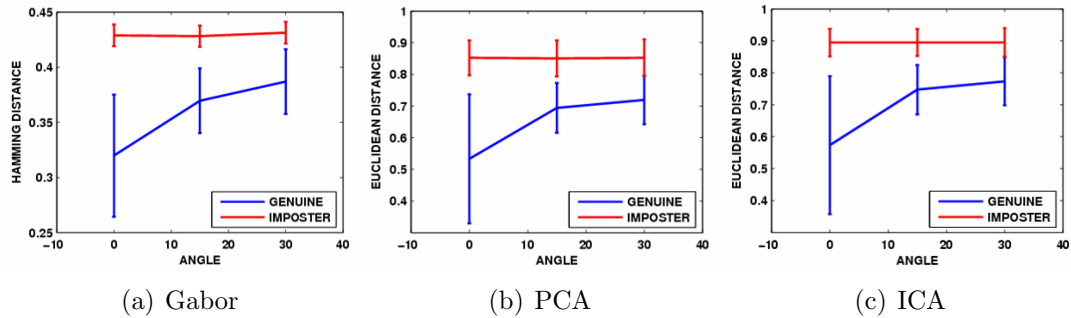
Figure 4.6: Off-Angle Results

4.5   Occlusion

Occlusion results from eyelashes, eyelids, camera orientation, hair, eye glasses, printed contact lenses and specular reflections that obscure iris texture. To compensate for this, traditional iris recognition systems mask out the occluded iris regions which can result in a reduction of informative iris texture hence degrading performance. Similar to this, we simulate eyelid occlusion by masking out portions of the iris region and evaluate the effect on recognition performance. Occlusion is simulated on upper, lower and combined iris regions at varying scales.

(a) Upper Eyelid Occlu-   (b) Lower Eyelid Occlu-   (c) Upper & Lower Eyelid
sion   sion   Occlusion
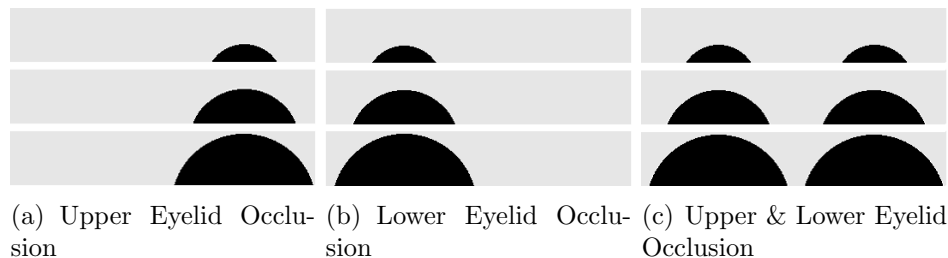
Figure 4.7: Simulated Occlusion

Fig. 4.7 displays some of the masks used to evaluate the effect of occlusion on three

encoding techniques. Fig. 4.7 (a) corresponds to occlusions resulting from the upper eyelid, 4.7 (b) corresponds to occlusions resulting from lower eyelid, and 4.7 (c) corresponds to occlusions resulting from both eyelids. The strengths for all three types of occlusion correspond to half-circle radius of the simulated eyelids. The corresponding radii are 20, 40, 60 pixels for all three types of occlusion in Fig. 4.7.

<div align="center">
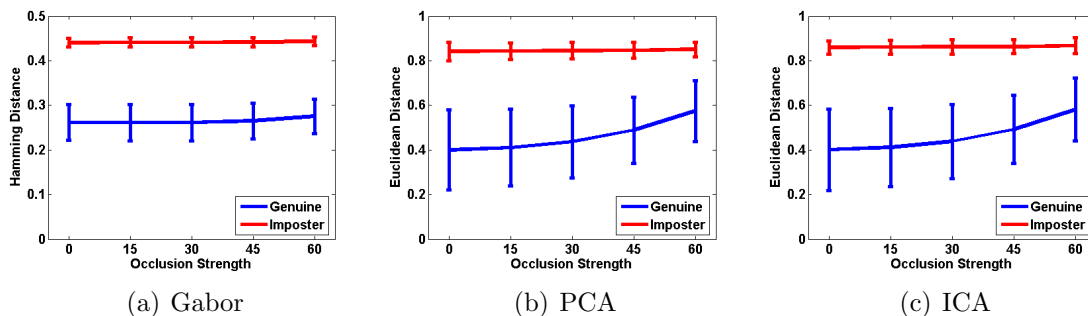
(a) Gabor   (b) PCA   (c) ICA

Figure 4.8: Occlusion Results

</div>

Fig 4.8 reflects the impact of occlusion on (a) Gabor, (b) PCA, and (c) ICA encoding techniques. Performance degradation is more pronounced for PCA and ICA encoding: as the occlusion increases the separation of genuine and imposter scores decreases. The Gabor encoding on the other hand is tolerant to occlusion. At the strongest occlusion levels, there is only a slight decrease in separation of genuine and imposter scores.

4.6   Lighting

Non-uniform or excessive lighting is a function of many sources but namely: granularity of the capturing system and user acclimation/cooperation. If the capturing system allows for variation in lighting, then there will be captures with variant lighting whether unintentional or not. A good example of this is the WVU dataset. Fig. 4.9 illustrates some of the variant lighting captures from the WVU dataset, which were captured from a hand held device. Along with that, capturing irises at a distance can also result in variant lighting conditions especially if the capturing system does not restrict user locality such as surveillance type applications.
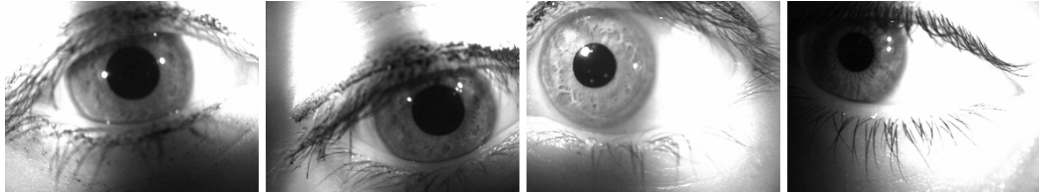
Figure 4.9: Variant Lighting Conditions from WVU dataset

In this study lighting is simulated by adding a constant (10, 30, 60, 90, 120, 150, 180) to specific iris regions in an attempt to simulate directional lighting. Fig. 4.10 (a) illustrates these regions and which part of the iris they correspond too. Fig. 4.10 (b) identifies these iris regions in pseudo-polar representation. Finally Fig. 4.10 (c) illustrates a real iris with lighting added to upper and lower regions (strength = 30).
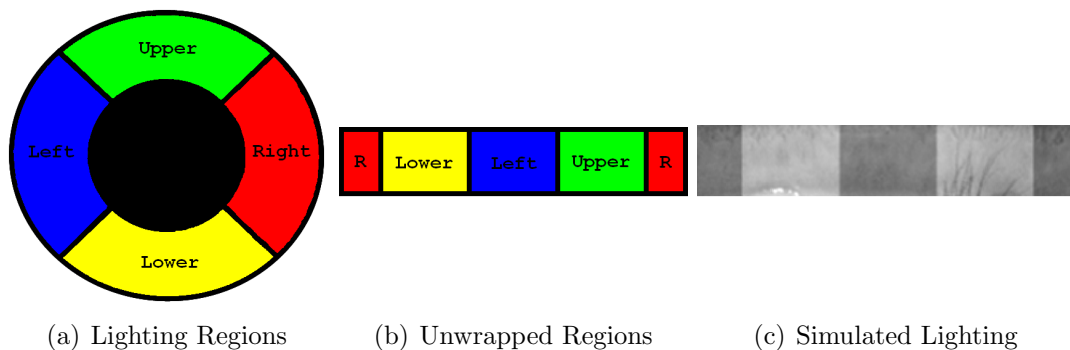
(a) Lighting Regions   (b) Unwrapped Regions   (c) Simulated Lighting

Figure 4.10: Simulated Lighting Regions

Overall lighting was simulated on left, right, left-right, upper, upper-right, upper-left, upper-lower, lower, lower-right, and lower-left regions. Results are shown for only the upper-lower region as this region was most influential on performance. Fig. 4.11 characterizes the effect of simulated lighting on the three encoding techniques. With all encoding techniques, there is minute change in genuine and imposter separation. This is consistent with [1] where Daugman illustrates the invariance of the Gabor encoding to variant lighting. PCA and ICA also demonstrate invariance to lighting. However this is more likely attributed to the background subtraction and contrast normalization done during enhancement rather than a result of the underlying techniques themselves.

