



**HAL**  
open science

# Adapting iris feature extraction and matching to the local and global quality of iris image

Sandra Cremer

► **To cite this version:**

Sandra Cremer. Adapting iris feature extraction and matching to the local and global quality of iris image. Other [cs.OH]. Institut National des Télécommunications, 2012. English. NNT : 2012TELE0026 . tel-03934515

**HAL Id: tel-03934515**

**<https://theses.hal.science/tel-03934515v1>**

Submitted on 11 Jan 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**THALES**

**DOCTORAT EN CO-ACCREDITATION  
TELECOM SUDPARIS ET L'UNIVERSITE EVRY VAL D'ESSONNE**

**Spécialité : Informatique**

**Ecole doctorale : Sciences et Ingénierie**

**Présentée par**

**M<sup>lle</sup> Sandra Cremer**

**Pour obtenir le grade de  
DOCTEUR DE TELECOM SUDPARIS**

**Adapting iris feature extraction and matching to the local and  
global quality of iris images**

**Soutenue le 9 octobre 2012 devant le jury composé de :**

**Rapporteur : Pr. Andreas Uhl**

**Rapporteur : Pr. Nicole Vincent**

**Examineur : Pr. Amine Nait-Ali**

**Examineur : Dr. Lorène Allano**

**Directeur de thèse : Pr. Bernadette Dorizzi**

**Encadrant de thèse : Pr. Sonia Garcia**

**Encadrant de thèse : Mme. Nadège Lempérière**

Thèse n° 2012TELE0026

**CONFIDENTIAL**

# Abstract

Iris recognition has become one of the most reliable and accurate biometric systems available. However this modality is very sensitive to the quality of the input iris images: if this quality is degraded, the performance of iris recognition systems is strongly diminished. One of today's challenges therefore is to build iris recognition systems that stay highly efficient if the acquisition conditions are loosened and the quality of the input images is deteriorated.

Generally iris recognition systems can be cut into four stages: segmentation, normalization, feature extraction and matching. Quality degradations of the input images can have repercussions on all of these stages. For instance, they make the segmentation more difficult which can result in normalized iris images that contain distortion or undetected artifacts. Such normalized images will be hard to handle for feature extraction and matching. Moreover, even when the segmentation is accurate in spite of image degradations, the feature extraction and matching have to overcome other difficulties, such as a reduced amount of useful information available in the normalized images. This can happen if the iris images are highly occluded.

In this thesis we propose methods to improve the feature extraction and matching stages of iris recognition, so that their robustness to degraded input images is increased, whether the previously carried out segmentation is perfectly accurate or not. We mainly work with two algorithms for feature extraction and matching. Both of them use convolution with two-dimensional Gabor filters to extract iris features. However, they exploit the resulting phase features differently for recognition: the first one is based on Daugman's approach and performs a bit to bit comparison of binary feature vectors, while the second one computes local cross-correlations between real-valued phase images.

The first part of our work is aimed at controlling the quality and quantity of information selected in the normalized iris images for feature extraction and matching, when using these two recognition algorithms. To this end we have defined a local quality metric that exploits a statistical model to measure the amount of occlusion and the richness of the iris texture in sub-images. We use this metric to choose, for each couple of normalized iris images to compare, the sub-images on which feature extraction will be performed. Consequently, features will be extracted on the regions of highest quality, leading to an enhanced performance. Moreover, we demonstrate that the quantity of features that are extracted and exploited for matching have a strong influence on recognition accuracy. The highest accuracy is achieved when this quantity is adapted to each couple of normalized irises to compare. We therefore define a global metric we prove related to the optimal features' quantity to automate the features' selection process.



In the second part, we study the link between image quality and the performance of the two recognition algorithms just described, that are based on different techniques for feature extraction and matching. We demonstrate that the second one is more robust to degraded images that contain artifacts, distortion or a poor iris texture, and that these characteristics are accurately measured by our global quality metric.

Finally, we propose a complete system for iris recognition that combines the use of the local and global quality metrics we have defined to select the optimal extraction and matching algorithms and optimize their performance. The global quality metric is used to determine, for each couple of iris images to compare, the quantity of features to extract as well as the algorithms to apply to maximize recognition performance. The local quality metric is used to select the best sub-images for feature extraction and matching. We exploit mainly two public databases, ND-IRIS-0405 and CASIA-IrisV3 to show that the resulting system is a good trade-off between accuracy and speed.

**CONFIDENTIAL**

# Remerciements

Tout d'abord, je voudrais remercier ma directrice de thèse, Bernadette Dorizzi, de m'avoir accueillie dans son équipe afin que je puisse mener ma thèse. Je la remercie pour son investissement dans cette thèse ainsi que pour son soutien, aussi bien sur le plan scientifique que sur le plan humain. Elle a joué un rôle clé dans chaque étape de ma thèse et est pour beaucoup dans son aboutissement.

Je remercie également Sonia Garcia qui a participé à mon encadrement au sein de l'équipe Intermedia. En particulier, elle m'a aidé à reprendre et à assimiler les travaux qui avaient été fait auparavant au sein de l'équipe sur la reconnaissance d'iris, ainsi qu' à trouver des idées innovantes pour poursuivre ces travaux.

Ensuite je remercie les trois personnes qui m'ont encadrée successivement chez Thales Communications & Security (TCS) : Valérie Letournel, Nadège Lempérière et Stéphane Bélardi. Chacun d'entre eux m'a permis d'avancer dans les différentes étapes de mon travail. Valérie a joué un rôle important dans le lancement de ma thèse et m'a souvent donné de précieux conseils durant ces trois dernières années. Nadège s'est fortement investie dans mon travail. Son apport technique ainsi que son soutien moral ont été essentiels pour moi. Finalement, Stéphane a contribué à l'avancement de ma thèse dès le départ, par le biais de nos brainstormings mensuels, et a suivi sa finalisation en prenant le temps de relire et corriger mon manuscrit.

Pour continuer, je souhaite remercier Philippe Robin, Pierre Vaures et Cyrille Bichon qui ont rendu cette thèse possible en lançant une thèse Cifre avec Telecom SudParis et ont fait en sorte qu'elle se déroule dans de bonnes conditions. Je remercie aussi Bertrand Sensfelder qui a veillé, avec Philippe Robin, à ce que ma thèse aboutisse et que mon travail soit valorisé au sein de Thales.

Ces remerciements ne seraient pas complets sans remercier toute l'équipe biométrie de TCS pour ses conseils, son soutien et sa bonne humeur durant ces trois dernières années. En plus des personnes nommées ci-dessus, je citerai d'abord ceux qui sont là depuis mes débuts (ou presque) : Florian Hadjadj, Sandra Marti, Thierry Lefevre, Nicolas Vervelle, Yves Legier, François Ribière, Michel Dieumegard, Mathieu Seillier et Aline Heuze (encore membre de l'équipe dans nos esprits). Puis ceux qui sont arrivés en cours de route : Vincent Monier, Nathalie Dumas, Silvia Gandy et Victor Leyva Sifuentes. J'en profite pour remercier spécialement Silvia pour avoir pris le temps de relire et corriger mon manuscrit de thèse. Je remercie aussi les deux stagiaires qui ont travaillé sur la reconnaissance d'iris au sein de cette équipe : Boris Tchorbadjian et Jean-Charles Bricolas, pour avoir participé au développement de certaines expériences présentées dans ma thèse.

Je remercie également mes collègues de l'équipe Intermedia, particulièrement

Nesma Houmani, Guillaume Sutra et Sanjay Kanade pour les moments que j'ai partagés avec eux quand je venais à Evry. Je remercie aussi Phan Viêt Anh, pour m'avoir aidé à m'approprier les codes OSIRIS et CORR qui ont servis de base à mon travail.

Je remercie Andreas Uhl et Nicole Vincent d'avoir accepté d'être rapporteur pour ma thèse, ainsi qu'Amine Nait-Ali et Lorène Allano pour avoir accepté d'être membres de mon jury de thèse.

Pour finir je voudrais remercier mon entourage proche. D'abord ma famille, et en particulier mes parents qui m'ont toujours soutenue pendant mes études et m'ont encouragée à faire une thèse. Je remercie mon père pour m'avoir initiée à Latex et surtout pour avoir pris le temps de relire et corriger ma thèse. Je remercie aussi mes amis pour leur soutien, notamment Clémentine et Daphné. Enfin, je remercie Florian pour son soutien quotidien durant ces trois dernières années. Avant les moments de réussites, il y a eu des moments de doutes et il a été présent pour m'encourager à persévérer... C'est donc aussi grâce à lui que cette thèse a abouti.

**CONFIDENTIAL**

# Acknowledgments

First of all, I would like to thank my supervisor, professor Bernadette Dorizzi, for giving me the opportunity of doing my doctoral work in her research team. I thank her for her devotion to this work, as well as for her scientific and moral support. She played a key role in the various steps of my thesis and is accountable for its success.

I would also like to thank professor Sonia Garcia for her co-supervision within the Intermedia team. In particular, she helped me to assimilate the work that had previously been carried out on iris recognition in this team and to find original ways of pursuing this work.

Next, I would like to thank the three people who successively supervised my work at Thales Communications & Security (TCS): Valérie Letournel, Nadège Lempérière and Stéphane Bélardi. Each one of them enabled me to move forward at different stages of my work. Valérie played an important role in the launching of my thesis and often gave me useful advice these last three years. Nadège was strongly devoted to the success of this thesis. Both her technical contribution and her moral support have been essential for me. Finally, Stéphane made contributions to my work from the beginning through our monthly brainstorming sessions. He also followed the completion of my work and took time to proofread my thesis.

To carry on, I would like to thank Philippe Robin, Pierre Vaures and Cyrille Bichon for making the collaboration between TCS and Télécom SudParis possible and ensuring that my doctoral work was carried out in mint condition. I would also like to thank Bertrand Sensfelder who made sure, with Philippe Robin, that my thesis could succeed and be promoted within Thales.

These acknowledgments would not be complete without thanking the entire biometrics team from TCS for their advice, support and good spirit during the last three years. In addition to the people I have already mentioned, I would like to start by naming the people that were part of this team from the beginning: Florian Hadjadj, Sandra Marti, Thierry Lefevre, Nicolas Vervelle, Yves Legier, François Ribière, Michel Dieumegard, Mathieu Seillier and Aline Heuze (who still kept the team's spirit). Then I would like to name the ones that arrived later on: Vincent Monier, Nathalie Dumas, Silvia Gandy and Victor Leyva Sifuentes. I would specially like to thank Silvia for taking the time to proofread my thesis. Moreover, I would like to thank the two trainees who worked on iris recognition within this team: Boris Tchorbadjian and Jean-Charles Bricolas. They contributed to the development of experiments presented in my thesis.

I would also like to thank my colleagues from the Intermedia team, and in particular Nesma Houmani, Guillaume Sutra and Sanjay Kanade for their company while I was at Evry. Also, thanks to Phan Viêt Anh for helping me

to assimilate how to work with the OSIRIS and CORR reference systems.

I would like to thank Andreas Uhl and Nicole Vincent for accepting to review my thesis, as well as Amine Nait-Ali and Lorène Allano for accepting to be juries for my defense.

Finally, I would like to thank my circle of friends and family. First, thanks to my family and particularly to my parents for supporting me through my studies and encouraging me to be a PhD candidate. Thanks to my dad for teaching me how to use Latex and especially for proofreading my thesis. Secondly, thanks to my friends for their support, particularly Clémentine and Daphné. Finally, thanks to Florian for his daily support for the last three years. Before the moments of success come the moments of doubts and he was there to encourage me to persevere... So he contributed to the success of my doctoral work.

**CONFIDENTIAL**

# Glossary

**EER:** Equal Error Rate

**FAR:** False Acceptance Rate

**FMR:** False Match Rate

**FNMR:** False Non Match Rate

**FRR:** False Rejection Rate

**FTC:** Failure To Capture

**FTE:** Failure To Enroll

**FTM:** Failure To Match

**GMM:** Gaussian Mixture Model

**ICE:** Iris Challenge Evaluation

**IOM:** Iris On the Move

**IR:** Infra-red

**IRIS:** Iris Recognition Immigration System

**NIR:** Near-infra-red

**NIST:** National Institute of Standards and Technology

**OSIRIS:** Open Source for Iris

**RER:** Retrieval Error Rate

**ROC:** Receiver Operating Characteristic

**SIFT:** Scale Invariant Feature Transform

**TCS:** Thales Communications & Security

**UIDAI:** Unique Identification Authority of India

**VASIR:** Video-based Automated System for Iris Recognition

**CONFIDENTIAL**

# List of Figures

1.1	Iris image . . . . .	1
1.2	Iris image acquired with the LG 2200 near-infra-red sensor . . . . .	2
1.3	Iris image acquired with a visible length sensor . . . . .	2
1.4	Examples of image degradations in the ICE 2005 database . . . . .	4
2.1	Inter-class and intra-class similarity score distributions . . . . .	12
2.2	FAR and FRR as functions of the threshold $th$ . . . . .	13
2.3	Example of a ROC curve and the associated EER . . . . .	14
2.4	Examples of images from the ND-IRIS-0405 database . . . . .	17
2.5	Examples of images from the CASIA V1.0 and CASIA-IrisV3-Interval databases . . . . .	18
2.6	Examples of images from the CASIA-IrisV3-Lamp dataset . . . . .	18
2.7	Examples of images from the CASIA-IrisV3-Twins dataset . . . . .	18
2.8	Statistics from the CASIA-IrisV3 datasets . . . . .	19
3.1	Illustration of the normalization process . . . . .	24
3.2	Illustration of the phase quantization . . . . .	25
3.3	Example of circles determined by an aberrant segmentation for an iris image and of the associated normalized image . . . . .	36
4.1	Normalized iris images from three different databases (1: ND-IRIS, 2: CASIA-IrisV3-Interval, 3: CASIA-IrisV3-Twins), before (a) and after (b) the pre-processing step . . . . .	42
4.2	Examples of sub-images extracted from three different databases (ND-IRIS, CASIA-IrisV3-Interval and CASIA-IrisV3-Twins) and the local quality values associated to them. Sub-images have been chosen to represent the different qualities encountered in the images. . . . .	44
4.3	Illustration of the phase quantization . . . . .	47
4.4	Distribution of the Hamming distances for inter-class and intra-class comparisons from the ND-IRIS-0405 database. . . . .	47
4.5	Application points contained in a $11 \times 51$ pixels sub-image . . . . .	49
4.6	Mean local Hamming distance for intra-class and inter-class comparisons of sub-images, for different quality values of these sub-images. . . . .	50
4.7	The difference $\delta = \text{mean}(\text{loc}HD_{inter}) - \text{mean}(\text{loc}HD_{intra})$ as a function of the minimum quality values . . . . .	50



4.8	ROC curves obtained when eliminating the 10% worst image couples according to different GMM-based quality measures on the ND-IRIS-0405 database. . . . .	54
4.9	ROC curves obtained when eliminating the 10% worst image couples according to a GMM-based or a segmentaion-based quality measures on the ND-IRIS-0405 database. . . . .	55
4.10	ROC curves obtained when eliminating the 10% worst image couples according to different quality measures on the CASIA-IrisV3-Lamp database. . . . .	55
4.11	ROC curves obtained when selecting the 10% best image couples according to different quality measures on the ND-IRIS-0405 database. . . . .	56
4.12	Distribution of the similarity scores computed with Krichen’s algorithm for inter-class and intra-class comparisons from the ND-IRIS-0405 database. . . . .	58
4.13	ROC curves obtained when selecting the 90% best image couples according to different quality measures on the ND-IRIS-0405 database and using Krichen’s algorithm for feature extraction and matching. . . . .	60
4.14	ROC curves obtained when selecting the 90% best image couples according to different quality measures on the CASIA-IrisV3-Lamp database and using Krichen’s algorithm for feature extraction and matching. . . . .	60
5.1	ROC curves of 3 iris recognition algorithms based on the same implementation of Daugman’s approach but using different formulas to compute the Hamming distance on the ICE 2005 database. . . . .	64
5.2	ROC curves of 4 iris recognition algorithms based on the same implementation of Daugman’s approach but using different formulas to compute the Hamming distance on the ICE 2005 database. . . . .	66
5.3	Global scheme of the method we propose for pertinent feature extraction using our quality measure. . . . .	68
5.4	ROC curves of 3 iris recognition algorithms based on the same implementation of Daugman’s approach but using different strategies for point selection for feature extraction on the ND-IRIS database. . . . .	70
5.5	ROC curves for different values of $N$ (number of selected points for feature extraction) and a set value of $M$ (initial number of points) for the ND-IRIS database. . . . .	71
5.6	ROC curves for different values of $N$ (number of selected points for feature extraction) and a set value of $M$ (initial number of points) for the CASIA-IrisV3-Interval database. . . . .	71
5.7	ROC curves for different values of $N$ (number of selected points for feature extraction) and a set value of $M$ (initial number of points) for the CASIA-IrisV3-Twins database. . . . .	72
5.8	False Rejection Rate (for $FAR = 10^{-4}$ ) for different databases and for different values of $N$ (number of selected points for feature extraction) given in function of $M$ (initial number of points). . .	72

