Measurement of the information for identification in iris images

Tatyana Dembinsky

A thesis submitted to the

Faculty of Graduate and Postdoctoral Studies

In partial fulfilment of the requirements

For the M.A.Sc. degree in Electrical Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering

School of Information Technology and Engineering

University of Ottawa

To my family

The undersigned hereby recommend to

The faculty of Graduate Studies and Research

Acceptance of this thesis,

Measurement of the information for identification in iris images

Submitted by

Tatyana Dembinsky

In partial fulfillment of

the requirements for the degree of

Master of Applied Science

In

Electrical Engineering

Andy Adler, Ph.D.

(Thesis Supervisor)

List of Acronyms

Acronyms	Details
CCD	Charge Coupled Device
EER	Equal Error Rate
FAR	False Accept Rate
FER	Failure to Enroll Rate
FMR	False Match Rate
FNMR	False Non Match Rate
FRR	False Reject Rate
GUI	Graphical User Interface
HD	Hamming Distance
IR	Infrared
ISO	International Organization for Standardization
IQM	Image Quality Measure
NaN	Not-a-Number
NIR	Near Infrared
NIST	National Institute of Standards and Technology
NSTC	National Science and Technology Council
osFAR	Observer Sample False Accept Rate
osFRR	Observer Sample False Reject Rate
PCA	Principal Component Analysis
psFAR	Predicted Sample False Accept Rate
psFRR	Predicted Sample False Reject Rate
PSNR	Peak Signal to Noise Ratio
ROC	Receiver Operating Characteristic
SVD	Singular Value Decomposition
WED	Weighted Euclidian Distance

List of Symbols

Symbols	Details
$D(p \parallel q)$	Relative entropy
d'	Decidability
E	Expectation
F	Number of features per sample
$F_0(x)$	Cumulative of $f_0(x)$
$F_n(x)$	Cumulative of $f_n(x)$
f_0	Centre frequency of filter
$f_0(x)$	Density distribution of single-comparison match scores
$f_n(x)$	Density distribution of n-comparison match scores
G	Number of independent features per sample
G_σ	Gaussian smoothing function
H	Entropy
$\mathbf{I}$	Identity matrix
$I(x, y)$	Iris image in Cartesian coordinates
$I(\rho, \phi)$	Normalized iris image in polar coordinates
M	Length of template
$MS_{i,j}$	Match score between samples i and j
N	Degrees of freedom
N_f	Number of filters
N_p	Number of samples per individual
N_q	Number of samples in database
n	Number of comparisons during matching
p	Mean of binomial distribution
$p(i)$	Probability mass function
$p(x)$	Feature distribution of one individual
Q	Quality score of sample
$q(x)$	Feature distribution of population
r	Radius
r	Pearson correlation co-efficient
$\mathbf{S}_p$	Regularized $\mathbf{\Sigma}_p$
$\mathbf{S}_q$	Regularized $\mathbf{\Sigma}_q$
$\mathbf{U}$	SVD transformation matrix
x	Feature vector
Y	Template of length M
y	Vector of independent features

Symbols	Details
α	Log-Gabor multiplicative factor
λ_n	Filter base wavelength
μ_D	Mean of impostor distribution
μ_p	Feature mean of individual
μ_q	Feature mean of population
μ_S	Mean of genuine distribution
$\mathbf{\Sigma}_p$	Feature covariance of individual
$\mathbf{\Sigma}_q$	Feature covariance of population
σ	Filter bandwidth
σ	Standard deviation
σ_D	Standard deviation of impostor distribution
σ_S	Standard deviation of genuine distribution
χ	Biometric system features
ω	Wavelet frequency

Abstract

Performance of iris recognition systems strongly depends on uniqueness of irises. One possible way to estimate the uniqueness is to measure the amount of identifying information contained in irises.

No proper identification can be done if the quality of involved samples is poor. In our study, we verified human ability to assess iris sample quality and concluded that their intuitive quality assessment may not be sufficient for biometric identification applications.

After image quality is assured, the biometric information measurement may be addressed. We studied the approach proposed by Daugman that measures the amount of information in irises. Experiments and statistical analysis showed that it does not provide a useful way to measure an amount of identifying information.

An alternative algorithm based on relative entropy was developed to measure information important for identification. We used it to detect the most informative regions of the iris and to understand dependencies within it.

Acknowledgements

I would like to thank all the people without whom this thesis would not be possible.

I wish to thank my supervisor Dr. Andy Adler for his guidance throughout my studies, for his valuable ideas and encouragement. I appreciate all the time he spent helping me in my research and then reviewing this thesis.

I am greatful to Canada Border Services Agency for motivating this study and for allowing me to use their iris camera to build the database of iris images for my experiments. Especially, I would like to thank Terri Webster who dedicated a lot of her time assisting me in image acquisition sessions.

I am indebted to Jens Roesner whose detailed online instructions helped me to dissassemble my digital camera and then successfully turn it into the IR imaging device.

Also, I would like to thank Camille Gomez for proofreading this document, Yednek Asfaw for being able to answer any question, Richard Youmaran for his resourcefulness, Thierry Metais for his big smile and Lee Xie for the emotional support.

Lastly, and most importantly, I wish to thank my family, for their constant encouragement and love. To them I dedicate this thesis.

Contents

List of Figures

Chapter 1

Introduction

Biometrics is a field of automatic person identification based on physiological and behavioral characteristics of individuals. Such characteristics include faces, fingerprints, irises, signatures, voices, etc. One of main advantages of using biometric systems is the fact that users no longer have to remember passwords or PIN numbers. This solves many related issues such as lost or stolen access codes, by making forgery a much more difficult task. Additionally, biometric systems allow for an easier and faster matching of people while not involving extensive manpower resources. For instance, traveler clearance may be significantly facilitated by using electronic passports (Bucknam & Christensen, 2003) for verifying the holder's identity and matching them against databases of criminals within seconds.

The main focus of this thesis is a subfield of biometrics — iris recognition. Iris recognition systems identify people by analyzing patterns of their irises which are claimed to be unique for every individual and stable over long periods of time (Wildes, 1997). As explained in (Kronfeld, 1968) and (Mann, 1950), the general structure of irises is determined genetically while small details depend on external factors, such as initial conditions of the embryo development. Thus, it is highly unlikely for two irises to be formed in the identical way.

A typical iris recognition process includes the following stages. Firstly, infrared (IR) iris cameras are employed to obtain high resolution eye images. Then, their quality is verified with the help of human operators or image processing algorithms. If the quality is

unacceptable, the sample is rejected and the acquisition is repeated. The samples are then processed and encoded in order to build iris templates containing identifying information. Finally, templates are matched to calculate differences between them. Based on the degree of difference, the person is either accepted or rejected.

Numerous iris recognition algorithms have been developed over a period of almost twenty years. The most acknowledged approach is the one proposed by Daugman (1994) where the author reported nearly zero recognition error rates. Until recently no thorough independent tests were performed to support this claim. In May 2005, International Biometric Group released results of iris technology evaluation sponsored by the US Department of Homeland Security. The study involved evaluating modern iris recognition hardware and software in terms of match rates, enrollment and acquisition rates and usability. Daugman's matching algorithm showed high resistance to impostor matches with false match rates varying from "once in every several hundred thousand to once in every several million comparisons" (International Biometric Group, 2005).

1.1 Thesis objective

In the modern world, biometric security applications see an increasing level of interest. They may be found in airports used for border control, in government institutions used for surveillance, in forensics as a part of criminal capture programs and in many private areas as a means of achieving higher security standards. Therefore, it is very important to ensure that the biometric characteristic used in the system provides enough identification power and will not cause the system to mistake one person for another.

In order to answer the question "How secure is a particular biometric system?", it is necessary to assess the biometric template uniqueness. One of the possibilities is to measure the biometric information content, or the amount of information contained in a person's characteristic that allows distinguishing them out of the whole population. This challenge motivated our research that addressed measurement of information for identification in irises.

No proper identification can be done if the quality of an involved sample is poor.

Image processing of the sample will give less accurate results, leading to imprecise features. Measuring the amount of information using such images will normally show a lesser biometric information content compared to high quality samples (Youmaran & Adler, 2006). In order to be able to detect images of poor quality, human operators are often involved. In our study, we attempted to verify human ability to correctly assess iris sample quality. For this project, we asked several participants to evaluate the quality of given biometric samples and then compared the results to quality scores obtained automatically.

After image quality is assured, the particular identification system precision may be estimated by measuring information content of the involved biometric based on produced templates. One of the existing approaches, proposed in (Daugman, 2003), introduced a term of discrimination entropy for measuring the amount of information in irises. In order to clarify the technique, we implemented the algorithm and applied it to the local data set. Then, we performed statistical analysis to determine how appropriate this approach is for calculating biometric information in irises.

Another method, presented in (Adler *et al.*, 2006), suggested using relative entropy to measure biometric information important for identification. We applied this technique to the database of iris images, in order to learn the nature of information content in irises. We addressed this issue by measuring relative entropy of single features, whole irises and their various regions.

This thesis is organized as follows. Chapter 2 gives an overview of main theoretic aspects involved in our work. Chapter 3 addresses an issue of human ability to evaluate biometric sample quality. In chapter 4, we introduce an open source iris recognition software by Masek & Kovesi (2003) that was used as a base of our experiments. Chapter 5 describes experimental work done to analyze the discrimination entropy introduced by Daugman and presents results of studying iris information content using the algorithm of Adler *et al.*. Finally, in chapter 6, conclusions and suggestions for future work are stated.

Chapter 2

Background

This chapter introduces main theoretic aspects involved in our work. Section 2.1 talks about the field of biometrics in general - its types, applications and main definitions. Section 2.2 discusses in detail a subfield of biometrics and the focus of this work - iris recognition. Main stages of the iris recognition process are explained and existing techniques are reviewed. Section 2.3 presents an introduction to information theory covering its main principles. Ways of measuring biometric information content are discussed in section 2.4. Finally, section 2.5 discusses quality issues applied to person identification and verification.

2.1 Biometrics

Biometrics is a collection of methods used for automatic person identification based on their physiological or behavioral characteristics, such as fingerprint, face, iris, signature, etc.

The history of biometrics dates back at least 31,000 years when prehistoric men left their handprints around cave paintings to serve as author signatures (Renaghan, 1997). In the 1800s, rapidly growing population of cities created a need for a formal system for keeping criminal records. The proposed solution was the Bertillion system that involved cards for storing body dimensions of people with an option for record sorting by any measured parameter. Later, the system evolved into the index of fingerprints that provided more precise identification (Wayman, 2004a). Nowadays, biometric systems substitute passports, pass-

words and PIN numbers facilitating numerous procedures such as border control, access to secured areas and resources, etc. Users of biometric systems are no longer required to remember passwords or carry IDs thus ensuring that a secret sequence will not be forgotten and a document will not be forged or stolen.

There exist two modes of using biometrics - identification and verification, or negative and positive identifications respectively (Mansfield & Wayman, 2002). During the identification process a biometric sample of the person to be identified is compared against the database of all the individuals previously enrolled in the system. The identity of the person is resolved if the input sample sufficiently matches one of the existing records. In the case of verification, the system determines if the person is the one they claim to be. For this purpose a biometric sample of an individual is obtained and compared to the sample of a claimed identity stored in the database. Normally, the process of verification is preceded by an enrollment stage when a person biometric sample is processed and reposited in global storage.

Typically, a biometric system consists of the following five stages (NSTC, 2006): 1) using a sensor to collect a biometric sample and to digitize it; 2) applying signal processing techniques for evaluating data quality and building a biometric template containing features of interest; 3) storing a newly obtained template in the database for later reference; 4) matching the current template to templates from the data storage; 5) making a decision about the person's identity. Figure 2.1 presents a diagram of this process.

Performance of biometric systems is normally measured in terms of the following error rates:

- **False Accept Rate (FAR):** "The probability that a biometric system will incorrectly identify an individual, or will fail to reject an impostor" (Mak & Thieme, 2003). For verification systems, this can be calculated as $\frac{\texttt{number of falsely accepted people}}{\texttt{number of all impostor attempts}}$.
- **False Match Rate (FMR):** "The rate for incorrect positive matches by the matching algorithm for single template comparison attempts" (Mak & Thieme, 2003). If the biometric system uses only one comparison attempt for making a decision, FMR is

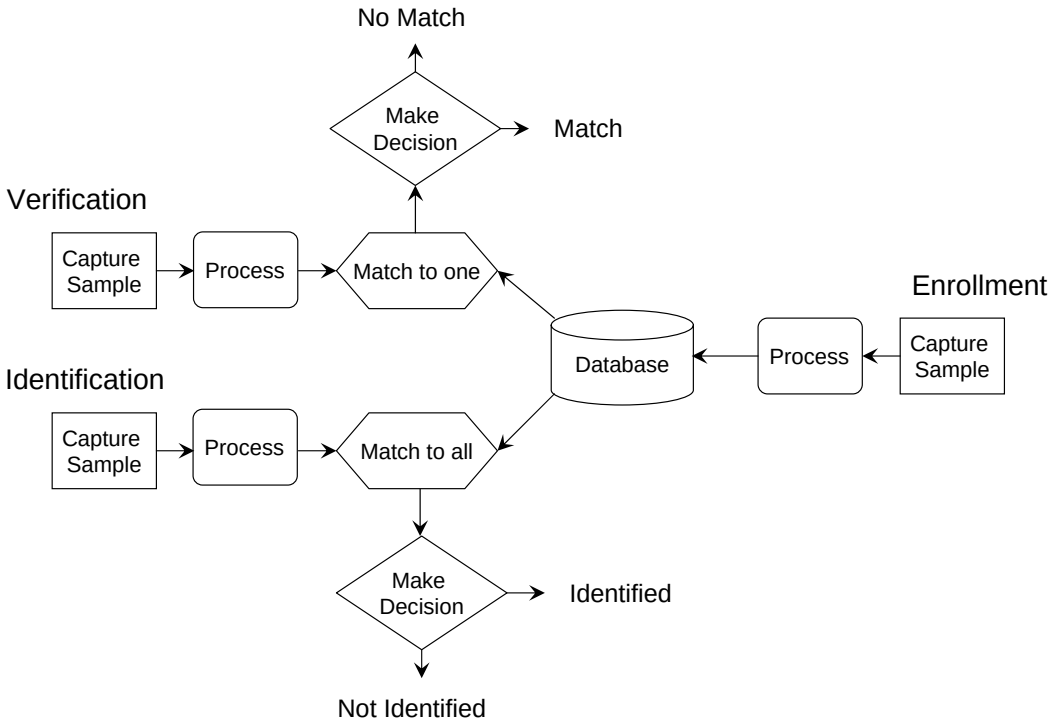


Figure 2.1: Components of a typical biometric system.

equal to FAR.

- **False Reject Rate (FRR):** "The probability that a biometric system will fail to identify a genuine enrollee" (Mak & Thieme, 2003). For verification systems, this can be calculated as $\frac{\texttt{number of falsely rejected people}}{\texttt{number of all genuine attempts}}$.

- **False Non Match Rate (FNMR):** "The rate for incorrect negative matches by the matching algorithm for single template comparison attempts" (Mak & Thieme, 2003). If the biometric system uses only one comparison attempt for making a decision, FNMR is equal to FRR.

- **Failure to Enroll Rate (FER):** "The proportion of the population for whom the system is unable to generate repeatable templates" (Mansfield & Wayman, 2002). As (Mansfield & Wayman, 2002) clarifies, this includes individuals who cannot provide

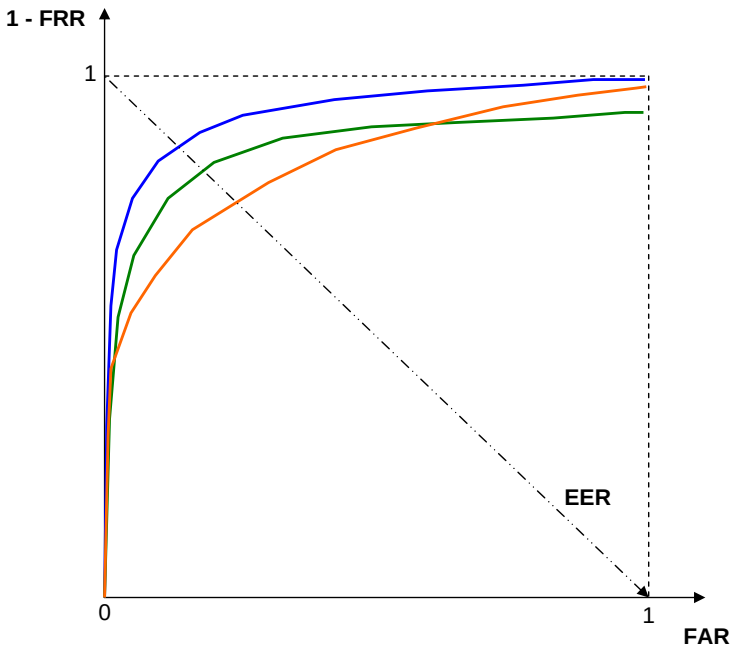


Figure 2.2: ROC curves of three systems demonstrating their FAR and FRR relationships. The top curve has the smallest EER value, thus, representing the most accurate system out of the three compared.

samples of sufficient quality for enrollment, who cannot present the feature at all and those whose samples fail to match enrolled ones.

In real-world applications a trade-off between FAR and FRR is normally used in order to achieve a desired performance. To analyze the relationship between these error rates, Receiver Operating Characteristics (ROC) curves are used, as shown in figure 2.2. The point of the ROC curve corresponding to equal FAR and FRR, or the equal error rate (EER), is typically selected to compare different identification systems, with lower values representing more accurate systems. In most cases EER is not a practical decision threshold. However, it is easy to understand and for this reason is widely used for comparisons (Waymann *et al.*, 2005). Figure 2.2 illustrates this process by showing three ROC curves that correspond to three different recognition systems. The top curve has the smallest EER value, thus, representing the most accurate system out of the three compared.

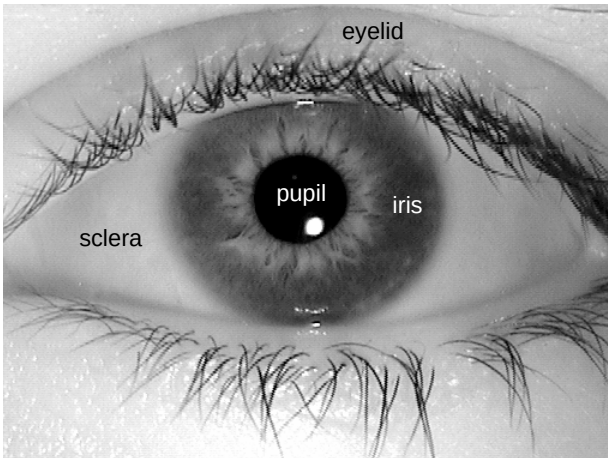


Figure 2.3: Human eye components.

2.2 Iris recognition

Iris recognition refers to identifying people by analyzing patterns of their irises. According to (Wildes, 1997), the iris "is a thin diaphragm stretching across the anterior portion of the eye". It is attached to the ciliary body of the eye and "opens into the pupil" (Wildes, 1997) (see figure 2.3). The iris is comprised of a number of tissue layers which contribute to the richness of its pattern, such as pigment cells, blood vessels and muscles. Moreover, it is believed that the iris is unique for every individual (Wildes, 1997). (Kronfeld, 1968) and (Mann, 1950) explained that even though the general structure of irises strongly depends on genetics, numerous small details are determined by external factors such as initial conditions of the embryo development. Thus, it is highly unlikely for two irises to be formed in the identical way. For instance, right and left eyes of the same person, as well as eyes of identical twins, do not show much correlation in their iris patterns (Adler, 1965).

Due to the iris pattern uniqueness, iris recognition is considered to be one of the most accurate biometric technologies. A number of other characteristics make it even more attractive for recognition systems:

- **External visibility.** Under normal conditions, the iris is a visible organ of the human body. Therefore, taking pictures of it is easy and does not raise ethical concerns. Moreover, irises may be used for surveillance purposes when advanced technologies are

involved to acquire high-resolution images from a distance.

- **Natural protection.** Eyes have a natural protection mechanism in the form of eyelids and tears that prevent this fragile organ from being damaged and keep it in a proper condition. Thus, at the moment of sample acquisition, condition related problems rarely occur. Eye irritations, caused by allergies and fatigue, normally do not worsen recognition, since they affect sclera only.

- **Stability over long periods of time.** It is clinically observed that irises do not change significantly over time (Wildes, 1997). This characteristic is particularly important for person identification because it allows recognizing an individual, who was once enrolled in the system, at any moment of their life. As an illustration, we can mention a famous story of an Afghan girl who was identified by comparing irises from her two photographs - one taken in 1984, another taken in 2002 (Daugman, 2002).

At the moment, the most well known iris recognition algorithm is the one proposed by Daugman (1994). Other notable systems were designed by Wildes (1997), Boles & Boashash (1998), and Lim *et al.* (2001). In following sections we introduce components that are common to most iris recognition system, using different existing approaches as examples.

2.2.1 Capturing a sample

The first, and a very important module of an iris recognition system, is capturing a sample. In the case of a low quality sample, any of subsequent stages may fail resulting in a faulty identification. It is commonly agreed that, in order to extract a sufficient amount of detail from the iris, its radius in the eye image should be no less than 70 pixels (Daugman, 2004). Monochrome CCD cameras with near infrared (NIR) illumination are used for this purpose. In order to remain noninvasive to humans and to minimize specular reflections, the illumination range is typically 700-900nm. To insure that participants position their eyes properly (centered and with the least possible tilt) a mirror or a video feedback is provided.

2.2.2 Iris segmentation

The first processing stage of the eye image obtained by the sensor is localization of the iris region. Normally, the process involves approximation of the pupil and the iris outer boundaries by two circles, detection of eyelids and eyelashes occluding the iris and removal of specular reflections. Commonly used techniques for determining iris location are 1) integrodifferential operator by Daugman and 2) Hough transform initially used by Wildes (1997).

Daugman (2004) describes the detection operator by means of the following equation:

$$\max_{(r,x_0,y_0)} \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r,x_0,y_0} \frac{I(x,y)}{2\pi r} ds \right| \tag{2.1}$$

where $I(x,y)$ is a monochrome image of an eye, r is a search radius, G_σ is a Gaussian smoothing function of scale σ and s is a circular arc defined by the radius r and a center (x_0, y_0). For each iteratively selected (x_0, y_0), a series of concentric circles with growing radiuses r is defined. The average intensity is then calculated along each circular contour. The point where the average intensity changes in the most significant way from one circle to another is considered to define a searched set of parameters. The blur function strength is gradually reduced in order to achieve higher precision. Firstly, the operator is applied to find the pupilary contour and the outer boundary of the iris. Then, the same procedure is used to localize eyelids by using an arc path instead of a circular one. Figure 2.4 shows an example of iris localization as reported in (Daugman, 2004).

The Hough transform is a classical digital image processing technique used for detecting lines in the image. Its more popular extension used in many imaging applications identifies locations of arbitrarily defined shapes. Many iris recognition algorithms, such as the ones proposed by Wildes (1997) and Ma *et al.* (2002), exploit the Hough transform for detecting components of the eye.

At the beginning, one of many existing edge detecting algorithms is used to build an edge map. Then, as described in (Wildes, 1997), for every set of parameters defining a curve that intersects at least one point on the edge map the Hough value is calculated. After that,

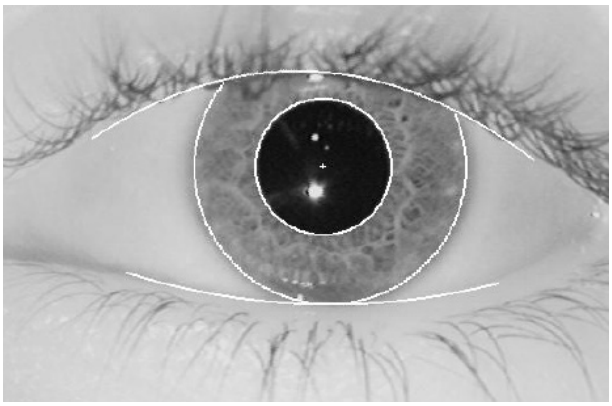


Figure 2.4: Iris localization and eyelid detection as performed by the algorithm of Daugman (reproduced from (Daugman, 2004)).

the values are collected to build a Hough space where the highest value corresponds to a curve that intersects the biggest amount of edge points. The set of parameters related to the maximum of the space is considered to define the best searched approximation. In the case of iris extraction, the approximated curve is a circle defined by a radius r and a center (x_c, y_c):

$$x_c{}^2 + y_c{}^2 - r^2 = 0 \tag{2.2}$$

Wildes uses a similar approach to detect eyelids. The difference is that in place of a circular curve the arc defined by the following equation is used:

$$(-(x - h_j)\sin\theta_j + (y - k_j)\cos\theta_j)^2 = a_j((x - h_j)\cos\theta_j + (y - k_j)\sin\theta_j) \tag{2.3}$$

where (h_j, k_j) is a parabola peak, a_j is a curvature parameter and θ_j is a rotation angle with respect to the x-axis.