(a) Gabor   (b) PCA   (c) ICA
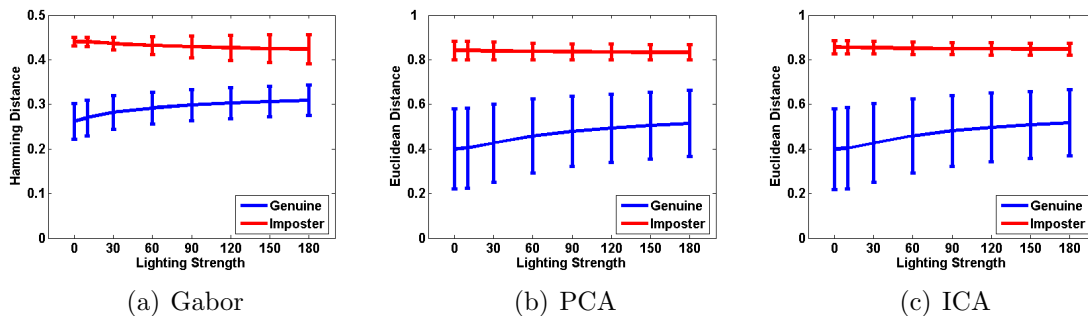
Figure 4.11: Lighting Results

4.7   Specular Reflection

Specular reflection is a result of light reflecting off the smooth surface of the eye lens back into the capturing device. Additional reflections may be caused by occlusions such as contact lenses and eye glasses. Specular reflection is similar to occlusion with respect to obscuring informative iris texture as well as negatively influencing segmentation performance. This factor was tested using a subset of the WVU dataset, which consisted of 50 users with 2 templates each. Fig. 4.12 (a) & (b) illustrate performance results for this factor.

(a) Gabor   (b) PCA   (c) ICA
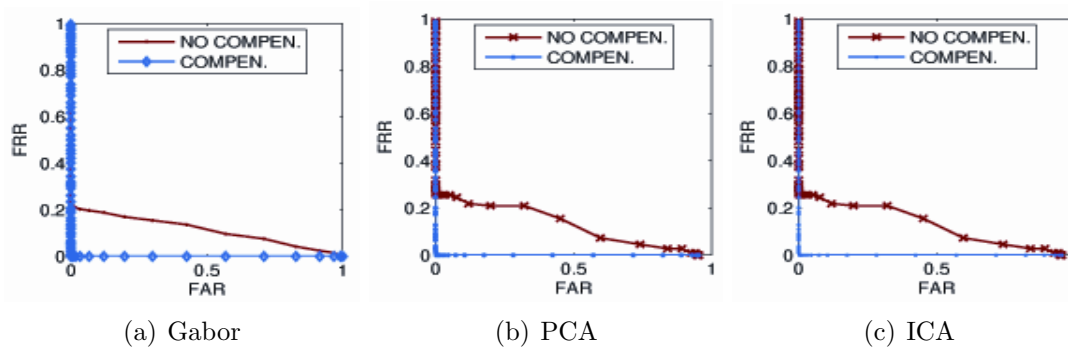
Figure 4.12: Specular Results

The resulting degradation in performance is due to failed localization of iris and pupil regions. However, performance improves for all encoding techniques once the reflections have been compensated. Compensation for specular reflections is done by

first localizing the reflections by use of hard thresholding. Then localized regions are median filtered (filter size = 16×16).

4.8 Pixel Counts

The amount of information within an iris is a function of its resolution. Capturing iris at a distance can result in variable resolutions. With that notion in mind, at which point does the information contained within the iris become unsuitable for distinguishing individuals? To simulate this factor we employ image downsampling on the normalized iris image by averaging with scales ranging from 2 to 14. For example, if the normalized image region is 64×360, downsampling with a scale of size 4 would reduce the image region to 16×90.

| (a) Gabor | (b) PCA | (c) ICA |
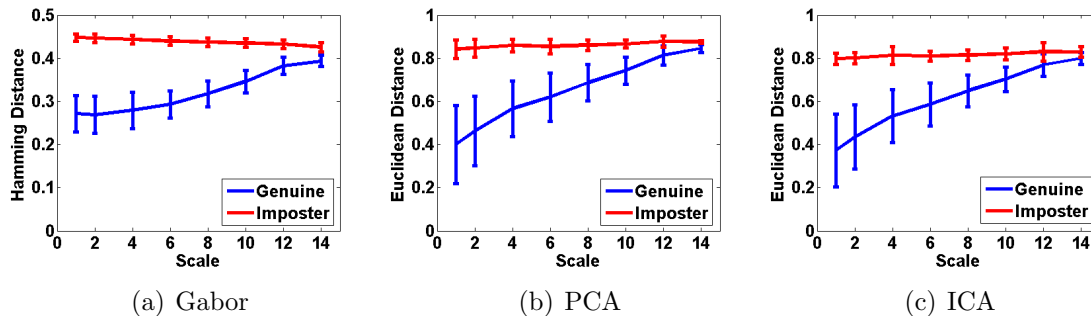
Figure 4.13: Pixel Count Results

Fig. 4.13 illustrates results from downsampling by averaging. For all encoding techniques performance degrades as the downsampling rate increases. The degradation for PCA and ICA encoding techniques is almost linear while Gabor encoding method is more tolerant to the downsampling.

4.9 Conclusions

The intent of this analysis was to study the impact of degrading factors on traditional Gabor , PCA, and ICA based iris recognition techniques. Since no public databases

exist that exhibit such factors, a subset from CASIA and WVU datasets was selected. The selected subset was then synthetically degraded in terms of defocus blur, motion blur, off-angle, occlusion, lighting, specular reflection and pixel counts at varying strengths. The impact on performance was quantified and illustrated in Fig. 4.2, 4.4, 4.6, 4.8, 4.11, 4.12, and 4.13.

From this analysis we notice that defocus blur, motion blur, and off-angle impact performance more acutely than occlusion, lighting, specular reflection, and pixel counts. This is not surprising as occlusions, lighting, and specular reflection in general impact segmentation more than directly effecting encoding techniques. Future studies will include the evaluation of these factors on segmentation performance to validate the aforementioned assessment.

Chapter 5

Factor Estimation and Evidence Fusion

5.1 Preprocessing

In this work, full iris segmentation is not required for evaluation of global image quality. Instead a pseudo "rough segmentation" method is used to evaluate local iris quality. This is accomplished by image down-sampling, then employing our own interpretation of Daugman's [1] or Wildes's [4] segmentation algorithms. Following this, estimated segmentation parameters are then re-scaled back with regard to the original image scale. Finally local analysis is performed on the normalized representations. The following sections describe estimation procedures for defocus, motion, occlusion, off-angle, specular, lighting, and pixel counts.

5.2 Defocus

Defocus primarily attenuates high spatial frequencies. Due to this relationship, defocus can be assessed by measuring high frequency content in the overall image or "roughly" segmented iris region. Daugman demonstrated this in [1] by proposing an (8x8) convolution kernel and measuring the total power in the response. This 2-D spectral power is then passed through a compressive non-linearity of the form:

$$f(x) = 100 * \frac{x^2}{(x^2 + c^2)} \tag{5.1}$$

in order to get a normalized score between 0 and 100. Here x is the total power spectrum measured by the (8x8) convolution kernel and c is the half-power of a focus score corresponding to 50%.

This spectral measure of focus works well when iris images are canonical about the iris as in Fig. 5.1 (a). However, when dealing with imaging not canonical about the iris, this spectral measure of focus can be misleading as in Fig. 5.1 (c).