**CONFIDENTIAL**

5.9	ROC curves of 3 iris recognition algorithms based on the same implementation of Daugman's approach but using different strategies for feature extraction's point selection, on images from the ND-IRIS database obtained with Lefèvre's segmentation. . . . .	73
5.10	ROC curves for different values of $N$ (number of selected points for feature extraction) and a set value of $M$ (initial number of points) for images from the ND-IRIS database obtained with Lefèvre's segmentation. . . . .	74
5.11	ROC curves for different values of $M$ (number of uniformly distributed application points) for the ND-IRIS-0405 database. . . . .	75
5.12	Evolution of the means of the normalized inter-class and intra-class score distributions with $N$ . . . . .	77
5.13	Normalized score distributions obtained when applying Daugman's feature extraction and matching algorithm with our region selection strategy or with Lefèvre's segmentation mask. . . . .	79
5.14	Example of an image for which the normalization failed . . . . .	79
5.15	Examples of (gallery, probe) normalized iris image couples for different global qualities. . . . .	81
5.16	ROC curves for 4 different values of $N/M$ on matches from the CASIA-IrisV3-Lamp database for which $0 < GlobQ_1 < 0.13$ . . . . .	82
6.1	Global scheme of the method we propose for pertinent feature extraction and matching according to Krichen's technique using our quality measure. . . . .	87
6.2	ROC curves for different values of $N$ (number of points selected for feature extraction) and a set value of $M$ (initial number of points) for the ND-IRIS-0405 database. . . . .	89
6.3	ROC curves for different values of $N$ (number of points selected for feature extraction) for the CASIA-IrisV3-Lamp database. . . . .	89
6.4	ROC curves for different values of $N/M$ for matches from the CASIA-IrisV3-Lamp database for which $0 \leq GlobQ_1 \leq 0.13$ . . . . .	90
7.1	ROC curve for the detection of off-angle images using the score defined in formula 7.2. . . . .	98
7.2	Zoomed ROC curve for the detection of off-angle images using the score defined in formula 7.2. . . . .	99
7.3	ROC curves for iris recognition algorithms using Krichen's or Daugman's feature extraction and matching technique and the quality based region selection strategy with $N/M=0.4$ on off-angle images. . . . .	99
7.4	Example of an off-angle and a non-off-angle iris image of the same subject and the associated normalized images . . . . .	100
7.5	Proposed strategy for iris recognition combining Krichen's and Daugman's algorithms as well as local and global quality measures. . . . .	102
7.6	Comparative ROC curves for image couples for which $GlobQ_1 < 0.4$ or $GlobQ_{off-angle} > th_{off-angle}$ . . . . .	106
7.7	Comparative ROC curves for image couples for which $GlobQ_1 \geq 0.4$ or $GlobQ_{off-angle} \leq th_{off-angle}$ . . . . .	106

**CONFIDENTIAL**

**CONFIDENTIAL**

# List of Tables

2.1	<i>Description of different NIR iris image databases . . . . .</i>	20
4.1	<i>Mean values of the local qualities corresponding to 500 sub-images from different categories, manually labeled . . . . .</i>	45
5.1	<i>Values tested for <math>N</math> for each value of <math>M</math> . . . . .</i>	76
5.2	<i>Values of <math>N</math> corresponding to the best performance . . . . .</i>	76
5.3	<i>Percentage of the full matching list for each category of <math>GlobQ_1</math> .</i>	81
5.4	<i>Best values for <math>N/M</math> for different databases and different values of <math>GlobQ_1</math> when using Daugman's algorithm . . . . .</i>	82
5.5	<i>Proportion of the full ND-IRIS-0405 database in each range of <math>GlobQ_1</math> and best value of the ratio <math>N/M</math> for each quality range .</i>	83
5.6	<i>Absolute and relative improvement of the FRR for <math>FAR=10^{-4}</math> for different global quality ranges . . . . .</i>	83
6.1	<i>Best values for <math>N/M</math> for different databases and different values of <math>GlobQ_1</math> when using Krichen's algorithm . . . . .</i>	91
7.1	<i>FRR for <math>FAR=10^{-3}</math> on the ND-IRIS-0405 database for different values of <math>GlobQ_1</math> and <math>N/M</math> when using the recognition process based on Daugman's and Krichen's techniques . . . . .</i>	95
7.2	<i>FRR for <math>FAR=10^{-3}</math> on the CASIA-IrisV3-Lamp database for different values of <math>GlobQ_1</math> and <math>N/M</math> when using the recognition process based on Daugman's and Krichen's techniques . . . . .</i>	95
7.3	<i>Best values for <math>N/M</math> for different values of <math>GlobQ_1</math> . . . . .</i>	104

**CONFIDENTIAL**

# Contents

<b>Abstract</b>	<b>i</b>
<b>Remerciements</b>	<b>iii</b>
<b>Acknowledgments</b>	<b>v</b>
<b>Glossary</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Iris acquisition and recognition . . . . .	2
1.1.1 Iris acquisition . . . . .	2
1.1.2 Iris recognition . . . . .	3
1.1.3 Iris recognition applications . . . . .	3
1.2 Today's challenges in iris acquisition and recognition . . . . .	4
1.3 Aim and content of the thesis . . . . .	6
<b>2 Performance evaluation in biometrics</b>	<b>9</b>
2.1 How to measure performance . . . . .	10
2.1.1 Definition of the error rates . . . . .	10
2.1.2 Performance curves and accuracy indicators . . . . .	13
2.1.3 Estimation of confidence intervals . . . . .	14
2.2 Available databases . . . . .	16
2.2.1 ICE and ND-IRIS-0405 . . . . .	16
2.2.2 CASIA-Iris . . . . .	17
2.2.3 Other iris datasets . . . . .	19
2.3 Benchmarks . . . . .	20
2.4 Conclusion . . . . .	21
<b>3 Literature review</b>	<b>23</b>
3.1 Major existing iris recognition systems . . . . .	24
3.1.1 Daugman's algorithm . . . . .	24
3.1.2 Alternate ways to generate binary codes . . . . .	27
3.1.3 Real-valued Feature vectors . . . . .	28
3.1.4 Fusion of different techniques . . . . .	29
3.2 Quality metrics . . . . .	29
3.2.1 Global quality metrics . . . . .	29
3.2.2 Local quality metrics . . . . .	30
3.3 Integration of quality metrics in recognition systems . . . . .	32
3.3.1 Global quality measures . . . . .	32

3.3.2	Local quality measures . . . . .	32
3.4	Reference systems . . . . .	34
3.4.1	Implementations of Daugman's algorithm . . . . .	34
3.4.2	Segmentation used to generate our input normalized images . . . . .	35
3.5	Conclusion . . . . .	37
<b>4</b>	<b>Elaboration of a local quality measure and analysis of its pertinence for matching</b>	<b>39</b>
4.1	Definition of our local quality measure . . . . .	40
4.1.1	Choice of a model . . . . .	40
4.1.2	Gaussian Mixture Models in general . . . . .	40
4.1.3	Detailed description of our Gaussian Mixture Model . . . . .	41
4.1.4	Computation of our local quality measure . . . . .	42
4.2	Examples and statistics . . . . .	44
4.2.1	Examples . . . . .	44
4.2.2	Statistics . . . . .	44
4.3	Pertinence of our quality measure for a Daugman type matching . . . . .	46
4.3.1	Description of the feature extraction and matching technique . . . . .	46
4.3.2	Local pertinence of our quality measure . . . . .	47
4.3.3	Global pertinence of our quality measure . . . . .	52
4.4	Pertinence of our quality measure for a Krichen type matching . . . . .	56
4.4.1	Description of the feature extraction and matching technique used . . . . .	57
4.4.2	Global pertinence of our quality metric . . . . .	58
4.5	Conclusion . . . . .	59
<b>5</b>	<b>Integration of our quality measure in OSIRIS</b>	<b>61</b>
5.1	Drawbacks of the state-of-the-art techniques . . . . .	61
5.1.1	Creation of a binary quality mask . . . . .	61
5.1.2	Weighted quality mask . . . . .	62
5.2	Proposed technique . . . . .	66
5.2.1	Description of the method . . . . .	66
5.2.2	Experiments and results . . . . .	67
5.3	Understanding of the best value of $N$ . . . . .	74
5.3.1	Importance of the ratio $N/M$ . . . . .	74
5.3.2	The impact of $N/M$ on score distributions . . . . .	77
5.4	Combination of global and local information . . . . .	80
5.4.1	Motivation . . . . .	80
5.4.2	Experiments and results . . . . .	80
5.5	Conclusion . . . . .	84
<b>6</b>	<b>Integration of our quality measure in Krichen's algorithm</b>	<b>85</b>
6.1	Integration of the region selection routine . . . . .	85
6.1.1	Description of the proposed technique . . . . .	85
6.1.2	Experiments and results . . . . .	86
6.2	Combination of global and local information . . . . .	88
6.2.1	Motivation . . . . .	88
6.2.2	Experiments and results . . . . .	90
6.3	Conclusion . . . . .	91

**CONFIDENTIAL**

---

<b>7</b>	<b>Adapting the technique used for matching to image quality</b>	<b>93</b>
7.1	Quality according to our GMM-based measure . . . . .	94
7.1.1	Matching performance per quality category . . . . .	94
7.1.2	Result interpretation . . . . .	95
7.2	Off-angle quality . . . . .	96
7.2.1	Off-angle quality measure . . . . .	96
7.2.2	Comparative matching performance on off-angle images . . . . .	97
7.3	Final strategy proposed for iris recognition . . . . .	100
7.3.1	Method description . . . . .	101
7.3.2	Experiments and results . . . . .	103
7.4	Conclusion . . . . .	105
<b>8</b>	<b>Conclusions and discussion</b>	<b>109</b>
<b>Appendix A: “Improving normalization circle detection for iris recognition using a variational approach”</b>		<b>113</b>
<b>Appendix B: “How a local quality measure can help improving iris recognition”</b>		<b>115</b>



**CONFIDENTIAL**

# Chapter 1

## Introduction

Reliable and secure identification of people is required in many situations and has therefore become a major concern. Most commonly, identification claims are verified with user names, passwords or identification cards. However, these methods can quite easily be bypassed by stealing the identification card or breaking the password. For this reason, growing attention has been given to biometric methods that use physical or behavioral specific characteristics for individual identification. Such characteristics cannot be lost or forgotten in the way passwords and identification cards can be, and some of them are shown to allow very high recognition accuracy.

The most commonly used biometric modalities are fingerprints, irises and face. They are exploited in many commercial systems and the recognition performance that can be achieved by using them are well quantified. However many other biometric features can be used such as the veins in the hand, the retina or the ear.

In this thesis, we work exclusively with the iris for recognizing people. This biometric modality is one of the most reliable modalities available, for it possesses a very rich pattern that is believed to be different between persons, and relatively constant throughout a lifetime. The iris is “the colored ring of tissue around the pupil through which light enters the interior of the eye” [64] as we can see in Figure 1.1. It is the only internal organ of the body that is externally visible.

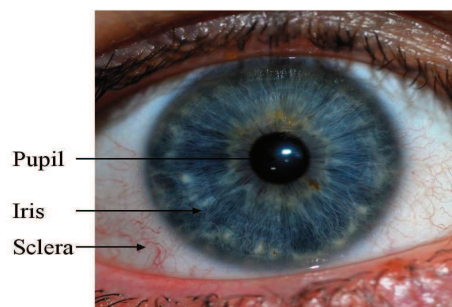


Figure 1.1: Iris image

In Section 1.1., we will present very briefly how iris acquisition and recogni-

tion work, and present some examples of iris-based systems. The major challenges for iris recognition will be put forward in Section 1.2. in order to present the motivations of this thesis in Section 1.3.

## 1.1 Iris acquisition and recognition

### 1.1.1 Iris acquisition

Acquiring iris images in which the iris pattern is clearly visible is not obvious, for the iris is relatively small (diameter approximately equal to 1cm) and is located behind the cornea which is highly reflective. To minimize the reflection on the cornea, most commercial iris acquisition systems use near-infra-red illumination for the acquisition. The used wavelength lies in the range of 700-900nm. Figure 1.2 represents an iris image acquired with the LG 2200 near-infra-red sensor.

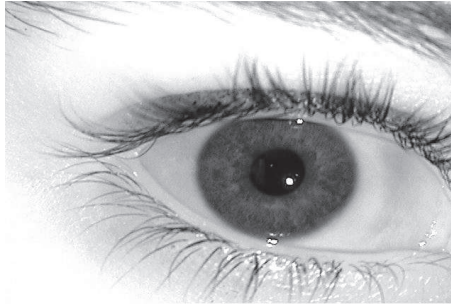


Figure 1.2: Iris image acquired with the LG 2200 near-infra-red sensor

Nevertheless, research has also been done on iris images that have been acquired with visible length illumination. For example, Proença and Alexander have acquired iris images using a Nikon E5700 camera and made the resulting database UBIRIS publically available [71]. Figure 1.3 presents an image from this database. The two images presented in Figures 1.2 and 1.3 show the typical quality difference between images from sensors that use near-infrared light and sensors that use light from the visible spectrum. The second one contains much more reflections on the cornea. Consequently, in this work, we have chosen to consider only iris images acquired with near-infra-red sensors, like the majority of commercial systems.



Figure 1.3: Iris image acquired with a visible length sensor

**CONFIDENTIAL**

Furthermore, in order for the iris image to be exploitable for recognition, the diameter of the acquired iris must be high enough. The ISO/IEC norm 19794-6 requires this diameter to be higher than 200 pixels. This adds constraints on the camera lenses that can be used, and usually imposes that the acquisition system's depth of field is very low. Consequently, acquiring focused iris images requires specific camera design and a precise subject to camera distance. This distance usually is around 50cm.

### 1.1.2 Iris recognition

Once the iris has been acquired, the image can be used for recognition. Daugman was the first one to propose a complete iris recognition system in 1993 [21]. His system, like most of today's commercial iris recognition systems, can be cut into four steps:

- **Segmentation of the iris:** localization of the iris in the image in order to separate it from other elements such as the pupil, the sclera, the eye-lashes or the eye-lids. Usually this is done by associating a binary mask to the image that covers the non-iris portions.
- **Normalization:** the iris is transformed into a rectangle of pre-set dimension. This step is necessary because iris dimensions change as the pupil dilation varies over time.
- **Feature Extraction:** extraction of the relevant information from the normalized iris image in order to obtain a reduced representation set of features.
- **Matching:** comparison of the features corresponding to different iris images, to determine whether or not the images represent the same iris.

### 1.1.3 Iris recognition applications

Iris recognition is currently used in several applications. For instance, it is used in some airports to replace passport presentation for frequent travelers. This system requires travelers to enroll. They can then present their iris at an airport terminal and it is compared to all the irises in the enrolled database. Such a system has been deployed at Schiphol Airport in the Netherlands, as well as in the UK airports where the Iris Recognition Immigration System (IRIS) was put into operation in January 2006. One million frequent travelers have been enrolled in this system, and there are currently almost 20 000 IRIS entries into the UK every week.

Iris recognition is also used for more general border control. For example, the United Arab Emirates border security system is based on iris recognition. It is deployed in 32 ports, comprising entry points via sea, land and air. Over 1 million iris templates are registered in a watch list and an average of 14 billion comparisons are performed each day.

Our final example is the Unique Identification Authority of India (UIDAI) project. Its deployment started in 2010 and is still in progress. The project's objective is to attribute a unique number (Aadhaar number) to every Indian citizen, so that they can all have access to services such as medical aid or opening

**CONFIDENTIAL**

a bank account. Each Aadhaar number is associated to a civil status and to biometric data (fingerprints and iris). One million new irises are enrolled every day with this system. It is the largest iris recognition system that has ever been deployed.

## 1.2 Today's challenges in iris acquisition and recognition

Since 1993 a great amount of research has been done on the topic of iris recognition. For the last five years, over 150 papers are published per year on this topic [15]. Nevertheless, many challenges persist.

One of the major challenges is to reduce the constraints on iris acquisition systems. Indeed, many applications would benefit from less intrusive and more fluid acquisition systems. For instance, such systems could reduce the waiting time at border control.

Consequently, researchers have gradually loosened iris acquisition conditions. From 2004 to 2005, the Computer Vision Research Lab at the University of Notre Dame collected iris images using the LG 2200 iris imaging system. They have inactivated the quality control associated to this imaging system. The quality of the acquired images is therefore degraded. For example, excessive blur appears in some images. In other images the iris is occluded by artifacts such as eyelids, eyelashes or specular reflection. Sometimes the iris even is off-angle or off-center. Figure 1.4. illustrates examples of such degraded iris images. All these images have been made publicly available for the 2005 Iris Challenge Evaluation (ICE 2005) [68]. We will give more details about this evaluation in this thesis.