2.2.3 Iris normalization

After the iris region is extracted, the normalization stage is necessary to eliminate dimensional inconsistencies between irises. These can be differences in iris magnifications, caused by changes in distance between the eye and the image capturing device; differences in pupil sizes, caused by variations in illumination; orientation changes resulting from head tilts, eye

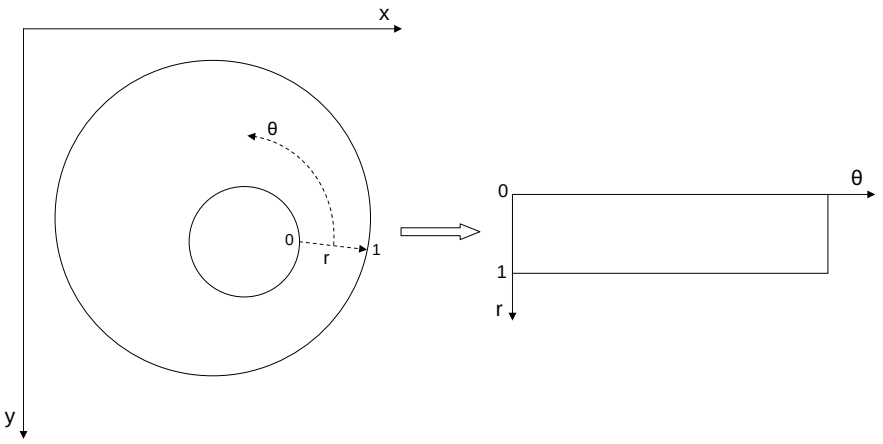


Figure 2.5: Iris normalization process: converting iris region from Cartesian coordinates into a polar representation (adapted from (Masek, 2003)).

or camera rotations. Normalization is needed to obtain two consistent sets of features from two dimensionally different eye images of the same person.

Daugman approaches this problem by suggesting a homogeneous rubber sheet model as seen in figure 2.5 adapted from (Masek, 2003). In this model each point of the iris (x, y) is transformed into a pair of dimensionless non-concentric coordinates (r, θ) where r is a value between 0 and 1 and θ is an angle between 0 and 2π. Daugman explains the remapping by means of the following formula:

$$I(x(r,\theta), y(r,\theta)) \rightarrow I(r,\theta) \tag{2.4}$$

where

$$\begin{aligned} x(r,\theta) &= (1-r)x_p(\theta) + rx_i(\theta) \\ y(r,\theta) &= (1-r)y_p(\theta) + ry_i(\theta) \end{aligned} \tag{2.5}$$

with $I(x, y)$ representing the iris image, and (x_p, y_p) and (x_i, y_i) being coordinates on the pupil and outer iris boundaries along the angle direction θ. The model takes into account inconsistencies in the iris size, pupil dilations and its non-concentricity, resembling a rubber

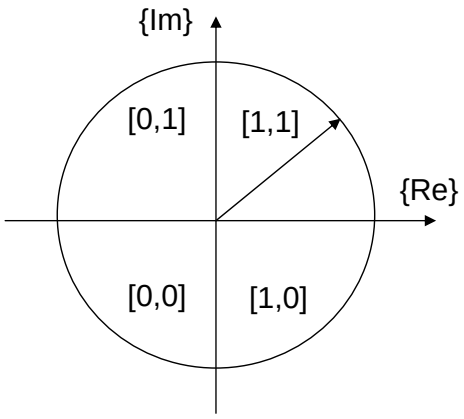


Figure 2.6: Quantization of the angle of each phasor to one of the four quadrants resulting in two bits of information (adapted from (Daugman, 2004)).

sheet that is connected to the iris outer border and is fixed at the center of the pupil. However, this presentation is not invariant to the iris orientation. The solution to orientation problems given by Daugman is described further in this chapter.

2.2.4 Feature encoding

The following stage of iris recognition performs extraction of iris features and building of templates for further comparison. In order to achieve the best identification result, the encoding algorithm should extract the most discriminating information out of the iris. Typically, techniques involving band-pass decomposition are used for building a biometric template (Masek, 2003).

Daugman in his work suggested using 2D Gabor wavelets for extracting localized frequency information. Firstly, complex-valued coefficients that represent phasors in a complex plane are calculated. Then, depending on their location, vector angles are quantized to one of the four segments of a unit circle (see figure 2.6 adapted from (Daugman, 2004)) contributing two bits into the template each.

This encoding procedure, involving both the demodulation and quantization, is de-

fined by Daugman as follows

$$h_{\{Re,Im\}} = \mathrm{sgn}_{\{Re,Im\}} \int_{\rho}\int_{\phi} I(\rho,\phi) e^{-i\omega(\theta_0-\phi)} e^{-(r_0-\rho)^2/\alpha^2} e^{-(\theta_0-\phi)^2/\beta^2} \rho d\rho d\phi \tag{2.6}$$

where $I(\rho, \phi)$ is an iris image in polar coordinates as obtained from the previous stage, (r_0, θ_0) are coordinates of the point for which the encoding is applied, α and β are size parameters of the multi-scale 2D wavelet and ω is a wavelet frequency. The resulting value $h_{\{Re,Im\}}$ corresponds to quantized real and imaginary parts of a phasor. After subsequent application of demodulation to every point of the normalized iris, 2048 bits of phase information are obtained. In parallel, the mask of iris occlusions is built, in order to localize and later ignore invalid regions during the matching process.

An alternative approach, described in (Field, 1987) and selected by Masek (2003) for his iris recognition system implementation, suggests using Log-Gabor filter which is Gaussian on a logarithmic scale with the following frequency response:

$$G(f) = \exp\left(\frac{-(\log(f/f_0))^2}{2(\log(\sigma/f_0))^2}\right) \tag{2.7}$$

where f_0 is a centre frequency of the filter and σ is its bandwidth.

2.2.5 Pattern matching

The final module of an iris recognition system is responsible for matching two iris templates. Its purpose is to measure how different the templates are and to decide whether they belong to the same individual or not.

The method designed by Daugman chooses using the Hamming Distance (HD) approach which is generally defined as

$$HD = \frac{1}{M}\sum_{j=1}^{M} X_j \otimes Y_j \tag{2.8}$$

and calculates the amount of different bits in binary sequences X and Y over total number

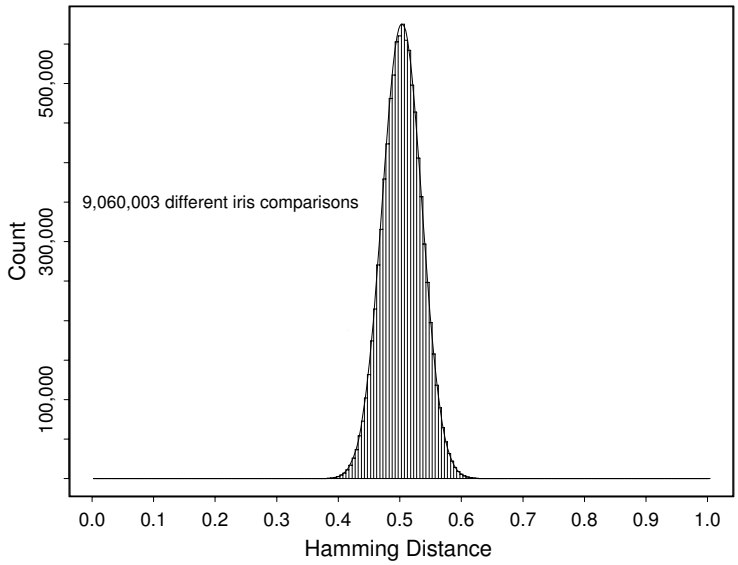


Figure 2.7: Distribution of Hamming Distances from 9.1 million non-genuine comparisons calculated using the algorithm by Daugman (adapted from (Daugman, 2004)).

of M bits.

Daugman extends this approach for iris templates:

$$HD = \frac{\| \ (templateA \otimes templateB) \cap maskA \cap maskB \ \|}{\| \ maskA \cap maskB \ \|} \tag{2.9}$$

where $templateA$ and $templateB$ are two iris templates being matched and $maskA$ and $maskB$ are corresponding masks specifying which bits of the templates belong to valid iris regions (not occluded with eyelids or eyelashes and not having specular reflections). In this way, a fractional HD is calculated with a value 0 for perfectly matching templates. Due to independence of templates from different irises, it is expected to obtain one half of bits that match and another half that do not match for any two irises pertaining to different eyes. An experimental justification of this theory may be seen in figure 2.7 adapted from (Daugman, 2004). It may be inferred that it is extremely improbable that less than a third of template bits will disagree if compared irises are different.

In addition to steps described above, Daugman modified the matching stage to compensate for possible iris rotations, which occur due to head tilts. To do this, one of the

compared iris templates is scrolled in a cyclic way. It may be easily shown that scrolling the template in polar coordinates is equivalent to iris rotation in Cartesian coordinates. The discussed algorithm performs matching of two templates several times while shifting one of them to n different locations. Finally, the smallest HD value is selected to be the matching result.

An alternative approach for matching templates, proposed by Zhu *et al.*, involved using the weighted Euclidian distance (WED):

$$WED(k) = \sum_{i=1}^{M} \frac{\left(f_i - {f_i}^{(k)}\right)^2}{\left({\delta_i}^{(k)}\right)^2} \tag{2.10}$$

where k is an iris that has several templates of size M in the database, f_i is the i^{th} bit in the unidentified template, ${f_i}^{(k)}$ and ${\delta_i}^{(k)}$ is the i^{th} bit in the database template and its standard deviation for the iris k.

2.3 Information theory

Information theory is a branch of mathematics founded by Shannon in the 1940s. The main target of this discipline is to define and measure information. Cover & Thomas list a number of fundamental questions from different fields that are addressed by the theory. Among them are: "What is maximal possible data compression and the biggest transmission rate of the channel?" (communication theory), "How can we characterize sequences of random variables? What are the estimates for rare event probabilities?" (mathematics) and other issues related to physics, economics, computer science, etc. Main applications are shown in figure 2.8 adapted from (Cover & Thomas, 1991). In the following sections we introduce fundamental concepts of information theory: entropy and relative entropy.

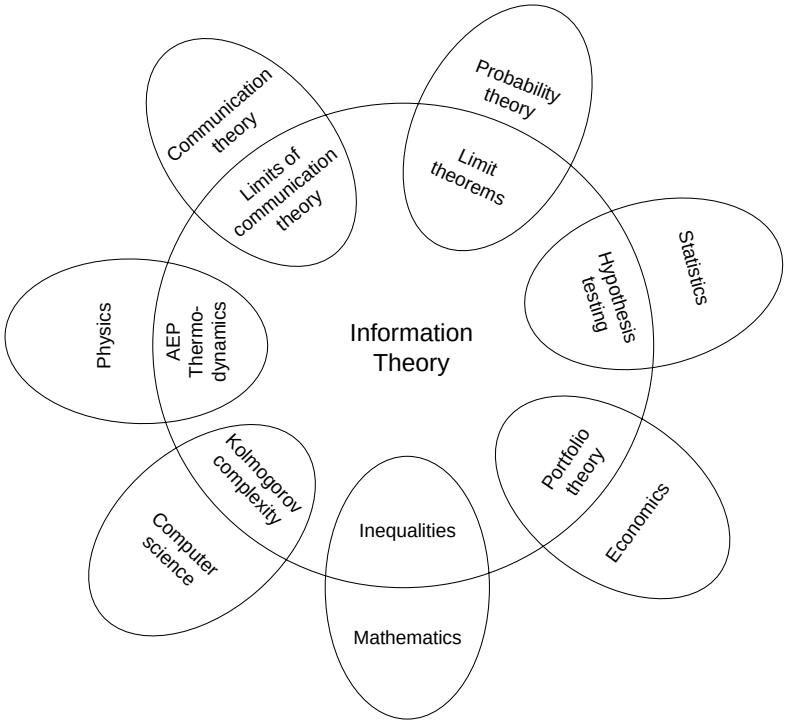


Figure 2.8: The relationship of information theory with other fields (adapted from (Cover & Thomas, 1991)).

2.3.1 Entropy

Entropy is a measure of randomness (or uncertainty) in a random event. This term was introduced by Shannon & Weaver (1949) and defined in the following way:

$$H(X) = -\sum_{i=1}^{n} p(i) \log_2 p(i) \tag{2.11}$$

where X is a discrete random variable with possible outcomes $1..n$, $p(i)$ is the probability mass function of X:

$$p(i) = Pr\{X = i\}, \quad i \in [1..n]. \tag{2.12}$$

When applied to communication theory, entropy gives a minimal channel capacity needed for a reliable transmission of an encoded information source. In other words, it is used to measure an amount of information in the message, or an average of bits needed to encode one symbol which is not necessarily an integer number (Cover & Thomas, 1991). For instance,

it was computed experimentally that entropy of an English text is between 1.1 and 1.6 bits of information per character depending on a text source (Shannon & Weaver, 1949).

It must be noted that entropy is a quantity defined in terms of probabilities of random variables and not their actual values. As an example, we may consider a source that repetitively generates only one character A. Its entropy will be equal to 0 since every next character is known in advance.

As explained in (Cover & Thomas, 1991), entropy satisfies a number of properties:

- Entropy value is always non-negative:

$$H(X) \geq 0 \tag{2.13}$$

- Events having zero probability do not contribute into final entropy value
- Uniform distribution of outcome probabilities maximizes entropy

$$H(p_1, p_2, ..., p_n) \leq H(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}) \tag{2.14}$$

2.3.2 Relative entropy

The relative entropy, or Kullback Leibler divergence, $D(p \parallel q)$ is a measure of distance from a true probability distribution p to an assumed distribution q (Cover & Thomas, 1991). In other words, it measures inefficiency of assuming that q takes place and not p. Typically, p represents precise calculation results or observations, whereas q is a theoretical model or an approximation. According to (Kullback & Leibler, 1951), the relative entropy is defined as

$$D(p \parallel q) = \sum_i p(i) \log_2 \frac{p(i)}{q(i)} \tag{2.15}$$

for discrete random variables and

$$D(p \parallel q) = \int p(x) \log_2 \frac{p(x)}{q(x)} dx \tag{2.16}$$

for the continuous case.

Main properties of relative entropy are given in (Kullback & Leibler, 1951) as:

- $D(p \parallel q)$ is often referred as a distance. However, it is not entirely true since the relative entropy is not symmetric and does not satisfy the triangle inequality $|x+y| \leq |x|+|y|$.
- The relative entropy is always non-negative with its value equal to zero if and only if $p = q$.

2.4 Measuring biometric information

Measuring the amount of information in a biometric, such as face, iris or fingerprint, is an issue gradually gaining interest because it can answer a number of questions related to security strength. First of all, there has always been an interest in finding out how unique different biometric types are, which ones are more reliable than others and under what conditions. Secondly, performance of different methods and technologies may be evaluated based on knowledge of information content of biometric samples and templates they produce. Finally, it becomes possible to say whether combining different biometric systems (as discussed in (Ross & Jain, 2003)) results in higher information content.

In an attempt to answer these questions several approaches were developed. Wayman (2004b) proposed a "cotton ball model" to measure separability of Gaussian feature distributions. Daugman (2003) suggested using discrimination entropy to calculate the information content of the iris by analyzing distributions of matching scores. The recently developed method by Adler *et al.* (2006) approaches measuring the amount of biometric information by using the concept of relative entropy. In the following sections we present a detailed overview of the latter two methods: discrimination entropy is introduced in section 2.4.1 and the relative entropy approach is explained in section 2.4.2.

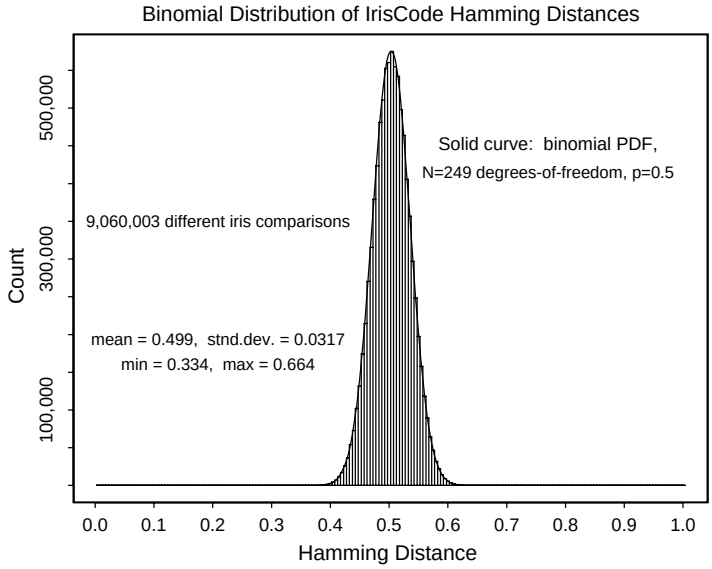


Figure 2.9: Distribution of impostor match scores forms a binomial distribution with 249 degrees of freedom and mean equal to 0.5 (reproduced from (Daugman, 2003)).

2.4.1 Discrimination entropy measure

In the approach described in (Daugman, 2003), the author addressed measuring the amount of information in the iris by analyzing the distribution of impostor match scores (the genuine distribution is not considered in this analysis). For this purpose, a database of 4,258 different iris images was used. It was acquired using the identification system designed by Daugman (2003) in different countries, including UK, USA, Japan and Korea. To obtain the HD distribution, the author performed 9.1 million comparisons between samples of different eyes. The analysis of the resulting plot showed the mean $p = 0.499$ and the standard deviation $\sigma = 0.0317$, as can be seen in figure 2.9 reproduced from (Daugman, 2003). The author then assumed that the data behaves as the binomial distribution and calculated its degrees of freedom:

$$N = p(1-p)/\sigma^2 = 249 \tag{2.17}$$

The curve of binomial distribution, described by the following fractional functional form

$$f(x) = \frac{N!}{m!(N-m)!} p^m (1-p)^{N-m} \tag{2.18}$$

with $x = m/N$, $p = 0.499$ and $N = 249$, revealed a close fit to the matching data. Daugman explained this by the fact that each comparison of two bits of iris templates may be seen as a Bernoulli trial, since a) irises are uncorrelated and b) every bit of the phase code has equal probability of being 1 or 0.

By analyzing the form of match score distribution, the author concluded that amount of difference between iris templates is distributed in the same manner as runs of 249 coin tosses, which is equivalent to saying that the iris contains 249 bits of information. This type of information was called by Daugman a discrimination entropy. Using typical dimensions of the eye, $r_{iris} = 5.5$mm and $r_{pupil} = 2.5$mm, the author calculated information density of the eye:

$$\frac{249}{iris\ area} = \frac{249}{\pi {r^2}_{iris} - \pi {r^2}_{pupil}} \approx 3.2 \text{ bits/mm}^2 \tag{2.19}$$

Another way to interpret this, as was explained by Daugman, is to say that the informativeness of the iris changes with the rate of 3.2 bits per mm of visible iris as person opens their eye more or less. The likelihood of any other person's iris matching this portion of the iris would be approximately the same as tossing a fair coin a number of times equal to 3.2 bits multiplied by the visible iris area and getting a fraction of "heads" that is the same as the HD when two templates are compared.

The previously done derivations were based on the assumption that iris templates are compared only once during the matching process. For the case of the modified matching algorithm, where two iris templates are compared several times for different iris orientations in order to find the best HD, Daugman (2003) proposed the following adjustments. If $f_0(x)$ is a density distribution of match scores obtained from single comparisons, as described in (2.18), then the cumulative of $f_0(x)$, $F_0(x)$, is a probability of getting a false match when the acceptance threshold is x:

$$F_0(x) = \int_0^x f_0(x)dx \tag{2.20}$$

Equivalently, $f_0(x)$ may be defined in terms of $F_0(x)$ as

$$f_0(x) = \frac{d}{dx} F_0(x) \tag{2.21}$$

Thus, the probability of not having a false match in one trial is $1 - F_0(x)$ and after n independent trials $(1 - F_0(x))^n$. Consequently, the probability of a false match after n tests using the acceptance threshold x is

$$F_n(x) = 1 - [1 - F_0(x)]^n \tag{2.22}$$

Finally, the corresponding density function may be written as

$$f_n(x) = \frac{d}{dx} F_n(x) = n f_0(x)[1 - F_0(x)]^{n-1} \tag{2.23}$$

Experimental measurements were performed for the same iris database that was used for the experiment showed in figure 2.9. 9.1 million pairs of templates belonging to different irises were compared by calculating the HD for 7 different rotations, in order to select the optimal match value. Finally, the best selected match scores were plotted as shown in figure 2.10 reproduced from (Daugman, 2003). As a result of multiple comparisons, the final distribution shifted towards a lower mean value $p = 0.458$. A curve defined by the theoretically derived formula (2.23) was then plotted using the same parameter for degrees of freedom, $N = 249$, and showed perfect fit to the experimental data.

2.4.2 Relative entropy measure

Adler *et al.* (2006) approached the problem of measuring the information content by, firstly, defining the biometric information as "the decrease in uncertainty about the identity of a person due to a set of biometric measurements". In other words, initially, the person is a part of the population and can be anybody. However, after biometric measurements are acquired, the information for identification is available and the uncertainty decreases.

As the approach motivation, Adler *et al.* presented the following example. Suppose, there is a biometric system identifying people based on their height and weight. Also assume that the height and the weight of the population are uniformly and independently distributed between 100-200 cm and 100-200 lb respectively. Since people's features are not perfectly

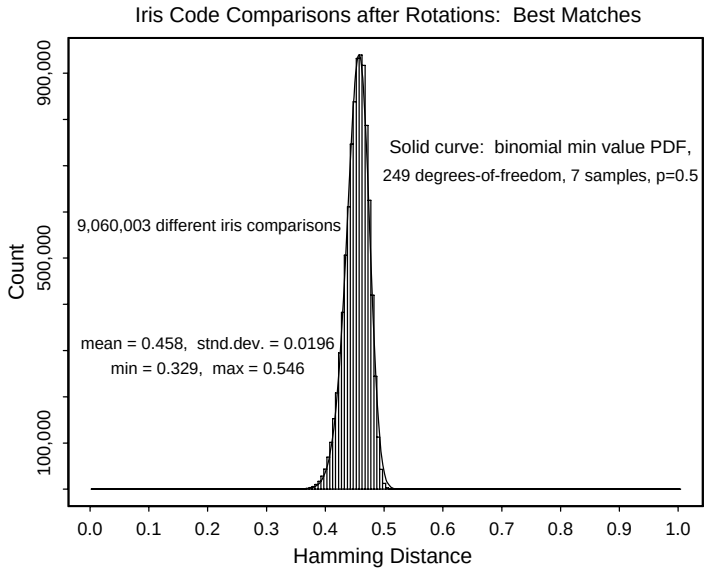


Figure 2.10: Distribution of iris match scores for the matching algorithm that employs 7 comparisons in order to find the best one (reproduced from (Daugman, 2003)).

stable and the measuring devices are not perfectly precise, it is impossible to measure the discussed characteristics with infinite accuracy and to identify a person uniquely. Let's say that observed errors are ±5 cm and ±5 lb. The system can resolve this by simply quantizing measured values to the grid [105, 115, ... , 195]. Thus, each feature can take on 10×10 equiprobable values, resulting in the following entropy (Shannon & Weaver, 1949):

$$H(X) = -\sum_{i=1}^{n} p(i) \log_2 p(i) = -\sum_{i=1}^{100} \frac{1}{100} \log_2 \frac{1}{100} = \log_2 100 \approx 6.6 \text{ bits} \tag{2.24}$$

From this, the authors concluded that biometric information strongly depends on the choice of features. Therefore, it is only possible to answer how much information provides a certain set of features and not the biometric characteristic itself. Additionally, redundant features shouldn't increase the overall information content (i.e. same data presented in different units).

The approach described in this section takes the mentioned aspects into account and develops a method that allows calculating the biometric information for a particular system. Attempts were made to stabilize the algorithm, in order to make it suitable for cases of

limited amounts of user samples.

(1) Theoretical measure selection.

Adler *et al.* selected relative entropy, $D(p \parallel q)$, introduced in section 2.4.2, as the most appropriate measure of biometric information:

$$D(p \parallel q) = \int_{\chi} p(x) \log_2 \frac{p(x)}{q(x)} dx \tag{2.25}$$

where χ is a set of all feature dimensions, $p(x)$ is the feature distribution for one individual, or intra-person distribution, and $q(x)$ is the feature distribution in the whole population, or inter-person distribution.

The relative entropy $D(p \parallel q)$ was preferred over the entropy $H(p)$ since it allows to determine the amount of information distinguishing one person, described by the distribution $p(x)$, from the total population, described by the distribution $q(x)$. Whereas, $H(p)$ reflects the biometric information contributed by all possible features whether they help in identification or not. This is the main difference between the $D(p \parallel q)$ based approach of Adler *et al.* and the technique of Daugman who measures iris information based on inter-person comparisons only without taking into account intra-person distributions.