(a) Canonical   (b) Filter Response   (c) Non-Canonical   (d) Filter Response
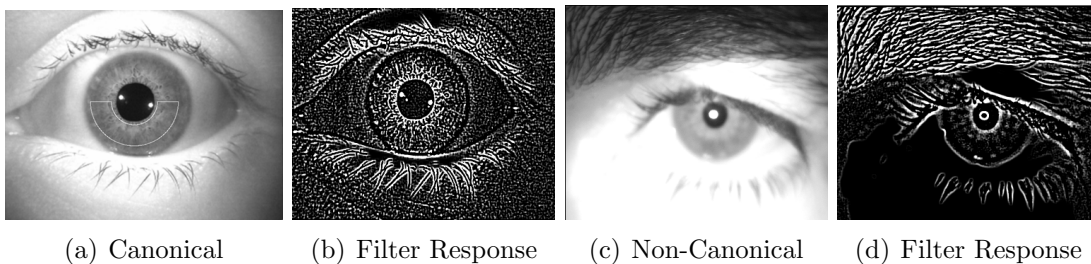
Figure 5.1: Sample Images from WVU dataset

Fig. 5.1 (b) and (d) display the responses of the band pass filtering when applied to Fig. 5.1 (a) and (c), respectively. Notice in Fig. 5.1 (b) that the iris region contains a significant amount of high spatial frequencies hence a highly focused image. In Fig. 5.1 (d) on the other hand, the iris region does not contain high frequency information. We conclude therefore that the iris region is defocused. However, introduction of in focus eyebrows in (d) results in a high focus global score hence the need for local focus assessment. To compensate for this we employ the same spectral measurement of focus but locally rather than globally. We also modify the compressive non-linearity to become:

$$f(x) = \frac{x^2}{(x^2 + P^2)}, \tag{5.2}$$

where x is the total power spectrum measured by an (8x8) convolution kernel and P is now the total power contained in the original image portion as defined by the local assessment region which is illustrated in Fig. 5.1 (a). This region was experimentally

chosen as the upper iris region is more likely to be occluded from the upper eyelid. The notion of P was introduced such that normalization is tolerant across datasets. It was experimentally found to give good results based on CASIA & WVU datasets.

5.3 Motion

Motion blur, as described previously, results from the relative motion between the object or camera during exposure time, which can result in linear and non-linear blur. Currently this work only includes estimation of linear motion.

Estimating linear motion blur is essentially estimating the primary direction in the image, along with the strength of this direction. To estimate the angle, we apply directional filters in Fourier space.
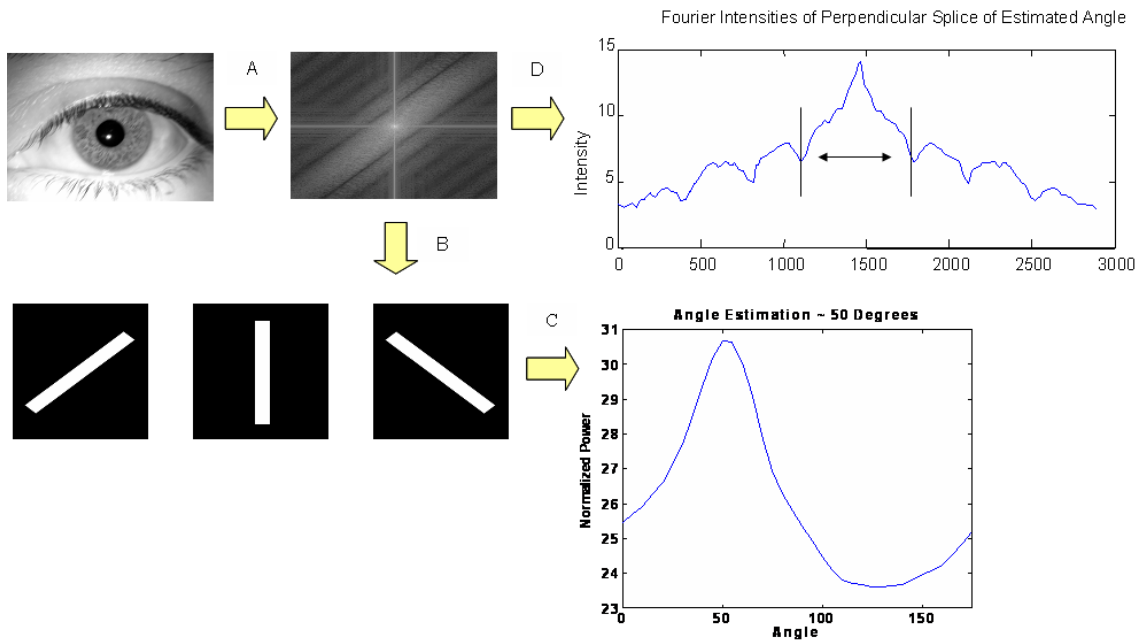
Figure 5.2: Motion Estimation Block Diagram

The input image is subjected to a Fourier transform as seen in Fig. 5.2 (arrow A). The dot product between the transformed input image and directional masks/filters similar to those shown in Fig. 5.2 (arrow B) (at 36 equally spaced orientations in the range (0, 180) degrees) is performed. The total power is calculated from each of

these responses. The response with most directional power gives an estimate of the angle as seen in Fig. 5.2 (arrow C). Let *I* be the image. Denote by *F(I)* the Fourier transform of *I*. To find the estimate of the motion blur angle, we apply directional filters of a given scale α. We denote the filter response at an orientation Θ by *H(Θ:α)*. The following equation expresses this process:

$$\hat{\Theta} = \arg \max_{\Theta \in [0:5:180]} \|F(I)H(\Theta : \alpha)\|^2 \tag{5.3}$$

Strength is estimated by analyzing a slice of Fourier coefficients perpendicular to the estimated angle of motion blur. Note the main "lobe" in Fig. 5.2 right above arrow B. The width of this main "lobe" is inversely proportional to the amount of motion blur strength. By measuring the power in the main "lobe" we can obtain an estimate of linear motion blur strength. Fig. 5.3 (a) is a plot of Fourier coefficients perpendicular to the estimated angle. In order to get the location of the main lobe the coefficients require smoothing. Fig. 5.3 (b) represents the smoothed coefficients by use of B-spline with least squares approximation. Once the coefficients are smoothed a gradient based approach is used to locate the main lobe.

(a) Noisy Fourier Coefficients   (b) Smoothed Coefficients
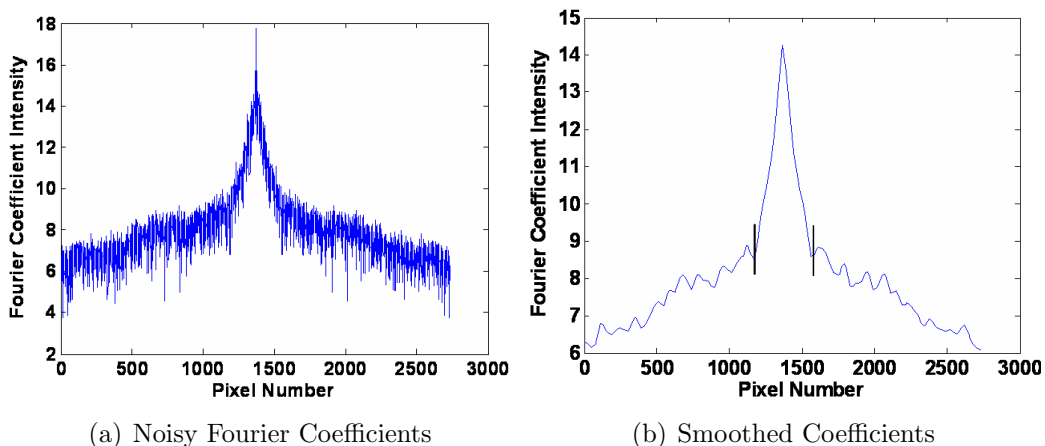
Figure 5.3: Perpendicular Fourier Coefficients