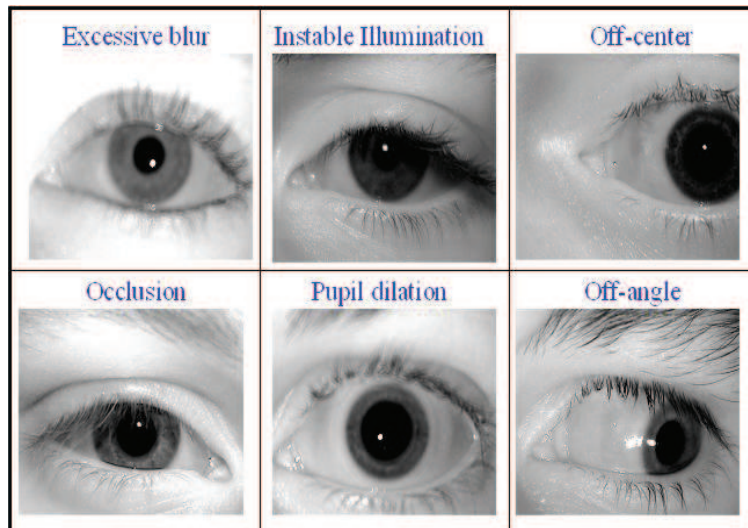


Figure 1.4: Examples of image degradations in the ICE 2005 database

Since then, researchers have pursued even more ambitious objectives. In

**CONFIDENTIAL**

2006, the Iris On the Move System (IOM) [60] was developed. It acquires the iris at a 3 meter distance, and can tolerate a subject moving at approximately one meter per second. The obtained image quality is significantly worse than the one presented here above. For instance, image resolution is much lower and the motion blur is stronger.

Today, there is only one commercial iris acquisition system that can operate at a camera distance of 2 meters from the iris. It is developed by AOptix [9]. All other systems require this distance to be lower than one meter. Moreover, all commercial systems necessitate a cooperation from the subject whose irises are to be acquired.

Generally, as soon as the constraints at the acquisition are loosened, the quality of the acquired images is degraded and iris recognition systems are less efficient for different reasons. First of all, the variability of the illumination and of the focus of the image makes the segmentation step more difficult. Segmentation errors then have repercussions on the performance of all the other stages. For instance, they can cause distortions in the normalized images that make the matching more difficult. They can also cause features to be extracted in regions of the normalized iris image that contain artifacts. This will lead to a noisy feature vector and diminish the chances that the matching be accurate. Secondly, even if the segmentation has succeeded in spite of the degraded images, the feature extraction and matching steps must overcome other difficulties. The quantity and the quality of information available after segmentation are reduced. For example in the case of highly occluded iris images, only a small portion of the iris can be used for the feature extraction and matching. In the case of excessive blur, the high frequencies of iris pattern are not collectible. These difficulties also tend to reduce recognition performance.

One line of research for handling these issues, related to acquisition conditions and image quality, consists in developing iris acquisition systems that generate good quality images even when the acquisition conditions are loosened. We can mention the work done by Venugopalan et al. in [84]. The system they propose is capable of capturing iris images up to distances of 8 meters with a resolution of 200 pixels across the diameter. It can also acquire irises from mobile subjects thank to a focus tracking module and velocity estimation.

Another line is to work on quality metrics in order to detect images of degraded quality at the acquisition and to repeat the acquisition if necessary. Researchers are also working on creating recognition algorithms that are more robust to image quality degradation. More details on these two research areas will be given in this thesis.

Apart from these issues related to acquisition conditions and image quality, there are many other challenges in the field of iris recognition. Since, we have chosen not to deal with them in our work, we will only mention them very briefly in this introduction.

One challenge is to take countermeasures against forgery (spoofing). To this end, researchers have worked on testing the liveness of eye images. The proposed solutions include testing the pupil's response to light [12], or using biophotonics to test the liveness of the iris tissue, since living tissues respond differently to different wavelengths of light [17]. Work has also been carried out to detect the

**CONFIDENTIAL**

presence of printed and patterned contact lenses on the eye [33].

Another challenge is to reduce the size of the data that is stored in iris databases. Indeed for applications such as the UIDAI, hundreds of millions of iris templates need to be stored. The objective is to find a trade-off between reducing the template size and keeping high recognition performance.

In link with the issue of data storage, the National Institute of Standards and technology (NIST) has conducted a study to define storage formats that would be inter-operable. So interoperability has also become a concern in the field of iris recognition.

Finally, a major issue is computation speed. In the UIDAI program, each time a new person is enrolled, their biometric data has to be compared to all the biometric data contained in the database to avoid identity duplication. Consequently, up to 150 trillion comparisons have to be done per day. Processing speed can be increased by performing indexing for such large databases instead of an exhaustive research, as proposed by Daugman in [32]. Only iris codes with “suspicious coincidences” would be considered as candidates for matching.

### 1.3 Aim and content of the thesis

The doctoral research presented in this thesis is the result of a collaboration between Thales Communications & Security (TCS), a French company which is a key player in identity systems worldwide, and the Intermedia research team of Telecom SudParis. One of the research interests of the Intermedia team is biometric authentication. In particular, they have worked on iris recognition in degraded conditions through the doctoral work of Emine Krichen [46]. Consequently, TCS has chosen to collaborate with the Intermedia team to benefit from this team’s experience and improve its own iris recognition technology.

In this context, two doctoral projects were initiated : the one of Thierry Lefèvre [52] and the one presented in this thesis. Their objective was to elaborate together an iris recognition system that stays highly efficient when the acquisition conditions are loosened and the quality of the iris images diminished. The aim was to handle images acquired when the subjects were still, but the constraints at the acquisition were limited. Consequently, the main challenges that were addressed were the ones linked to input iris images of degraded quality, when the level of the degradations is close to the ones presented in Figure 1.4. However two complementary approaches were chosen to do so.

In [52] Lefevre has chosen to work on the segmentation and normalization steps in order to improve their robustness to degraded images. The objective was to reduce the amount of distortions and non-detected artifacts in the normalized iris images. On the other hand, we have chosen to concentrate on the two last steps of iris recognition, namely the feature extraction and the matching. Consequently, in this thesis, we will take as input, images that represent the iris after segmentation and normalization. Since we have no information on the quality of the segmentation that has been previously carried out on the iris images we consider, we will try to propose a solution that can be applied independently of the segmentation precision.

In line with what we have explained in Section 1.2, the objective of our doctoral work is to propose a solution for feature extraction and matching that is robust to the following elements:

**CONFIDENTIAL**

- distortions in the normalized images caused by an inaccurate segmentation
- artifacts contained in the normalized image that have not been detected at the segmentation step
- a reduced amount of non-occluded iris texture available for matching whether the segmentation is accurate or not

Consequently, our first challenge is to detect normalized iris images that are affected by one of these elements. To this end, we have defined a local quality metric aimed at measuring the amount of artifacts, as well as the richness of the texture in the usable iris area, in sub-images from the normalized iris images. It is based on a Gaussian mixture model (GMM). Our idea was to exploit this measure to select the regions in the normalized iris images used for feature extraction and matching.

We will see in this thesis, that only a small number of authors have worked on measuring the quality of iris images locally. Moreover, the method we chose for exploiting our local measure differs from theirs in that it enables to keep control on the amount of information exploited for recognition. Indeed, we have analyzed the impact of the quantity of information available for matching on recognition performance, and have proposed to adapt this quantity to each pair of irises to compare, when we select the exploited regions with our local quality measure.

We have applied our method with two state of the art algorithms. The first one is described by Daugman in [21] and is used in most commercial systems. Its performance was evaluated on very large databases and was demonstrated to be very high in terms of accuracy and speed. In our work we have used the OSIRIS implementation of this algorithm [49], less optimized than the commercial versions. The second one was described by Krichen et al. in [48]. It is presented as being more robust to degradations of the image quality than Daugman's system.

In order to be robust to degradations of iris images, as well as to compromise between accuracy and speed when performing recognition, we have chosen to adapt feature extraction and matching to each pair of irises to compare. To this end, we have defined a global quality metric associated to each pair of irises. It gives a good indication of the amount of highly textured, non-occluded and overlapping regions in the pair of normalized images. We analyze how the optimal amount of information extracted in the iris images and exploited for matching is linked to our global quality measure. We also study how our measure can be used to predict the performance of Daugman's and Krichen's algorithms.

The final system we propose for iris recognition integrates our method to select the regions of the normalized iris images exploited for feature extraction. This method does not require any segmentation mask, but uses our local quality measure to avoid selecting regions that contain artifacts and to privilege the most highly textured regions for feature extraction. It also exploits our global quality measure to adapt the proportion of the initial normalized images that is selected for feature extraction and matching, to each couple of images. This enables us to handle the variations of the quantity of information available from one image to another. Moreover we use global quality measures to adapt the techniques applied for feature extraction and matching to each image couple in

**CONFIDENTIAL**



order to optimize performance in terms of accuracy and speed.

This thesis is organized as follows:

In Chapter 2, we will describe the methods and tools that can be used to evaluate biometric systems. We will use them throughout our thesis to evaluate the performance of the iris recognition algorithms we propose. In this chapter we will also present the public iris image databases which were the basis of our evaluation, and describe the major benchmarks that were conducted for iris recognition.

Chapter 3 will present a literature review. More specifically, we will present the most important existing iris recognition systems, the quality metrics for iris images that have been defined and the way they have been integrated in recognition systems. From this, we will explain the novelty of our work.

In Chapter 4, we will explain how we have elaborated our own local quality metric. More precisely, we will describe the statistical model we have build to measure image quality locally in terms of occlusion and amount of texture. We will present the performance of our quality metric through examples and statistics. We will also define a global quality metric. Then, we will demonstrate that our local and global quality measures are correlated to the matching performance, whether the method used for matching is Daugman's or Krichen's. We will deduce the pertinence of combining the usage of our quality measures with these matching algorithms.

This will lead us to propose, in Chapter 5, a method to integrate our local quality measure in an iris recognition system based on Daugman's algorithm. It consists in selecting the regions in the normalized iris images used for feature extraction and matching. We will present the advantages of our method compared to the ones presented in the literature for integrating local quality in Daugman's algorithm and evaluate our method's performance on different databases. We will explain the results we have obtained and demonstrate that our global quality measure can be used to optimize our method so that each image couple to compare is treated appropriately.

In Chapter 6, we will apply the same method to an iris recognition system based on Krichen's algorithm and show that the results obtained in the previous chapter can be extended to this algorithm.

In Chapter 7, we will compare the iris recognition systems based on Daugman's and Krichen's algorithms when they are used with the region selection method presented in Chapters 5 and 6. For this purpose, we will consider different categories of images build by taking into account different image quality characteristics. This will lead us to presenting our complete algorithm for iris recognition that combines Daugman's and Krichen's algorithms as well as our local and global quality measures.

Finally, we will conclude in Chapter 8 and put our work into perspective.

**CONFIDENTIAL**

## Chapter 2

# Performance evaluation in biometrics

As explained in Chapter 1, biometric systems are used in a large range of applications that are often linked to security issues. A failure of these systems can have severe consequences. It is therefore very important to evaluate these systems before deploying them, to determine the types of failures that can occur, and to measure the frequency of these failures.

As explained by Phillips et al. in [69] there are three different ways of evaluating biometric systems:

- technology evaluation,
- scenario evaluation,
- operational evaluation.

The technology evaluation only takes into account the performance of the system's algorithms without considering any application aspects. Other aspects related to the scenario can be evaluated: such as the performance linked to specific sensors, environments or populations. This constitutes the scenario evaluation. Finally, the operational evaluation tests a complete biometric system in a specific application environment with a specific target population. In this thesis, our goal is to improve iris recognition systems from an algorithmic point of view. We have therefore exclusively used technology evaluation. We have used pre-acquired iris images from public databases to carry out this evaluation.

In Section 2.1 we will present in detail how we measure the performance of biometric recognition algorithms. Such an evaluation can only be done if the algorithms can be tested on given datasets, publicly available and containing a large number of biometric data of various quality levels. We will describe the databases we have used for evaluating our iris recognition algorithms in Section 2.2. Finally, two recognition systems can only be rigorously compared if they are used with the same input data and the same protocol. Evaluation campaigns have been organized for this purpose. We will present the major evaluation campaigns in the field of iris recognition in Section 2.3.

## 2.1 How to measure performance

The aim of a biometric recognition system is either to verify an identity claim, or to identify people, based on their biometric data. To this end, it compares the acquired biometric data to one or many other biometric samples, previously acquired. Consequently, such a system performs comparisons of two biometric samples in order to decide whether they match or not. More precisely, a matching score is calculated to quantify the similarity between the two samples. A decision is then made by comparing this similarity score to a threshold value  $th$ . If the score is smaller than  $th$ , then the decision is taken that the samples do not match and vice-versa.

It is important to note that a person's identity is associated to each biometric sample. Therefore a match of biometric samples comes to accepting that the two samples are associated to the same identity.

### 2.1.1 Definition of the error rates

#### Failure rates

Prior to the comparison of two biometric samples, a recognition system can decide that it is not capable of carrying out this comparison. This can be due to three factors:

- the system has failed to automatically capture the biometric modality when it is presented to the sensor.
- the system has failed to enroll a user i.e., to retrieve the associated biometric sample
- the system considers that the quality of one of the biometric samples is insufficient to be used for matching

In consequence, three failure error rates are associated to recognition systems: the Failure To Capture (FTC) rate, the Failure To Enroll (FTE) rate and the Failure To Match (FTM) rate.

#### Accuracy rates

Let us consider two biometric samples of the same biometric modality (for example the right iris) that have passed these rejection tests. These samples verify one of the following hypothesis:

- $H_0$ : they belong to different persons
- $H_1$ : they both belong to the same person

The comparison of these two samples by a recognition system has two possible outcomes:

- $D_0$ : the system decides that the samples do not match
- $D_1$ : the system decides that the samples match

Given these definitions, it is easy to understand that the recognition system makes an error when it decides:

**CONFIDENTIAL**

- $D_0$  when  $H_1$  is true : False Non Match or False Rejection
- $D_1$  when  $H_0$  is true : False Match or False Acceptance

Note that, a match leads to accepting that the two biometric samples belong to the same person while a non match leads to rejecting this hypothesis. Therefore, a False Match can also be referred to as a False Acceptance, whereas a False Non Match can be referred to a False Rejection.

The objective when building a biometric system is to minimize simultaneously the occurrence of the False Acceptances and the False Rejections. We therefore define the False Acceptance Rate (FAR) and the False Rejection Rate (FRR) that are the probabilities of these two types of errors:

$$\begin{aligned} FAR &= P(D_1|H_0 = true) \\ FRR &= P(D_0|H_1 = true) \end{aligned}$$

In practical terms, these rates are evaluated by calculating the similarity scores for a set of matches that verify  $H_0$ , also called inter-class matches or impostors, and a set of matches that verify  $H_1$  (intra-class matches or authentic). This leads to the computation of the inter-class and intra-class distributions:  $p(s|H_0 = true)$  and  $p(s|H_1 = true)$ . For a given threshold  $th$  to which the similarity scores are compared in order to decide between  $D_0$  and  $D_1$ , we can then define the FAR and FRR as follows, assuming that the similarity scores have been normalized in order to be between 0 and 1:

$$FAR = \int_{th}^1 p(s|H_0 = true) ds$$

$$FRR = \int_0^{th} p(s|H_1 = true) ds$$

Figure 2.1 illustrates the computation of the FAR and FRR when  $th = 0.35$ . As we can see, for a given set of inter-class and intra-class distributions, changing the value of  $th$  would modify the values of the FAR and the FRR inversely. More specifically, a reduction of the FAR leads to an increase of the FRR and vice-versa. So the only way to improve recognition performance is to modify the system so that the overlapping between the two distributions is less important. In the case of an ideal recognition system, these two distributions would be completely separated.

Biometric recognition systems can be used in two types of scenarios: for verification and identification. We will now see how the error rates we have defined are used in each type of scenario.

### Verification

In the case of a verification, a person claims a particular identity and the biometric system is used to verify or reject the claim. The system compares a biometric sample acquired at the time of the claim to a sample that has been

**CONFIDENTIAL**

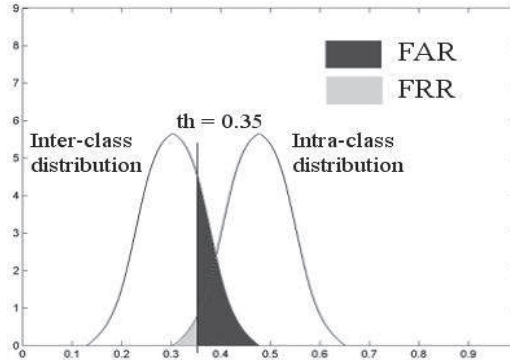


Figure 2.1: Inter-class and intra-class similarity score distributions

previously enrolled and is associated to the claimed identity. So the verification scenario comes down to exactly one comparison of two biometric samples, which is what we have dealt with here above. Therefore the FAR and FRR can directly be used to measure the errors made by a verification system. The FAR will evaluate the probability that the system accepts an identity claim that is false and the FRR represents the probability that it rejects a claim that is true.

### Identification

In an identification scenario, a biometric sample is acquired without any associated identity claim. The objective is to determine if the sample matches any of the  $N$  biometric samples contained in a database that are known and have been previously enrolled. So the system performs  $N$  comparisons of biometric samples.