(2) Feature distribution model.

Adler *et al.* present in their work the following model. A typical biometric system acquires N_p biometric samples for each person and then calculates a vector x of F features for each sample. The number of samples of all the users in the system is denoted as N_q. If a vector x is a value of the random variable X, then the population feature mean μ_q is defined as a vector

$$\mu_q = E_q[X] = \frac{1}{N_q} \sum_{i=1}^{N_q} x_i \tag{2.26}$$

with the individual feature mean μ_p defined in the similar way. The population feature covariances, $\mathbf{\Sigma}_q$, is then written as

$$\mathbf{\Sigma}_q = E_q \left[(X - \mu_q)^t (X - \mu_q) \right] = \frac{1}{N_q - 1} \sum_{i=1}^{N_q} (x_i - \mu_q)^t (x_i - \mu_q) \tag{2.27}$$

where μ_q and μ_p are vectors of size $F \times 1$ and feature covariances are matrices of size $F \times F$. The individual feature covariances $\mathbf{\Sigma}_p$ is determined similarly.

In order to avoid instabilities of $\log_2(p(x)/q(x))$ occurring due to insufficient data, the authors decided to approximate intra- and inter-personal feature distributions using a Gaussian model. It is commonly used for approximating real-world distributions and is convenient for analytical derivations. Based on this model, Adler *et al.* (2006) defined the population feature distribution $q(x)$ as

$$q(x) = \frac{1}{\sqrt{|2\pi\mathbf{\Sigma}_q|}} \exp\left(\frac{1}{2}(x-\mu_q)^t {\mathbf{\Sigma}_q}^{-1}(x-\mu_q)\right) \tag{2.28}$$

and the individual feature distribution $p(x)$ in the same way substituting q with p. Finally, the relative entropy $D(p \parallel q)$ was derived

$$\begin{aligned} D(p \parallel q) &= \int p(x)(\log_2(p(x)) - \log_2(q(x)))dx \\ &= -k\left(\ln|2\pi\mathbf{\Sigma}_p| + \mathbf{I} - \ln|2\pi\mathbf{\Sigma}_q| - E_p\left[(x-\mu_q)^t {\mathbf{\Sigma}_q}^{-1}(x-\mu_q)\right]\right) \\ &= k\left(\ln\frac{|2\pi\mathbf{\Sigma}_q|}{|2\pi\mathbf{\Sigma}_p|} + trace\left((\mathbf{\Sigma}_p + \mathbf{T}){\mathbf{\Sigma}_q}^{-1} - \mathbf{I}\right)\right) \end{aligned} \tag{2.29}$$

where $\mathbf{T} = (\mu_p - \mu_q)^t(\mu_p - \mu_q)$ and $k = \log_2\sqrt{e}$.

(3) Resolving numerical instability.

Adler *et al.* explained that the algorithm instability is caused by the fact that in the set of F features some of them strongly correlate. Therefore, in order to eliminate the problem, it is necessary to find a way for extracting G uncorrelated features ($G \leq F$). As a solution, the authors chose to use principal component analysis (PCA) (Draper *et al.*, 2003). It allowed creating a mapping from the original feature space X of size F to the new space of independent features Y of size G:

$$\mathbf{U}^t : X \to Y \tag{2.30}$$

The mapping was done using the Singular Value Decomposition (SVD) (Alter *et al.*, 2000) of the covariance matrix of the population:

$$\mathbf{U}\mathbf{S}_q\mathbf{U}^t = svd(cov(X)) = svd(\mathbf{\Sigma}_q) \tag{2.31}$$

Adler *et al.* decided to apply PCA to the covariance matrix of the population, $\mathbf{\Sigma}_q$, since it is built based on a sufficient amount of biometric samples and is a better estimate than $\mathbf{\Sigma}_p$. The authors used $\mathbf{S}_q$ as an indicator of feature uniqueness: the smaller $[\mathbf{S}_q]_{j,j}$ is, the less significant is the feature j in the set. In order to eliminate correlated features, the matrix $\mathbf{S}_q$ was truncated to size $G \times G$ using the threshold $[\mathbf{S}_q]_{j,j} < 10^{-10}\,[\mathbf{S}_q]_{1,1}$. $\mathbf{U}$ was then reduced to size $F \times G$. Finally, the obtained base $\mathbf{U}$ was applied to each individual covariance matrix $\mathbf{\Sigma}_p$, in order to transform them into the new feature space Y:

$$\mathbf{S}_p = \mathbf{U}^t\mathbf{\Sigma}_p\mathbf{U} \tag{2.32}$$

After Adler *et al.* incorporated the PCA regularization into the initial equation (2.29), the relative entropy measure took the following form:

$$D(p \parallel q) = k\left(\ln\frac{|\mathbf{S}_q|}{|\mathbf{S}_p|} + trace\,\mathbf{U}\left((\mathbf{S}_p + \mathbf{S}_t)\mathbf{S}_q{}^{-1} - \mathbf{I}\right)\mathbf{U}^t\right) \tag{2.33}$$

where $\mathbf{S}_t$ is a regularized form of $\mathbf{T}$.

(4) Resolving instabilities caused by insufficient data.

The authors of the method realized that solving a problem of feature dependencies using the SVD method may not be enough. The algorithm still can be unstable due to the fact that some individuals are not represented by a sufficient number of biometric samples. In the case when number of images per person, N_p, is smaller than the amount of features, G, $\mathbf{\Sigma}_p$ is singular which causes $D(p \parallel q)$ to diverge towards infinity. To resolve this common problem in practice, Adler *et al.* proposed to develop a lower bound estimate. Firstly, the following two assumptions were made:

1) Diagonal elements of the regularized covariance matrix $[\mathbf{S}_p]_{i,i}$ are valid for any i.

2) Non-diagonal values of $\mathbf{S}_p$, $[\mathbf{S}_p]_{i,j}$ where $i \neq j$, are only valid for the most significant L features, when $L < N_p$.

The rest of entries of $\mathbf{S}_p$ were set to zeros using the mask $\mathbf{M}$ defined as follows:

$$\mathbf{M} = \begin{cases} 1 & \text{if } i = j \text{ or } (i < L \text{ and } j < L) \\ 0 & \text{otherwise} \end{cases} \tag{2.34}$$

Adler *et al.* explain that the choice of L is a trade-off between using more data and stability. If L is selected to be large, more features are taken into account providing more accurate computations, while risking the covariance matrix stability. On the other hand, small L insures stable calculations by reducing the amount of data and consequently the measure precision.

Finally, the process of regularization of the covariance matrix using both introduced above techniques was written as:

$$[\mathbf{S}_p]_{i,j} = (\mathbf{M}_{i,j})\left[\mathbf{U}^t \mathbf{\Sigma}_p \mathbf{U}\right]_{i,j} \tag{2.35}$$

(5) Information content of the biometric system.

Equation (2.29) allows the calculation of biometric information for one specific individual. Since this measure varies from person to person, it is important to average this value throughout the population in order to obtain the biometric information content of the system. If desired, the personal measures may be weighted for creating preference of one group of people over another.

2.5 Biometric sample quality measures

Biometric technology highly depends on the quality of samples that are acquired from users and are further processed by algorithms. Intuitively, the poorer the image sample is, the

more difficult it is to identify the person. In statistical terms, we can say that the sample quality and the amount of information in it are directly connected. This was demonstrated by the work of Youmaran & Adler (2006) where dependency between the level of blur in the image and its information content was explored.

Many aspects may affect the sample quality, such as insufficiently advanced sensors, poor environmental conditions of the system setup or lack of user cooperation. Until recently, there were no standards allowing to score biometric sample quality in an objective way. Recently, INCITS released a document on biometric sample quality standards offering definitions of quality metrics. This section presents an overview of main ideas reflected in (INCITS, 2005).

Firstly, the authors defined terminology that was used to describe biometric systems and their performance:

- **Observer Sample False Accept Rate (osFAR):** The frequency with which a sample of an individual falsely matches samples stored in the database.
- **Observer Sample False Reject Rate (osFRR):** The frequency with which a sample of an individual falsely fails to match with the samples of the same person in the database.
- **Predicted Sample False Accept Rate (psFAR):** The frequency with which a sample of an individual is expected to falsely match those stored in the database.
- **Predicted Sample False Reject Rate (psFRR):** The frequency with which a sample of an individual is expected to falsely fail to match with the samples of the same person in the database. It must be noted that the terms defined above depend on quality of all the samples involved and on the matching algorithm used.
- **Peak Signal to Noise Ratio (PSNR):** A measure used to quantify pixel wise difference between two digital images, normally introduced by using compression algorithms.
- **Performance:** An assessment of error rates of the biometric system.

		Fidelity	
		Low	High
Utility	Low	Low `fidelity` and low `utility` indicate that sample `utility` and system performance could be improved with a new sample with higher `fidelity`, so recapture is recommended.	High `fidelity` and low `utility` indicate poor source `character`. The sample `utility` cannot be significantly improved with a new sample, and an additional biometric type may be recommended.
	High	Samples with poor `fidelity` demonstrate inconsistent `utility` and reduce system performance. They typically will not demonstrate high `utility`.	Samples with high `fidelity` and high `utility` indicate successful capture.

Table 2.1: Relationship between character, fidelity and utility (reproduced from (INCITS, 2005)).

- **Source:** A body part that is used for biometric sampling.

It is difficult to give a single definition to a term "quality" because in biometric systems, it is used in different connotations. However, it is possible to distinguish three main quality components that were described in (INCITS, 2005) as follows:

- **Character:** Character is a quality measure of a sample dependant on inherent features of the source. Some people may provide samples of lower character due to injuries or medical conditions, such as scars on fingers or faces, droopy eyelids, aniridia (underdeveloped irises), etc.

- **Fidelity:** Sample fidelity reflects the degree of accuracy with which a biometric sample represents the source. This normally depends on fidelity of several intermediate steps, such as sample capturing and further image processing. Typical examples of low fidelity are samples of low resolution and samples with high level of blur or noise.

- **Utility:** Utility of a sample in a biometric system reflects the effect (positive or negative, observed or predicted) of the sample on the system performance. Typically, this

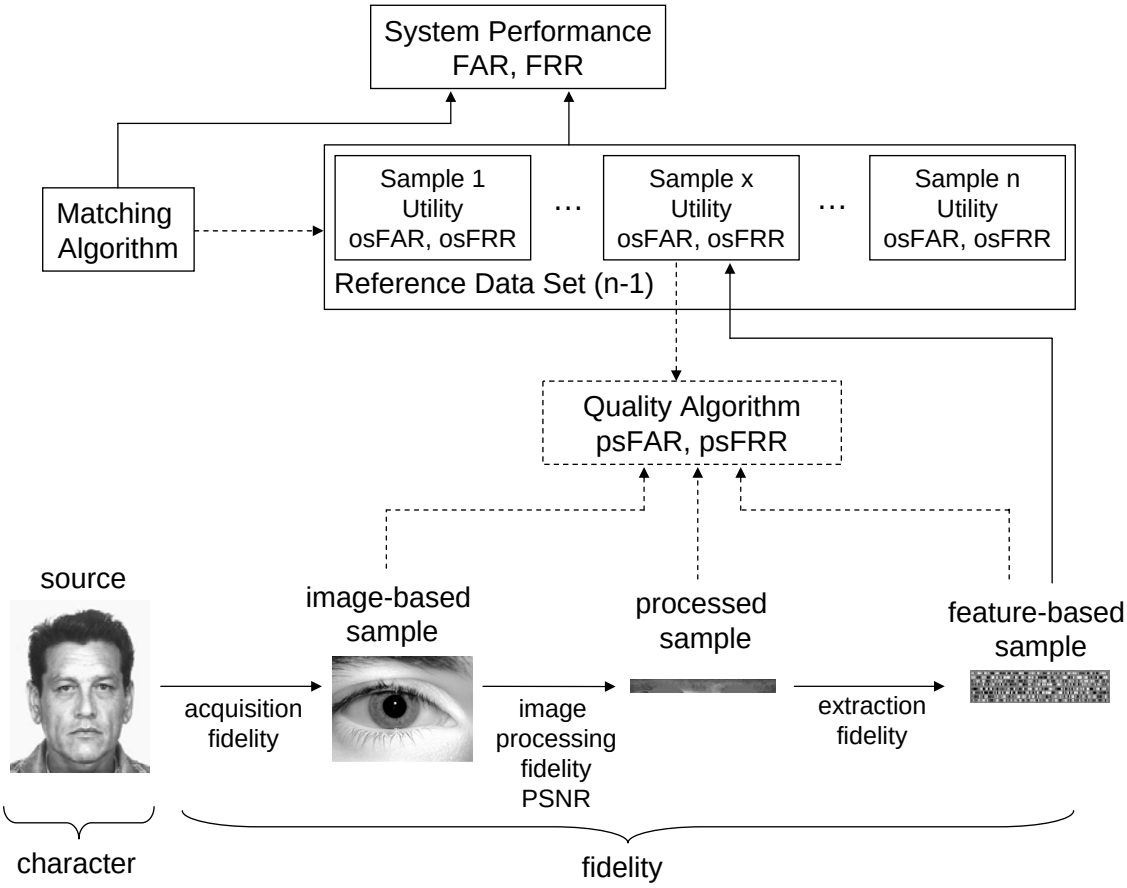


Figure 2.11: Quality reference model illustrating character, fidelity and utility terms (adapted from (INCITS, 2005)).

measure depends on the quality of both the reference dataset and the sample.

The introduced terms are closely connected with each other. Their relationship is illustrated in table 2.1 reproduced from (INCITS, 2005).

Finally, authors offered a reference model in order to help distinguish different aspects of biometric quality, using previously defined terminology. The model is shown in figure 2.11 adapted from (INCITS, 2005). It allows to calculate quality scores using two different approaches - "bottom-up" and "top-down".

"Bottom-up" approach assumes that the sample character is verified and the relationship between sample fidelity and utility is known and is used to predict the utility of the sample in the system, given the fidelity score. In the case when the relationship between

fidelity and utility is impossible to establish, the fidelity score may be still used to find out what elements contributing into the sample utility are responsible for its low value. This could be a poor character of the sample, indicating that a different biometric may be required; or a poor fidelity of the capturing stage, in which case adjustment of parameters may be needed followed by the sample recapturing; etc.

"Top-down" approach calculates relationships between sample features and utility and then estimates psFAR and psFRR. Utility is calculated as a "function of 1) the character of the source, 2) the fidelity of the sample (and its components), 3) the cumulative fidelity of the reference dataset" (comprised of samples, their utility and osFAR/osFRR values), "and 4) the matching application" (INCITS, 2005).

Chapter 3

Biometric Sample Quality Assessment

3.1 Introduction

One of the most important aspects of biometric systems is sample quality. Poor image quality may easily affect an overall biometric system performance by the lack of accuracy in extracted features and cause high error rates (FAR and FRR). Therefore, it is crucial to be able to detect samples of poor quality before they are processed and initiate an additional sample capture or apply techniques for automated quality improvement. Recently, the International Standards Organization (ISO) attempted to standardize quality measures by releasing a biometric sample quality draft standard (INCITS, 2005). In this document, as was introduced in chapter 2.5, measures contributing to performance of a biometric system were defined - character, fidelity and utility. It was agreed that sample utility should be the most important factor which assigns higher quality to samples that allow better person identification, meaning better separation of genuine and impostor scores.

It is traditionally believed that human assessment of biometric data is the most precise way to evaluate sample quality. For instance, this assumption was used by Tabassi *et al.* (2004) for designing biometric quality standards. Moreover, human evaluators of biometric samples are involved in many government institutions (i.e. immigration and passport agencies).

In this work, we attempted to evaluate the ability of humans to adequately assess

biometric sample quality by answering following three questions: 1) do all participants agree about sample quality? 2) are algorithmic quality measures consistent with each other? 3) do human quality scores and algorithmic scores match?

3.2 Methods

Firstly, we needed to gather quality scores from human participants and from a number of automatic quality measures. In order to obtain results of higher significance, two different biometrics were used - irises and faces. Unlike fingerprints, irises do not have a well defined expert community that could argue that the experiments should take into account the level of expertise of each evaluator.

3.2.1 Image quality evaluation by human participants

To allow human participants to evaluate biometric samples, the web based quality assessment environment, as shown in figure 3.1, was designed. It enabled a user to evaluate quality of two data sets, iris and face images, for a biometric identification application. For each image, the participant was asked to assign a quality score, from 1 to 5, by clicking a corresponding button of the form. The scores were collected and stored in the database.

In order to help participants learn the range of image quality variation and by thus ensure a uniform evaluation, a training section was incorporated into the environment. During the learning stage, users were given an opportunity to review both data sets without assessing them.

3.2.2 Automatic quality measurement

The developed algorithm for automatic quality measurement was based on the biometric score distribution. The goal was to incorporate the sample utility aspect into the model, so that ability of samples to influence separation of genuine and impostor scores would be reflected.

Figure 3.1: Web based evaluation form for human evaluation of biometric image quality. Participants assigned each image a quality level from 1 to 5.

Many sources report that sample quality has little or no effect on impostor distribution. On the other hand, the genuine distribution gets more precise and narrow with increasing biometric sample quality. Therefore, we decided to model the impostor distribution as a constant. The first step of the study involved modeling a term of sample quality:

$$E[MS_{i,j}] = E[Q_i Q_j] \tag{3.1}$$

where $MS_{i,j}$ is a matching score between two biometric samples i and j of the same person calculated by the identification algorithm and defined in the range from 0 to 1, and Q_i and Q_j are biometric quality values of the same samples i and j respectively ($0 \leq Q \leq 1$). In other words, the model defined quality scores to be proportional to the matching score of involved samples.

Values of Q were then calculated by minimizing

$$\sum_{i} \sum_{j, i \neq j} (MS_{i,j} - Q_i Q_j)^2 \tag{3.2}$$

for all comparisons between samples of the same person. It must be noted that most identification softwares result in a match score of 1.0 when comparing identical biometric samples ($i = j$), irrespective of the image quality. Therefore, such comparisons were excluded from the model. Finally, in order to simplify the minimization problem, (3.2) was modified in the following way:

$$\sum_{i,j} (\log MS_{i,j} - \log Q_i - \log Q_j)^2 \tag{3.3}$$

As an illustration of the minimization process described above, let us consider the following example. For a set of four samples of one person matching scores are calculated. Solving the following least square matrix equation should provide us with quality scores for each sample used:

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \log Q_1 \\ \log Q_2 \\ \log Q_3 \\ \log Q_4 \end{bmatrix} = \begin{bmatrix} \log MS_{1,2} \\ \log MS_{1,3} \\ \log MS_{2,3} \\ \vdots \end{bmatrix} \tag{3.4}$$

Thus,

$$\log Q = (H^t H)^{-1} H^t (\log MS) \tag{3.5}$$

where Q is a vector of N biometric quality scores (one per each sample), MS is a vector of M available match scores between given samples obtained from identification algorithms, H is an $M \times N$ matrix where every row k specifies which samples were compared to calculate a matching score MS_k by setting corresponding values to 1-s and keeping the rest as 0-s.

The described procedure allowed us to calculate a quality estimate for each biometric sample for each recognition algorithm separately. The only requirement was to have at least three biometric samples per individual. Otherwise, the equation (3.5) becomes singular.

Another automatic approach used for measuring image quality was the Image Quality

Measure (IQM) algorithm by MITRE (2005). The software is based on the method that evaluates quality of digital images by the power spectrum of their spatial frequencies. This is a freeware software that can be applied to a wide range of images, such as aerial, medical or forensic images. We used it to have an alternative set of quality scores for the image collections of irises and faces.

3.3 Results

For running experiments, datasets of irises and faces were used. 84 iris images were selected out of the internal database captured by the L.G. iris camera. They included 6 samples per eye of 7 different individuals. As a source of face images, the Mugshot Identification Database (NIST, 2002) was used. 98 frontal face images were chosen that included 3-4 samples of 29 different people. An effort was made to select face images with minimal variation in age, in order to minimize matching score differences which are unrelated to the image quality.

8 people were invited to participate in the manual quality evaluation stage. Among them there were 7 males and one female, between the ages 20-40. All of the participants were graduate students and researchers in electrical engineering at the University of Ottawa.

During the stage of automated evaluation, iris and face recognition algorithms were applied. As an iris identification software, we decided to use an open-source implementation developed by Masek (2003). The obtained Hamming Distances (HD) were distributed between 0.2 and 0.4, with lower scores for better matches. In order to make these results suitable for use with the quality calculation model, the scores were inverted using a simple formula $1 - HD$ and normalized to spread over the range [0-1]. Six different face recognition algorithms were used to generate matching scores of the face database. Software implementing these methods was released by three different vendors over the period of 6 years, 1999-2005. Finally, the propsed model was used to generate biometric quality scores.

The experimentally obtained data (human participant evaluation scores, calculated quality scores for each recognition algorithm, IQM results) was then compared by calculating the Pearson r correlation co-efficient. Analyzing these results allowed us to answer questions

	Mean Algorithm	**IQM**
Mean Human	0.234	0.159
Mean Algorithm		0.003

Table 3.1: Correlation of biometric image quality measures for face images.

	Mean Algorithm	**IQM**
Mean Human	0.175	0.458
Mean Algorithm		-0.036

Table 3.2: Correlation of biometric image quality measures for iris images.

posed at the beginning of the study:

Are human evaluators consistent with each other? Yes. Human participants gave similar quality scores to the experimental images which was justified by an average correlation co-efficient of $r = 0.723$ (iris), $r = 0.613$ (face), $p < 0.001$.

Are algorithms consistent with each other? Yes. Face recognition algorithms resulted in strongly correlated quality measurements reflected by $r = 0.534$, $p < 0.001$. It was impossible to perform similar analysis of iris recognition algorithms since only one implementation was used.

Are human evaluators consistent with measures from biometric algorithms? No. On average, human assessments of quality do not correlate or show very low correlation which may be clearly seen from resulting correlation co-efficients: $r = 0.175$, p not significant (iris) and $r = 0.234$, $p < 0.05$ (face).

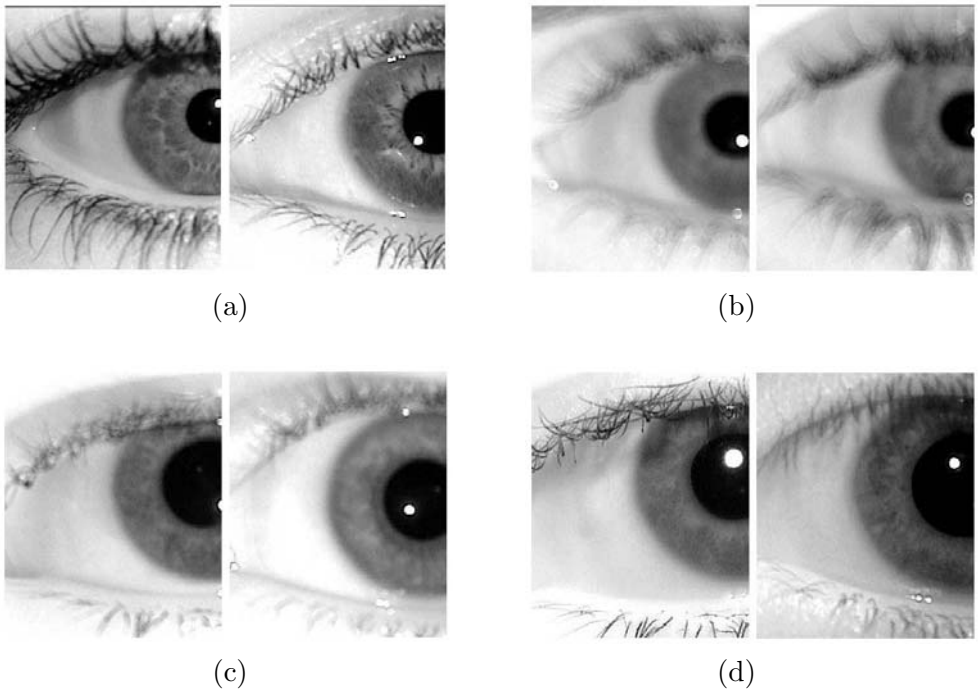
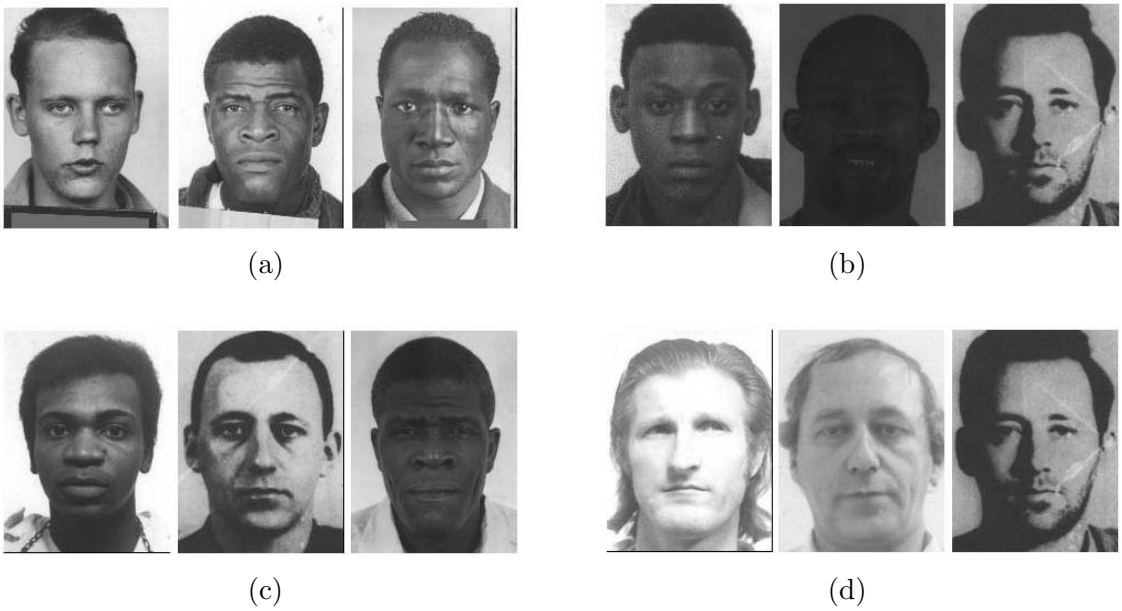


Figure 3.2: (a) The average highest quality irises selected by humans; (b) the average lowest quality irises selected by humans; (c) the average highest quality irises selected by biometric algorithms; (d) the average lowest quality irises selected by biometric algorithms.