5.4 Off-Angle

The system processes off-angle iris images by estimating the gaze direction through the application of a projective transformation to bring an iris image into a frontal view image. The general block-diagram of the estimation system is shown in Fig. 5.4 which was reproduced from [7]. Below is a brief description of the estimation process. To estimate the angle of rotation we assume that a rough initial estimate of the angle is available.
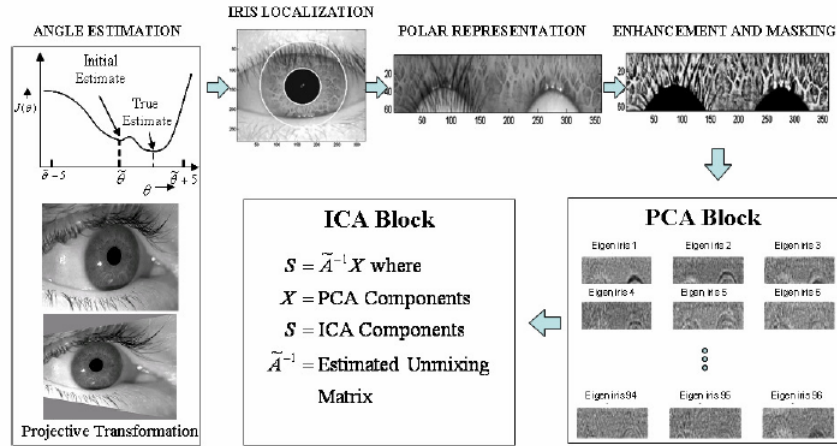
Figure 5.4: Off-Angle Estimation Block Diagram

The best estimate would be obtained by exhaustively searching all possible angles for roll and pitch. We use one objective function to refine the estimate: Daugman's integro-differential operator (see [1] for details) as a measure of iris circularity. We pick the estimates that maximize the value of the integro-differential operator. To be more specific, let Ψ_1 and Ψ_2 be two rotational angles and $J(\Psi_1, \Psi_2)$ be an objective function that has to be optimized. For each pair of (Ψ_1, Ψ_2) in the range $\Psi_1 \in [\Psi_{1,\min}, \Psi_{1,\max}]$ and $\Psi_2 \in [\Psi_{2,\min}, \Psi_{2,\max}]$, (i) the off-angle iris image is rotated by using the projective transformation and (ii) the objective function $J(\Psi_1, \Psi_1)$ is calculated. Once the angles are estimated we apply the projective transformation using the optimal angles estimated using the above procedure to rotate the off-angle image into a frontal view image. After this step, any iris recognition algorithm that operates on frontal view iris images can be applied.

5.5   Occlusion

Occlusion, specular reflection and lighting variation are estimated sequentially. This is to reduce spurious measurements introduced from eyelid occlusion, when estimating specular reflection. In order to estimated light variation accurately it is prudent to remove specular reflections since the resulting influence will acutely impact lighting variation.

With regard to that, occlusion is measured first. For eyelid occlusion measurement, we make the assumption that the sclera region and eyelid region are of differing intensities. This assumption allows us to adopt a gradient based approach to finding the edges of upper an lower eyelid occlusion on a "stretched" normalized iris image. To include portions of the sclera in the normalized image, we expand it by approximately 1.1 times the size of the estimated iris radius. The expansion was experimentally chosen based on evaluation of CASIA and WVU datasets. The sclera portion of the normalized image is smoothed by averaging. Next a horizontal gradient is calculated along the sclera portion to locate 4 points, two for each eyelid. Finally a half circle is fit to these points and a mask is generated (radius of each circle is dependant upon the Euclidean distance between the estimated points for each eyelid). Fig. 5.5 (a), (b), and (c) illustrate this process.

(a) Iris Illustration   (b) Horizontal Gradient   (c) Generated Mask
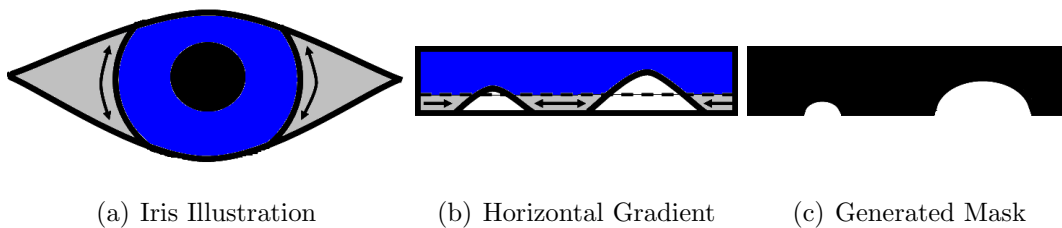
Figure 5.5: Occlusion Estimation

5.5.1   Specular Reflection

Once eyelid occlusions are estimated, occlusions resulting from specular reflection are estimated on the remaining iris portion unaffected by the eyelids. This factor is

estimated by hard thresholding. Based on evaluation of CASIA and WVU datasets, a threshold of 240 experimentally gave good results.

5.5.2   Lighting Variation

After estimating occlusions from eyelids and specular reflection, the remaining unoccluded iris portion is split into four regions. (illustrated in fig 4.10). The mean in each region is calculated and the variance of the means is used for our estimate of lighting.

5.5.3   Pixel Counts

This ISO Iris Image Standard [17] specifies that good quality iris images should have an iris diameter of 200 pixels. That being said, both WVU and CASIA datasets have an iris diameter of at least 200 pixels but both datasets do not consist of entirely high quality data.

Our measure for pixel counts is calculated as the ratio of iris pixels to occluded pixels within the iris region. It is expressed as follows:

$$\text{Pixel Counts} = \frac{X_{estimated}}{X_{estimated} + X_{occluded}} \tag{5.4}$$

where $X_{occluded}$ represents the number of pixels occluded from eyelids and specular. $X_{estimated}$ represents the number pixels estimated from rough segmentation.

5.6   Dempster Shafer Theory

To generate an overall global/local quality of iris images based on the estimated individual factors, we adopt a Dempster-Shafer theory approach [18] to information fusion. This approach was proposed as a solution to a number of problems in the field of artificial intelligence, software engineering, and pattern classification.

Dempster Shafer theory differs from Bayesian theory in following ways:

1. **No need to specify priors and conditionals**.

2. **Specify degrees of ignorance in the place of priors, called uncertainty**.

3. **Belief assignment is based on evidential reasoning**.

The belief for propositions (events in Bayesian theory) start at 0, with uncertainty equal to 1. Based on incoming evidence, belief assignments are updated, hence decreasing the uncertainty. In DS theory, belief models are built on a finite boolean algebra of mutually exclusive propositions known as the frame of discernment Θ. The belief in a proposition $Bel(A)$ is a measure of certainty that A is true. Shafer gives the following expressions for assigning and measuring beliefs.

If Θ is a frame of discernment, then a function $m : 2^\Theta \rightarrow [0, 1]$ is called a basic probability assignment when:

1. $m(\emptyset) = 0$

2. $\sum_{A \subset \Theta} m(A) = 1$.

To measure the belief of a proposition A, one must add up the belief in all subsets B belonging to A:

$$Bel(A) = \sum_{B \subset A} m(B) \qquad (5.5)$$

Dempsters rule of combination is used to combine beliefs over the same frame of discernment that are from distinct sources of evidence. This is measured by computing the orthogonal sum of all belief functions m which results in a new belief function based on the combined evidence.

$$m(A) = \frac{\sum_{A_i \cap B_j = C} m_1(A_i) m_2(B_j)}{1 - \sum_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j)} \tag{5.6}$$

Dempster's rule makes the following assumptions about evidence:

1. Independence.

2. Combination order is unimportant.

The problem with these assumptions lies in the fact that we do not have a good understanding of the dependencies between the quality factors and to assume independence between them is unreasonable (since our evidence is from the same source). In light of this, Murphy [19], [20] modified Dempster's rule such that it is suitable to use information from the same source as seen in (5.7) & (5.8).

$$m(A) = \frac{\sum_{A_i \cap B_j = C; C \neq \emptyset} f(m_1(A_i) m_2(B_j))}{\sum_{A_i \cap B_j} f(m_1(A_i) m_2(B_j))}, \tag{5.7}$$

where

$$f(m_1(A_i) m_2(B_j)) = [m_1(A_i) m_2(B_j)]^n, n \in [0, 1]. \tag{5.8}$$

Murphy characterizes n as a method to weight evidence. She explains that choosing $n > 0.5$ will give more weight when combining new evidence, while choosing $n < 0.5$ will give less weight when combining new evidence [20]. Other proponents of Murphy's rule characterize n as governing correlation between evidence [21]. It is explained in [21], that choosing $n > 0.5$ assumes more independence between the evidence while choosing $n < 0.5$ assumes correlation. In light of both views choosing $n = 0.5$ is considered neutral and equal weight is applied to all evidence during their integration.