For this scenario, we define the following error rates:

- $FRR_N$ : False Rejection Rate for the  $N$  comparisons
- $FAR_N$ : False Acceptance Rate for the  $N$  comparisons

We assume that a single biometric sample is present in the database for each user and that no indexing or retrieval mechanism is used. Then the  $FRR_N$  measures the chances that the decision  $D_0$  is taken when the acquired biometric sample is compared to the single sample in the database that verifies  $H_1$ . The  $FAR_N$  measures the chances that  $D_1$  is taken for any of the  $N - 1$  or  $N$  comparisons with the biometric samples that verify  $H_0$ . Consequently, we have the following relation:

$$\begin{aligned} FRR_N &= FRR \\ FAR_N &= 1 - (1 - FAR)^N \end{aligned}$$

Indeed, given that a single biometric sample is present in the database for each user, there is at most one comparison out of the  $N$  in total that verifies the hypothesis  $H_1$ , so  $FRR_N = FRR_1 = FRR$ .

On the other hand, there can be  $N$  comparisons that verify  $H_0$ . So the complementary of  $FAR_N$  is the probability that all the impostors are rejected

**CONFIDENTIAL**

which is the same as  $(1 - FAR)^N$  since  $1 - FAR$  is the probability of rejecting an impostor. Therefore  $FAR_N = 1 - (1 - FAR)^N$ .

Note that we can make the following approximation if the FAR is small:  $FAR_N \cong N \cdot FAR$ .

To the contrary, if we assume that the biometric samples in the database have been indexed or classified then a limited part of the database is searched during the identification. The penetration rate  $P$  represents the average percentage of the database that is selected at the indexing step and searched during identification.

In this case, the expressions of the  $FRR_N$  and the  $FAR_N$  are different. Indeed, to compute the  $FRR_N$  we must take into account the fact that the biometric sample that matches the newly acquired sample may have been eliminated at the indexing step and therefore is not one of the  $NP$  samples that the test sample will be compared to. The probability for this to happen is measured by the Retrieval Error Rate (RER). So the  $FRR_N$  can be expressed the following way:

$$FRR_N = RER + (1 - RER) \cdot FRR$$

Since the test sample is compared to  $NP$  samples from the database:

$$FAR_N = 1 - (1 - FAR)^{NP}$$

### 2.1.2 Performance curves and accuracy indicators

We have seen that two types of errors can occur in biometric recognition systems and that they can be evaluated by computing the False Rejection Rate (FRR) and False Acceptance Rate (FAR) as a function of the threshold  $th$  that separates the accepted similarity scores from the rejected ones. Therefore a first way of representing the performance of such systems is to plot  $FRR(th)$  and  $FAR(th)$  as it is done in Figure 2.2.

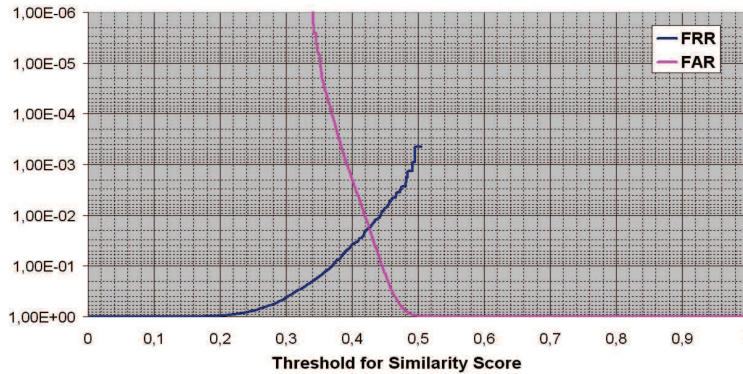


Figure 2.2: FAR and FRR as functions of the threshold  $th$

This performance can also be summarized in a Receiver Operating Characteristic (ROC) curve, that plots the FRR as a function of the FAR. For an ideal

**CONFIDENTIAL**

system i.e., a system for which the inter-class and intra-class distributions do not overlap, the ROC curve takes a unique shape. As in this case it is possible to achieve simultaneously  $FAR=0$  and  $FRR=0$ , this curve will coincide with the X and Y axis.

Figure 2.3 represents the ROC curve associated to Figure 2.2.

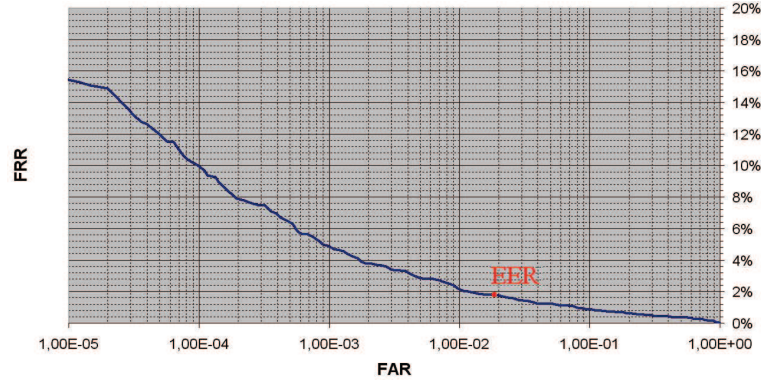


Figure 2.3: Example of a ROC curve and the associated EER

Accuracy indexes are also used to evaluate the performance of a recognition system:

- Equal-Error Rate ( $EER$ ) is the value of the ROC curve where  $FAR=FRR$
- $SetFAR$  is the FRR for a set value of the FAR
- $SetFRR$  is the FAR for a set value of the FRR

Usually the accuracy index that is used depends on the application for which the system is destined. For instance, for high security access applications the priority is to avoid allowing access to an impostor, so we will set the FAR at a very low value. The index used to evaluate the system will therefore be the  $SetFAR$ . On the other hand, for forensic applications the priority is to avoid falsely rejecting a biometric sample when comparing it to a blacklist which would lead to the failure to identify a criminal. Therefore we will set the FRR at a very low value for this application and use the  $SetFRR$  as an accuracy index.

### 2.1.3 Estimation of confidence intervals

We have seen that the performance of a biometric recognition system is mainly evaluated on the basis of its False Acceptance Rate (FAR) and False Rejection Rate (FRR). These two rates are defined as functions of the inter-class and intra-class similarity score distributions. In practical terms these distributions are unknown. They are estimated with the similarity scores resulting from a finite number of inter-class and intra-class comparisons. Therefore, the computation of the FAR and FRR only gives us an estimation of these rates' true values.

The more inter-class and intra-class comparisons we have at our disposal, and the better they represent the diversity of the target population, the closer

**CONFIDENTIAL**



the estimated values of the FAR and FRR will be to the true ones. Either way, it is important to be able to associate our estimated values of the rates to confidence intervals before relying on them.

There are mainly two ways of estimating the confidence intervals for the FAR and FRR: one is parametric [75] and the other one non-parametric [45]. They both suppose that the inter-class and intra-class similarity scores are generated from independent and identically distributed variables.

### Parametric estimation of confidence intervals

In this section we will consider the FRR. However, the same arguments can be used for the FAR.

Let us consider  $N_{intra}$  intra-class comparisons associated to  $N_{intra}$  similarity scores. As explained before, each similarity score will be compared to a threshold  $th$  and a decision is made:

- if  $s > th$ , we decide  $D_1$ , which leads to a True Acceptance (TA)
- if  $s \leq th$ ,  $D_0$ , which leads to a False Rejection (FR)

We can assume that for a set value of  $th$ , the  $N_{intra}$  decisions can be modeled by independent Bernoulli trials. So we can consider the binomial distribution defined by the probability  $p$  of a False Rejection, estimated by  $\frac{\text{number of FR}}{N_{intra}} = FRR$ . So:

$$E(p) = FRR$$

$$Var(p) = \sqrt{\frac{FRR(1-FRR)}{N_{intra}}}$$

Under certain conditions on  $N_{intra}$  and  $p$ , the binomial distribution can be approximated by a normal distribution  $\mathcal{N}(FRR, \sqrt{\frac{FRR(1-FRR)}{N_{intra}}})$ . In this case, the FRR can be seen as the mean of this normal distribution and we can use the properties of normal distributions to give the 90% confidence interval of the FRR:

$$FRR \in [FRR - \alpha \sqrt{\frac{FRR(1-FRR)}{N_{intra}}}, FRR + \alpha \sqrt{\frac{FRR(1-FRR)}{N_{intra}}}] \quad (2.1)$$

where  $\alpha = 1.645$ . In the same way, we can show that:

$$FAR \in [FAR - \alpha \sqrt{\frac{FAR(1-FAR)}{N_{intra}}}, FAR + \alpha \sqrt{\frac{FAR(1-FAR)}{N_{intra}}}] \quad (2.2)$$

where  $\alpha = 1.645$ . To obtain a 95% confidence interval, the value  $\alpha$  takes the value  $\alpha = 1.96$

These equations establish the following results:

- To be 90% confident that the true error rate is within  $\pm 30\%$  of the observed error rate, there should be at least 30 errors.

**CONFIDENTIAL**



- To be 90% confident that the true error rate is within  $\pm 10\%$  of the observed error rate, there should be at least 260 errors.

This means, that we will need to perform at least 300 000 inter-class matches to be able to evaluate an FAR around  $10^{-4}$ .

This estimation of the confidence interval is parametric in that we need to make assumptions on the distributions of the error rates (the FRR and the FAR) to use the percentiles of the normal distribution.

### Non-parametric estimation of confidence intervals

Another method to estimate the confidence intervals is the bootstrap technique [45] that has the advantage of being non parametric. The idea of this technique is to re-sample the test samples in order to have  $K$  sets of intra-class and inter-class similarity scores. The FRR and FAR are calculated for each set, which results in  $K$  values for the FRR and  $K$  values for the FAR. From this we can compute the mean value of the error rates as well as the standard deviation which gives us confidence intervals.

## 2.2 Available databases

We have seen in Section 2.1 that it is important to have a large number of biometric sample comparisons in order to evaluate the performance of a recognition system with an acceptable precision. To meet this need, several public databases have been made available.

In the case of the iris, the databases acquired in the near-infra-red (NIR) with stationary subjects that are the most often used by researchers to evaluate their iris recognition systems are: ICE [5], ND-IRIS-0405 [70] and CASIA-Iris [18]. These databases are interesting for us because they contain a high number of iris images of various quality levels. Moreover, since they are often used by authors, evaluating our algorithms on these databases enables us to compare our performance to those presented in literature. We will also mention some of the other databases that exist and explain why we did not exploit them in this thesis.

### 2.2.1 ICE and ND-IRIS-0405

The ICE databases [5] were collected by the Computer Vision Research Lab at the University of Notre Dame. The sensor used to acquire the iris images was the LG 2200. ICE 2005 contains 2953 images of 244 irises and ICE 2006 contains 60 000 images of 480 irises.

The ND-IRIS-0405 [70] is a superset of these iris image datasets. It contains 64 980 images corresponding to 356 unique subjects and 712 unique irises. The age range of the subjects is 18 to 75 years old. 250 of the subjects are Caucasian, 82 are Asian and 24 are of other ethnicities. The images in this database contain a large variety of artifacts such as occlusion, motion-induced interlacing, out-of-focus images and artifacts resulting from contact lenses.

Figure 2.4 shows examples of images from the ND-IRIS-0405 database.

**CONFIDENTIAL**

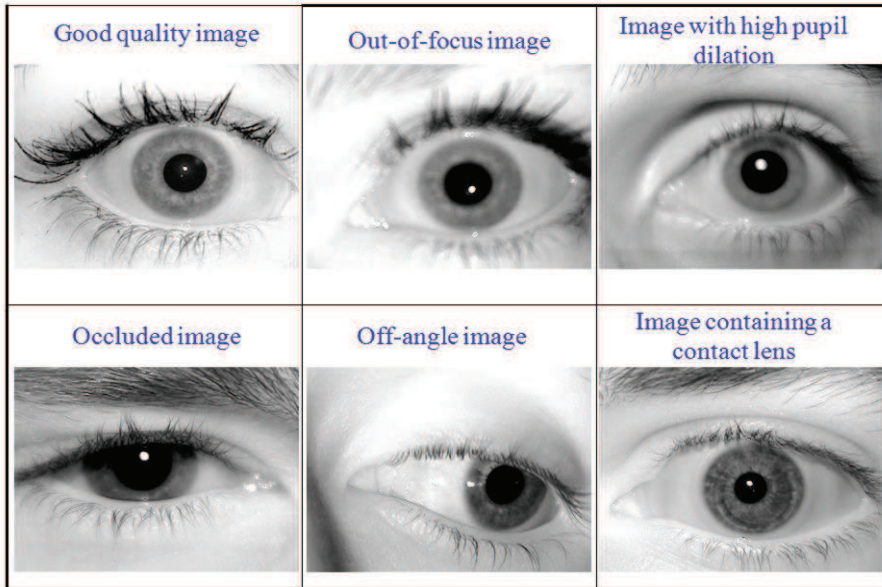


Figure 2.4: Examples of images from the ND-IRIS-0405 database

### 2.2.2 CASIA-Iris

CASIA V1.0 and CASIA V3.0 databases were collected by the Chinese Academy of Sciences Institute of Automation (CASIA) [18]. Almost all subjects are Chinese except a few in CASIA-IrisV3-Interval. The images from CASIA V1.0 were captured with a camera that was developed at the CASIA. The database includes 756 images from 108 irises. For each iris, 7 images were captured in two sessions : three samples were collected in the first one and four in the second one. The pupil regions were replaced by a circular region of constant intensity to mask out the specular reflections from the NIR illumination.

CASIA V3.0 contains three subsets : CASIA-IrisV3-Interval, CASIA-IrisV3-Lamp and CASIA-IrisV3-Twins. CASIA-IrisV3-Interval is a superset of CASIA V1.0 and contains 2639 irises from 249 subjects. Unlike CASIA V1.0, in CASIA-IrisV3-Interval the pupil regions are not replaced with a circular region of constant intensity. The image artifacts in this database are limited to occlusions. Globally the image quality is good with extremely clear iris texture details. Figure 2.5 represents an example of images from CASIA V1.0 and CASIA-IrisV3-Interval.

CASIA-IrisV3-Lamp was collected using OKI's hand-held iris sensor. A lamp was turned on and off close to the subject, so the images contain non linear deformation due to variations of visible illumination. The database contains 16213 iris images from 411 subjects.

CASIA-IrisV3-Twins was collected outdoors with the same sensor as CASIA-IrisV3-Lamp. It contains 3183 iris images from 100 pairs of twins and is the first publicly available twins' iris dataset. Figures 2.6 and 2.7 present images respectively from the CASIA-IrisV3-Lamp and the CASIA-IrisV3-Twins databases.

**CONFIDENTIAL**

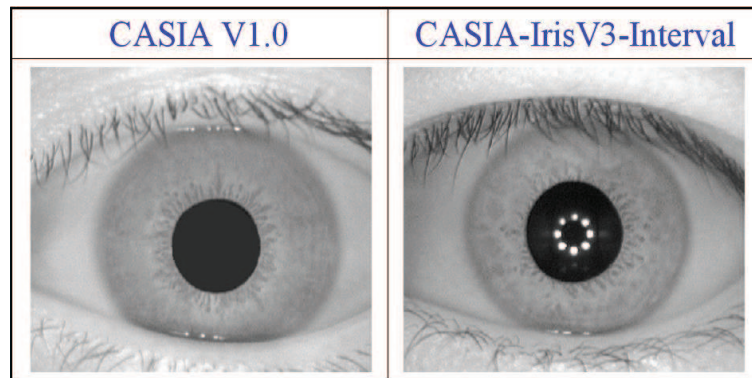


Figure 2.5: Examples of images from the CASIA V1.0 and CASIA-IrisV3-Interval databases

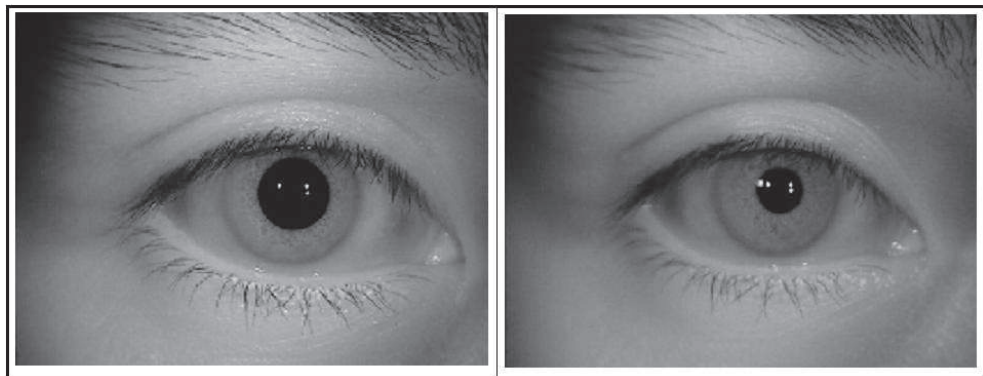


Figure 2.6: Examples of images from the CASIA-IrisV3-Lamp dataset

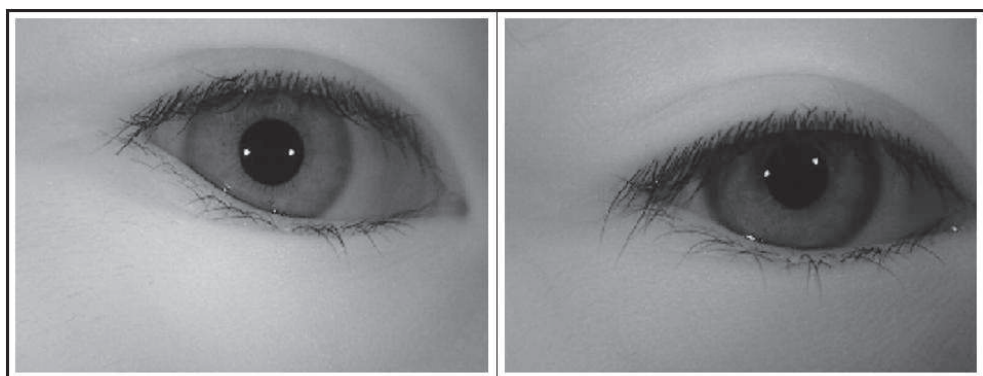


Figure 2.7: Examples of images from the CASIA-IrisV3-Twins dataset

**CONFIDENTIAL**

The image quality of the Twins and Lamp subsets is clearly less good than the one of Interval, with particularly many occlusions and non linear deformation due to variations of visible illumination.