Tables 3.2 and 3.1 show degree of correlation between different types of quality measurement for irises and faces correspondingly. Application of IQM software gave results that did not correlate with any other evaluation type, except for humans vs. IQM for iris images ($r = 0.458$, $p < 0.001$).

In order to illustrate and to gain a better understanding of human and algorithm quality decisions, samples of images from different groups (best and worst images, as selected by human participants and algorithms) were selected (see figures 3.2 and 3.3). In the case of iris images, it may be seen that people have a clear preference for sharp in-focus images. This explains agreement of their scores with scores given by the IQM algorithm that is based on spectrum power analysis. However, as stated in (Daugman, 2004), iris recognition algorithms are not very sensitive to image sharpness since blur does not affect significantly enough the phase information inside the iris.



(a) (b)

(c) (d)

Figure 3.3: (a) The average highest quality faces selected by humans; (b) the average lowest quality faces selected by humans; (c) the average highest quality faces selected by biometric algorithms; (d) the average lowest quality faces selected by biometric algorithms.

3.4 Discussion

In this chapter we attempted to evaluate how reliable humans may be in assessing quality of biometric samples used for identification purposes. In order to do so, experimentally obtained human scores were compared to measurements performed by the IQM software and by the alternative method based on match scores calculated by automated iris and face recognition algorithms. The results showed that, in general, human evaluators are consistent with each other, as well as different algorithms are consistent within their group. However, comparing different types of assessment showed low or no correlation. The only case where two different quality measures showed similar results was quality assessment of iris images done by the IQM and humans. This phenomenon may be explained by their preference for sharper images.

These results indicate that intuitive quality assessment done by human evaluators may not be sufficient for biometric identification applications since algorithm decisions are often based on features seemingly insignificant or not visible for humans.

Chapter 4

Experimental Software

In this chapter, we introduce an open-source implementation of the iris recognition system created by Masek & Kovesi (2003). This software, developed in the Matlab® environment, was used by us as a base component of our experiments for building iris templates of different configurations and for calculating match scores. Here, we present main elements of the implementation and its performance based on our results.

The main purpose of the work by Masek & Kovesi was to provide independent researchers with a means for verifying uniqueness of the iris and for evaluating effectiveness of using the iris for person identification. Masek carried out an elaborate analysis of existing algorithms before choosing the most appropriate elements for the implementation. The detailed description of involved methods was given in section 2.2.

4.1 CASIA database

Masek & Kovesi used the CASIA database, publicly provided by Chinese Academy of Sciences (2003), for developing and testing their iris recognition software. The database contains 756 grayscale eye images with resolution of 320×280 pixels, captured by means of an IR iris camera. 108 unique eyes are represented with 7 samples each. The images were obtained from two sessions separated by one month. Most of participants were of Asian origin.

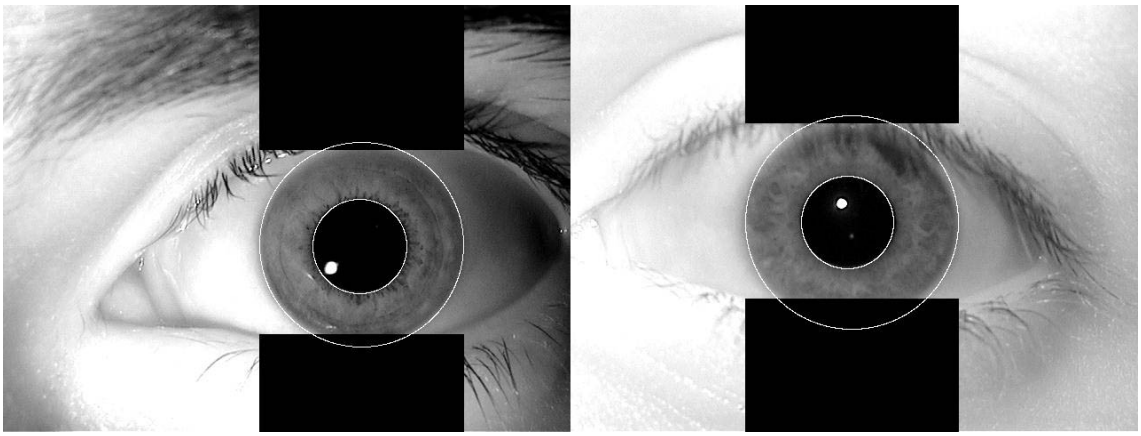


Figure 4.1: Iris segmentation using software by Masek & Kovesi (2003) . Eyelids are approximated by horizontal lines. Eyelashes are not detected.

4.2 Main components of the implementation

4.2.1 Iris segmentation

For the first stage of iris segmentation, the authors decided to use the circular Hough transform algorithm (see section 2.2.2) preceded by Canny edge detection. This part of the algorithm requires manual setting of the possible range of iris and pupil radii. For our eye database captured using the L.G. iris camera, we chose to use the following values: the pupil radius range of [50 - 90] pixels and the iris range of [100 - 150] pixels.

Eyelid detection was done using linear Hough transform to approximate each eyelid by a straight line (as a simplification). Then, a horizontal line was drawn through the closest to the pupil point of intersection of the Hough transform eyelid approximation and the iris boundary. Finally, eyelash detection was performed by a simple thresholding. However, Masek & Kovesi could use this technique only in the case when eyelashes were dark and when no other iris details had low intensities. For our database it was impossible to apply thresholding due to fair eyelashes. A typical result of iris segmentation using the software by Masek & Kovesi (2003) is shown in figure 4.1.

The authors performed the implementation testing using CASIA (Chinese Academy of Sciences, 2003) and LEI (Barry & Ritter, 2003) databases of iris images. For the CASIA database, the segmentation was reported to have a success rate of 83% and for the LEI

database it was only 62% due to lower quality of images. We did not perform success rate testing since it was not the objective of the work. Images were visually analyzed to detect correct segmentation. Approximately 20% of all images, for which the algorithm failed to extract irises correctly, were manually excluded from subsequent calculations.

4.2.2 Iris normalization

Masek & Kovesi used Daugman's rubber sheet model (see section 2.2.3) for implementing the iris normalization stage. At this point two manually set parameters were introduced: the angular resolution (number of iris points sampled along each circular line) and the radial resolution (number of circular lines located inside the iris and around the pupil). For the general application we used the default values chosen by the authors: 240 for the angular resolution and 20 for the radial resolution.

The iris normalization stage produces two 2D arrays with horizontal and vertical dimensions corresponding to the angular and the radial resolutions respectively. As can be seen in figure 4.2, the first array represents the unwrapped iris region and the second one is created to mark regions that contain reflections or are occluded by eyelids. Pixels set to 1 in the validity mask are excluded from the calculations at later stages of encoding and matching. The corresponding invalid pixels of the normalized iris are set to the average value of pixels in the iris in order to minimize the noise effect on features extracted later.

Masek & Kovesi reported normalization to have successful results. However, they also noticed that iris images with significantly varying pupil diameters showed some variations in normalized irises. The phenomenon was explained by claiming that pupil dilations trigger small changes in the pattern of the iris.

4.2.3 Feature encoding

The authors chose to use 1D Log-Gabor wavelets, as introduced in section 2.2.4, for the feature encoding stage. The normalized iris in its 2D representation was broken into separate rows which were treated as 1D signals. Each row was then convolved with 1D Log-Gabor

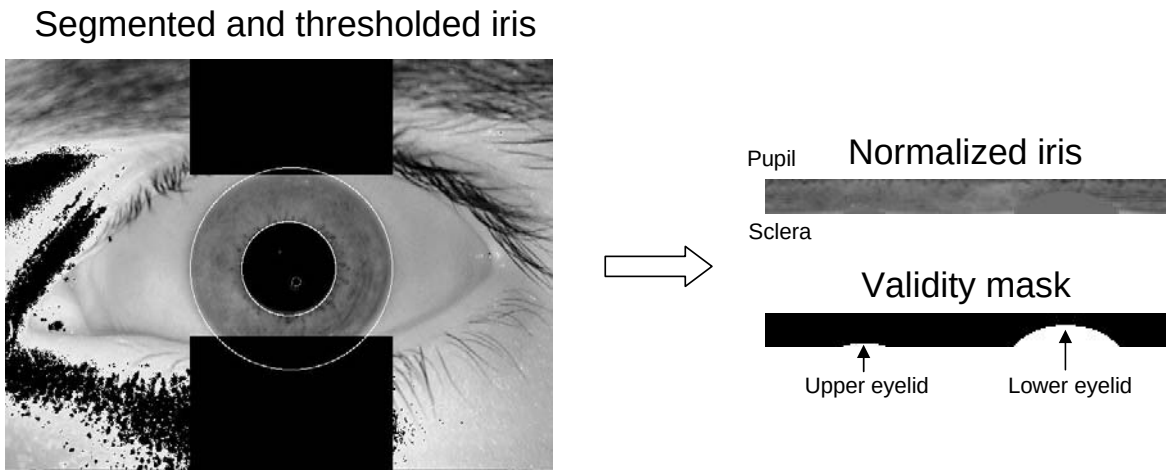


Figure 4.2: Iris normalization using software by Masek & Kovesi (2003). After the iris is segmented and reflections are thresholded out, two 2D vectors are built - the normalized iris with pixels sampled out of the original iris region and the validity mask marking reflections and occlusions.

wavelets. The rows, or circles inside the iris, were preferred over the columns containing radial information since the angular direction provides more independent information (Masek, 2003).

As a next step, the output of the convolution was quantized using Daugman's approach, as described in section 2.2.4. Each phasor was quantized to one of the four quadrants resulting in two bits of code. In order to minimize the number of disagreeing bits for slightly misaligned templates, grey code was used (Masek, 2003), such that the difference between adjacent quadrants is one bit only. Regions that produce phasors of amplitude close to zero were marked as invalid in the validity mask, in order to eliminate insignificant information.

Masek & Kovesi explained that the encoding performance depends on selection of a number of parameters: number of filters, N_f, their base wavelengths, λ_n, bandwidths of filters, σ, and the multiplicative factor between centre wavelengths of successive filters, α. The optimal parameter values were selected by experimenting with LEI and CASIA databases. The performance for each set of parameters was tested in terms of decidability d', that authors defined in the following way:

$$d' = \frac{|\mu_S - \mu_D|}{\sqrt{\frac{(\sigma_S{}^2 + \sigma_D{}^2)}{2}}} \tag{4.1}$$

where μ_S and μ_D are means of the intra-class and the inter-class Hamming Distance (HD) distributions respectively and σ_S and σ_D are their standard deviations. In other words, d' is a measure of distance between genuine and impostor distributions. The higher the value of decidability is, the better the separation of the intra-class and the inter-class distributions is.

Masek & Kovesi performed thorough testing and showed that using multiple filters does not improve decidability. Therefore, using one filter was recommended, which additionally results in more compact templates reducing computation times. For the CASIA database the optimum centre wavelength was found to be 18 pixels and σ - 0.5.

Another set of parameters that contributes to the algorithm performance is angular and radial resolutions used during the normalization stage. Using the decidability factor as a performance indicator the authors found out that an angular resolution of 240 and a radial resolution of 20 offer the best trade-off between accuracy and efficiency of the identification.

The size of the resulting template may be calculated as `angular resolution` × `radial resolution` × 2 × `number of filters`. If optimal parameter values mentioned above are used, 9600 bits of phase information are obtained per iris.

For our experiments, we decided to use default values for all filter parameters and template dimensions discussed above, because characteristics of our database are close to those of CASIA database, due to usage of the similar image acquisition technology.

4.2.4 Template matching

In order to perform bit-wise comparison of iris templates, Masek & Kovesi decided to use the HD measure as proposed by Daugman (2004) and described in section 2.2.5. For each iris two 2D arrays were used - the result of encoding and the mask of non-valid bit locations. The partial HD was calculated by counting non-matching bits of two templates in the iris regions where both validity masks had zeros, meaning that both templates had valid information at

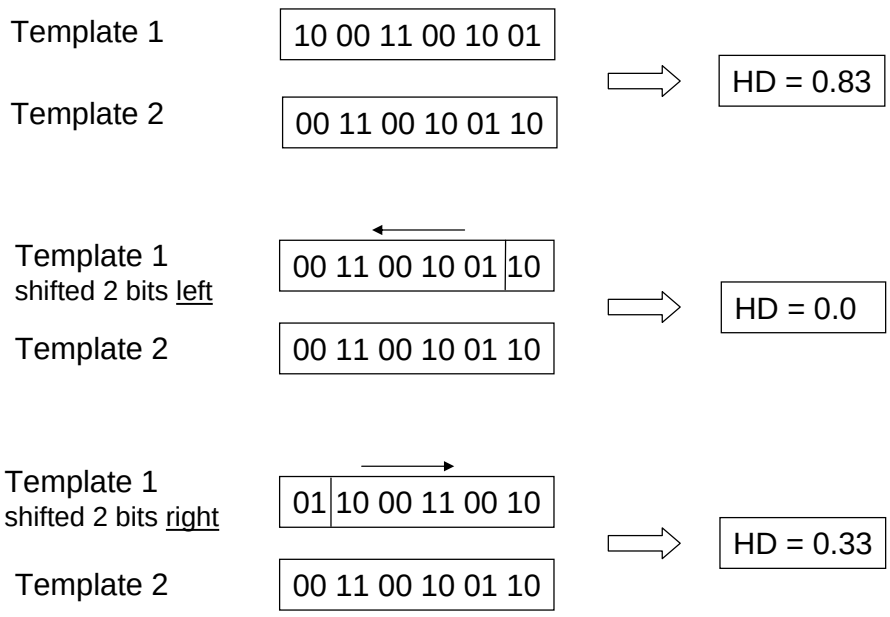


Figure 4.3: An example of template shifting adapted from (Masek, 2003). The reference template (Template 1) is shifted 2 bits right once and shifted 2 bits left once in search of the best match. Shifting left gave the perfect match with HD=0.

these locations.

To compensate over possible iris rotations resulting from head tilts, the authors implemented a template shifting scheme proposed by Daugman (2004) (see section 2.2.5). One of the matched templates was shifted in a circular way left and right using increasing number of shifts to find the best match. At the end of the process, the smallest HD was chosen to be the result of the comparison. The size of one shift in bits was selected to be the amount of phase information extracted from one pixel of normalized iris. This value is defined by the number of used filters times 2 (bits obtained by applying one filter). As was described in the previous section, Masek & Kovesi chose to use only one filter for the encoding. Therefore, the shift size was chosen to be 2 bits. One shift, as defined by the authors, included horizontal scrolling of the template to the left and then to the right using the same amount of bits. An illustration of one step of shifting may be seen in figure 4.3 adapted from (Masek, 2003).

In order to find the optimal number of shifts for the best possible person identification, the authors analyzed the intra-class HD distributions for increasing number of shifts. With

every new shift the rotational inconsistencies were decreasing and the distribution mean and standard deviation were converging to a constant value. It was found experimentally that for the CASIA database it requires 8 shifts (or 17 comparisons) for the intra-class distribution to converge.

4.3 Author results

Masek & Kovesi performed experiments using two databases of iris images, CASIA and LEI. We present their results related to the CASIA database only, since it was captured using an infrared iris camera and, therefore, has characteristics similar to our local database.

Application of the segmentation algorithm to 756 eye images of the CASIA database resulted in 624 successfully extracted irises, which is equal to a success rate of 83%. This allowed for 1,679 intra-class and 192,699 inter-class comparisons.

Distributions of obtained match scores, as reported by Masek, are shown in figure 4.4 reproduced from (Masek, 2003). The results were obtained using default parameter values: number of filters = 1, centre wavelength = 18 pixels, σ = 0.5, template size = 20×480, number of shifts = 8 (or 17 comparisons). Figure 4.4 specifies that there exists an overlap in distributions. The authors suggested using the decision threshold of 0.4 which resulted in FAR and FRR equal to 0.005% and 0.238% respectively.

Note that FMR and FNMR of the presently discussed software are equal to FAR and FRR, because in order to make a decision only one HD value is considered.

4.4 Our results

In this section we present match score distributions calculated by the software of Masek & Kovesi for the local database. Our data set contained 360 images: 12 different eyes with 30 samples per each eye. Participants involved in the database acquisition included people of Asian, African and Caucasian origins.

During the process of cross-comparisons, 5,220 intra-class and 59,400 inter-class match

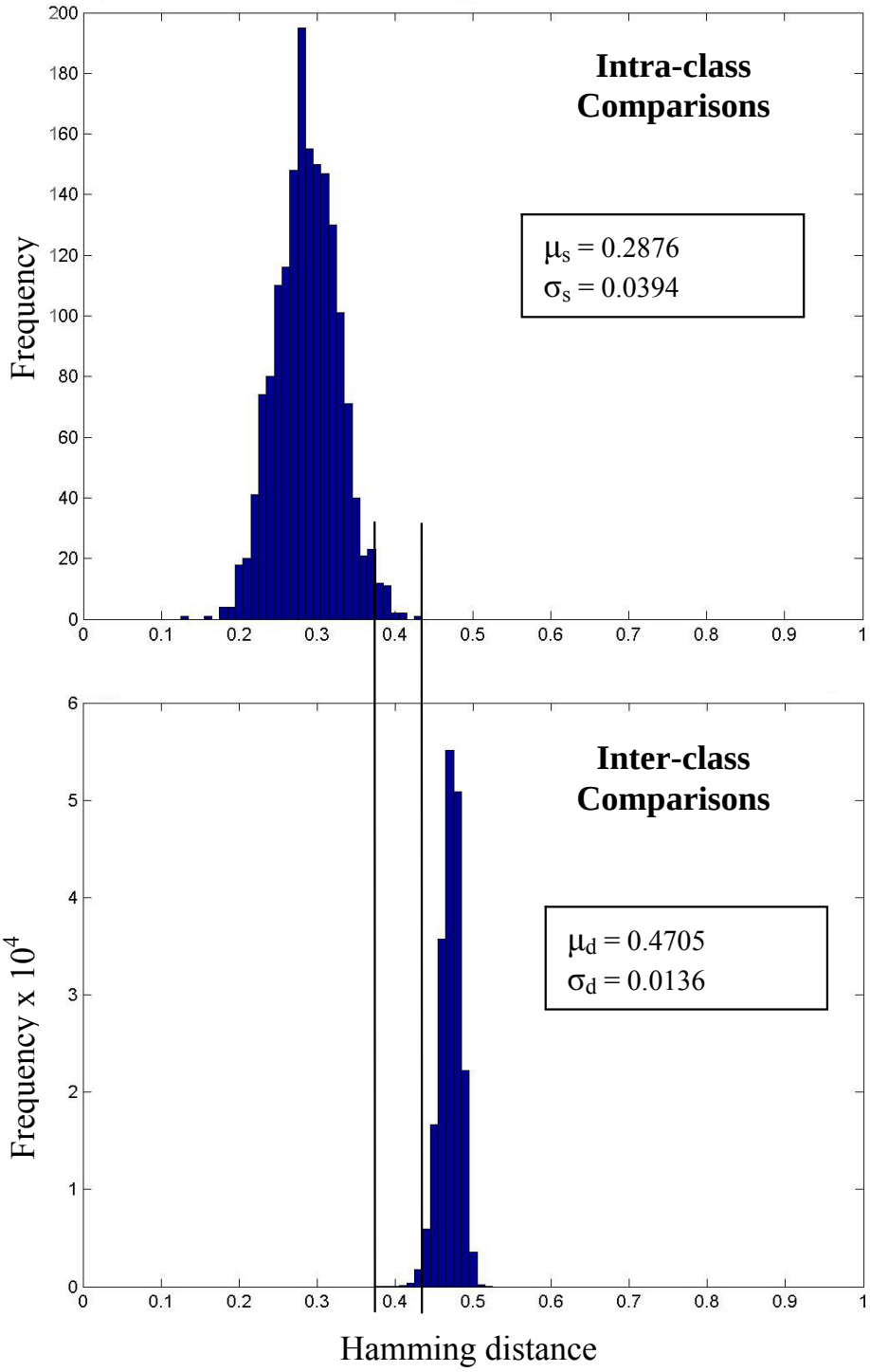


Figure 4.4: Genuine and impostor distributions for the CASIA database as reported by Masek (2003) (reproduced from (Masek, 2003)). The used parameters are: number of filters = 1, centre wavelength = 18 pixels, $\sigma = 0.5$, template size = 20×480, number of shifts = 8 (or 17 comparisons). Vertical lines mark the intersection region of two distributions.

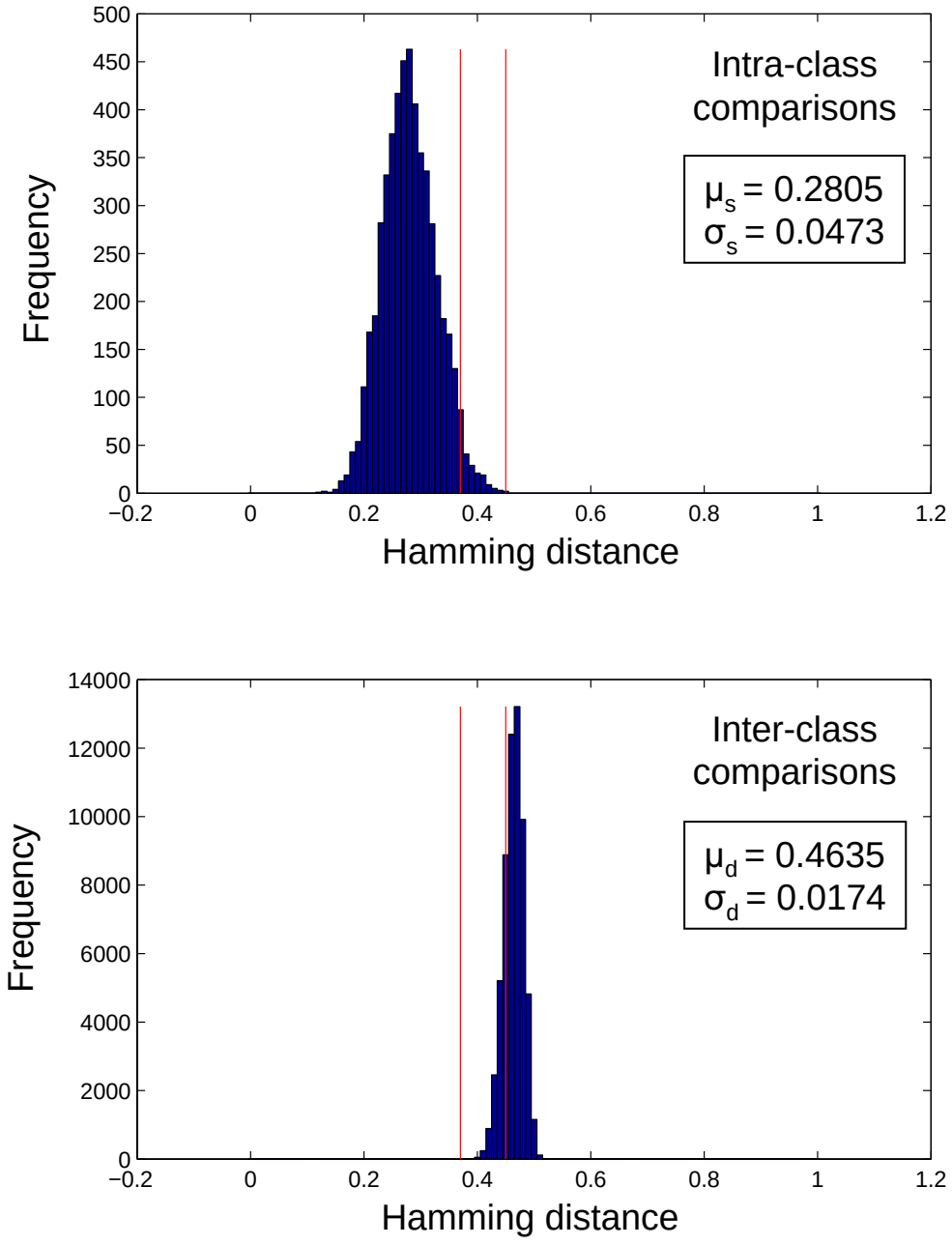
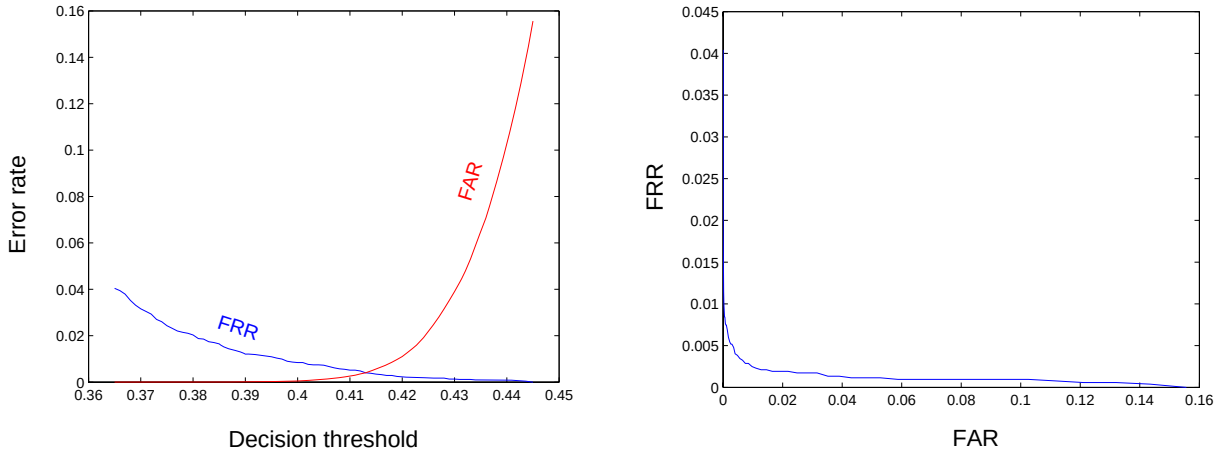


Figure 4.5: Genuine and impostor distributions for the local database using the software by Masek & Kovesi (2003) and default algorithm settings. Red vertical lines mark the intersection region of two distributions.