5.6.1 Dempster Shafer Theory Applied to Quality Assessment

We adopt a frame of discernment containing two propositions which represent opposite beliefs:

1. A - Image quality is bad (Our belief that quality is bad).

2. B - Image quality is good (Our belief that quality is good).

The normalized values for each quality factor are assigned as beliefs to proposition A. Since these propositions represent opposite beliefs, the assigned belief to $B = \bar{A}$. We adopt Murphy's rule of combination to combine beliefs with parameter $n = 0.5$ for all evidence. Equation (5.9) is a generalized expression for combining beliefs from k quality factors m_1 to m_k.

$$m_i(A) = \frac{(m_{i-1}(A) \cdot m_i(A))^n}{((m_{i-1}(A) \cdot m_i(A))^n + (m_{i-1}(B) \cdot m_i(B))^n}, i = 2, .., k \qquad (5.9)$$

where $m_i(B) = 1 - m_i(A)$ since our propositions are complements of each other. Murphy has shown that different orderings result in different results for combined beliefs [20]. Since we have seven quality factors, that will result in 7! combinations. Our goal is to attain the orderings that result in the minimum and maximum values. These values provide valuable information about the global quality assessment of the iris image. Mladenovski in [21] has proved that by sorting the beliefs in ascending order with $n = 0.5$ for all belief combinations, a maximum value can be obtained. Similarly, if sorted in descending order a minimum value can be obtained. The following section illustrates some fusion results of real data from WVU and CASIA datasets.

5.6.2 Evidence Fusion Examples Based on Murphy's Rule

The sample iris images in Fig. 5.6 are from CASIA and WVU datasets. Image (a) represents a good quality from CASIA and (c) represents a good quality image from WVU (based on visual evaluation). Images (b) and (d) represent degraded quality images which are effected by occlusion (b) and motion in (d). The estimated angle for Fig. 5.6 (d) is 85.

| (a) Casia | (b) Casia | (c) WVU | (d) WVU |
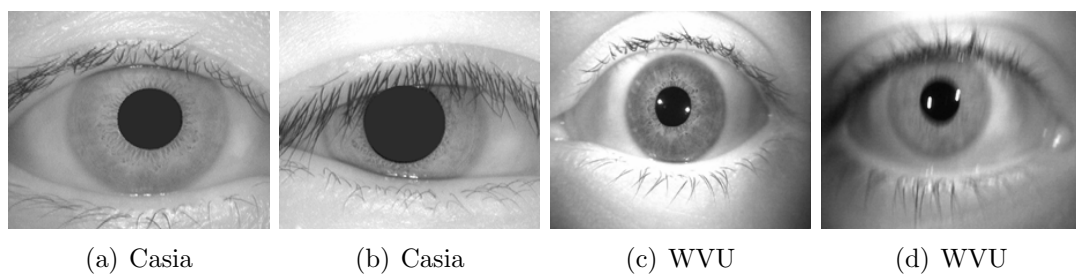
Figure 5.6: Sample Images from WVU and CASIA datasets

Table 5.1 lists the estimated factors (factors are between 0 and 1, with 1 implying heavy degradation) for these images and the combined quality for them. The quality column represents the lower bound (minimum value attained from fusion of all factors) on image quality. The estimated factors validate the assessment made above.

Image	Defocus	Motion	Occlusion	Specular	Lighting	Pixel Count	Quality
(a)	0.22	0.01	0.10	0.00	0.00	0.10	0.89
(b)	0.23	0.01	0.44	0.00	0.14	0.44	0.69
(c)	0.01	0.01	0.01	0.00	0.01	0.01	0.97
(d)	0.27	0.66	0.04	0.00	0.21	0.05	0.63

Table 5.1: Estimated Factors for images in Fig. 5.6

Chapter 6

Results

Evaluation of our quality metric was tested using WVU and CASIA datasets. Quality comparison is illustrated for three encoding techniques: Gabor, PCA, and ICA. Table 6.1 represents the rough segmentation results from WVU and CASIA datasets. In order to establish a baseline for our quality metric, all images that failed rough segmentation were excluded from the forthcoming analysis. Determination of failed segmentation was based on visual inspection.

Dataset	Number of Images	Failed Segmentations	Performance
CASIA	756	18	98%
WVU	2495	370	85%

Table 6.1: Rough Segmentation Performance

6.1 Quality Characterization

The mean factor scores for both datasets are illustrated in Table 6.2. Based on the statistics in this table we notice that the CASIA dataset is degraded primarily by occlusion and pixel counts while the WVU data suffers from occlusion, pixel counts, and lighting. The following sections will further detail the results about each factor.

Dataset	Defocus	Motion	Occlusion	Specular	Lighting	Pixel Count	# of images
CASIA	0.16	0.03	0.25	0.00	0.16	0.24	738
WVU	0.13	0.05	0.30	0.01	0.30	0.31	2125

Table 6.2: WVU and CASIA Mean Quality Factor Scores

6.1.1 Quality Bounds

Fig. 6.1 and 6.2 represent distributions of upper and lower quality bounds attained from using Dempster Shafer theory for information fusion for both CASIA and WVU datasets. Clearly the upper quality bounds for both WVU and CASIA in Fig. 6.1 are not discriminating in terms of quality. The majority of the distribution for both datasets falls on the high tail end. CASIA and WVU datasets both have an upper bound $mean = 0.99$.

(a) CASIA   (b) WVU
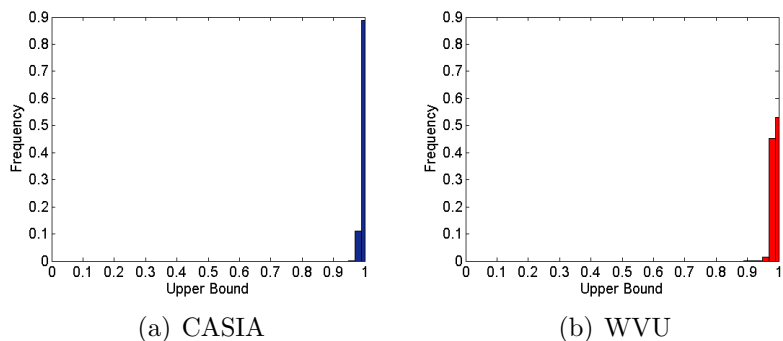
Figure 6.1: Upper Quality Bound Frequencies for CASIA and WVU

On the other hand, the lower quality bounds do provide discriminating information in terms of quality as seen in the histograms of Fig. 6.2 (This will be verified in the performance section). From Fig. 6.2 we notice that the mean of the distribution for CASIA is 0.79 while the WVU dataset has a greater spread, with the $mean = 0.65$. Fig. 6.3 is a scatter plot of the lower bounds for CASIA and WVU datasets. The lower quality bounds are used in the performance section to divide the datasets for performance prediction.

(a) CASIA   (b) WVU
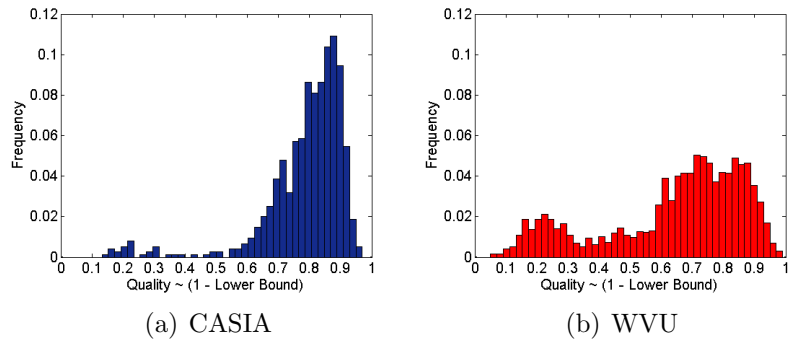
Figure 6.2: Lower Quality Bound Frequencies for CASIA and WVU