The characteristics of the three subsets of CASIA-IrisV3 are summarized in Figure 2.8.

Subset \ Characteristics	CASIA-Iris -Interval	CASIA-Iris -Lamp	CASIA-Iris -Twins
Sensor	CASIA close-up iris camera	OKI IRISPASS-h	OKI IRISPASS-h
Environment	Indoor	Indoor with lamp on/off	Outdoor
Session	Two sessions for most iris images	one	one
Attributes of subjects	Most are graduate students of CASIA	Most are graduate students of CASIA	Most are children participating Beijing Twins Festival
No. of subjects	249	411	200
No. of classes	395	819	400
No. of images	2,639	16,212	3,183
Resolution	320*280	640*480	640*480
Features	Cross-session iris images with extremely clear iris texture details	Nonlinear deformation due to variations of visible illumination	The first publicly available iris image dataset of twins
Total	A total of 22,034 iris images from more than 700 subjects and 1500 eyes.		

Figure 2.8: Statistics from the CASIA-IrisV3 datasets

### 2.2.3 Other iris datasets

The databases presented above are the ones in the NIR domain that are the most commonly used for evaluating iris recognition systems. However, other NIR wavelength iris image databases exist: MMU [4], UBATH [6] and WVU. Table 2.1 presents these databases. UBATH is not very interesting for us because it contains almost exclusively high quality iris images. As to MMU, each of its subsets contains less than 1000 iris images. Finally, WVU contains a higher number of images, but it is less exploited by researchers than the ones we have worked with. Table 2.1 describes these NIR iris image databases.

Note that visible wavelength iris image databases also exist. UBIRIS [71] and UPOL [24] are such databases. Databases acquired with mobile subjects

**CONFIDENTIAL**

Table 2.1: Description of different NIR iris image databases

Database	Number of irises	Number of images	Camera used	Image quality
MMU1	90	450	LG IrisAccess	deteriorated
MMU2	199	995	Parasonic BM-ET100US Authenticam	deteriorated
UBATH	800	16 000	ISG LightWise LW-1.3-S-1394	high quality
WVU	488	3099	OKI irispas-h	deteriorated

or a high camera distance from the iris exist as well. We can name MBGC [3] and QFire [41]. However in this thesis we have chosen to work only with images acquired with NIR illumination and stationary subjects, so we have not exploited these databases.

## 2.3 Benchmarks

After describing how to evaluate the performance of a given biometric recognition system and some iris image databases, we will present in this section a few benchmarks for comparing the performance of different iris recognition systems.

Two recognition systems can only be rigorously compared if they are used with the same input data and the same protocol. For this reason, benchmarks have been organized. They are in general public and organized by an independent organism. The benchmarks linked to iris recognition have mainly been coordinated by the National Institute of Standards and Technology (NIST).

The first one was the Iris Challenge Evaluation (ICE) of 2005. Its results were published in March 2006. This benchmark used the ICE 2005 database described in Section 2.2. For most subjects, both the left and right irises were captured. The images were acquired with the LG 2200 sensor.

The ICE 2005 benchmark was the first to provide a large publicly available dataset, as well as a common protocol for measuring algorithm performance. In all, 16 different algorithms were compared. The performances range from  $FRR = 0.2$  to  $FRR = 30$  for a  $FAR = 10^{-3}$ . More details about this benchmark can be found in [68].

This benchmark was closely followed by the Iris Challenge Evaluation (ICE) of 2006 of which the results are presented in [70]. Its goal was to measure state-of-the-art iris recognition algorithms and establish a baseline on which to measure future evaluations. In opposition to the ICE 2005 benchmark, ICE 2006 measured performance with sequestered data (data not previously seen by the researchers or developers). The dataset used for this evaluation contained 60 000 images of 480 different irises.

The Noisy Iris Challenge Evaluation (NICE) took place from 2008 to 2009. It exploited noisy images from the UBIRIS database [71] and the evaluation was divided into two parts. The first part [72] focused exclusively on the seg-

**CONFIDENTIAL**



mentation and noise detection stages of iris recognition, while the second part [7] considered only the feature extraction and matching stages of the degraded visible wavelength iris images that had been segmented according to the first part of the evaluation.

More recently, NIST has initiated the Iris Exchange (IREX) evaluations that are aimed at completing three activities:

- IREX I: addressing standards, formats and compression for data interchange
- IREX II: defining and measuring iris image quality
- IREX III: measuring the accuracy and the speed of iris identification algorithms

The reports of these three evaluations can be found on the website [2].

It is interesting to note that both ICE and NICE use datasets with degraded image quality. This shows that the robustness of the iris recognition algorithms to degraded qualities has been an issue for several years now.

In line with this, IREX II was recently conducted to improve the general knowledge on iris image quality and to understand how it is linked to recognition performance. The results of this benchmark are presented in [80] and will be discussed in more detail in Chapter 3 with the rest of our literature review on iris image quality. One part of IREX III also is to assess the utility of quality values in terms of its relation to recognition outcome. This evaluation demonstrates that the most capable quality assessor assigns low quality values to only 23.6% of image pairs involved in the poorest 2% of false negative outcomes (i.e. the highest dissimilarity scores). For other implementations, this figure is as low as 2.5%. These figures show that rejection based on low image quality values would not prevent more than three fourth of these identification errors from happening.

## 2.4 Conclusion

In this chapter, we presented the different methods that are commonly used to evaluate biometric recognition systems. We demonstrated that such an evaluation requires a large number of inter-class and intra-class comparisons of biometric samples. Following this, we have presented the iris image databases that have been made publicly available to enable authors to evaluate their algorithms. In particular we have described the ICE 2005, ND-IRIS-0405 and CASIA-IrisV3 databases, used in our work. Finally, we presented the benchmarks that have taken place in the field of iris recognition.

It is interesting to note, that many of those benchmarks address the issue of the robustness of iris recognition algorithms to degradations of the input images' qualities. This shows that this issue has been one of the major challenges in iris recognition for several years now.

In this thesis, we also address the issue of the impact of degradations of the image quality on the performance of iris recognition systems. As explained in Chapter 1, our aim is to improve the feature extraction and matching steps of iris recognition so that it stays highly efficient if the acquisition conditions

**CONFIDENTIAL**

are loosened and the quality of the iris image diminished. In the next chapter, we will present a literature review to demonstrate that iris image quality has become a major line of research during the last years.

**CONFIDENTIAL**

## Chapter 3

# Literature review

In this chapter we present a literature review on iris recognition. A large part of it is based on the detailed survey that Bowyer et al. have written on iris recognition in 2008 [14]. In addition, we emphasize on the methods authors have proposed to handle the degraded quality of iris images for recognition. This will lead us to point out what contribution this doctoral work makes in the management of the issue of degraded iris images.

The idea of using the iris as a biometric modality was suggested by A. Bertillon in 1885 [11], long before the first biometric system was built. Then in 1987, two ophthalmologists, Flom and Safir proposed a conceptual design of an automatic iris biometric system and obtained a patent [30]. An investigation of the feasibility of iris biometrics was conducted a few years later by Johnston at Los Alamos National Laboratory in 1992 [42] and it was concluded that iris biometrics held potential for both verification and identification.

Daugman was the first author to propose an operational iris recognition system in 1993 [21]. In this system, Gabor filters are used to generate a binary code representing the iris. Irises are then compared by an efficient comparison of their binary codes using bitwise operations. Later on, many authors have looked at alternate ways to generate a binary code. Others have chosen to represent the iris by a real-valued feature vector. In Section 3.1 we will describe Daugman's algorithm in some detail, as well as some of the alternate ways to generate the binary code and methods to represent the iris based on real valued feature vectors.

Furthermore, it is interesting to note that whatever system is used, the recognition performance has been shown to depend on the quality of the input images. Therefore, work has been undertaken to define relevant quality metrics for iris images and to integrate them in iris recognition systems. This work is presented in Sections 3.2 and 3.3.

Finally, in Section 3.4, we present some reference systems, as well as the algorithm we have used for the segmentation and normalization of the iris in order to work with normalized iris images.



## 3.1 Major existing iris recognition systems

### 3.1.1 Daugman's algorithm

#### Initial algorithm

Daugman was the first author to elaborate an operational iris recognition system. The system he proposed in 1993 [21], was innovative on many levels. He cut the problem into the four following steps: segmentation, normalization, feature extraction and matching, and proposed an original solution for each one of them.

In the segmentation step, Daugman approximates the iris outer and inner boundaries by circles of center  $(x_0, y_0)$  and radius  $r$ , and detects these circles by an integro-differential operator that searches the parameter space. His operator is given in formula 3.1 where  $*$  represents the convolution operator.

$$\max_{r, x_p, y_0} |G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r, x_p, y_0} \frac{I(x, y)}{2\pi r} ds| \quad (3.1)$$

Then he normalizes the iris in order to overcome the fact that the size of a given iris can vary from one image to another. To do this he assumes that the iris stretches linearly when the pupil dilates and contracts. The iris is transformed into a rectangle of pre-set dimensions: its radial coordinates are represented along a vertical axis and the angular coordinates on the horizontal axis. More precisely the transformation is done according to the following formula:

$$\begin{cases} x(r, \theta) = (1 - r)x_p(\theta) + rx_s(\theta) \\ y(r, \theta) = (1 - r)y_p(\theta) + ry_s(\theta) \end{cases}$$

This normalization technique is illustrated in Figure 3.1. It is often referred to as the rubber-sheet model of the iris.

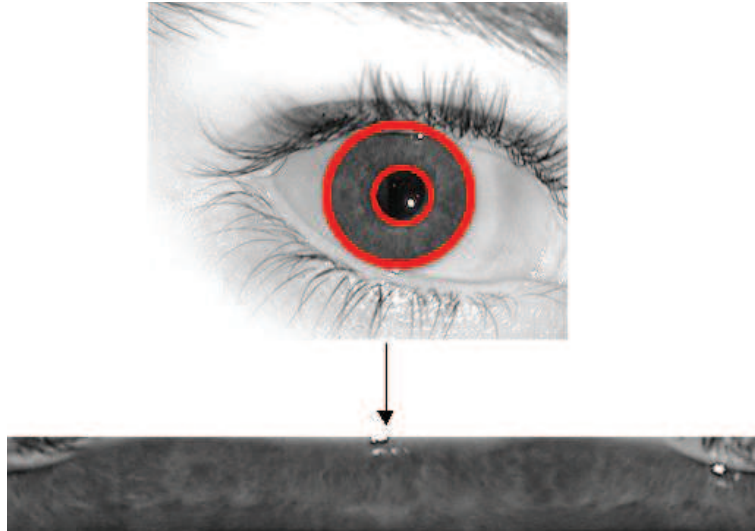


Figure 3.1: Illustration of the normalization process

**CONFIDENTIAL**

In the next step, Daugman has used convolution with two dimensional Gabor filters to extract the texture from the normalized iris image. These filters analyze the iris texture at different resolutions and orientations. He then uses the phase of the resulting complex coefficients to represent the iris. More precisely, in order to compress the information, the phases are quantized: each coefficient is assigned a different pair of bits depending on the sign of its real part and imaginary part as we can see in Figure 3.2.

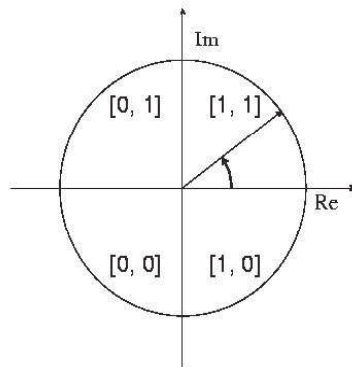


Figure 3.2: Illustration of the phase quantization

Daugman chose to keep only the phase information for encoding the iris because it is less dependent on imaging contrast or illumination than the amplitude information; it is also much more discriminating. Moreover, the phase quantization he uses enables us to capture the information of wavelet zero-crossings, since it is directly linked to the signs of the real and imaginary parts of the complex coefficients resulting from the Gabor filtering.

As a result to this encoding scheme, each iris image is represented by a binary code, which Daugman has called “iris code”. Its length depends on the number of Gabor filters that are applied. Daugman chose to use a 2048 bit binary code.

Finally, the matching of two irises is carried out by a bitwise comparison of the associated binary codes. More specifically, Daugman uses the Hamming distance to evaluate the dissimilarity between two iris codes. This distance represents the ratio of the number of bits that disagree on the total number of bits. It is equal to 0 when the two iris codes are identical and equal to 1 when all bits disagree. Since each bit of any iris code has equal odds of being 0 or 1, the probability that any pair of iris codes match is equal to 0.5. So the distribution of observed Hamming distances between independent iris codes should be a binomial distribution centered at 0.5. Therefore, the result of the matching of two irises can be seen as the result of a statistical test of independence between the two iris codes.

If each of the 2048 bits of the iris code were independent, then the inter-class distribution would be a binomial distribution with  $p = 0.5$  and  $N = 2048$ . However, the 2-D Gabor filters introduce an intrinsic correlation between the code’s bits. Moreover, there is a spatial correlation between the pixels of the

**CONFIDENTIAL**

iris image. Daugman demonstrates that the consequence of these two sources of correlation is that there are only 173 independent binary degrees of freedom in his iris code. So the inter-class distribution can be modeled by a binomial distribution with  $N = 173$  and  $p = 0.5$ . In consequence, the likelihood of the Hamming distance between two inter-class irises being equal to zero, meaning that the iris codes are rigorously identical, is roughly one in  $2^{173} \simeq 10^{-52}$ .

In order to take into account the variations of rotation of the iris from one image to another, the Hamming distance is computed for several different orientations when comparing a pair of iris images. Each orientation corresponds to a circular permutation of the iris code in the angular coordinate. The minimum computed Hamming distance is assumed to correspond to the correct alignment of the two images.

Allowing different degrees of head tilt generates biased inter-class and intra-class distributions: they will be shifted towards the lower Hamming distances, since we favor the lowest distance out of several trials.

### Latest improvements

Daugman has written several papers after this one to present techniques that can be associated to his first iris recognition algorithm in order to improve performance. In [20] he improves the segmentation of the iris. He adds the detection of curvilinear edges to localize the upper and lower eyelid boundaries to the existing detection of the iris' inner and outer boundaries. In consequence, a mask can be associated to each iris code to eliminate the bits that were generated from occluded regions of the iris. Daugman therefore changes the formula to compute the Hamming distance between iris codes to:

$$HD = \frac{\|(codeA \otimes codeB) \cap maskA \cap maskB\|}{\|maskA \cap maskB\|} \quad (3.2)$$

where  $\otimes$  represents the exclusive OR.

Daugman also demonstrates that the performance of his algorithm stays high when it is applied on a database of several thousand eye images acquired in trials in Britain, the USA, Japan and Korea that allows 9.1 million comparisons.

In his latest paper on iris recognition [23], Daugman presents four additional improvements for his system. Two of them are related to the segmentation of the iris: the first one is a method based on active contours that aims at modeling the inner and outer boundaries of the iris more faithfully, while the second one is a statistical inference method for detecting and excluding eyelashes. Another one is related to off-angle images. He describes a Fourier-based method to detect such images and rotate the eye into orthographic perspective. Finally, he presents a new way of normalizing the Hamming distance in order to take into account the amount of iris data that is available in images and the required scale of database search.

The raw Hamming distance  $HD_{raw}$  is calculated using the formula 3.2. In consequence, the number of bits taken into account varies from one comparison

**CONFIDENTIAL**

to another, since it depends on the masks associated to each image. If a big portion of an image is masked, then only very few bits will be used for the computation of the raw Hamming distance. This distance can therefore be biased. Therefore it is important to take into account the variations of the number of available bits. Consequently, Daugman decides to normalize the raw Hamming distance according to the following formula:

$$HD_{norm} = 0.5 - (0.5 - HD_{raw})\sqrt{\frac{n}{911}} \quad (3.3)$$

where  $HD_{raw}$  is calculated using formula 3.2,  $n$  is the number of available bits for the specific iris code comparison that is considered, and 911 is the mean of the number of available bits for all the comparisons of the database.