(a) (b)

Figure 4.6: Error rates for the local database after applying the software by Masek & Kovesi (2003): (a) FAR and FRR curves as a function of decision threshold; (b) FAR and FRR relationship.

scores were calculated. Comparisons of identical eye samples were ignored. It was decided to use default algorithm parameter values proposed by the authors of the software, since they provided sufficient performance. The resulting match score distributions may be seen in figure 4.5. Intra-class comparisons were distributed with the mean of 0.2805 and the standard deviation of 0.0473. The distribution of inter-class match scores had a mean of 0.4635 and a standard deviation of 0.0174. As may be seen in figure 4.5, there was an overlap between two HD sets, appearing between score values of 0.37 and 0.45. FAR and FRR rates for different decision threshold selections are shown in figure 4.6. For the decision threshold of 0.4 the algorithm provided FAR = 0.11% and FRR = 0.73%.

Compared to the results obtained by Masek & Kovesi using the CASIA database (FAR = 0.005% and FRR = 0.238%), our application showed slightly bigger error rates. This is likely due to the differences between the databases.

Chapter 5

Information Content of Iris Templates

The information content of biometric images has important applications in many uses of biometric technology, such as: determining the uniqueness of biometric features, establishing inherent limits to biometric template size requirements, clarifying feasibility of biometric encryption and estimating performance limits of biometric matchers. Such a mesaure is also useful to understanding biometric fusion, novel biometric features and quantifying privacy issues (Adler *et al.*, 2006).

In this chapter we present results of experiments that were done in order to gain a better understanding of the information content of iris images. This was addressed by studying existing techniques for measuring entropy of biometric samples, such as (Daugman, 2003) and (Adler *et al.*, 2006), and applying them to iris templates of our local database.

The experimental database is presented in section 5.1. Section 5.2 discusses the work done to explore the information measurement approach offered by Daugman. Firstly, results of applying the algorithm to iris templates of various dimensions are presented and analyzed. Then, a detailed statistical analysis of the approach is performed, in an attempt to explain the nature of the discrimination entropy. In section 5.3, we present the implementation of the algorithm by Adler *et al.* applied to irises. Later in this section, we talk about measurements of biometric information that were performed in different modes: entropy of features, entropy of the whole iris and entropy of certain iris regions.

5.1 Experimental data set

The main image data set used for the most experiments presented in this work was obtained using L.G. iris camera, provided by the staff of the Canada Border Services Agency (CBSA). The database was built using 640×480 grayscale eye samples of 7 people of different origins - Caucasian, Asian and African. The involved participants were graduate students at the university of Ottawa and the CBSA staff, 3 of them were female and the rest were males. In total, 657 samples of 14 eyes were collected in one session. Out of the original data set, 535 images successfully passed the stage of iris segmentation performed using the software by Masek & Kovesi.

For the final experimental database, 12 eyes with 30 best samples each were kept, resulting in a set of 360 images. Having an equal amount of images per eye was important for this study. The fact of equally represented eyes in the database allowed for the most accurate statistical analysis.

Additionally, as described in appendix A, we attempted to build an IR image capturing system for gathering our own database of eye images. Unfortunately, at this stage, we were unable to achieve an acceptable performance. The further investigation needs to be done in order to make the system suitable for acquiring iris samples.

5.2 Discrimination entropy

In section 2.4.1, we presented the approach for measuring information content of the iris as was proposed by Daugman (2003). The distribution of inter-class match scores was modeled using a binomial distribution, and the amount of information was measured in terms of its degrees of freedom.

5.2.1 Experimental results

In this section we present results of applying the model by Daugman to our database of iris images in order to determine the information content of iris templates. Templates of different

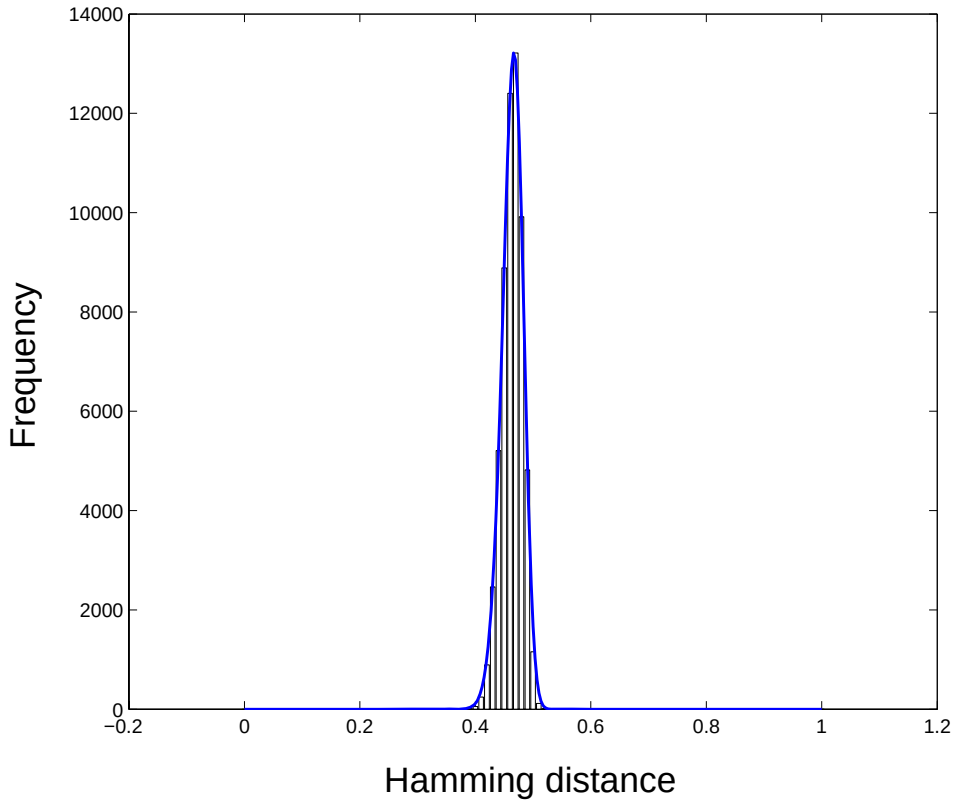


Figure 5.1: Fitting a binomial distribution to the impostor match score distribution of $20 \times (240 \times 2)$ templates. Solid line represents the curve of binomial distribution defined by (2.23) with 301 degrees of freedom, mean value of 0.5 and 6 comparisons.

dimensions were obtained by applying iris recognition software of Masek & Kovesi.

The Matlab® environment was used for implementing a procedure of fitting a curve of binomial distribution to a given distribution of match scores. The curve was modeled according to the definition given in 2.23 for approximating the distribution of Hamming Distances (HD) using multiple comparisons. In order to find the best fit, an iterative procedure involving different curve parameters was executed. Each iteration involved the following steps:

1) Define the binomial distribution curve for a given set of parameters: N_i (degrees of freedom), p_i (mean) and n_i (number of circular template rotations during the matching stage).

2) Transform the curve to be defined over the interval [0,1], since originally binomial

distribution is not defined for non-integer numbers and is only set for values [0, 1, 2, ... N_i-1].

3) Normalize the binomial curve by scaling it to fit the data distribution vertically.

4) Calculate the error between the data and the approximation by summing squared distances between every pair of corresponding points.

5) If the resulting error is smaller than all the errors computed in previous iterations, set it to be a minimum value and store the corresponding parameters N_i, p_i and n_i.

At the end of all iterations the set of parameters that produced the approximation with the smallest error is used to define the binomial distribution curve that fits match scores the best. The Matlab® implementation of the described procedure is given in appendix B.

An example of approximated binomial curve may be seen in figure 5.1. For the distribution of 59,400 inter-class match scores calculated using templates of size $20 \times (240 \times 2)$, the following parameters resulted in the best fit of the binomial curve: $N = 301$, $p = 0.5$ and $n = 6$. According to (Daugman, 2003), this means that the iris template of these dimensions contains 301 bits of information.

It was decided to use the number of template comparisons as an additional fitting parameter n, even though the matching algorithm was always performed using a constant number of shifts, 8, resulting in 17 comparisons per pair of iris templates. The need for this parameter was discovered experimentally by seeing that it was impossible to find an acceptable binomial fit for the match score distribution otherwise. Varying n allowed us to obtain very close data approximations.

The next stage of experiments involved determining the number of degrees of freedom for HD distributions obtained using iris templates of different configurations. The purpose of this was to study the relationship between templates of different dimensions and the discrimination entropy. An attempt was made to discover the effect of the following parameters: angular and radial resolutions and number of filters used for template encoding.

For each combination of parameters mentioned above, we used the software by Masek

& Kovesi (2003) to obtain a corresponding set of iris templates. Next, cross-comparisons were performed to build the distribution of impostor and genuine match scores. Finally, the fitting of the binomial distribution curve was done in order to determine the amount of information in the iris template of given characteristics, as was defined by Daugman (2003). The obtained results are presented in table 5.1.

We attempted to calculate information content for both impostor and genuine distributions, in order to gain better understanding of Daugman's approach. Intra-class comparisons revealed a significantly smaller number of degrees of freedom compared to inter-class distribution, which is an expected behaviour indicating that there are more dependencies between templates of the same eye than between templates of different eyes. Additionally, in all cases, the procedure of binomial curve fitting into intra-class Hamming Distances resulted in a bigger number of rotations. The amount of information obtained by analyzing genuine comparisons did not show significant changes for templates of varying sizes. As may be seen from the result table 5.1, templates of size $8 \times (128 \times 2)$ resulted in 30 degrees of freedom, while significantly bigger $20 \times (270 \times 2)$ templates showed only 39 degrees of freedom. Graphical presentation of number of degrees of freedom calculated for genuine distributions is given in figure 5.2. Each curve represents templates of the same radial resolution. The graph shows little consistency. Therefore, we omit genuine match scores from further experiments.

The experimental results obtained from applying the information measure technique of Daugman to the distribution of inter-class comparisons are presented in figure 5.3. 5.3(a) shows the behaviour of all obtained degrees of freedom with more detailed analysis of templates with radial resolution of 20 points. 5.3(b) zooms in the area of 5.3(a), marked by the dotted contour, where all radial resolutions are present allowing better view on relationship between template dimensions and amount of information.

Figure 5.3(a) shows that number of degrees of freedom increases proportionally to the angular resolution. When increase in the angular resolution does not provide any more new iris pixels, due to the limited eye image resolution, Daugman's measure starts converging to a constant value. The average iris area in the local eye image database is equal to approximately

Radial resolution	**Angular resolution**	**Number of filters**	**Impostor** N	**Impostor** n	**Genuine** N	**Genuine** n
8	128	2	93	6	30	16
8	64	4	26	5		
12	150	2	128	6	31	13
12	150	4	128	6	31	13
12	150	8	128	6	31	13
15	160	2	171	10	32	11
15	200	2	206	6	30	9
15	200	4	206	6	30	9
20	180	2	176	6	30	9
20	180	4	176	6	30	9
20	210	2	230	6	33	9
20	240	2	301	6	36	9
20	270	2	383	6	39	9
20	300	2	505	11	38	7
20	330	2	578	6		
20	360	2	677	6		
20	390	2	781	6		
20	420	2	880	6		
20	480	2	988	5		
20	510	2	1104	5		
20	540	2	1218	5		
20	700	2	1460	4		
20	1000	2	2374	4		
20	1500	2	3084	4		
20	2000	2	3260	4		
22	300	2	484	6	39	8
25	288	2	480	11	38	7
30	260	2	367	6	36	8
30	280	2	429	6	38	8
40	224	2	270	6		
40	240	2	311	6		
40	254	2	352	6		
40	270	2	408	6		
40	284	2	453	6		

Table 5.1: Discrimination entropy, N, and a number of rotations, n, calculated for impostor and genuine match score distributions, as a function of radial and angular resolutions and a number of filters.

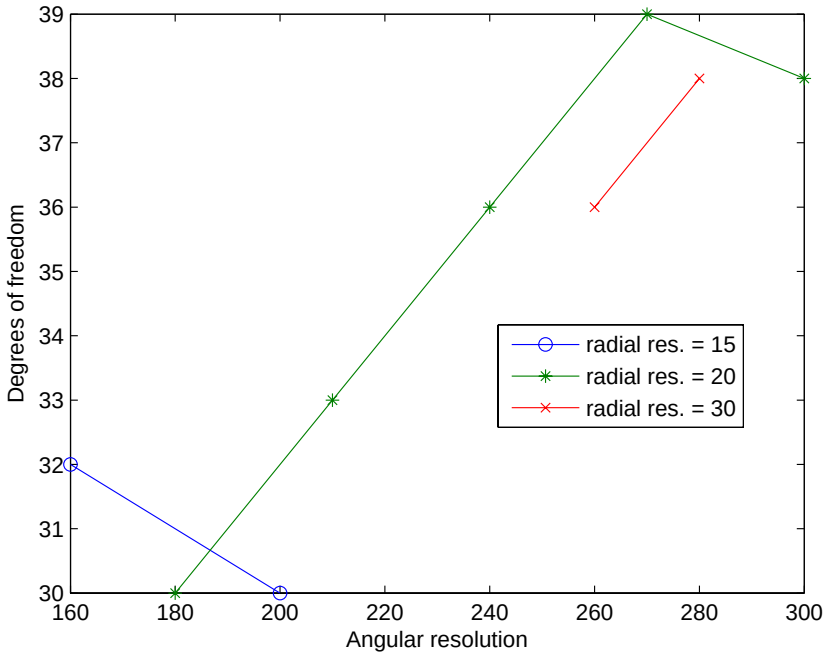
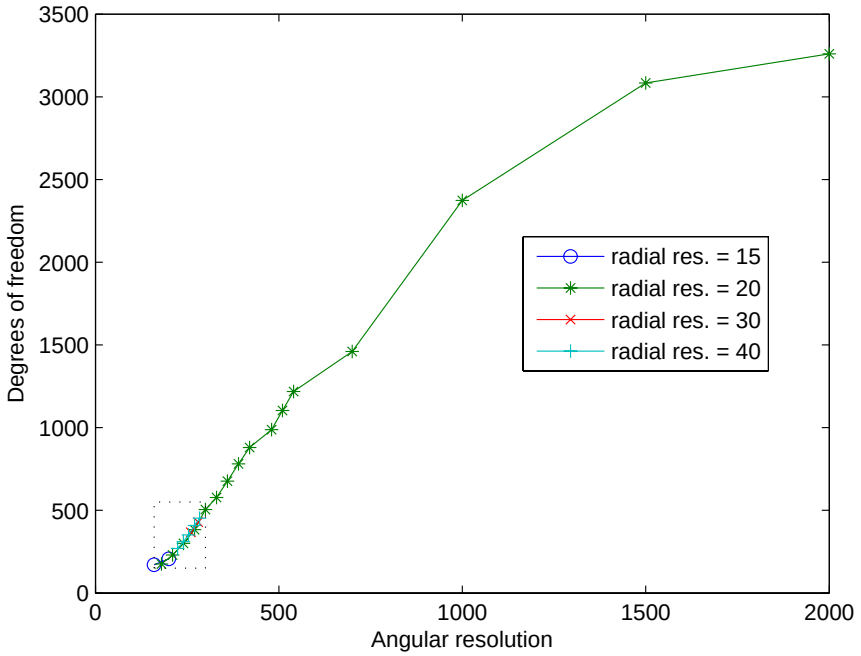


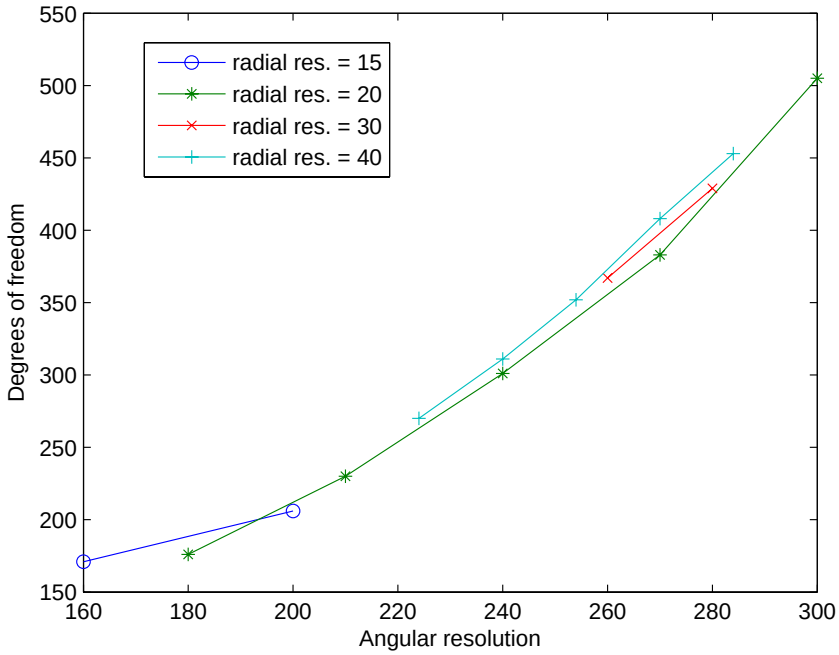
Figure 5.2: Degrees of freedom for genuine match score distributions calculated using approach from (Daugman, 2003). Each curve represents templates of equal radial resolution.

35,000 pixels, therefore, we expect that any template having size bigger than $35,000 \times$ `number of filters` $\times$ 2 will not provide additional biometric information about the iris. However, even though converging of Daugman's discrimination entropy occurs, as can be seen from figure 5.3(b), it happens for much bigger than expected template dimensions.

Figure 5.3(b) shows in more detail the region of figure 5.3(a) where the behaviour of information content in templates of different radial resolutions is presented. Due to the computational complexity of the experiment, we could not obtain results for a large amount of different template sizes. Nevertheless, it is still clear that increasing radial resolution provides more degrees of freedom. However, this increase is insignificant compared to the effect created by growing angular resolution. Thus, for instance, templates of sizes $20 \times (240 \times 2)$ and $40 \times (240 \times 2)$ have an information difference of 10 degrees of freedom, while $40 \times (254 \times 2)$ templates provide 41 more degrees of freedom than $40 \times (240 \times 2)$ templates.



(a)



(b)

Figure 5.3: (a) All collected experimental data related to degrees of freedom for impostor match score distributions obtained using Daugman's discrimination entropy approach. Each curve represents templates of equal radial resolution. The longest curve represents behavior of information content for radial resolution of 20 and increasing angular resolution; (b) Zoomed area of (a), marked by a dotted contour, showing relationship between templates of different radial resolutions.

5.2.2 Statistical analysis

In this section we attempt to gain better understanding of the nature of Daugman's approach for measuring information content. The author introduced a new term — discrimination entropy. However, few explanations were provided in order to clarify its characteristics. It was claimed that degrees of freedom specify the amount of identifying iris information, which may resemble the definition of relative entropy. On the other hand, the algorithm is based solely on the distribution of inter-class comparisons without taking into account knowledge of variations in individual samples. This contradiction was the primary motivation for performing the statistical analysis presented below.

In order to understand what type of information is described by the discrimination entropy it was decided to learn the relationship between Daugman's measure and a known information criterion — entropy. Our choice was based on the fact that both measures rely on one distribution only.

To achieve this goal, the following model was deployed. First of all, we defined the term of an iris feature. Feature X_i is a bit of the iris template located at its i-th position. For our model, we defined a vector of M features:

$$\{X_1, X_2, ..., X_M\} \tag{5.1}$$

where each feature X_i can have a value of either 0 or 1.

As a next step, we assigned probabilities to all the possible combinations of M features that comprised 2^M different binary strings. In order to simplify the notation, we defined a random variable Y to represent binary strings of length M. Different patterns of probability assignment are presented later in this section together with experimental results.

The entropy of Y was then calculated according to the formula which was discussed in section 2.3.1:

$$H(Y) = -\sum_{j=1}^{2^M} p(y_j) \log_2 p(y_j) \tag{5.2}$$

where y_j is a j-th binary string. This gave us the amount of information in the modeled

template of length M.

From the other side, biometric information content may be measured using the approach offered by Daugman (2003) approximating HD distribution of the given set of binary templates by the binomial distribution. In order to do this for the model we presented, it was required to calculate differences between every pair of binary strings y_k and y_l before forming the distribution. This was identical to finding the probability of occurrence of every specific HD. If HD between bit strings y_k and y_l is defined as

$$HD_{k,l} = \frac{1}{M}\sum_{i=1}^{M}(y_{k_i} - y_{l_i})^2 \tag{5.3}$$

with y_{k_i} and y_{l_i} being two corresponding bits of compared templates, then the probability of HD having a value of h, $P(HD = h)$, is calculated as follows:

$$P(HD = h) = \sum_{i,j,H_{i,j}=h} p(y_i)p(y_j) \tag{5.4}$$

As an illustration, let us consider the case for $M = 2$. Possible HD values are 0, 1/2 and 1 with probabilities

$$\begin{aligned} P(HD = 0) &= p(00)^2 + p(01)^2 + p(10)^2 + p(11)^2 \\ P(HD = \frac{1}{2}) &= 2\Big(p(00)p(01) + p(00)p(10) + p(01)p(11) + p(10)p(11)\Big) \\ P(HD = 1) &= 2\Big(p(00)p(11) + p(01)p(10)\Big) \end{aligned} \tag{5.5}$$

where $p(b_ib_j)$ is an occurrence probability of a binary string b_ib_j.

Finally, the procedure for fitting a curve of binomial distribution was applied to the obtained HD distribution, resulting in parameters of the binomial distribution curve that provide the best fit into the data. One of the calculated values, the number of degrees of freedom, determined the bits of information in the template, according to the work by Daugman. After both information measures the standard entropy and discrimination entropy were obtained, their relative behaviour was studied to help us understand the nature of the

Daugman's measure.

The next step after designing the experimental procedure was developing a scheme for assigning probabilities to 2^M binary strings of length M, that served as a model of iris templates. We divided this stage into two different experiments. During the first test we decided to study the effect of scattering of probabilities by introducing the scatter parameter s which was defined over the interval $[0, 0.5]$. $s = 0.5$ corresponded to the highest spreading of the probabilities, while $s = 0$ defined the case when all the probabilities were equal. The motivation behind this was to see how information content behaves if the ideal case of equal probabilities for all templates does not occur. The assignment was done in following steps:

- Select 2^M random values out of the interval [0,1].
- Equalize the set so that the minimal value is 0 and the maximum value is 1.
- Multiply the values by s restricting them to the interval [0,s].
- Centralize the value set by adding to it $(0.5 - s/2)$.
- Ensure that resulting probabilities sum up to 1 by dividing every value by their total sum.