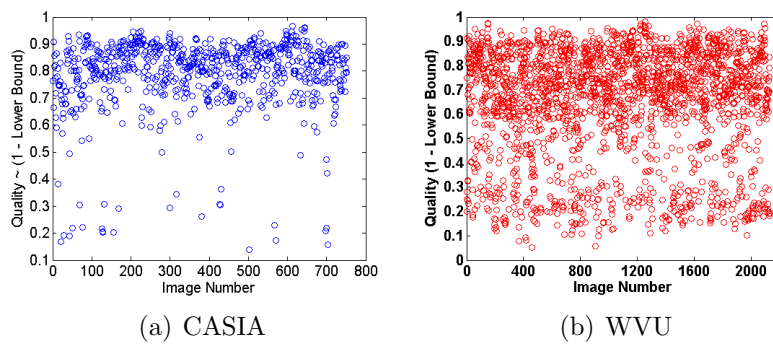(a) CASIA   (b) WVU

Figure 6.3: Overall Quality Scatter plots for CASIA and WVU

6.1.2 Factor Distributions for CASIA and WVU datasets

This section contains histograms of the quality factors for CASIA and WVU datasets. The x-axis for all plots in section represents the strength of the estimated factor. Fig. 6.4 (a) and (b) are histogram plots of the defocus estimates for both datasets. We notice that with CASIA there are relatively low defocus scores with a $mean = 0.16$. The WVU dataset has a wider range of defocus scores which slightly models an exponential distribution.
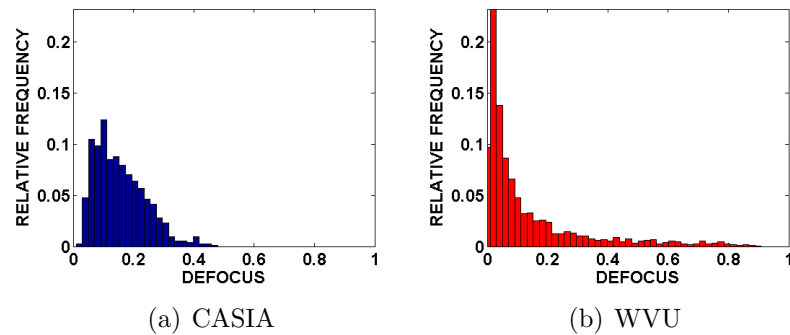

(a) CASIA   (b) WVU

Figure 6.4: Defocus Frequencies for CASIA and WVU

Fig. 6.5 (a) and (b) represent histograms of the motion blur estimates. From these plots we can easily notice that neither dataset contains a significant amount of motion blur.
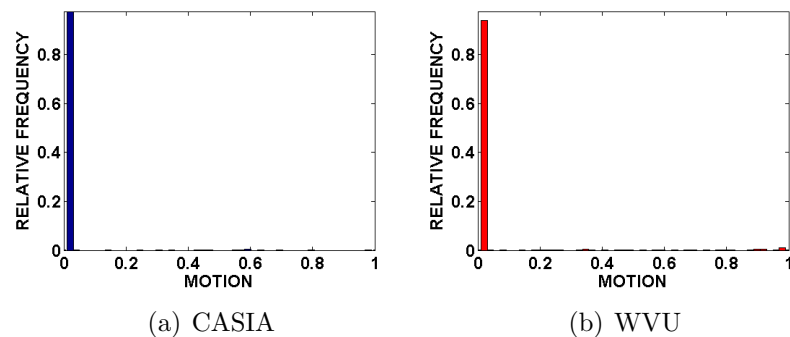

(a) CASIA   (b) WVU

Figure 6.5: Motion frequencies for CASIA and WVU

This is consistent with CASIA, as this dataset was collected in an ideal environment. While the WVU dataset is not ideal, based on our metric we do not notice significant amounts of motion blur.

(a) CASIA   (b) WVU
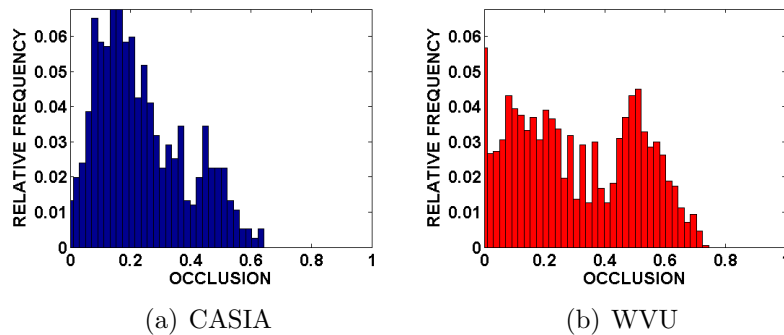
Figure 6.6: Occlusion frequencies for CASIA and WVU

Fig. 6.6 (a) and (b) are histogram plots of occlusion estimates. For both datasets there are significant amounts of occlusion. CASIA has a mean occlusion estimate of 0.25. From visual inspection of this dataset one may notice eyelid and especially the eyelash occlusions. The WVU dataset has mean occlusion estimate of 0.30, hence it is degraded more than CASIA with respect to this factor. Although not the only factor, it is indeed noticeable when visually inspecting the dataset.

(a) CASIA   (b) WVU

Figure 6.7: Lighting Variation frequencies for CASIA and WVU

Fig. 6.7 (a) and (b) represent the plots for lighting variation. For CASIA, the majority of the distribution is on the low tail end, with the $mean = 0.16$. We notice the same

tail like distributions with WVU, but with a significant portion on the high end. The WVU dataset has a mean lighting score of 0.30.

(a) CASIA   (b) WVU
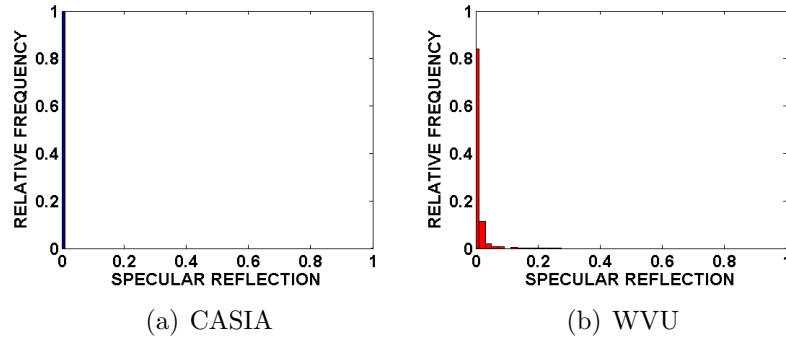
Figure 6.8: Specular frequencies for CASIA and WVU

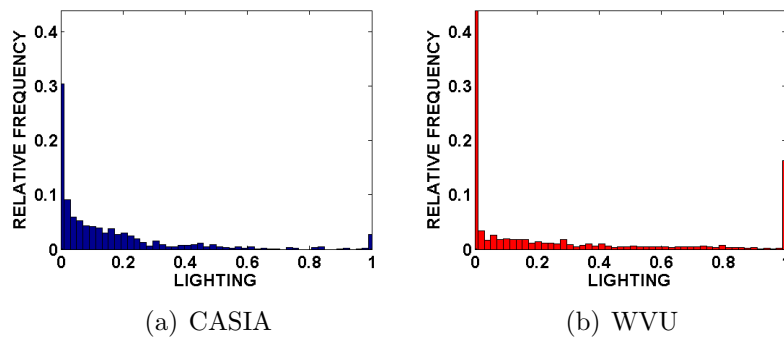Fig. 6.8 (a) and (b) characterize the distributions for the specular reflection estimates. The specular reflections occuring were masked out in the CASIA dataset and hence the lack of specular reflection in (a) is not surprising. CASIA has a mean specular reflection score of 0. The WVU data in (b), although not ideal, lacks specular reflections as well but not to the extent of CASIA. The mean of this quality factor with respect to WVU data is 0.01.