As shown in [20], Daugman's method allows high speed matching, since the computation of the Hamming distance only takes  $10\mu s$ . It also allows very high performance in terms of accuracy. When the criteria on the Hamming distance is set to 0.327, then there are no false acceptances and no false rejections for the 9.1 million comparisons that are carried out. However, he shows that the acquisition conditions of the iris have an impact on the intra-class distribution. If these conditions are non-ideal (images acquired at different distances, and by different optical platforms), this distribution is shifted towards the higher Hamming distances. In [20], the intra-class is not shifted enough to increase the error rates. However, this suggests that a degradation of the acquisition conditions can eventually lead to a reduction of the recognition accuracy.

Moreover, Daugman points out that it is important to mask the artifacts in the iris image so that they do not influence the iris comparison. This suggests that the non-detection of some artifacts can also reduce the recognition accuracy.

### 3.1.2 Alternate ways to generate binary codes

While many authors have kept the idea of representing the iris by a binary code which allows high speed matching, they have tried to test different techniques to extract the iris texture.

Some have generated binary codes from the iris by using Gaussian filters [77], dyadic wavelet transforms [57] or the discrete cosine transform [62]. In [81], Thornton et al. have used seven types of filters for extracting features from the iris and have compared the recognition performance for the different filters. The filters that were tested were the Daubechie's wavelet, the Bi-orthogonal wavelet, the Coiflet wavelet, the Symlet wavelet, the Haar wavelet, circular symmetric filters and the Gabor wavelet. They demonstrate that the best performance is achieved with Gabor wavelets. They also point out that it is crucial to tune the filter's parameters to optimize performance.

Huang et al. also stress the importance of choosing wisely the filters parameters and the scales at which to apply them. The aim of their work [37] is to adjust the scales for analysis specifically to different images.

Sun et al. have also worked on adjusting the filter's parameters for optimal performance in [34]. As in [77] they apply Gaussian filters to obtain ordinal features used for recognition. They generate a large amount of ordinal features

**CONFIDENTIAL**

by applying 708 filters to 224 regions of the iris in order to provide an over-complete iris feature. Then they use a Similarity Oriented Boosting algorithm to learn which filters and regions generate the most discriminating information.

### 3.1.3 Real-valued Feature vectors

While some researchers have focused their work on finding the best way to generate a binary code that summarizes the iris information, others have chosen to use a real-valued feature vector representation of the iris for comparison. Wildes was the first author to propose an iris recognition system based on a real-valued feature vector [86]. He applies Laplacian of Gaussian filters on the normalized iris images at four different resolutions. The filtered images are then directly used to form a feature vector, and the normalized correlation is calculated between the vectors to determine their similarity. Wildes' system is more complex than Daugman's and it leads to a less compact representation of the iris. However, it might be capable of finer distinctions between irises since it uses more data for comparison.

Another early system based on real-valued feature vectors is Boles and Boashah's one [13]. They work with 1D signals (concentric circular bands of the iris region), on which they apply a wavelet transform and extract a zero-crossing representation. They evaluate the dissimilarity between two irises by measuring globally the difference in energy between the two corresponding zero-crossing representations.

Following these two first approaches, a great variety of alternate real-valued feature vector based systems were proposed. Some generate the feature vector using wavelet transforms, such as Gabor filters [8, 56], Daubechies wavelets [8, 40, 31] or a Discrete Cosine Transform [8], while others use statistical methods, such as independent component analysis [38] or principal component analysis [26].

Kumar et al. have worked on generating specific correlation filters for each individual in [85]. Their method requires a training database containing several images for each iris in order to generate the associated correlation filter. This filter is then applied to a test iris image to determine if the iris matches the one in the training set. The filter generation is carried out by using the Optimal Trade-off Synthetic Discriminant Function (OTSDF). The goal is to minimize the Average Correlation Energy (ACE) and the Output Noise Variance (ONV). This way, the correlation peak will be narrow and will not depend on the image blur.

When working with images of degraded quality, systems that use real-valued feature vector can be interesting. Because they exploit more information for matching, they could be more robust to image degradation. Krichen et al. have proposed such a system in [48]. They use Gabor-filtered phase images as feature vectors and then compare irises by computing local normalized cross-correlations between different sub-images of the vectors. They use both the values of the different correlation peaks and their position to determine the similarity between the irises. We will describe their algorithm in detail in Chapter 4 where we use it to conduct experiments for evaluating our quality measure. They show that his system is more robust to image degradations than the OSIRIS implementation of Daugman's system.

**CONFIDENTIAL**

In the same way as Krichen, Du et al. [27] keep the Gabor filtering step but choose to exploit the Gabor phases in a different way than Daugman. More specifically, they apply a Scale Invariant Feature Transform (SIFT) in which they use the Gabor phases associated to key points to build the related descriptors. The descriptors of each iris are then compared two by two with the Euclidean distance. Different combinations are tested so that different transformations between the two irises can be compared. As Krichen's method, this technique is presented as being more robust to image degradations and specifically to local distortions.

### 3.1.4 Fusion of different techniques

Whatever technique is used for feature extraction, it has its advantages and drawbacks. For instance, Daugman's method allows a compact representation of the iris and high speed matching. Other systems, such as Wildes', use more data for comparison so they achieve finer distinction between irises even if they generate a slower matching.

For this reason, researchers have tried to combine different feature types in order to exploit complementary characteristics. Most often, a "local" feature extraction (for instance the one proposed by Daugman) is combined to a "global" one.

A first strategy is to elaborate a "cascaded" system that first calculates the similarity score by using one of the two feature extraction techniques (usually the less time-consuming one). Then, if this score is in an uncertainty interval, it uses the second feature extraction technique to calculate a new similarity that will enable to decide whether or not the irises match. Such a strategy is described by Sun et al. in [78] and by Zhang et al. in [87].

Another strategy, is to use both types of feature extraction techniques independently to calculate two different matching scores. Then the results are combined by performing score level fusion. This can be done, for instance, by calculating a weighted sum of the scores obtained with different matching methods. Vatsa et al. [83] and Park and Lee [65] apply this approach.

## 3.2 Quality metrics

### 3.2.1 Global quality metrics

So far we have described how researchers have tested a great range of techniques for the extraction of features in iris images and their matching. Whatever technique is used, many researchers have pointed out that the quality of the input images is a crucial determining factor of recognition performance. In [43], Kalka et al. point out 7 factors that reduce recognition performance: defocus, motion blur, off-angle images, eyelid or eyelash occlusion of the iris, specular reflection, illumination variations and low image resolution. Bowyer et al. show in [16] that pupil dilation and contact lenses also have a negative impact on performance. These researchers quantify each of these factors, by defining an associated quality metric and demonstrate the correlation of the different metrics and the performance.

**CONFIDENTIAL**

Many different quality metrics have been defined, most of which measure focus [29, 20, 23, 44], occlusion [54] or pupil dilation [54]. Belcher and Du measure “feature information” by calculating the relative entropy of the iris texture when compared with a uniform distribution in [10].

Researchers have also proposed to define fused quality measures that take into account several of these factors [43, 28].

The National Institute of Standards and Technology (NIST) has carried out an Iris Quality Calibration and Evaluation (IREX II IQCE), already mentioned in Section 2.3, which is aimed at defining and quantifying iris image properties that are influential on performance of iris recognition [80]. It lists various global quality metrics and compares their influence on recognition accuracy. Most of these metrics are related to the following standard quality components:

- Gray scale spread
- Iris size
- Dilation
- Usable iris area
- Iris-sclera contrast
- Iris-pupil contrast
- Iris shape
- Pupil shape
- Margin
- Sharpness
- Motion blur
- Signal to noise ration
- Gaze angle
- Interlace

It is demonstrated that the metric with the highest impact on recognition performance is the usable iris area, followed by iris-pupil contrast, pupil shape, iris-sclera contrast, gaze angle and sharpness.

All of these quality metrics can be seen as global quality metrics in the sense that they quantify the quality of an entire image.

### **3.2.2 Local quality metrics**

On the other hand, researchers have pointed out that quality can be defined in a more local way. A given iris image can indeed include both low quality regions that contain artifacts and high quality regions in which high textured parts of the iris are clearly visible. Consequently, some authors have defined

**CONFIDENTIAL**

local quality metrics to quantify the quality of different regions in a given iris image.

Chen et al. [19] were the first to propose this approach. They use a wavelet transform based quality metric that measures the energy in a given concentric band of the iris. Krichen proposed in his thesis [46] a probabilistic local (pixel-level) quality measure relying on a Gaussian Mixture Model (GMM) that can quantify the levels of blur or occlusion. In [39], Li and Savvides propose the same type of model, namely a GMM based approach to localize pixels that have any type of occlusion. Note that they use two GMMs, which are trained respectively on good quality and low quality images, while Krichen used only one GMM learned on the good quality texture. Krichen therefore did not need any database of noisy images and can easily adapt his technique to any type of new artifacts.

The local quality measures presented here above are directly linked to the quality of regions of the iris image. In principle, it should be pertinent to exploit them whatever technique is used for encoding and matching the iris. On the other hand, Hollingsworth et al. chose in [35] to define a local quality measure that is specific to Daugman's recognition technique for it is not determined from the image of the iris but from its iris code. They distinguish consistent and inconsistent bits. If a bit is stable in iris codes resulting from different images of the same iris, then it is consistent. Otherwise it is inconsistent. So they measure the quality of a bit by computing its stability.

Since the bits are obtained with the phase quantization process that is described in Section 3.1.1 and that follows the Gabor filtering, an unstable bit corresponds to an output Gabor coefficient for which the sign of the real or imaginary part changes. This can mainly be due to two factors:

- The fact that the output Gabor coefficient has a real or imaginary part that is close to zero. This coefficient varies slightly from one image to another of the same iris, so those values can flip between positive or negative values.
- The fact that the output Gabor coefficient is associated to a portion of the image that commonly contains some type of occlusion. Its value will therefore change completely whether the occlusion is present or not.

In practical terms, they put forward different ways of measuring the consistency of a bit. One way uses a training process that requires several registered images of the same iris. They consider  $n$  iris codes in the same class to determine for each bit the number of times,  $m_1$ , that it is equal to 1 and the number of times,  $m_0$ , that it is equal to 0. The stability of the bit is then measured by:

$$s = \frac{|m_1 - m_0|}{m_1 + m_0} \quad (3.4)$$

Dong et al. have also worked on a bit stability measure in [25]. They use the same training process as Hollingsworth et al. but compute the bit stability using a different formula:

$$s = 2 \frac{m_1^2 + m_0^2}{(m_1 + m_0)^2} - 1 \quad (3.5)$$

**CONFIDENTIAL**



They update this measure when the number of training images increases.

We will see in the next section that these local quality metrics can be very useful for iris recognition. They are the key to localizing the parts of the iris image or the parts of the iris code that will be the most useful for matching.

### 3.3 Integration of quality metrics in recognition systems

#### 3.3.1 Global quality measures

As explained in Section 3.2, a pertinent global quality measure for iris images will be correlated to the performance of recognition. The report of the IREX II evaluation [80] describes different ways of using global quality metrics to enhance recognition performance. They can be used to screen out poor quality images before recognition, initiate the acquisition of new data or influence a multi-modal biometric fusion process.

One of the results presented in the IREX II report is the improvement of the FRR when 3% of the images are screened out according to different quality components. This experiment was performed on the ICE 2006 database. The baseline FRR (without screening out any images) is  $FRR=0.1$ . The highest improvement is obtained when rejecting the images with the lowest usable iris area. In that case  $FRR=0.09011$ . Then comes the rejection according to the iris/pupil contrast, which enables to achieve  $FRR=0.09085$ . The lowest improvement is obtained when rejecting the images with the worst Signal to Noise Ratio. In that case,  $FRR=0.0999$ .

Finally, a scalar quality metric was considered that combined all or some of the quality components presented in Section 3.2.1. When rejecting the 3% images with the lowest scalar quality the achieved FRR was  $FRR=0.892$ .

Global quality metrics can also be used to adapt the recognition algorithm to image quality. For instance, Schuckers et al. use in [76] a global quality measure to detect off-angle iris images and adapt the recognition algorithm to the deviation of the eye.

#### 3.3.2 Local quality measures

Local quality metrics such as the ones described in Section 3.2.2 can be very useful for iris recognition. Indeed, we have seen that errors made at the segmentation step can leave us with normalized iris images containing mixed quality regions. Furthermore, Daugman points out in [20] that the highest recognition performance are obtained when the feature extraction is done in regions of the normalized iris image that are of good quality, namely well textured parts of the iris that do not contain artifacts such as eyelid or eyelash occlusion, specular reflection and excessive blur. At first, authors systematically excluded some regions of the iris images to avoid taking into account low quality regions. Then they have started to use local quality metrics to identify high quality regions and exploit them for matching. We describe these two approaches in this subsection.

**CONFIDENTIAL**

### Systematic region exclusion

Before using local quality metrics researchers systematically excluded the regions that are considered the most likely to contain artifacts. For example, in [21] Daugman chose to exclude two portions of the iris: one at the top that is often occluded by the upper eyelid, and another at the bottom where specular reflection is common. Other authors exclude iris regions near the limbic boundary [61, 58, 82]. Pereira and al. showed that some specific combinations of concentric bands of the iris region improve the reliability of recognition [67, 66], and therefore excluded the remaining bands. More recently, Murugan and Savithiri [63] demonstrate that they improve the reliability of iris recognition when they extract the features only from half of the iris (either bottom half or the inner half).

The drawback of such an exclusion strategy is that it does not take into account the specificity of each iris image. Even though the excluded regions are those that are the most likely to contain artifacts, these can unpredictably be found in other regions. This is particularly common when the acquisition conditions have been loosened, which is the case of most of the public databases available.

### Different ways of using local quality measures

The local quality metrics we have presented earlier can be a solution to this difficulty. Therefore, some researchers have integrated them into their recognition systems to improve performance. However, very few work has been done in this direction, since much more attention has been given to global quality metrics in order to screen out degraded images.

In [46], Krichen localizes pixels that have any type of occlusion by using his local quality measure, and then masks these pixels before matching with Daugman's algorithm. So he generates a mask that is specific to each iris image and that makes a deterministic decision between what is noise and what is not. Li and Savvides apply the same strategy in [39]. They show that the integration of such masks in their recognition system improves considerably the performance of recognition. However, when applying a mask, the unmasked regions will all be considered equally at the extraction step. This can be a problem for two reasons: first of all, masks are rarely perfect and some parts of the unmasked regions may still contain artifacts. Secondly, some regions are more informative than others because they are highly textured, so it would be wise to privilege the extraction in those regions.

Chen et al. proposed to integrate his local quality measure in a different way in [19]. When calculating the Hamming distance, they attribute a weight to different concentric bands of the iris depending on their quality. Consequently, they compute a weighted Hamming distance to compare two iris codes *codeA* and *codeB*, following formula 3.6.

$$HD_w = \frac{1}{L} \frac{\sum_{i=1}^L \sqrt{E_{g(i)}^A \times E_{g(i)}^B} \times (\text{code}A_i \otimes \text{code}B_i)}{\sum_{i=1}^L \sqrt{E_{g(i)}^A \times E_{g(i)}^B}} \quad (3.6)$$

In 3.6,  $g(i)$  is the index of the band that contains the  $i$ -th bit of the iris code. The symbols  $E_{g(i)}^A$  and  $E_{g(i)}^B$  are the associated local quality measures of the

**CONFIDENTIAL**

$g(i)$ -th band in iris images A and B, respectively.

Their weighting scheme is such that regions with high quality in both iris images A and B contribute more to the Hamming distance than regions with poor quality. They show that doing this improves the EER of their Daugman-based system.

In his thesis [46], Krichen also integrates a local quality measure in the calculation of the similarity score of his correlation-based algorithm: the local correlations of sub-images are weighted by the minimum quality of the sub-images in the computation of the global similarity score. He also uses his local quality measure to compute a weighted Hamming distance, similarly to Chen et al.

Hollingsworth et al. have elaborated different ways of using their bit consistency measure. One way, presented in [35], is to define a threshold that separates consistent bits from inconsistent ones, and mask all the inconsistent ones. So they use a mask similarly to Li and Savvides in [39], but use a different criteria (bit consistency) to decide which bits should be masked. Another way, described [36], is to compare the positions of the stable bits when matching two iris codes. The idea is that the location of consistent bits should be more or less the same in two iris codes from the same iris. Consequently, they use the information on the position of the stable bits to compute a similarity score that they fuse with the Hamming distance.

Dong et al. use their bit consistency measure in another way in [25]. They generate a weight map from this measure: the weight associated to each bit measures its stability with the formula 3.5. This weight is then incorporated in the Hamming distance in order to compute a weighted Hamming distance. They demonstrate that their strategy “is effective for iris matching and greatly improves the performance of iris recognition systems”.

Finally, Ring and Bowyer describe in [74] a technique to eliminate low quality regions *a posteriori*. They notice that the presence of low quality regions can reduce the similarity between authentic matches. They work with a system based on Daugman’s approach, and explain that many bits corresponding to the low quality regions will disagree which will increase the Hamming distance. In the case of independent irises, the percentage of bits that disagree should be more or less the same from one region of the iris to another and close to 50% . On the other hand, in the case of an authentic match in which there are some low quality regions, the percentage of bits that disagree will be abnormally high in the low quality regions compared to the other ones. Therefore, Ring and Bowyer compute local Hamming distances for different pairs of sub-images. If the local Hamming distance is abnormally high for certain pairs of sub-images compared to others, they consider that these sub-images correspond to low quality regions and then leave them out to recalculate the global Hamming distance.