For each scatter parameter s, we performed 20 information calculations, each time using a new set of randomly selected template probabilities. The results of comparing discussed information measures for $M = 2$ may be seen in figure 5.4. The x-axis reflects the increase in probability scattering, while the y-axis shows the resulting bits of information. Mean values of information measurements are plotted along with corresponding standard deviations. The blue line represents the amount of degrees of freedom in the HD distribution for the selected probabilities. The green line is a standard entropy for the same set of probabilities. It may be seen that both curves behave in a similar way, converging to the maximum entropy value of M when the probabilities are equal, or the spreading is insignificant. Increasing probability spreading causes the entropy value to decrease. However, Daugman's measure becomes increasingly unstable for bigger scatter values. Its values remain bigger than the entropy and sometimes even exceed the maximal entropy, marked by the dotted line.

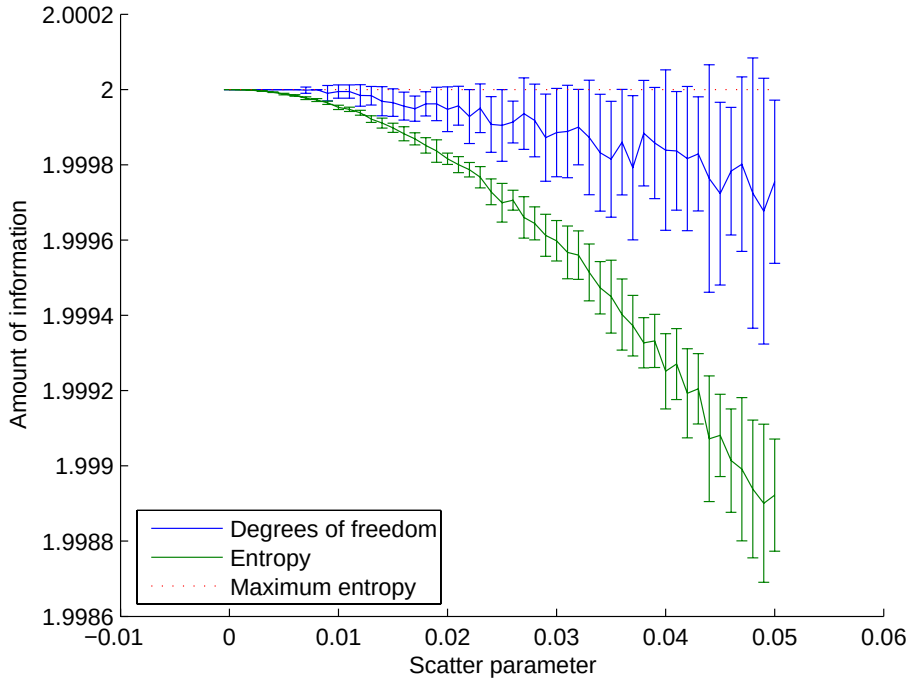


Figure 5.4: Comparing entropy and degrees of freedom for differently spread probabilities for binary strings of size $M = 2$. 20 rounds of experiments are performed. Mean values are plotted along with corresponding standard deviations. The blue line, representing Daugman's measure, and the green line of the entropy show similar behaviour for close to equal probabilities. Discrimination entropy measure becomes unstable for bigger scatter values. Its values remain bigger than the entropy and sometimes even exceed the maximal entropy (the dotted line).

The second experiment of comparing two discussed information measures involved analysis of templates with dependencies between their bits. In order to simplify the model, we chose to consider binary templates of size $M = 2$. The dependency was defined as a parameter d that changed between 0 and 1. The middle point $d = 1/2$ corresponded to the case when there was no dependency between different features, while $d = 0$ and $d = 1$ reflected extreme situations when the template bits were never the same or were always the same respectively. Firstly, the 3 cases for $d = 0, 1/2, 1$ were studied to help us build a general model:

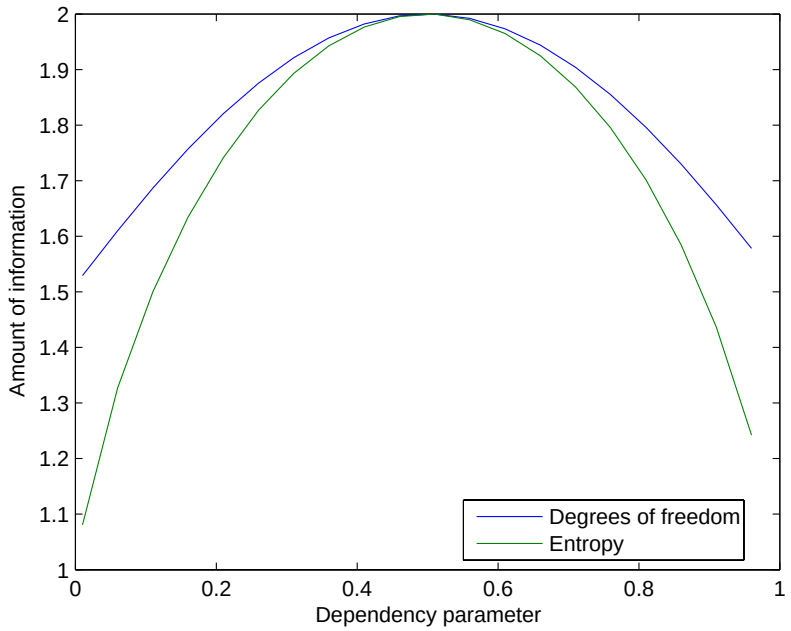


Figure 5.5: Comparing entropy and degrees of freedom for templates with dependencies for binary strings of size $M = 2$. The blue line, representing Daugman's measure, and the green line of entropy both show the maximum value of 2 bits at the point where probabilities are independent ($d = 1/2$). Increasing dependency affects amount of information in both cases. However, discrimination entropy always shows more bits of information compared to the entropy.

d	0	$\frac{1}{2}$	1
$\begin{bmatrix} P(00) & P(01) \\ P(10) & P(11) \end{bmatrix}$	$\begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}$	$\begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix}$	$\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$

From observing the sets of probabilities displayed above we concluded that the general relationship between 2D binary vectors probabilities and the dependency parameter d was the following:

$$\frac{1}{2}\begin{bmatrix} d & 1-d \\ 1-d & d \end{bmatrix} \tag{5.6}$$

Based on this relation, we built a number of probability sets for different dependency values and then performed information measure comparisons as described earlier in this section. The

results may be seen in figure 5.5. The x-axis shows changes in the dependency parameter, starting from $d = 0$ when template bits are never equal, through $d = 1/2$ for independent template content, and ending with $d = 1$ for the case when only templates 00 and 11 are possible. The green curve of the entropy and the blue curve of Daugman's information measure both reach their maximum at $d = 1/2$ when template bits are independent, while gradually decreasing for growing values of dependency. However, the discrimination entropy always remains bigger than the entropy.

5.2.3 Discussion

In previous sections we presented experiments that approached the task of studying the discrimination entropy introduced in (Daugman, 2003) from two different angles: by applying the method to the local database of iris images and by comparing the measure to the standard definition of entropy.

Calculating degrees of freedom

In the first part of the experiments we built a number of template sets for the same database using different template dimensions. After calculating HD distances, we used a curve fitting procedure of the binomial distribution with the match score distribution in order to determine amount of information provided by templates of different sizes.

During the process of experimental testing, a number of questions arose. The ones that we could answer, helped us to understand better the nature of the discrimination entropy and of the iris information content in general. The questions that remained unanswered will serve as a base for the future work.

First of all, we noticed that using a fixed value for the parameter n of the binomial distribution, corresponding to the number of comparisons used during the matching, did not provide an acceptable binomial curve fit. When we permitted n to be one of variable parameters of the searching routine, the fitting algorithm consistently resulted in n values that were smaller than actual number of comparisons (4-11 compared to 17). A possible explanation is that, due to internal dependencies of templates, the experimentally obtained

match scores cannot precisely fit the theoretical model developed by Daugman. Observed HD distributions were always wider than the model curves.

In our experiments we also used genuine distributions for performing information content measurements. This was done in an attempt to learn better the investigated technique. As expected, the results showed significantly smaller degrees of freedom, due to high similarities of compared templates. Additionally, fitted binomial curves had a higher parameter n, compared to results of experiments on impostor distributions. This may be explained by the fact that when templates are initially oriented in a similar way (no significant head tilts), high enough number of comparisons (in our case 17) involving different template rotations almost always result in a very narrow HD distribution. Thus this is sufficient for finding the best possible relative orientation.

The calculation of degrees of freedom for iris templates of growing sizes showed that the information content converges for bigger templates. This happens because iris codes of higher resolution are calculated using densely sampled iris pixels. As a result more dependencies are introduced, since neighboring pixels normally have related intensities.

Another relationship between template dimensions and the resulting bits of information was discovered. Changes in angular resolution influenced discrimination entropy much stronger than changes in radial resolution. The possible explanation is that dependencies along radial lines are significantly higher than those along the circular lines. On the other hand, this phenomenon may be explained by differences in iris circular length and its radial length. When radial resolution increases, pixels are sampled with higher density introducing more dependencies than in the case of angular resolution.

Another observation made while running experiments was related to using multiple filters for feature encoding. We noticed that increasing the number of filters does not change characteristics of distributions. This may be explained by a mistake in the software of Masek & Kovesi, or possibly another reason which will require future investigation.

Comparing discrimination entropy to standard entropy

The second part of experiments included comparing behaviors of Daugman's discrim-

ination entropy and the standard entropy, typically denoted as H. In order to do this, we used a simplified model of 2-bit templates with predefined probabilities that allowed us to calculate exact H values and then compare them to corresponding results of Daugman's algorithm. The simple model was sufficient for our experiments, because it allowed us to learn discrimination entropy behavior without involving complicated theoretical derivations.

The experiment, based on sets of template probabilities of varying scattering, attempted to model a real world scenario when some templates are more likely to occur than others. In the case of small scattering values, both investigated measures showed similar behaviour, hinting that in the ideal situation of nearly equiprobable templates, Daugman's information measure behaves as a standard entropy H. When template probabilities were allowed a higher degree of variance, we observed an increasingly unstable behavior of the discrimination entropy. This may be an indication of the fact that the model suggested in (Daugman, 2003) assumes that iris templates all have equal probabilities, resulting in undefined behavior for other cases.

In the second experiment of this stage we modeled dependencies between bits of templates. This was done to simulate a real world situation involving dependencies between neighboring pixels of irises. Again, both investigated measures showed identical behavior in the ideal case when no dependencies were introduced. This supported our previous assumption that Daugman's measure and H are identical under ideal conditions. The trend of both curves was also similar when stronger internal dependencies were introduced into modeled templates.

During this stage of experiments, another important observation was made. Discrimination entropy appeared to be bigger than the standard entropy when the modeled conditions were different from the ideal case. The more scattered, or dependant, were template probabilities, the bigger the difference between entropies became. Moreover, in the first experiment, Daugman's information measure sometimes resulted in slightly higher than maximal entropy values. In future work we will attempt to explain this phenomenon.

Conclusion

The discussed above experiments showed that the technique for measuring information content proposed in (Daugman, 2003) does not necessarily provide a useful way of determining amount of biometric information in the iris. Firstly, the proposed theoretical model is too idealistic and does not closely fit the experimental measurements, as in the case of approximating HD distribution by a binomial distribution curve when we had to introduce more flexibility into the algorithm in order to achieve satisfying results. Secondly, the model becomes unstable when templates are given different occurrence probabilities in an attempt to model a real world scenario. Finally, the proposed criterion measures the overall amount of information in the iris but does not measure the amount of biometric information used for person identification, because it is only based on inter-person comparisons and does not involve variations of features of the same individual which is one of the main components of relative entropy. This statement was supported by experimental results that showed correlation between the standard entropy H and the discrimination entropy of Daugman.

5.3 Relative entropy

In this section we present experiments in which we used the approach by Adler *et al.* (2006) for measuring the amount of biometric information in irises. Three kinds of information measurement were performed: for iris features, for the whole iris and for iris regions. Certain preparation steps and algorithm modifications were required which are discussed below.

Feature definition.

First of all, experiments required a definition of the iris feature. Since we used the iris recognition software by Masek & Kovesi (2003) for all our experiments, it was decided to employ its feature encoding mechanism based on 1D Log-Gabor wavelets. In order to make features more appropriate for relative entropy calculations, their non-quantized form was used, giving us two real valued features for each pixel of normalized iris corresponding to real and imaginary parts of the phasor.

Feature extraction.

Feature extraction was done by modifying the program by Masek & Kovesi to output an additional array at the encoding stage. After the Log-Gabor filter was applied to a row of normalized iris and before the results were quantized, real and imaginary parts of phasors were stored aside in a separate array. Then, using the validity mask, the feature set was updated to have Not-a-Number (NaN) values (special values that signify undefined expressions) at non valid locations to give us an indication where features were not defined.

In this way, features were extracted for each iris image in the local database of 360 eye images and stored as additional files. As a result, after the software by Masek & Kovesi was applied to an eye image, 3 output files were obtained: the iris template, the validity mask and the feature set.

Feature database building.

After features were extracted and stored aside, it was possible to build the database of iris features. The biometric content algorithm by Adler *et al.* is based on analyzing the behaviour of each feature throughout the whole collection of iris templates. However, the way the iris encoding was performed did not ensure that corresponding bits in different iris templates described the same feature due to rotational inconsistencies. Therefore, the alignment stage was required to adjust feature sets so that locations k for all the templates corresponded to the same feature x_k.

We designed the alignment procedure that cyclically rotated iris templates until feature locations matched. It was comprised of the following steps:

- For each eye in the eye image database, select one template to serve as a reference.
- For the rest of templates of the individual, match each template against the selected reference using the matching algorithm. Keep the rotation value that resulted in the best HD.
- Rotate iris templates, validity masks and feature sets using corresponding misalignment values from the previous step.

- Perform cross comparisons without using the shifting scheme for the set of rotated templates to validate the alignment. Keep the set that gives the best HD distribution for further experiments.

- Repeat the steps above using different reference templates.

After the described procedure was repeated for every eye, the feature sets were gathered into the feature database. They were grouped by their belonging to different eyes, resulting in the list of feature matrices, where each matrix represented one eye and had as many rows as there were samples for the eye, with the number of columns defining the number of features used. In the case of the local eye database that was used for our experiments, the feature database contained 12 feature matrices of size 30×9,600.

5.3.1 Feature entropy

In the first round of experiments, we used the method from (Adler *et al.*, 2006) to determine the amount of information in each iris feature.

As a preparation step, it was decided to remove features that have NaN values in one or more templates, to ensure that calculations uniformity was not affected. Finally, we had 3,768 out of 9,600 features left - those that were valid throughout the whole database.

The distribution of information values calculated for 3,768 features and averaged over all eyes is shown in figure 5.6. The values spread between 0.2804 and 1.6268 bits, with mean of 0.8535 bits and standard deviation of 0.2129.

In order to help understand the relationship between feature information content as a function of the corresponding iris pixel location, we display relative entropy of features graphically as shown in figure 5.7. The data is presented in the form of the iris template where each point reflects the amount of information in the feature at that location. Brighter points correspond to features with higher information content, while darker ones show less significance. Points with zero intensity indicate features that were not included in calculations due to occurrences of NaNs among their values. It must be noted, that each pair of horizontally adjacent points of data in figure 5.7 represents one pixel of the normalized iris,

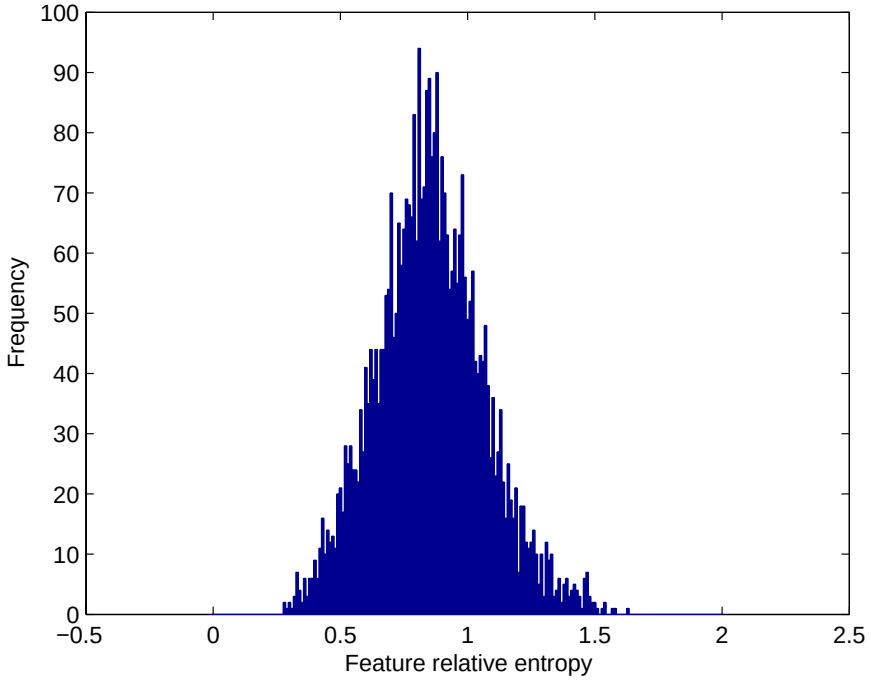


Figure 5.6: Distribution of feature information content, averaged over all eyes, as calculated for 3,768 iris features that do not have NaN values in any template of the database.

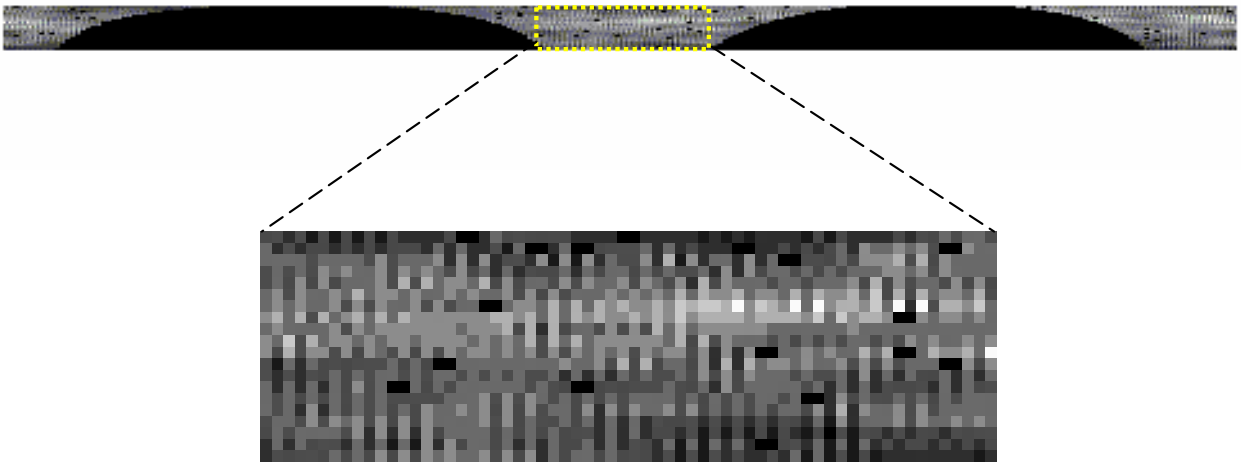


Figure 5.7: Spatial distribution of feature information in the iris template, averaged over all eyes. Brighter points correspond to features with higher information content, while darker ones show less significance. Black points reflect locations where feature relative entropy was not calculated due to non-valid values in one or more iris templates.

due to the fact that encoding provides us with two features - real and imaginary parts of the phasor.

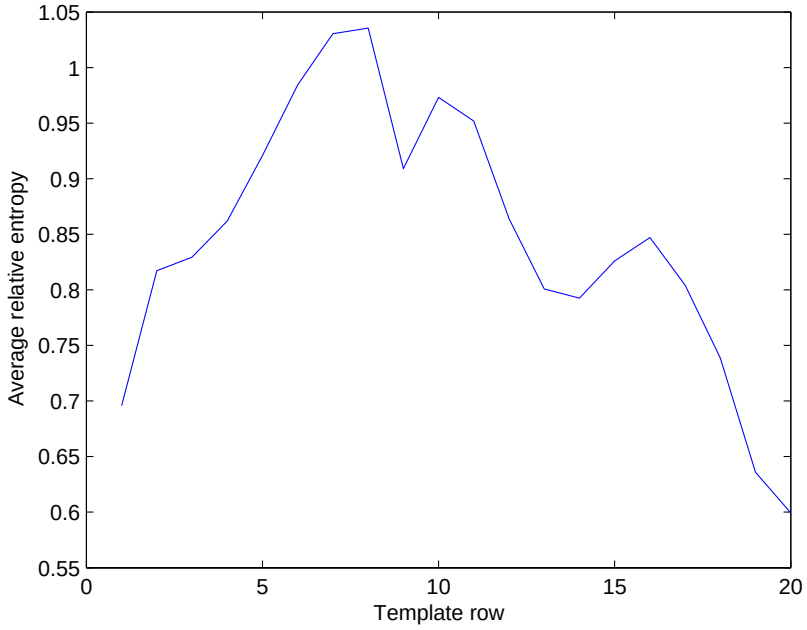
Figure 5.8 shows another interpretation of experimental results. The curve points in 5.8(a) were calculated by averaging the amount of feature information along rows of the template, or along circles inside the iris. The higher the row number is, the further the circle is from the pupil. The discussed experiments were based on 20×480 templates (radial resolution = 20, angular resolution = 240). It may be seen that features extracted from iris pixels located along rows 7 and 8 are the most informative, while features closer to the outer iris boundary are less important.

Similarly to 5.8(a), 5.8(b) presents feature information as a function of the radial angle inside the iris. This was done by averaging information values along template columns. Whenever there were no valid features in the column for performing calculations, the information value was set to zero as an indication of the case. Features resulting from real and imaginary parts of the phasor are presented by blue and green curves respectively. This is a more accurate way of data plotting which allowed us to study the relationship between two types of features. Figure 5.8(b) shows slightly higher information content in features extracted from the lower half of the iris. Another observation was made by studying the relative behaviour of two curves. It may be seen that whenever the curve of real features has a local maximum, the curve of imaginary features has a local minimum, and vice versa. This effect can also be seen in 5.7 where neighboring columns have contrasting intensities.

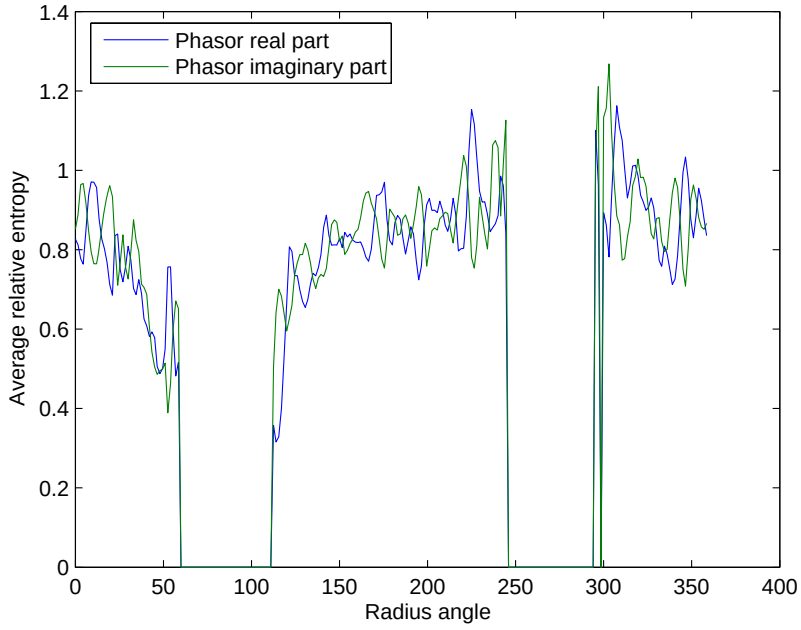
5.3.2 Iris template entropy

In this section we discuss experiments that were done to calculate the amount of biometric information contained in each of 12 irises in our database.

Iris features, as they were defined in this work, have variations of a very small degree. This fact caused instabilities when a large number of features were selected for relative entropy calculations. The situation did not improve significantly even for smaller regularization parameter L values (see chapter 2.4.2 for the description of L). Therefore, it was decided to



(a)



(b)

Figure 5.8: (a) Information content averaged along template rows, or circles inside the iris, where smaller row numbers correspond to circles closer to the pupil. The inner half of the iris seems to provide features with higher information content, with rows 7 and 8 having the most informative features. (b) Feature relative entropy averaged along template columns, or radial angles of the iris. Features from the lower part of the iris show slightly higher information content. Features obtained from real and imaginary parts of the phasor are presented by blue and green curves respectively. Columns with zero information values indicate that there were no valid features to perform the calculation.

limit the number of features involved in the calculation.