(a) CASIA   (b) WVU
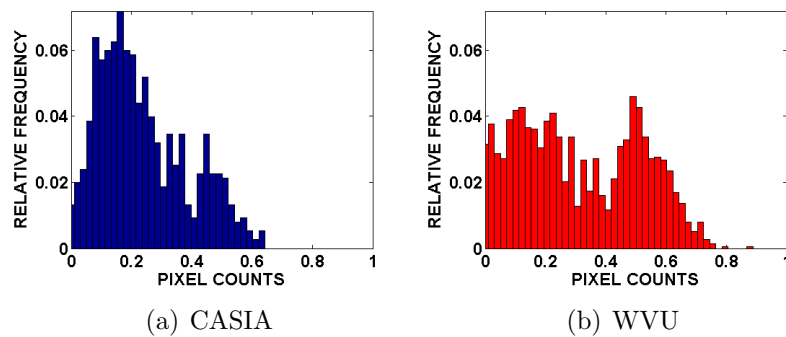
Figure 6.9: Pixel Count frequencies for CASIA and WVU

Fig. 6.9 are the distributions for our estimates on pixel counts. Similar to occlusion, both datasets suffer from this quality factor. CASIA has a mean pixel count score of 0.24 and WVU has mean pixel count score of 0.31.

6.2   CASIA and WVU Performance

Performance of biometric systems is typically characterized by error rates such FAR, FRR, and EER as described in Chapter 2. To evaluate the performance of our quality metric we divide CASIA and WVU datasets up into three intervals based on the lower quality bound. Interval one consists of the entire dataset. Interval two corresponds to those images pertaining to $quality \geq 0.75$ and interval three corresponds to those images pertaining to $quality \geq 0.85$. Equivalence classes are tallied up for all the images at each interval. If no class can be established, i.e only one image from a user, then no class is established hence it is removed. Performance is then evaluated at each interval for Gabor, PCA, and ICA encoding techniques.

We experiment with two different types of training and testing is performed for PCA and ICA for the CASIA dataset. In the first training scenario, the first template from each user is used to train while the remaining templates are used for testing. In the second experiment three templates from each user are used for training, while the rest are used for testing. For the WVU dataset, only the first scenario is performed for PCA and ICA encoding. The reason lying in the fact that the WVU dataset is large and including more training samples becomes computationally unfeasible. While reducing the size of the dataset would solve this problem, performance of just a subset leaves something to be desired considering the non-ideality of the data.

6.2.1 Gabor Encoding

Fig. 6.10 characterizes the performance at each interval. Based on the statistics in Table 6.3 we notice a performance increasing trend: as quality of the data increases so does performance. Table 6.3 lists EER, d', and mean quality values for each interval along with the number of images pertaining to each interval. The entire CASIA dataset (except for those images that failed rough segmentation) performs well, in general. However, by using our quality metric we are able to select images which can achieve the desired level of performance, with the last interval attaining an EER of 0.11 and dprime of 3.13.

(a) CASIA Performance Prediction

(b) WVU Performance Prediction
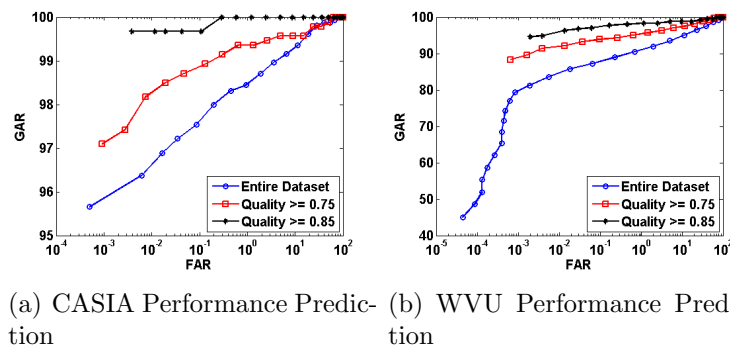
Figure 6.10: Verification Performance Prediction

With respect to the WVU results in Fig. 6.10 (b) and Table 6.4, we again see a performance increasing trend. The last interval achieves an EER of 1.17 and dprime of 3.53 which are comparable to what was achieved for CASIA.

Interval	EER %	Dprime	Quality	Image Count
All	1.30	2.63	0.79	738
Quality $\geq$ 0.75	0.63	2.79	0.85	556
Quality $\geq$ 0.85	0.11	3.13	0.89	273

Table 6.3: CASIA Results

Interval	EER %	Dprime	Quality	Image Count
All	5.07	2.53	0.65	2125
Quality $\geq$ 0.75	3.20	2.96	0.84	841
Quality $\geq$ 0.85	1.17	3.53	0.89	393

Table 6.4: WVU Results

6.2.2   Global PCA Encoding

Fig. 6.11 (a) and (b) represent the ROC curves for the PCA encoding technique using two different scenarios for the CASIA dataset. Similar to Gabor performance, we notice as quality interval increases so does PCA performance for both scenarios. Tables 6.5 and 6.6 list the performance statistics for both scenarios respectively. Fig. 6.11 (c) and Table 6.7 represents the ROC curves and statistics for the WVU data using scenario 1.

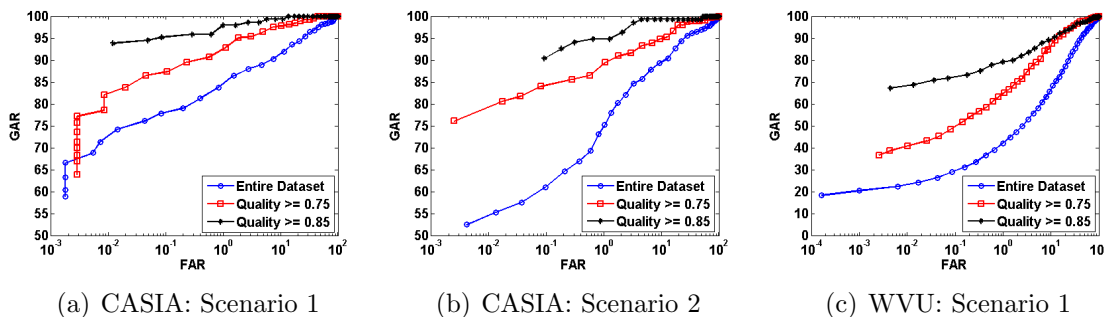| (a) CASIA: Scenario 1 | (b) CASIA: Scenario 2 | (c) WVU: Scenario 1 |

Figure 6.11: PCA Performance Prediction

For CASIA scenario 1, a maximum EER of 1.65 and dprime of 2.57 is achieved. In scenario 2 performance also increases with the quality, although not as acutely as scenario 1. This could result from the increase in training template, resulting in overtraining. For scenario 2, a maximum EER of 2.10 and dprime of 2.38 is achieved.

Interval	EER %	Dprime	Training	Testing
All	7.51	1.74	108	631
Quality $\geq$ 0.75	3.58	2.14	102	445
Quality $\geq$ 0.85	1.65	2.57	63	186

Table 6.5: CASIA: PCA Scenario 1

Interval	EER %	Dprime	Training	Testing
All	9.42	1.58	321	308
Quality $\geq$ 0.75	6.17	1.90	234	173
Quality $\geq$ 0.85	2.10	2.38	75	45

Table 6.6: CASIA: PCA Scenario 2

Interval	EER %	Dprime	Training	Testing
All	20.30	1.60	338	1787
Quality $\geq$ 0.75	11.03	2.22	202	580
Quality $\geq$ 0.85	9.05	2.33	101	220

Table 6.7: WVU: PCA Scenario 1

6.2.3 Global ICA Encoding

Fig. 6.12 (a) and (b) illustrate the ROC curves for the ICA encoding technique using two different scenarios for the CASIA dataset. Similar to Gabor and PCA performance, we notice as quality interval increases so does ICA performance for both scenarios. Tables 6.8 and 6.9 list the performance statistics for both scenarios respectively.