## 3.4 Reference systems

### 3.4.1 Implementations of Daugman’s algorithm

Since Daugman’s iris recognition has had its performance validated on very large databases and is used in most commercial systems, it is used by many researchers as a reference: they compare the performance of their own systems

**CONFIDENTIAL**

to the one of Daugman's. To make this comparison easier, Open Source versions of Daugman's algorithm have been made available. However, only very few of these Open Source versions exist.

The first one was developed by Masek in Matlab [59]. It is cut into 4 modules: segmentation, normalization, feature extraction and matching. The feature extraction module is slightly different from the one proposed by Daugman, for it applies 1D log-Gabor filters to the iris instead of 2D Gabor filters. Liu et al. rewrote it in C language [55], and the National Institute of Standards and Technology (NIST) in C++ [1].

NIST provides a Video-based Automated System for Iris Recognition (VASIR) [50]. This implementation of Daugman's iris recognition algorithm was designed to work both on conventional iris images and iris images collected at a distance. It can be cut into seven parts:

- Image acquisition
- Eye region detection/Extraction
- Quality measures and selection of the best quality image
- Segmentation
- Normalization
- Feature extraction and encoding
- Similarity matching templates

Our research group has also developed an Open Source version of Daugman's algorithm: the Open Source for Iris (OSIRIS) [49]. It consists of the same type of modules as Masek's. The segmentation module uses the Hough transform to model the iris boundaries by two circles. These circles are then used by the normalization module to "unfold" the iris image according to Daugman's rubber sheet model. At the feature extraction step, 2D-Gabor filters are applied around application points in the normalized iris image. The Gabor filters as well as the application points are input parameters of the system, so they can be chosen by the user. Finally matching is done by computing the Hamming distance between the binary codes that represent each iris.

Modifications of the segmentation module have led to a second version of this system : OSIRIS-V2. In this version, an algorithm based on active contours is used for segmentation. The latest version of this system, OSIRIS-V4 will be available on-line soon. It integrates the last improvements of the segmentation module presented in [79].

### 3.4.2 Segmentation used to generate our input normalized images

The reference system we have used throughout this work includes the normalization, the feature extraction and the matching modules of OSIRIS-V2. However we have not used the OSIRIS-V2's segmentation module. Even though this module performs much better than the segmentation module of OSIRIS-V1, its parameters have been defined in order to work well on the ICE 2005 database.

**CONFIDENTIAL**

Consequently, its performances are not optimal on databases that have been acquired with a different sensor such as the CASIA databases.

As explained in Chapter 1, our objective is to work on the feature extraction and matching steps of iris recognition systems so that they become robust to image degradations, whether the segmentation is accurate or not. Nevertheless if the segmentation is completely incongruous, the output segmented image is not exploitable for matching, whatever method is used. More specifically, the objective of the segmentation module is twofold:

- determine the parameters of the two circles, ellipses or curves that are the closest to the inner and outer boundaries of the iris. The resulting circles, ellipses or curves will then be used by the normalization module to “unfold” the iris.
- generate a binary segmentation mask that separates the iris texture from artifacts such as eyelids, eyelashes or specular reflections.

If the first point fails completely, then the normalized iris image will not be exploitable as illustrated in Figure 3.3.

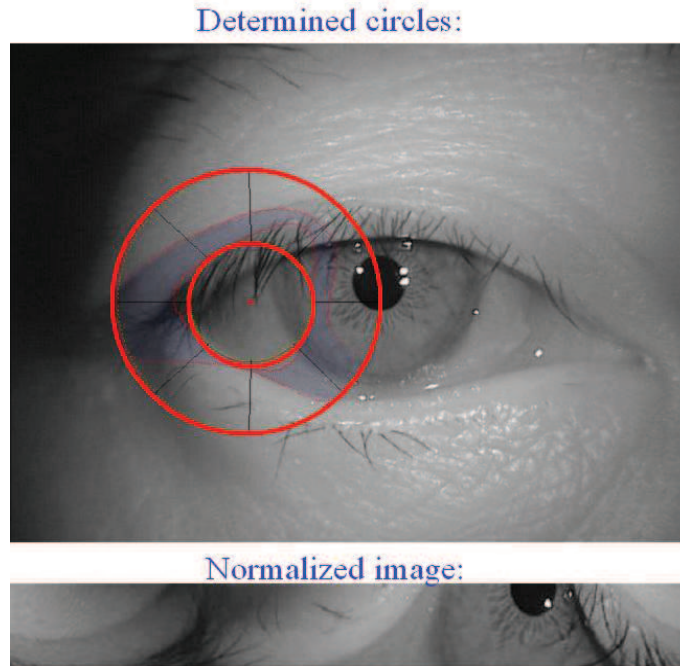


Figure 3.3: Example of circles determined by an aberrant segmentation for an iris image and of the associated normalized image

So even if our aim is to be robust to segmentation inaccuracies, we require that the circles or ellipses used for normalization and determined by the segmentation module are only aberrant for a very small percentage of images in the database. Since the segmentation module from OSIRIS-V2 does not fulfill this requirement on the CASIA databases, we have not exploited it for modeling the

**CONFIDENTIAL**

boundaries of the iris.

At the beginning of our doctoral research, we have worked with images that had been normalized using the output of a manual segmentation. More precisely, the center and the radius of the two circles that fit best to the inner and outer boundaries of the iris have been determined manually for a given number of images in the ND-IRIS-0405, CASIA-IrisV3-Interval and CASIA-IrisV3-Twins databases. These images have then been normalized with OSIRIS-V2's normalization module. Moreover the circles determined manually were used to initialize an active contour algorithm that generated a binary segmentation mask for each iris image.

Later on, we have exploited a segmentation module that was elaborated by a member of our team, Thierry Lefèvre, as part of his PhD work [52]. This module determines the circles that model the inner and outer boundaries of the iris with a much higher accuracy than the segmentation module of OSIRIS-V2. The process applied to determine these circles is based on a parametric variational approach and is described in [51] and presented in Appendix A. Moreover the accuracy of the binary segmentation mask that is generated is also improved compared to OSIRIS-V2. It is obtained by estimating the histograms of the iris and of the sclera and maximizing the difference between these two histograms.

Throughout this thesis we will work almost exclusively with iris images that have been previously segmented according to one of the two techniques described here above (1: manual determination of circles + active contour for binary mask or 2: Lefèvre's segmentation module) and normalized with OSIRIS-V2's normalization module. For each experiment we will specify which of the two segmentation techniques has been used to generate our input normalized iris images. We only exploit the initial iris images (before segmentation and normalization) in Chapter 7 to detect off-angle images.

### 3.5 Conclusion

In this Chapter we presented a literature review on iris recognition. More specifically we presented the major iris recognition algorithms, the quality metrics for iris images that have been defined and the way they have been integrated in recognition systems.

In this thesis, we will define our own metrics to measure the quality of iris images locally and globally and integrate them in two existing iris recognition systems described in Sections 3.1.1 and 3.1.3. The first one is based on Daugman's system and the second one is based on Krichen's. We will demonstrate that the method used for integrating our metrics is novel and allows a significant improvement in recognition performance. Moreover, following an approach often used in the literature, we have chosen to combine a "local" feature extraction and matching algorithm (Daugman's) to a "global" one (Krichen's). However, the way of doing this is original since it exploits our global quality metrics.

Unlike the quality metrics defined in IREX II, our measures are applied on the iris images after normalization. Their purpose is to detect segmentation inaccuracies, as well as to evaluate the usable iris area and the richness of its texture. The usable iris area is a quality component that has also been studied

**CONFIDENTIAL**

in IREX II. However, segmentation inaccuracies are not measured *a posteriori* in IREX II. Instead, IREX II evaluates criterias, such as the iris/pupil contrast, that can impact the accuracy of the segmentation *a priori*. Moreover, many criterias considered in IREX II are meant to be measured on the original iris image and not on the normalized one. In this work, we will compare our global quality metric to the usable iris area, that is shown as the most highly related to recognition performance in IREX II.

Finally, we presented in this chapter the reference systems that exist for iris recognition. In our work we will use the reference system OSIRIS, an implementation of Daugman's algorithm. More precisely, we will compare the feature extraction and matching algorithms we propose to the OSIRIS' feature extraction and matching modules. Since our work is aimed at improving these two steps of iris recognition, we will work with iris images that have previously been segmented and normalized. The segmentation has either been done by determining manually the circles that model best the iris boundaries followed by an active contour algorithm or by applying Lefèvre's segmentation algorithm. The normalization is done with OSIRIS's normalization module.

**CONFIDENTIAL**

## Chapter 4

# Elaboration of a local quality measure and analysis of its pertinence for matching

The goal of this thesis is to contribute to the elaboration of an iris recognition system that stays highly efficient if the acquisition conditions are loosened and the quality of the iris images diminished. More precisely our aim is to improve the feature extraction and matching steps so that these perform well, even when confronted to two major difficulties:

- errors at the segmentation step,
- a limited amount of information available.

Therefore, the first challenge is to detect locally, in the normalized iris images, regions that present one of these two difficulties. This will enable us to adapt the feature extraction and matching to this type of difficulties. To this end, we defined a local quality measure that can:

- discriminate the noisy iris regions from the good ones, to detect segmentation errors,
- distinguish highly textured regions from poorly textured ones, to quantify and localize the highly discriminative information.

This chapter presents the local quality measure we defined to fulfill these requirements in Section 4.1, as well as the experiments we conducted to study its properties in Section 4.2. Finally, we will prove in Sections 4.3 and 4.4 that the regions of highest quality according to our measure are the regions that should be exploited to achieve the best recognition performance, when using respectively Daugman's and Krichen's algorithms. In these sections, we will also show that our local quality measure can be used to define a global quality that is correlated to recognition accuracy.



## 4.1 Definition of our local quality measure

### 4.1.1 Choice of a model

We wish to localize regions of the normalized iris images that contain artifacts, as well as highly textured regions that are free from occlusion. In other words, we wish to classify different regions of the normalized iris images. We have chosen to use supervised classification for this, which requires labeled samples. Two types of models can be used for such a classification: a discriminative model or a generative model.

The objective of a discriminative model based classification is to learn the separation between different types of classes. The learning process requires labeled samples from each class. So if we were to use this type of model for the classification of iris regions, we would need to label regions containing each type of artifact (eyelash occlusion, eyelid occlusion, specular reflection...), poorly textured regions, highly textured regions etc.

On the other hand, a generative model based classification allows to model a single class using only labeled samples of this class. Therefore no training needs to be done on other classes.

Since our aim is to isolate the non-occluded and highly textured parts of the iris from the other regions, we have chosen to use a generative model. This way, we only need to label such high quality iris regions and avoid explicitly defining all the other classes of sub-images. Doing this would be difficult, for the other classes are numerous and diverse. Furthermore, modeling the borders between all these other classes would be complex.

In particular, we do not explicitly model statistically poor textured or noisy sub-images. We will see in Section 4.2 that both of these 2 types of sub-images will be characterized by observations that are different from those of our good quality model. Furthermore, the characteristics of images that are poorly textured but free from occlusion are closer to the highly textured ones than occluded images are. Consequently, a generative model allows us to distinguish those two types of regions from one another by analyzing how close they are to the high quality regions.

We have chosen a Gaussian Mixture Model (GMM) to characterize high quality sub-images, for it enables to model easily and accurately the probability distribution related to a class of sub-images. In Section 4.1.2 we will explain briefly what Gaussian Mixture Models are. Then we will present how we have build a GMM adapted to our purpose in Section 4.1.3. Finally, we will describe how we have used this GMM to define a local quality measure in Section 4.1.4.

### 4.1.2 Gaussian Mixture Models in general

The first Gaussian Mixture Model (GMM) was elaborated by Reynolds [73] for speech recognition. It was presented as a simplified version of classical hidden Markov models.

The GMM models a probability distribution associated to a class  $\lambda$  by a weighted sum of  $M$  Gaussian distributions. This is expressed in formula 4.1.

$$p(x|\lambda) = \sum_{i=1}^M p_i \cdot b_i(x) \quad (4.1)$$

**CONFIDENTIAL**

In 4.1,  $x \in \mathfrak{R}^D$  is a vector of dimension  $D$ ,  $b_i$  is the  $i$ -th  $D$ -dimensional Gaussian distribution and  $p_i \in [0; 1]$  is the associated weight. Each Gaussian distribution is defined by:

$$b_i(x) = \frac{1}{2\pi^{D/2}|\Sigma_i|^{1/2}} e^{-0.5 \cdot (x-\mu_i)^T \Sigma_i^{-1} (x-\mu_i)} \quad (4.2)$$

In 4.2,  $\mu_i \in \mathfrak{R}^D$  is the mean value and  $\Sigma_i \in \mathfrak{R}^D \times \mathfrak{R}^D$  is the covariance matrix. Each Gaussian distribution is entirely defined by these two elements  $\mu_i$  and  $\Sigma_i$ . Therefore, when a probability distribution is modeled with a GMM, it is entirely defined with the  $M$  sets of  $(p_i, \mu_i, \Sigma_i)$ .

The objective of the training of the GMM is to determine these  $M$  sets of elements on the training dataset. This is done iteratively in three steps: initialization, re-estimation and checking if the halt criterion is verified. We will describe these steps in more detail in the next sub-sections.

### 4.1.3 Detailed description of our Gaussian Mixture Model

We have used a GMM to characterize high quality sub-images, similarly to what is done in [47]. These sub-images come from normalized iris images. However the sub-images with which we trained the GMM are not only free from occlusion and well-focused, but are also highly textured. Moreover, we have enriched the model by adding local observations, measured in a  $5 \times 5$  neighborhood of each pixel, as the input vector of our GMM. We use the local mean, local variance and local contrast in addition to the values of the pixel gray-level. Note that we define the local contrast in a  $5 \times 5$  neighborhood as the difference between the maximum gray-level value and the minimum gray-level value in the neighborhood. So the dimension  $D$  of our observation vector is  $D = 4$ .

Given the 4 parameters we have chosen, our model is sensitive to global changes of gray-levels in images, which depend themselves strongly on the acquisition conditions. Consequently, if we were to learn our model on images that were acquired in certain conditions and apply it on images acquired in different conditions, the output probabilities might be biased. To avoid this, we have pre-processed all images so that their mean gray-level be close to a constant value. Figure 4.1 presents normalized iris images before (column *a*) and after (column *b*) this pre-processing.

We have done this for all the images with which we have trained the GMM, as well as for all the test images on which we apply the GMM. This way, we can apply our GMM on images that have been acquired with different illuminations.

We have used  $K \cdot T_k$  observation vectors to train the GMM, where  $K$  is the number of sub-images used for training and  $T_k$  is the number of pixels in the  $k^{th}$  sub-image with  $k \in [1; K]$ . Specifically, we have chosen 50 sub-images for training that all have the same dimension of  $11 \times 51$  pixels, so  $K = 50$  and  $\forall k \in [1; K], T_k = 561$ . These sub-images were selected manually in 50 different pre-processed normalized iris images of dimension  $64 \times 512$  that come from different databases, namely CASIA-IrisV3 and ND-IRIS-0405. They were chosen so as to be free from occlusion, well focused and highly textured. Moreover we

**CONFIDENTIAL**

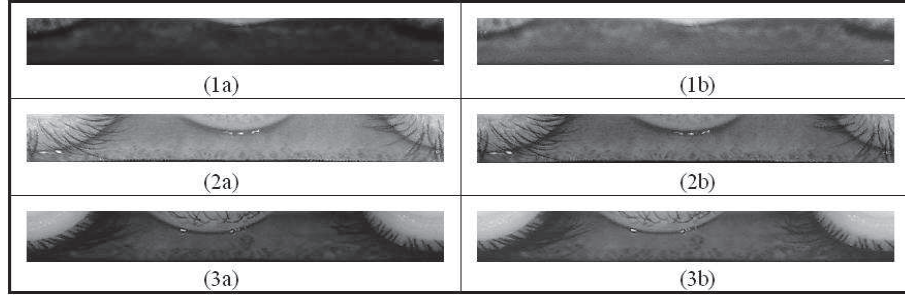


Figure 4.1: Normalized iris images from three different databases (1: ND-IRIS, 2: CASIA-IrisV3-Interval, 3: CASIA-IrisV3-Twins), before (a) and after (b) the pre-processing step

have chosen to use a GMM containing  $M = 10$  Gaussian functions. The values of all these parameters have been set empirically, after conducting tests to analyze the influence of their values on the training of the GMM. The selected values compromise between maximizing accuracy and minimizing computation complexity.