Similar to the experiment of feature relative entropy calculation, we firstly removed all the features that appear to have NaN values in one or more iris templates. As a next step, we used remaining valid features to randomly select a smaller subset for the calculations. The fact that features were selected randomly helped us to ensure that the iris template was represented in a uniform way.

Another means of making the algorithm more stable for the case of iris templates was feature normalization. Since absolute feature values were very small with mean of approximately 0.007 and standard deviation of also 0.007, we chose to increase their absolute values by multiplying them by 100. For our purposes, this operation did not affect final results, because entropy is not calculated based on elements' values but rather based on their probabilities.

After features were selected and normalized, the rest of the algorithm was applied unchanged. Taking the number of features to use and the regularization constant L as input parameters, the procedure outputted a vector of information content values - one element per eye. After that, the relative entropy values were averaged to obtain an estimate of the biometric content in the iris template in general. Alternatively, the results can be viewed without averaging to compare information content in samples of different eyes. The Matlab® implementation of the described above procedure is given in appendix C.

For all the experiments that we performed, the constant regularization parameter of 29 was used. This was the maximum possible value because for each eye in the database we had 30 samples. The reason we used the highest setting for L was that smaller values reduce information content while not contributing significantly into the stability issue.

As an experiment, we performed iris relative entropy calculations for varying feature amounts to see how this fact affects the overall information content of the template. The selected number of features changed from 30 to 270 with the step of 10. After each iteration entropy values of eyes were averaged for obtaining a general information estimate. The resulting curve is shown in figure 5.9. It may be seen that the more features are used the

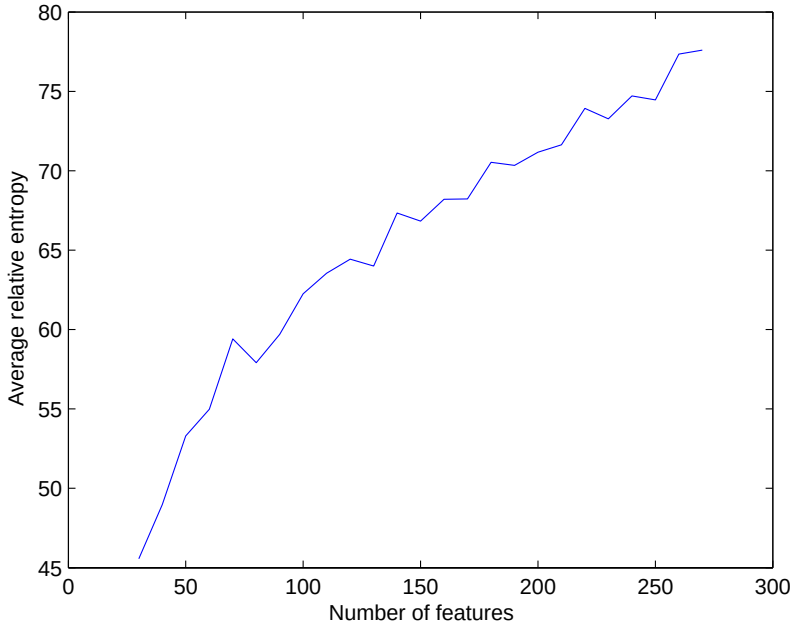
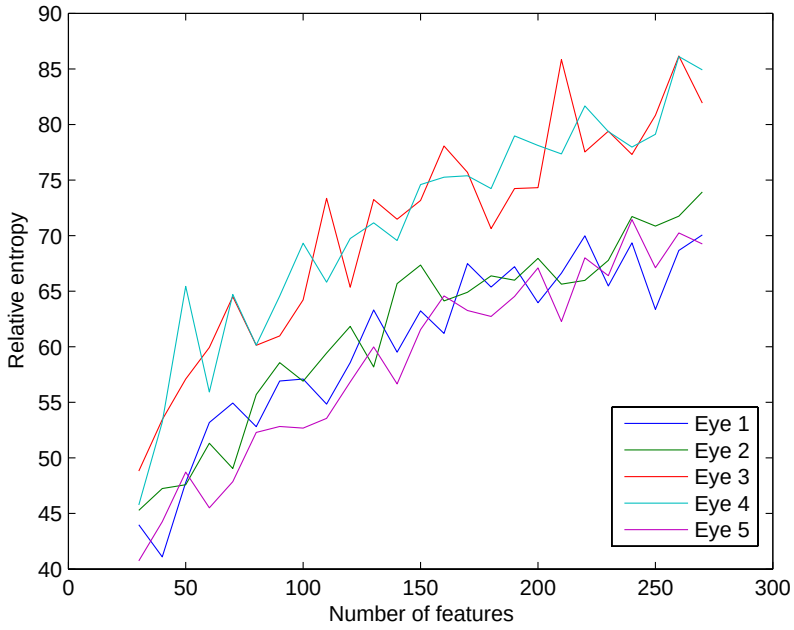


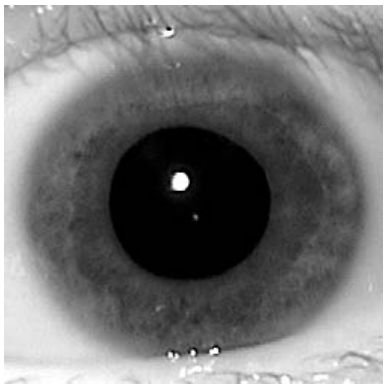
Figure 5.9: Averaged information content calculated for different feature amounts. The regularization parameter L is set to the constant value of 29.

more information is provided by the template. However, it is hard to analyze the curve trend since it was impossible to use more than 300 features out of available 3,768 valid ones due to computational instabilities.

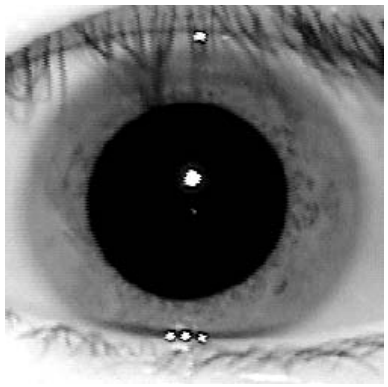
The partial results of the experiment described above are shown in figure 5.10(a) where information content is presented for five eyes separately, without averaging, as a function of number of features used. It may be clearly seen that eyes 3 and 4 provide iris templates of significantly higher information content than the rest. Additionally, an interesting observation was made: irises that demonstrated higher entropy belong to Caucasian participants, whereas irises with a lesser amount of information belong to Asian and African participants. Figures 5.10(b) and 5.10(c) present eyes 3 and 5 having higher and lower information content correspondingly. This issue should be investigated more thoroughly in the future by involving more people of different origins.



(a)



(b)



(c)

Figure 5.10: (a) Subset of experimental results from 5.9 corresponding to information content of five different eyes. Each curve represents relative entropy of one particular eye as a function of number of used features. Eyes 3 and 4 provide iris templates of significantly higher information content than the rest. (b) Eye 3 from (a) having higher information content. (c) Eye 5 from (a) having lower information content.

5.3.3 Iris region entropy

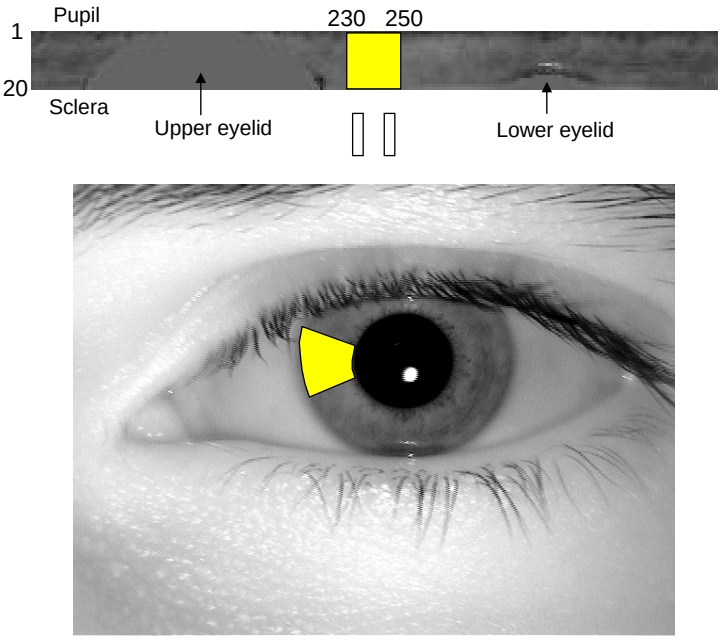
As a final stage of experiments, it was decided to introduce a possibility for restricting areas of the iris where features were selected. This way it became possible to learn which parts of the template were more informative than others.

Similarly to the previous stage of using the approach by Adler *et al.* for finding biometric information content of complete iris templates, we randomly selected a certain amount of features for the processing. However, this time we chose them from inside a defined rectangular region, passed as an input parameter to the algorithm. The Matlab® implementation of this algorithm is given in appendix C.

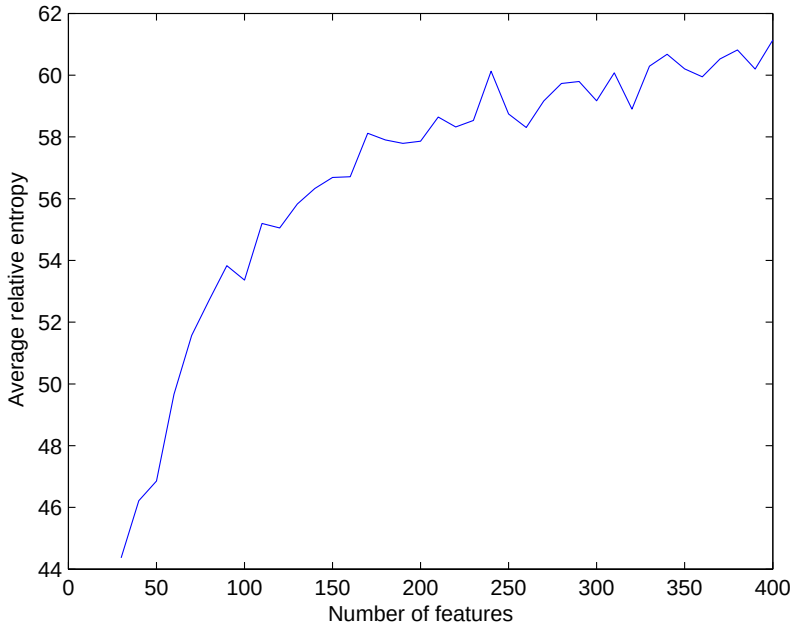
Firstly, we chose to experiment with a template region defined between columns 230 and 250, and over all the rows, as shown in figure 5.11(a). We performed a number of relative entropy calculations for different amount of features, from 30 to 400 with the step of 10. This allowed us to study information content behaviour of the template when features are sampled with varying density, avoiding instability issues as was mentioned in previous section. Since the amount of valid features inside the discussed region was 408, it was possible to calculate the information content for the case when features had the most possible density. For every iteration we obtained 12 entropy values, one for each eye, which were then averaged and plotted as showed in 5.11(b). This time we can see that for higher saturation levels the information content of the template converges. For the region we selected, this value is approximately 60 bits.

For the next experiment, we repeated the attempt to compare information content in iris template rows and columns, as was done for feature relative entropy calculations (see section 5.3.1). This time, instead of averaging feature information, we defined rectangular regions of a unit width - one for each row or a column - and then performed information measurement for each region separately.

In the case of row comparison, we chose to use 70 randomly selected features for each iteration of calculations. The relative entropy results are displayed in figure 5.12(a) where information provided by template rows, or circular regions of the iris, is shown as a function



(a)



(b)

Figure 5.11: (a) Selecting a rectangular iris region between the template locations (230,1) and (250,20). The corresponding area is marked on the original eye image. (b) Analyzing information content for a rectangular region shown in (a) for different feature amounts. For high feature saturation, the amount of information converges to approximately 60 bits.

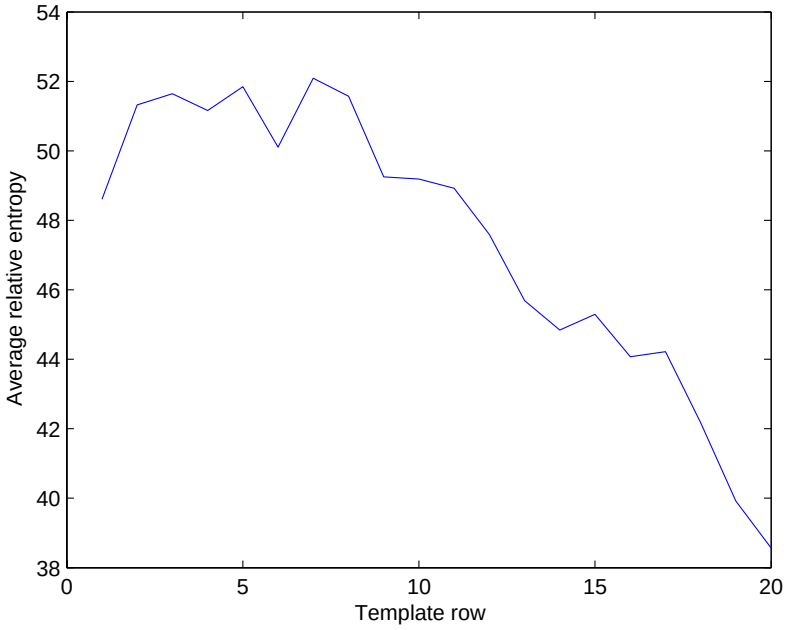
of the row location. Similar to results reflected in figure 5.8(a), the most of biometric content is located in the parts of iris closer to the pupil, while the further regions have significantly less information. Rows with the highest entropy values are again 7 and 8.

Finally, when template columns, or different radial angles, were compared, the calculations were performed using 10 randomly selected features. In the case when a column showed no sufficient amount of valid features to choose from, the relative entropy was set to 0 to indicate angles for which the value was not calculated. In a similar to 5.8(b) manner, we divided the gathered entropy data into two sets - one for each type of features. Information measurements of regions calculated based on real and imaginary phasor components is plotted with a blue and a green curves respectively. Again, we observe the effect of opposing behaviour in the results of two types. Whenever there is a local information increase in one curve, the other curve shows a decrease, and vice versa. Comparisons of entropy values as a function of the radial angle show very little difference. It may be argued that the lower part of the iris shows more information content, however, it is marginal.

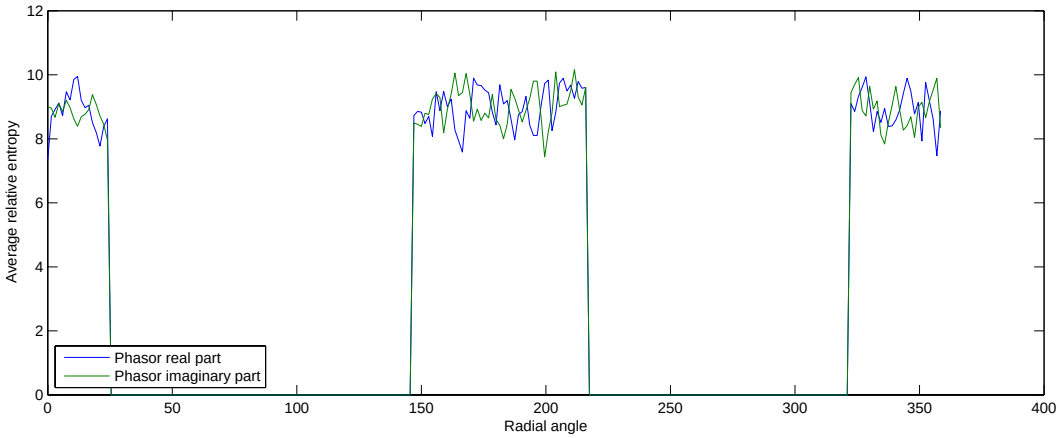
5.3.4 Discussion

In previous sections we presented experiments aimed at measuring biometric information content of the iris using the method developed by Adler *et al.*. A number of experiments were designed to understand the information distribution in the iris. We attempted to perform measurements for single features, for whole templates and for certain iris regions. Due to limitations of the used data and of the algorithm itself, we were unable to fulfill some of the set goals. However, many useful results were obtained and many important observations were done, thus, helping us to clarify the studied subject.

Initial stage of experiments involved measuring biometric content of single features. The results showed that feature information varies from 0.28 to 1.63 bits with the mean of 0.85 bits depending on the feature location. It was shown experimentally that pixels located along central circles of the iris, closer towards pupil, provide the highest amount of biometric information. This experiment helped us to understand better what pixels of the iris are more



(a)



(b)

Figure 5.12: (a) Information content as calculated for each template row separately by defining iris regions of a row size. The measurements were done using 70 randomly selected features. (b) Information content as calculated for each template column separately, by defining iris regions of a column size. The values are gathered into two sets - real phasor component features and imaginary phasor component features, represented by blue and green curves respectively. The measurements were done using 10 randomly selected features in each column. Zero values correspond to columns (radial angles) for which relative entropy was not calculated due to insufficient amount of valid features.

informative than others. Calculating the information content of the whole iris was impossible at this moment. A naive approach of estimating the total biometric content by summing up all feature information values would result in tremendous amount of bits — 3,241.8. This is an erroneous approach that does not account for dependencies inside the iris that unavoidably minimize the amount of biometric information.

The next experiment approached measuring information content of the whole template. Measurements were done for increasing amount of features, in order to find a value that relative entropy converges to. This value would have described the information content of the whole iris. Convergence of the entropy is explained by earlier mentioned dependencies between neighboring pixels of the iris. The more features are used for the calculation, the more dependencies are introduced, because of increasing density of involved iris pixels. Unfortunately, the algorithm of Adler *et al.* applied for this type of iris features demonstrated unstable behavior by being unable to perform calculations involving more than 300 features. As a result, we were only able to observe an insignificant convergence of relative entropy that stopped at the value of 77 bits.

Another experiment, designed to calculate relative entropy of certain iris regions, allowed us to fully observe the phenomenon of iris pixel dependencies. By selecting a 20×20 template region we were able to observe a clear relative entropy convergence, occurred at the point when no significant information was introduced by added features.

Additionally, we repeated an experiment done before using single feature information for distinguishing regions of higher importance in the iris. The results from both experiments showed similar curve trends, assigning inner iris circles, those that are closer to the pupil, more significance; even though features of inner circles are sampled more densely than features of outer circles and have more dependencies between them. This result supported theoretically an observation that was made while inspecting the database of iris images - it may be seen by a bare eye that this is a region where irises have the most reach patterns.

One of the observed behaviors of the relative entropy of the iris remained unexplained. We noticed that information content calculated for radial lines of the template had depen-

dency between radial lines containing features obtained from the real part of the pixel encoding and radial lines corresponding to the imaginary parts of the same encoding. Whenever the curve corresponding to real features had a local maximum, the curve corresponding to imaginary ones had a local minimum, and vice versa. One of possible reasons may originate from the nature of 1D Log-Gabor wavelets. This issue will have to be explored in more detail in the future work.

Conclusion

The discussed above experiments employed the approach for measuring biometric information content proposed in (Adler *et al.*, 2006). This approach measures relative entropy of biometric templates which is a more appropriate measure of biometric information than H, because it helps to assess how identifying a certain biometric is. The performed experimental work helped us study the biometric information distribution inside the iris and the relationship between pixels of the iris. Due to limitations of the features used and the algorithm design, at this stage we were unable to calculate the relative entropy of the whole iris template. The future work will attempt to address this matter.

Chapter 6

Conclusion

The objective of this thesis was the measurement of information for identification in irises.

Before approaching the issue of biometric information measurement, we addressed the field of biometric sample quality. No proper identification can be done if a quality of an involved sample is poor. Measuring the amount of information based on images of low quality will normally show a lesser biometric information content, compared to high quality samples (Youmaran & Adler, 2006). In our study, we investigated human ability to assess biometric sample quality. Several participants were asked to evaluate quality of given biometric samples. Then, human evaluation was compared to quality scores obtained automatically.

The results showed that, in general, human evaluators are consistent with each other, as well as different algorithms are consistent within their group. However, comparing different types of assessment showed low or no correlation. Thus, we concluded that intuitive quality assessment done by human evaluators may not be sufficient for biometric identification applications since algorithm decisions are often based on features seemingly insignificant or not visible for humans.

After image quality is assured, the particular identification system precision may be estimated by measuring information content of the involved biometric based on produced templates. One of existing approaches, proposed in (Daugman, 2003), introduced a term of discrimination entropy for measuring the amount of information in irises. In order to clarify the technique, we, firstly, applied it to sets of iris templates of different dimensions calculated

for the local database; secondly, we performed comparison of the discrimination entropy to the standard definition of entropy, H.

The study revealed that discrimination entropy does not necessarily provide a useful way of determining the amount of biometric information in the iris. Firstly, we discovered that the proposed theoretical model was too idealistic and did not fit the experimental measurements sufficiently. In the case of approximating the match score distribution by a binomial distribution curve, we had to introduce more flexibility into the algorithm in order to achieve good agreement. Secondly, the approach demonstrated instability when we modeled conditions different from ideal ones. Finally, we concluded that the proposed criterion measures the overall amount of information in the iris but does not measure the amount of biometric information used for person identification, because it is only based on inter-person comparisons and does not involve variations of features of the same individual which is one of the main components of relative entropy. This statement was supported by experimental results that showed correlation between the standard entropy H and the discrimination entropy of Daugman.

As a final step, we approached the thesis objective by using the algorithm of Adler *et al.* for the local database of iris images. Using this method, we attempted to calculate information content in single features, in the whole iris and in its various regions which helped us to learn the degree of informativeness of different iris areas.

The approach from (Adler *et al.*, 2006) measures relative entropy of biometric templates which is a more appropriate measure of biometric information than H, because it helps to assess identification power of biometric templates. The performed experimental work helped us to detect the most informative regions of the iris — inner iris circles that are located closer to the pupil. This knowledge may help to improve efficiency of future recognition systems. Additionally, we discovered that the relative entropy converges after a certain density of used iris pixels is reached. This phenomenon is explained by dependencies between closely located pixels of the iris. Due to limitations of selected features and the algorithm design, at this stage we were unable to calculate the relative entropy of the whole

iris template.

For the described above experiments, we used a set of iris images obtained using L.G. iris camera that was provided by the staff of the Canada Border Services Agency (CBSA). The database contained 360 640×480 grayscale samples of eyes captured from 7 participants of different ethnicities.

6.1 Future work

As a subject of information content of irises was researched, many questions and related issues arose. Some of them were answered and presented in this thesis, others may be included as a part of the future work.

- Firstly, the issue of instabilities of the algorithm by Adler *et al.* should be investigated in order to allow full iris information measurements. The problem solution could involve a different choice of features or modifications to the actual algorithm.

- Another important issue requiring attention is the further analysis of the behavior of the discrimination entropy as mentioned in section 5.2.3, in order to understand why it results in higher than H values and sometimes even reaches levels above the maximal entropy.

- During our experiments we noticed that increasing the number of 1D Log-Gabor filters for feature encoding did not affect characteristics of Hamming Distance (HD) distributions. This may be caused by a mistake in the software of Masek & Kovesi, or another reason which will require future investigation.

- Experiments measuring information content of the iris showed that information content may depend on the ethnicity of the person (see section 5.3.2). This subject may be further explored by using bigger iris databases where many ethnicities are represented by a sufficiently large number of people.

- Additionally, in the future work, we should consider investigating dependencies between image quality and information content of irises.

- Finally, more work may be done in order to improve performance of the IR camera presented in appendix A.

Bibliography

Adler A, Youmaran R and Loyka S (2006). Towards a measure of biometric information.

Adler F (1965). *Physiology of the Eye*. St. Louis, MO: Mosby.

Alter O, Brown P and Botstein D (2000). Singular value decomposition for genome-wide expression data processing and modeling. *Proc Natl. Acad. Sci.*, **97**, 10101–10106.

Barry C and Ritter N (2003). Database of 120 greyscale eye images. Lions Eye Institute, Perth Western Australia.

Boles W and Boashash B (1998). A human identification technique using images of the iris and wavelet transform. *IEEE Trans. Signal Processing*, **46**.

Bucknam R and Christensen R (2003). Protecting our nation: Biometrics in government. *Biometric Consortium Conference, Arlington, USA*.

Chinese Academy of Sciences (2003). Database of 756 greyscale eye images. Institute of Automation.

Cover T and Thomas J (1991). *Elements of Information Theory*. New York: Wiley-Interscience.

Daugman J (1994). U.S. Patent No. 5,291,560: Biometric personal identification system based on iris analysis. Issue Date: 1 March 1994.