(a) CASIA: Scenario 1   (b) CASIA: Scenario 2   (c) WVU: Scenario 1
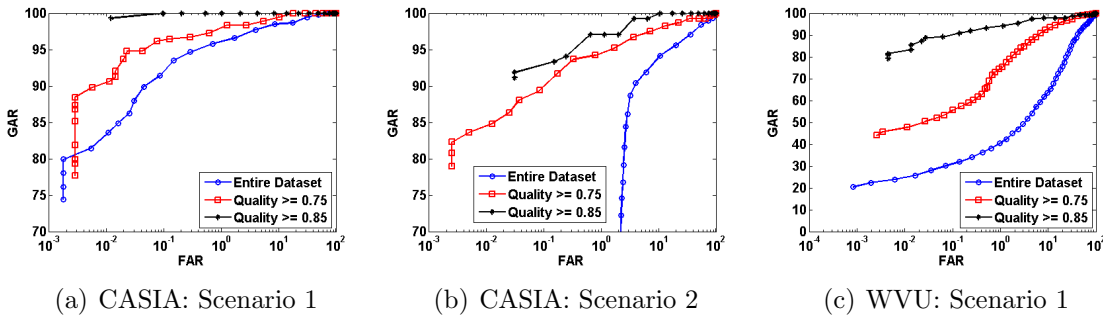
Figure 6.12: ICA Performance Prediction

For scenario 1, a maximum EER of 0.01 and dprime of 2.29 is achieved. In scenario 2 performance also increases with the quality, although not as acutely as scenario 1. We notice the same trend here as we did with PCA; incorporating more training samples reduces performance. For scenario 2, a maximum EER of 2.01 and dprime of 1.99 is achieved. Results for WVU are characterized in Fig. 6.12 (c) and Table 6.10.

Interval	EER %	Dprime	Training	Testing
All	2.29	1.91	108	631
Quality $\geq$ 0.75	1.28	2.23	102	445
Quality $\geq$ 0.85	0.01	2.68	63	186

Table 6.8: CASIA: ICA scenario 1

Interval	EER %	Dprime	Training	Testing
All	6.05	1.37	321	308
Quality $\geq$ 0.75	3.28	1.69	234	173
Quality $\geq$ 0.85	2.01	1.99	75	45

Table 6.9: CASIA: ICA scenario 2

Interval	EER %	Dprime	Training	Testing
All	21.78	1.59	338	1787
Quality $\geq$ 0.75	7.76	2.29	202	580
Quality $\geq$ 0.85	2.73	2.55	101	220

Table 6.10: WVU: ICA Scenario 1

Chapter 7

Conclusion & Future Work

7.1 Conclusions

Quality assessment is very important in biometric systems and image processing applications. In this work we studied the impact of seven factors such as defocus, motion, occlusion, off-angle, lighting variation, specular, and pixel counts on Gabor, PCA, and ICA encoding techniques. To get an understanding of how these factors affect performance and gain insight on estimating them, we took a subset of good quality images from CASIA and WVU datasets and synthetically degraded image quality. From these studies we concluded that defocus blur, motion blur, and off-angle significantly effect performance more than the other factors.

Next we estimate these factors given a single image. Defocus and motion blur are estimated using power based metrics. Off-angle is estimated by projectively transforming the image over roll and pitch angles to find a maximum of the integro-differential operator. A gradient based approach is adopted for occlusion estimation. Specular reflections are measured by hard-thresholding. To estimate lighting variation, the normalized iris region is divided up into four blocks. The variance of the means of each block is used for our measure of lighting variation. Finally pixel-counts is estimated as the ratio of occluded pixels to the number of estimated pixels. The estimated factors are then fused by use of Dempster-Shafer theory using Murphy's

rule to gain quality bounds (minimum and maximum values from fusing when using Murphy's rule) on image quality. Using the lower quality bound, we conducted further experiments on performance prediction for Gabor, PCA, and ICA encoding techniques. Our experiments showed that we are able to reliably predict recognition performance based on our quality metric for the three encoding techniques.

7.2 Future Work

Given the nature of this work, there are many areas that should be studied further such as:

1. Thorough study of the affect of quality factors on segmentation performance.

2. Evaluation of other quality factors such as pupil dilation and SNR.

3. Conduct further studies on Murphy's rule.

4. Evaluate other datasets.

5. Improve current estimation techniques.

The synthetic studies only tested performance on encoding block of an iris recognition system. It would be invaluable to conduct further experiments to see how these factors effect the segmentation block. Insight gained from this study could result in reliable non-ideal segmentation methodologies. Studying other quality factors is also of interest. The factors considered in this work are by no means exhaustive. Studying other factors may lead to better performance in terms of prediction. Following that, the further study of n parameter in Murphy's rule is of interest. Currently we are assigning all quality factors the same weight by leaving $n = 0.5$. Conducting research to evaluate the correlation between the different factors could prove useful in terms of predicting performance.

Bibliography

[1] J.G. Daugman, "How iris recognition works," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 21–30, January 2004.

[2] L. Masek, "Recognition of human iris patterns for biometric identification," B.S. Thesis, The University of Western Australia, http://www.csse.uwa.edu.au/pk/studentprojects/libor/, 2003.

[3] V. Dorairaj, Y.N. Shah, and N.A. Schmid, "Exploring and developing processing techniques for iris recognition system," September 2004.

[4] R.P. Wildes, "Iris recognition: An emerging biometric technology," *Proceedings of the IEEE*, vol. 85, no. 9, pp. 1348 – 1363, September 1997.

[5] Y. Du, B. Bonney, R. Ives, and D.M. Etter, "Analysis of partial iris recognition," in *Proc. of the SPIE 2005 Symp. on Defense and Security*, Orlando, FL., March 2005, vol. 5779, pp. 31–40.

[6] V. Dorairaj, N.A. Schmid, and G. Fahmy, "Performance evaluation of non-ideal iris based recognition system implementing global ica encoding," *Proc. of the IEEE International Conference on Image Processing (ICIP), Genova, Italy*, 2005.

[7] V. Dorairaj, N.A. Schmid, and G. Fahmy, "Performance evaluation of iris based recognition system implementing pca and ica techniques," *Proc. of the SPIE 2005 Symp. on Defense and Security*, pp. 51–58, March 2005.

[8] Z. Wang, A.C. Bovik, and L. Lu, "Why is image quality assessment so difficult," *Proc. of 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2002)*, pp. 3313–3316, May 2002.

[9] E. Tabassi, C. L. Wilson, and C. I. Watson, "Fingerprint image quality," Tech. Rep. NISTIR 7151, NIST, 100 Bureau Drive, Stop 8940, Gaithersburg, MD 20899, August 2004.

[10] M.Y. Yao, S. Pankanti, N. Haas, N. Ratha, and R. Bolle, "Quantifying quality: A case study in fingerprints," *Proc of IEEE Conference on Automatic Identification Advanced Technologies (AutoID)*, pp. 126–131, March 2002.

[11] B.J. Kang and K.R. Park, "A study on iris image restoration," *Proc. Audio and Video-Based Biometric Person Authentication*, vol. 3546, pp. 31–40, July 2005.

[12] L. Ma, T. Tan, Y. Wang, and D. Zhang, "Personal identification based on iris texture analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1519–1533, December 2003.

[13] G. Zhang and M. Salganicoff, "Method of measuring the focus of close-up image of eyes," Tech. Rep. 5953440, United States Patent, 1999.

[14] X. Zhu, Y. Liu, X. Ming, and Q. Cui, "A quality evaluation method of iris images sequence based on wavelet coefficients in 'region of interest'," *Proc. of the 4th Int'l Conf. on Computer and Information Technology*, pp. 24–27, September 2004.

[15] Y. Chen, S. Dass, and A. Jain, "Localized iris quality using 2-d wavelets," *Accepted for presentation at ICBA*, 2006, Hong Kong, China.

[16] "Casia iris image database (ver 1.0)," *http://www.sinobiometric.com/casiairis.html*.

[17] International Organization for Standardization and International Electrotechnical Comission, "Biometric data interchange formats - part 6: Iris image data," 19794-6, ISO/IEC, March 2004, Draft Version.

[18] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, 1976.

[19] R. Murphy, "Adaptive rule of combinations for observations over time," *Fusion and Integration for Intelligent Systems*, pp. 125 – 131, December 1996.

[20] R. Murphy, "Dempster-shafer theory for sensor fusion in autonomous mobile robots," *IEEE Trans. On Robotics and Automation*, vol. 14, no. 2, pp. 197–206, April 1998.

[21] M. Mladenovski, "Information fusion schemes for real time risk assessment in adaptive control systems," M.S. thesis, West Virginia University, 2004.

[22] L. Guo, *Software Quality and Reliability Prediction Using Dempster-Shafer Theory*, Ph.D. thesis, West Virginia University, Morgantown, 2004.

[23] A.K. Jain, R. Bolle, and S. Pankanti, *Biometrics Personal Identification in Networked Society*, Kluwer Academic Publishers, 1999.