Daugman J (2002). How the afghan girl was identified by her iris patterns. Retrieved from *http://www.cl.cam.ac.uk/ jgd1000/afghan.html*.

Daugman J (2003). The importance of being random: Statistical principles of iris recognition. *Pattern Rec.*, **36**, 279291.

Daugman J (2004). How iris recognition works. *IEEE Trans. Circuits Syst. Video Technol.*, **14**, 21–30.

Draper B, Baek K, Bartlett M and Beveridge J (2003). Recognizing faces with pca and ica. *Computer Vision and Image Understanding*, **91**, 115–137.

Field D (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am.*, **4**, 2379–2394.

Frankel K, Sousa S, Cowan R and King M (2004). Concealment of the warfighter's equipment through enhanced polymer technology. *25th Army Science Conference, Orlando, USA*.

INCITS (2005). Biometric sample quality standard draft (revision 4). M1/06-0003, Information Technology Industry Council (ITI).

International Biometric Group (2005). Independent testing of iris recognition technology. Retrieved from *http://www.biometricscatalog.org/itirt/itirt-FinalReport.pdf*.

Kronfeld P (1968). The gross anatomy and embryology of the eye. *The Eye*, **1**, 1–66.

Kullback S and Leibler R (1951). On information and sufficiency. *Annals of Mathematical Statistics*, **22**, 7986.

Lim S, Lee K, Byeon O and Kim T (2001). Efficient iris recognition through improvement of feature vector and classifier. *ETRI Journal*, **23**, 61–70.

Ma L, Wang Y and Tan T (2002). Iris recognition using circular symmetric filters. *National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences*.

Mak M and Thieme M (2003). Study project proposal. Evaluating multi-modal biometrics systems: Concepts of operation and methods of performance evaluation. M1/03-0141, International Biometric Group.

Mann I (1950). *The Development of the Human Eye*. New York: Grune and Stratton.

Mansfield A and Wayman J (2002). Best practices in testing and reporting performance of biometric devices. NPL Report, version 2.01.

Masek L (2003). Recognition of human iris patterns for biometric identification. Requirement for the Bachelor of Engineering degree, University of Western Australia.

Masek L and Kovesi P (2003). 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.

The MITRE Corporation (2005). Image Quality Measure (IQM) v7.1. Retrieved in Apr., 2006 from *http://www.mitre.org/tech/mtf/*.

National Institute of Standards and Technology (NIST) (2002). NIST special database 18: Mugshot identification database (MID). Retrieved from *http://www.nist.gov/srd/nistsd18.htm*.

NSTC (2006). Biometrics overview. NSTC Subcommittee on Biometrics.

Renaghan J (1997). Etched in stone. *Zoogoer*.

Roesner J (2004). How to disassemble an olympus 2040. Retrieved in 2005 from *http://www.jr-worldwi.de/photo/index.html?mod_oly_dis.html*.

Ross A and Jain A (2003). Information fusion in biometrics. *Pattern Rec. Letters*, **24**, 2115–2125.

Shannon C and Weaver W (1949). *The Mathematical Theory of Communication*. Univ. of Illinois Press.

Tabassi E, Wilson C and Watson C (2004). Fingerprint image quality. NISTIR 7151, National Institute of Standards and Technology (NIST).

Wayman J (2004a). Biometrics now and then: The development of biometrics over the last 40 years. In *H. Daum (ed.) Biometrics in the Reflection of Requirements: Second BSI Symposium on Biometrics 2004*, SecuMedia, Bonn.

Wayman J (2004b). The cotton ball problem. *Biometric Consortium Conference, Washington DC, USA*.

Waymann J, Jain A, Maltoni D and Maio D (2005). *Biometric Systems: Technology, Design and Performance Evaluation*. Springer.

Wildes R (1997). Iris recognition: an emerging biometric technology. *Proc. IEEE*, **85**, 1348–1363.

Youmaran R and Adler A (2006). Measuring biometric sample quality in terms of biometric information. *Biometric Consortium Conference, Baltimore, USA*.

Zhu Y, Tan T and Wang Y (2000). Biometric personal identification based on iris patterns. *Proc. 15th Intern. Conf. Pattern Rec.*, **2**.

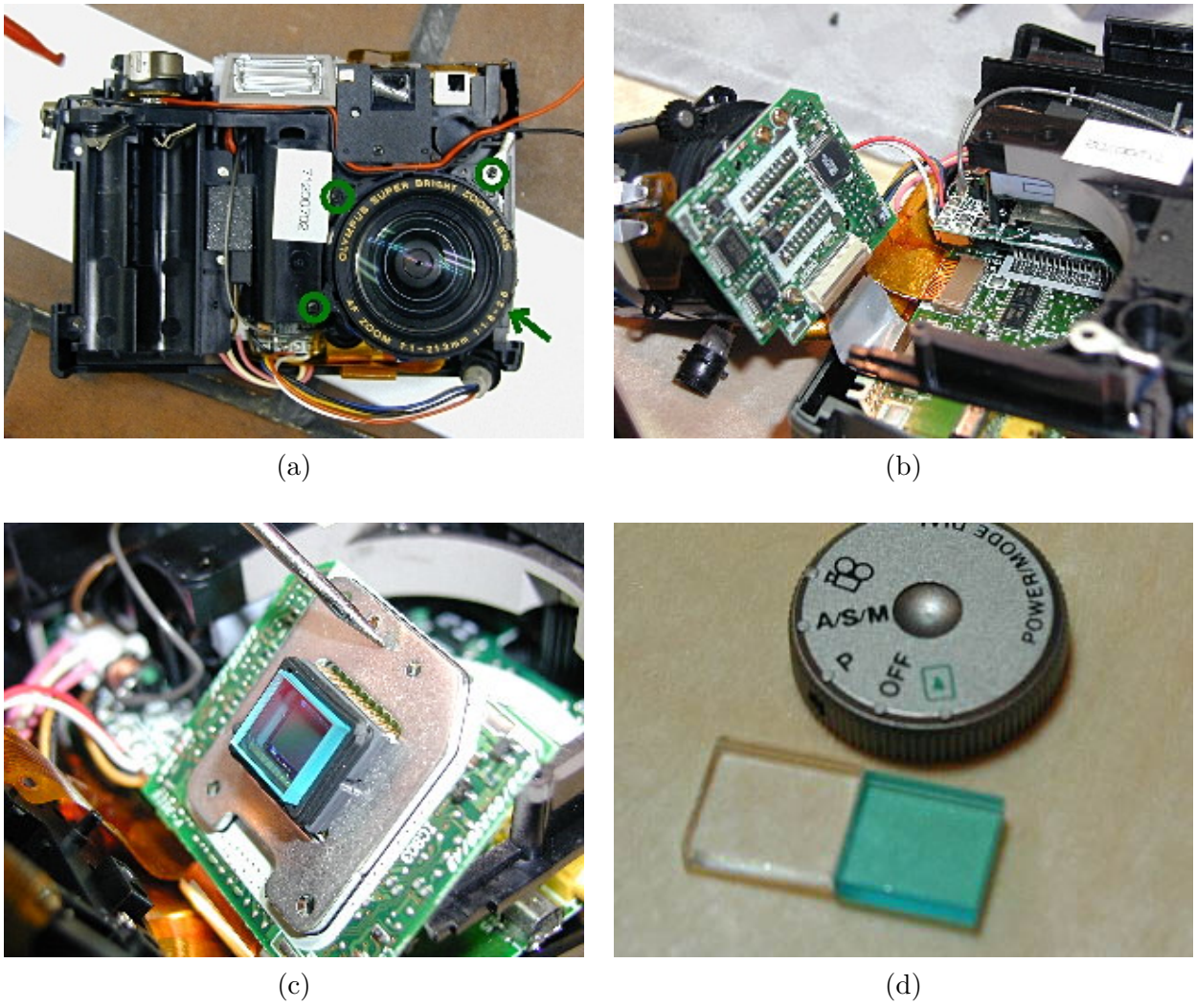
Appendix A

IR Image Capturing System

During the study we attempted to build an IR image capturing system for gathering our own database of eye images suitable for iris recognition. The motivation of this work was an absence of a large enough public data set with a well represented population of different origins and ages, taken in several sessions over a long period of time and under varying conditions. Having a database like this could have been a significant help in testing iris recognition technology and in developing new methodologies. Therefore, it was decided to, firstly, design a set up for capturing iris images and then to organize a number of iris capturing sessions involving students on the campus of the University of Ottawa. Due to a number of problems described below, we could not successfully use the system; however, it showed a lot of potential and can be improved in the future.

A.1 IR iris camera

The first step involved building the core element of the system - an IR iris camera. Most contemporary digital cameras have CCD sensors capable of capturing IR light. However, typically, they have a built-in IR cut filter for preventing color distortion. Therefore, we decided to purchase a used digital camera and to modify it by removing the filter. To do so, we used an online manual by Roesner (2004) giving detailed step-by-step instructions for disassembling an Olympus C-2040 Zoom camera or any other device of this series. The



(a) (b)

(c) (d)

Figure A.1: The process of Olympus C-2040 Zoom digital camera disassembly in order to remove the IR cut filter (images presented by permission of Roesner (2004)). (a) The digital camera with the front cover removed; (b) detached CCD board; (c) IR cut filter in its original position; (d) the replacement glass plate.

camera used for modification was Olympus C-4000 Zoom. As may be seen in figure A.1 from snapshots downloaded from the website of Roesner presented here with the author's permission, after the device front cover was removed and the CCD board detached, we were able to reach the IR blocking filter and to substitute it with a clear glass of matching size.

As a next step, we acquired a Hoya 58mm R72 IR filter in order to limit the light range entering the camera to IR frequencies. The filter was attached externally in front of the camera lens with the help of a lens adapter and a number of step up rings.

Since eyes do not emit a sufficient amount of heat, an additional source of IR illumination was required. This was accomplished by building a collar around the camera lens with 6 IR light-emitting diodes (LEDs) of 890nm wavelength attached to a power source.

A.2 Image capturing system

In order to facilitate the image acquisition process, we decided to develop a software system that would allow one to operate the modified Olympus camera via the USB connection and to store resulting eye images on the hard drive.

To achieve this, we used the Olympus Software Development Kit (SDK) that includes an interface written in the C programming language for controlling the camera through software applications. As a next step, we decided to use the Python programming language and the wxPython library containing graphical user interface (GUI) tools to develop an application that could provide us with the required functionality for building a database.

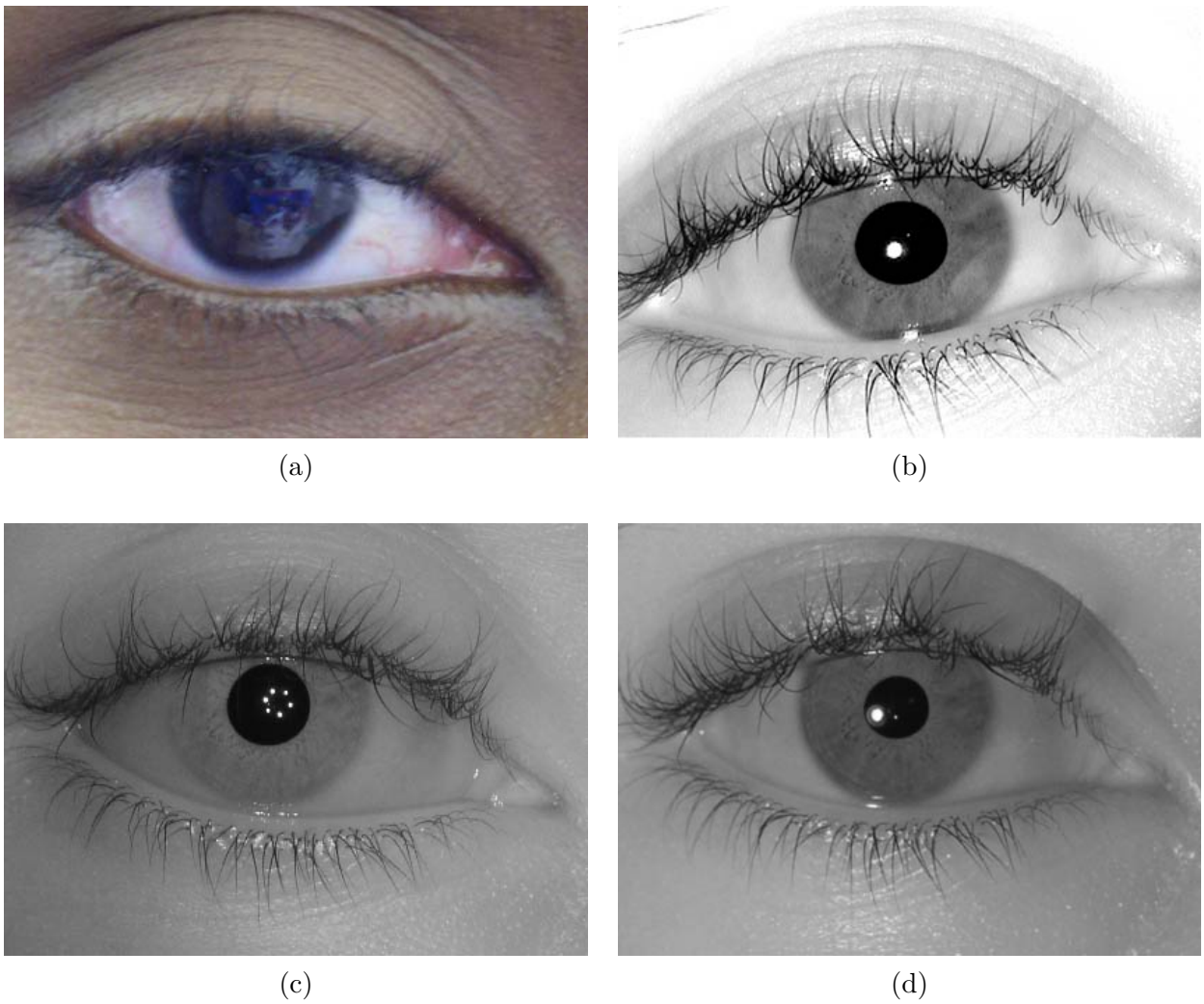
The final program allowed us: 1) to create user profiles for storing samples of their eyes on the hard drive, 2) to use existing profiles in case an additional image acquisition session is performed and 3) to capture eye images by operating the iris camera through the GUI.

A.3 Results

Before the modified Olympus C-4000 Zoom camera was integrated into the image acquisition system, it was tested on a number of simple objects. One of testing photographs is shown in figure A.2. It was taken during the day in the summer using daylight as the only illumination source. An interesting effect may be seen - trees, which normally appear dark on common photographs, acquired a bright glow. This is explained by the fact that foliage is strongly reflective of the near IR light (Frankel *et al.*, 2004). On the other hand, the sky appears significantly darker because the emitted waves of wavelengths between 400 and 500nm are mostly blocked by the IR filter.

Figure A.2: Outdoor IR image illustrating different reflectivity of objects under IR light. Foliage is very bright due to its ability to reflect most of IR light, while the sky becomes dark because the emitted waves of 400-500nm are blocked by the IR filter.

The process of eye image acquisition revealed several problems. First of all, the system was highly sensitive to subject movements due to the long exposure values of the camera. A possible future solution may involve building a head rest for helping participants minimize head movements. Secondly, the amount of iris detail in resulting images appeared to be not sufficient for successful iris recognition. A possible explanation is that the illumination used does not provide light of high enough frequencies able to capture small details of the iris



(a) (b)

(c) (d)

Figure A.3: Different iris imaging techniques as applied to the same eye: (a) original Olympus C-4000 Zoom camera; (b) professional L.G. iris camera; (c) modified Olympus C-4000 Zoom with 6 IR LEDs of 890nm wavelength; (d) modified Olympus C-4000 Zoom camera with a light bulb as an illumination source.

texture. Figure A.3 compares several images obtained using different imaging techniques. Original image taken using regular digital camera is presented in A.3(a). Due to specular reflections and a dark coloring of the iris no texture details are visible. A.3(b) is an eye image captured with an L.G. iris camera used in a number of commercial iris recognition applications in different areas. In A.3(c) we show an IR image as captured by modified Olympus camera from a distance of about 15cm using 6 IR LEDs emitting light of 890nm wavelength. The reflection of the LED camera collar may be easily seen in the pupil. This eye image showed significantly smaller amount of fine details in the iris compared to A.3(b).

An attempt to improve this result was made using a regular light bulb that provides wide range of wavelengths including near IR light. The obtained image, presented in figure A.3(d), showed some improvement in detail contrast. However, this still was not enough for successful iris recognition. Match scores calculated for the presented images were in the range 0.48-0.5 resulting in a failed recognition.

Therefore, we had to temporarily terminate the project. Instead, the database described in section 5.1 was chosen for the research experiments. In future, we intend to investigate the matter further and to improve the system performance so that captured eye images are suitable for iris recognition applications.

Appendix B

Binomial Distribution Fitting Code

This appendix lists the Matlab® code used for fitting the curve of binomial distribution into a given distribution of match scores. The detailed description of the procedure is given in section 5.2. The function `binomfun` defines the binomial distribution curve for given parameters and returns the distance between the curve and the data. The function `find_min` iteratively calls `binomfun` in order to find a parameters that result in the best approximation.

```
%--------------------------------------------------------------------------
% BINOMFUN defines the binomial distribution curve for given parameters and
% returns the distance between the curve and the data
%
% Args:     params  - parameters defining a binomial distribution
%           data_x  - input data scale
%           data_y  - distribution values
%
% Returns:  sse     - distance between the approximation
%           curve_x - approximated curve scale
%           curve_y - approximated curve values
%
% (C) 2006, Tatyana Dembinsky
%--------------------------------------------------------------------------

function [sse, curve_x, curve_y] = binomfun(params, data_x, data_y)

% get parameters
N = params(1);
p = params(2);
n = params(3);

% build a curve
f0 = binopdf(0:N-1, N, p);
F0 = binocdf(0:N-1, N, p);
curve_y = n.*f0.*(1-F0).^(n-1);

% transform the distribution to fit the interval [0:1]
% interpolate both sets of data to share all the points
data_points = length(data_x)*N;
new_x_range = 0:1/(data_points-1):1;
curve_x = 0:1/(N-1):1;
curve_y_interp = interp1(curve_x, curve_y, new_x_range);
```

```
data_y_interp = interp1(data_x, data_y, new_x_range);

% scale the curve vertically to give the best fit
if max(curve_y_interp) == 0
    scale = 1;
else
    scale = max(data_y_interp)/max(curve_y_interp);
end
curve_y_interp = scale*curve_y_interp;
curve_y = scale*curve_y;

% calculate the error value
error_vector = curve_y_interp - data_y_interp;
sse = sum(error_vector .^ 2);

%--------------------------------------------------------------------------
% FIND_MIN finds the best approximation of the given data by a binomial
% distribution curve
%
% Args:     data_x  - input data scale
%           data_y  - distribution values
%           N_range - range for looking up degrees of freedom
%           p_range - range for looking up binomial distribution mean value
%           n_range - range for looking up value for number of shifts
%
% Returns:  params_min  - a vector of values [N, p, n] (degrees of freedom,
%                         mean, number of shifts) providing the best
%                         approximation
%
% (C) 2006, Tatyana Dembinsky
%--------------------------------------------------------------------------

function params_min = find_min (data_x, data_y, N_range, p_range, n_range)

% perform iterative search for different parameters to
% find the best combination giving the minimal error
sse_min = realmax;
params_min = [0,0,0];
for N = N_range
    for p = p_range
        for n = n_range
            params = [N, p, n];
            [sse, curve] = binomfun(params, data_x, data_y);
            if sse < sse_min
                sse_min = sse;
                params_min = params;
            end
        end
    end
end
```

Appendix C

Iris Region Relative Entropy Code

This appendix lists the Matlab® code used for calculating relative entropy of a specific iris region. The description of the algorithm is given in sections 2.4.2, 5.3.2 and 5.3.3. The function `calc_eye_rel_entropy` calculates relative entropy of the iris region defined by an input parameter. The code calculates the entropy of the whole iris when the region parameter specifies full template dimensions and the entropy of one feature if 1×1 region is specified.

```
%---------------------------------------------------------------------------
% CALC_EYE_REL_ENTROPY calculates relative entropy of a specific iris
% region for each eye in the database
%
% Args:     db      - database of iris features given as a cell array of
%                     matrices - one matrix per each eye where each line
%                     represents one template of the eye
%           ftr_num - number of features to use for calculations
%           L       - number of actual features used
%                     (regularization parameter)
%           row_num - number of rows in the template
%           col_num - number of columns in the template
%           active_region - the region of the iris template to use for
%               the relative entropy calculation (everything outside it is
%               ignored). The region should be presented as a matrix
%               [top_left_corner_X top_left_corner_Y; bottom_right_corner_X
%               bottom_right_corner_Y]
%
% Output:   eye_rel_entropy - array of relative entropy of the certain iris
%                region for each eye in the database. The order is preserved.
%
% (C) 2006, Tatyana Dembinsky
%---------------------------------------------------------------------------

function eye_rel_entropy = calc_eye_rel_entropy(db, ftr_num,
                                    L, row_num, col_num, active_region)

% remove empty matrices out of the database (remember their locations)
no_ftr_ind = [];
for i = 1:length(db)
    if length(db{i}) == 0
        no_ftr_ind(end+1)=i;
    end
```

```
end
db(no_ftr_ind) = [];

% build a matrix of all the population
min_Np = ftr_num;
pop = [];
for i = 1:length(db)
    eye_ftrs = db{i};
    sample_num = length(eye_ftrs(:,1));
    pop = [pop; eye_ftrs];
    if sample_num < min_Np
        min_Np = sample_num;
    end
end

if L >= min_Np
    error('L should be smaller than %d', min_Np);
end

% build a list of valid feature indices (no NaNs)
templ_length = row_num * col_num;
valid_ind = [];
for i = 1:templ_length
    if sum(isnan(pop(:,i))) == 0
        valid_ind(end+1) = i;
    end
end
% build a list of indices of features from the selected active region
active_ftr_ind = [];
for i=active_region(1,1):active_region(2,1)
    active_ftr_ind = [active_ftr_ind row_num*(i-1)+
                                (active_region(1,2):active_region(2,2))];
end
% restrict valid features to the ones from the specified region
temp = [];
for i=1:length(active_ftr_ind)
    if length(find(valid_ind==active_ftr_ind(i))) ~= 0
        temp = [temp active_ftr_ind(i)];
    end
end
valid_ind = temp;

% select features for processing
valid_ftr_num = length(valid_ind);
if ftr_num > valid_ftr_num
    error('Available features in the region: %d. Requested features: %d.',
                                                valid_ftr_num, ftr_num);
end
% select features randomly out of the collection of available ones
ftr_ind = [];
for i = 1:ftr_num
    rand_val = ceil(rand(1,1)*length(valid_ind));
    ftr_ind = [ftr_ind valid_ind(rand_val)];
    % remove the feature index so it won't be repeated
    valid_ind(rand_val) = [];
end
```

```
% feature normalization parameter
feat_norm = 1e2;

% calculate mean and covariance matrix for each eye
% reduce feature number per eye
eyes_means = [];
eyes_covars = {};
eyes_num = length(db);
for i = 1:eyes_num
    eye_ftrs = db{i}(:,ftr_ind) * feat_norm;
    if length(eye_ftrs) ~= 0
        eyes_means = [eyes_means; mean(eye_ftrs)];
        eyes_covars{i} = cov(eye_ftrs);
    end
end

% normalize features and reduce their number for the whole population
% calculate mean and covariance
pop = pop(:,ftr_ind) * feat_norm;
pop_mean = mean(pop);
pop_covar = cov(pop);
ftrs_num = length(pop_mean);

% apply SVD method
[U pop_S V] = svd(pop_covar);
ff = find(diag(pop_S) > 1e-5*pop_S(1,1));
U = U(:,ff);
V = V(:,ff);
pop_S = pop_S(ff,ff);
if norm(U-V)<1e-10
    U = (U+V)/2;
    V = U;
else
    error('U ~= V');
end
nG = size(U,2);

% regularize each eye covariance matrix to eliminate singularity
eyes_S = {};
M = ones(nG);
% decrease the feature number
M(L+1:nG,:) = 0;
M(:,L+1:nG) = 0;
M = M|eye(nG);
for i = 1:eyes_num
    eyes_S{i} = M.*(V'*eyes_covars{i}*U);
    if( det(eyes_S{i})<=0 )
        error('problem with eyes_S{%d}',i);
    end
end

% calculate entropy
eye_rel_entropy = [];
k = log2(exp(.5));
det_pop_S = det(pop_S);
```

```
for i = 1:eyes_num
    T = (eyes_means(i,:)-pop_mean)'*(eyes_means(i,:)-pop_mean);
    St = V'*T*U;
    covar_det = log(abs(det_pop_S/det(eyes_S{i})));
    tr = trace(U*((eyes_S{i}+St)*inv(pop_S)-eye(nG))*V');
    eye_rel_entropy(i) = k*(covar_det+tr);
end
```