Once the number of Gaussians has been set, we have build our GMM in the following way:

- **Initialization** using the K-mean algorithm. At first, the means of the Gaussians are initialized randomly. We associate each observation vector  $O_t^k$ , where  $k \in [1; K]$  and  $t \in [1; T_k]$ , to the Gaussian for which the mean is the closest to this observation vector  $O_t^k$  in terms of euclidean distance. Then we compute the mean vectors  $\mu_i$  and the covariance matrices  $\Sigma_i$  for each of the 10 groups. The weight  $p_i$  associated to each Gaussian is equal to the proportion of observation vectors in each group.
- **Re-estimation** using the Expectation Maximization (EM) algorithm. This algorithm re-estimates the parameters  $\mu_i$ ,  $\Sigma_i$  and  $p_i$  in order to maximize the log-likelihood  $\overline{LL}_{train}$  defined in 4.3 where  $P(O_t^k|\lambda)$  is defined as in 4.1.

$$\overline{LL}_{train} = \frac{1}{K} \sum_{k=1}^K \frac{1}{T_k} \sum_{t=1}^{T_k} \log(P(O_t^k|\lambda)) \quad (4.3)$$

- **Halt criteria:** the iteration process is stopped when the following criterion is verified:

$$\left| \frac{\exp(\overline{LL}_{train}^{before}) - \exp(\overline{LL}_{train}^{after})}{\exp(\overline{LL}_{train}^{before})} \right| < \varepsilon \quad (4.4)$$

where  $\varepsilon = 10^{-2}$

#### 4.1.4 Computation of our local quality measure

In the previous paragraph we have described how we have trained a GMM that characterizes high quality sub-images. In this paragraph, we will present how we have exploited this GMM to define a local quality measure.

**CONFIDENTIAL**

The GMM we have trained can assign a likelihood to each observation vector associated to a pixel in an image. This likelihood is between 0 and 1. The higher is its value, the closer the tested observation vector is to the observation vectors on which the GMM has been trained. There are different ways of exploiting the likelihood given by our model.

One way would be to exploit it for categorical classification. This would imply setting a threshold for the likelihood. Then each pixel would be classified into one of the two categories: high quality pixel and non-high quality pixel. However, this would mean binarizing the real-valued likelihood and consequently, losing information. We have therefore chosen to use our GMM for probabilistic classification, instead. More precisely, we will associate to each sub-image the likelihood that it is of high quality. We will show in the next sub-section that the likelihood given by the model to sub-images containing artifacts or poorly textured sub-images will be lower than the one obtained on the high quality images. Furthermore, the characteristics of images that are poorly textured but free from occlusion are closer to the highly textured ones than occluded images are. Therefore poorly textured images should tend to have a higher likelihood than occluded images. In consequence, the probabilistic classification should not only allow us to distinguish high quality sub-images from other types, but it should enable us to determine which sub-images within the other types are closest to the high quality ones.

In practical terms we have defined the quality measure associated to a given sub-image  $w$  by the formula 4.5.

$$Q(w) = \exp\left(-\frac{1}{a \cdot b} \sum_{i=1}^{a \cdot b} |\log(P(x_i|\lambda)) - \overline{LL}_{train}|\right) = \exp\left(-\frac{1}{a \cdot b} \sum_{i=1}^{a \cdot b} d_{ll_i}\right) \quad (4.5)$$

In 4.5,  $a$  and  $b$  are the width and the height of the sub-image  $w$  respectively, so  $w$  contains  $a \cdot b$  pixels. The  $x_i$ ,  $i \in \{1, \dots, a \cdot b\}$ , are the input vectors of our GMM. Each one of them is associated to a pixel indexed  $i$  and takes into account 4 local observations: the gray-level of the pixel  $i$ , the local mean, the local variance and the local contrast of the gray-levels of all the pixels in a  $5 \times 5$  region around the pixel  $i$ .

$P(x_i|\lambda)$  is the likelihood given by the GMM to the input vector  $x_i$  and  $\overline{LL}_{train}$  is the mean log-likelihood on the training set defined in formula 4.3. We subtract this mean log-likelihood from the log-likelihood given by the GMM and consider the absolute value of the result. This is done to normalize the log-likelihood with respect to the training database. We call this value the log-likelihood distance. The lower this distance is, the closest the vector  $x_i$  is to the training set.

We then calculate the mean of these distances for all the pixels of the sub-image  $w$ . Once again, the lower this mean distance is, the closer the sub-image  $w$  is to the sub-images in the training set (namely sub-images that are free from occlusion and highly textured).

Finally, we use a negative exponential to obtain a result with values in  $[0; 1]$ . The closer its value will be to 1, the higher are the chances that the sub-image  $w$  is of good quality.

**CONFIDENTIAL**

## 4.2 Examples and statistics

For these experiments we used images that have been segmented manually and normalized with OSIRIS' normalization module presented in Chapter 3. They have also been pre-processed so that their mean gray-level be close to a constant value. Moreover, these images are not the ones that were used to train our GMM.

### 4.2.1 Examples

In this section we use the GMM we have trained as described in Section 4.1.3 to evaluate the quality of test sub-images. Figure 4.2 presents different sub-images (localized by a white or black rectangle) from three databases (ND-IRIS, CasiaV3-Interval and CasiaV3-Twins) and the local quality measures  $Q(w)$  associated to them, calculated using the formula 4.5. We can see that the lowest values are given to occluded sub-images, whether the occlusion comes from eyelashes, eyelids or specular reflections. Regions of the iris that are free from occlusion but very lowly textured are given intermediate values. The highest values are given to highly textured sub-images that are free from occlusion.






	Occluded sub-images	Quality
Eyelash Occlusion (ND-IRIS-0405)		$Q(w) = 0.01$
Eyelid Occlusion (Casia-IrisV3-Twins)		$Q(w) = 0.01$
Specular Reflection (Casia-IrisV3-Interval)		$Q(w) = 0.19$
	Non occluded sub-images	Quality
Lowly textured (Casia-IrisV3-Interval)		$Q(w) = 0.41$
Highly textured (ND-IRIS-0405)		$Q(w) = 0.89$

Figure 4.2: Examples of sub-images extracted from three different databases (ND-IRIS, CASIA-IrisV3-Interval and CASIA-IrisV3-Twins) and the local quality values associated to them. Sub-images have been chosen to represent the different qualities encountered in the images.

### 4.2.2 Statistics

We have confirmed the results presented in these examples by measuring the local qualities of 500 sub-images of  $11 \times 51$  pixels issued from normalized iris im-

ages of the ND-IRIS, CASIA-IrisV3-Interval or CASIA-IrisV3-Twins database and that have been labeled manually as belonging to one of the following categories:

- sub-images occluded by eye-lids
- sub-images occluded by eye-lashes
- sub-images containing a specular reflection
- sub-images that are free from any artifact and poorly textured
- sub-images that are free from any artifact and highly textured

For each one of these categories we have computed the mean local quality. This is presented in Table 4.1.

Table 4.1: *Mean values of the local qualities corresponding to 500 sub-images from different categories, manually labeled*

Categories of sub-images	Mean local quality
Eye-lid occlusion	0.07
Eye-lash occlusion	0.14
Specular reflection	0.05
No artifact, iris poorly textured	0.55
No artifact, iris highly textured	0.75

These statistics confirm the observations drawn from the examples presented in Section 4.2.1: the lowest quality values are given to occluded sub-images, whether the occlusion comes from eyelashes, eyelids or specular reflection, while the highest values are given to highly textured sub-images that are free from occlusion. Regions of the iris that are free from occlusion but very lowly textured are given intermediate values.

Consequently, we can conclude that the local quality measure we have defined in this section fulfills the objective we had set, for it enables us to:

- discriminate the portions of the iris that contain artifacts from the artifact-free ones, which will enable us to detect segmentation errors

**CONFIDENTIAL**

- distinguish highly textured regions from poorly textured ones, which will enable us to quantify and localize highly discriminative information

### 4.3 Pertinence of our quality measure for a Daugman type matching

The aim of this section is to prove that the regions of highest quality according to our measure presented in Section 4.1 are the regions that should be exploited to achieve the best recognition performance with Daugman’s algorithm [21]. We also wish to show that our local quality measure can be used to define a global quality that is correlated to recognition accuracy.

We chose Daugman’s algorithm for it was demonstrated in [22] that it enables to achieve high recognition performance and is computationally very efficient. Moreover, it was shown that its performance depends highly on the quality of the regions selected for feature extraction. Consequently, we believed that exploiting our quality measure in OSIRIS [49] is a promising way to improve its performance strongly on degraded images.

We wish to demonstrate that there is a link between the image quality according to our measure and the performance of the iris recognition system described here above.

Throughout this section, the normalized iris images we consider have been obtained with Lefèvre’s segmentation algorithm and the normalization module from OSIRIS presented in Chapter 3. They have also been pre-processed as explained in Section 4.1.

#### 4.3.1 Description of the feature extraction and matching technique

In this section, we use the feature extraction and matching technique proposed by Daugman in [21] to compare normalized iris images. More specifically, 2D Gabor filters will be applied around  $M$  application points uniformly distributed in the normalized iris images. The resulting Gabor phases will then be quantized so that each phase is represented by two bits, the values of which depend on the quadrant in which the phase is located, as shown in Figure 4.3.

As a result, each iris is represented by a binary code and two iris codes can be compared bit to bit. This is achieved by computing the Hamming distance between the binary codes of the irises A and B according to formula 4.6, in which  $\otimes$  represents the exclusive OR and  $L$  is the length of  $codeA$  and  $codeB$ .

$$HD = \frac{\|codeA \otimes codeB\|}{L} \quad (4.6)$$

Since the rotation of the iris can differ from one image to another, there can be a translation between the normalized iris images we wish to compare. We take this into account by computing the Hamming distance for several circular permutations of  $codeB$  in the angular coordinate (along the X axis in the normalized image) and by keeping the minimum computed Hamming distance to compare irises A and B.

**CONFIDENTIAL**



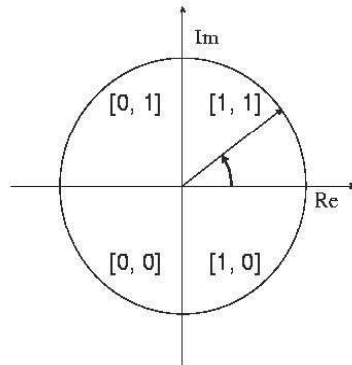


Figure 4.3: Illustration of the phase quantization

If the two normalized iris images we compare come from two different irises, the corresponding binary codes should be independent, and there is a 50% chance that two compared bits match. Consequently, the mean Hamming distance for inter-class comparisons should be 0.5. However, since we test several Hamming distances and keep the minimum, the distribution of the Hamming distances kept for comparison is biased towards the lower values. Figure 4.4 presents the normalized distributions of Hamming distances for inter-class and intra-class comparisons on the ND-IRIS-0405 database.

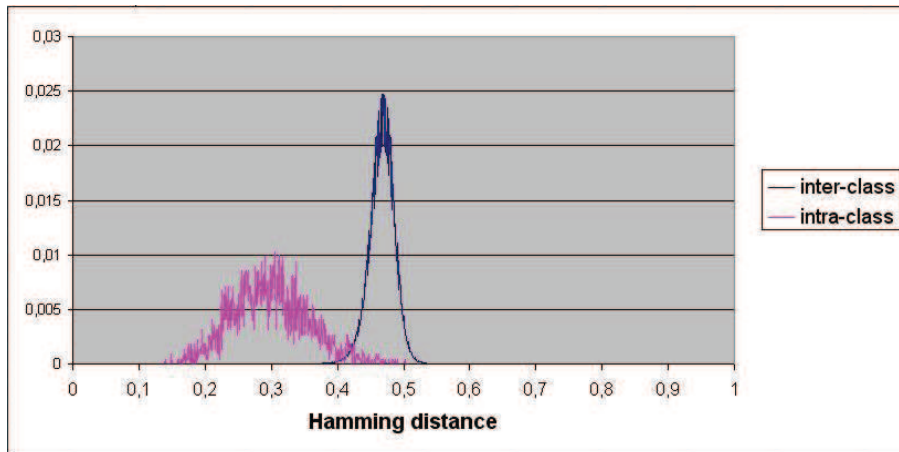


Figure 4.4: Distribution of the Hamming distances for inter-class and intra-class comparisons from the ND-IRIS-0405 database.

### 4.3.2 Local pertinence of our quality measure

#### Motivation

A pertinent quality metric for iris recognition should be linked to the recognition performance: the higher the quality according to this metric, the higher the performance should be.

**CONFIDENTIAL**



The recognition performance of Daugman’s algorithm is directly linked to the Hamming distance computed to compare two irises. This distance gives an indication on how well images match globally. As explained in Section 4.3.1, it is computed as a bitwise comparison of the binary codes corresponding to two normalized iris images A and B. Since each bit was obtained by applying a Gabor filter around a point in the normalized iris image, every bit is associated to a position in the image. Consequently, a bit to bit comparison is a local comparison and the Hamming distance can be seen as the mean of distances that measure the local dissimilarity.

It therefore makes sense to measure the recognition performance locally for a pair of normalized images with Daugman’s algorithm. To this end, we define the local Hamming distance. It is computed by considering the reduced binary codes  $codeA_{reduc}$  and  $codeB_{reduc}$ , associated to a pair of sub-images  $(w_A, w_B)$ , located at the same position in the normalized iris images A and B. These reduced codes contain only the bits associated to the  $m$  application points located in the given sub-images  $w_A$  and  $w_B$ , out of the total  $M$  application points uniformly distributed in each normalized iris image. The local Hamming distance is defined as in formula 4.7 where  $L_{reduc}$  is the length of  $codeA_{reduc}$  and  $codeB_{reduc}$ .

$$locHD = \frac{\|codeA_{reduc} \otimes codeB_{reduc}\|}{L_{reduc}} \quad (4.7)$$

Such a local Hamming distance can be computed for sub-images extracted from normalized iris images from the same iris or from different irises. We will call them respectively the intra-class local Hamming distance ( $locHD_{intra}$ ) and the inter-class local Hamming distance ( $locHD_{inter}$ ).

We wish to determine if the local Hamming distance between a couple of sub-images  $(w_A, w_B)$  and the local quality measure of sub-images  $w_A$  and  $w_B$  as defined in Section 4.1 by formula 4.8, are correlated.

$$Q(w) = \exp\left(-\frac{1}{a \cdot b} \sum_{i=1}^{a \cdot b} |\log(P(x_i|\lambda)) - \overline{LL}_{train}|\right) = \exp\left(-\frac{1}{a \cdot b} \sum_{i=1}^{a \cdot b} d_{li}\right) \quad (4.8)$$

To do this we consider the minimum between  $Q(w_A)$  and  $Q(w_B)$ , in order to associate a single quality value to the couple of sub-images  $(w_A, w_B)$ . We can then compare this quality value to the local Hamming distance between  $w_A$  and  $w_B$ . We have chosen to select  $\min(Q(w_A), Q(w_B))$  because one of the two sub-images having a low quality is sufficient to degrade performance. To illustrate this, let us consider the image quality in terms of occlusion. If  $w_A$  and  $w_B$  are an intra-class comparison then the resemblance of  $w_A$  to  $w_B$  will be less obvious if  $w_A$  is occluded whatever the quality of  $w_B$ , than if  $w_A$  and  $w_B$  are both free from occlusion.

### Experiments and results

We have considered 900 intra-class and 900 inter-class couples of normalized iris images selected randomly in the ND-IRIS-0405, CASIA-IrisV3-Interval and CASIA-IrisV3-Twins databases. The dimension of the normalized iris images is

**CONFIDENTIAL**

64x512 pixels. We consider 72 sub-images of 11x51 pixels per image and perform Daugman’s feature extraction, to obtain a reduced binary code for each sub-image. For this experiment, we have chosen  $M = 8 \times 127 = 1016$  application points located on a grid in the normalized iris images. The interval between the points is equal to 4 pixels along the X axis and 8 pixels along the Y axis. So there are  $m = 1 \times 13$  application points in a sub-image, as illustrated in Figure 4.5.

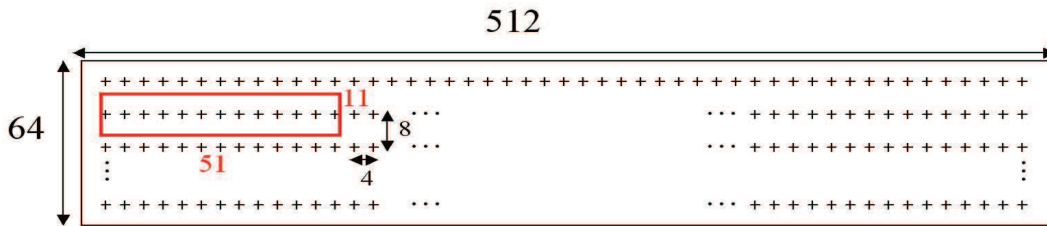


Figure 4.5: Application points contained in a  $11 \times 51$  pixels sub-image

For each of the 1800 couples of irises we calculate 72 local Hamming distances from the reduced binary codes corresponding to the sub-images. Additionally, we compute the quality of all the sub-images we consider and associate to each couple of sub-images the minimum value of the two qualities.

Note that we have previously computed the the global Hamming distances between the 1800 couples of irises. This was done by taking into account several circular permutations of the iris code and selecting the minimum global Hamming distance. It enabled us to keep, for each pair of irises, the values of the permutation’s parameters corresponding to the minimum global Hamming distance. This way, we could register each couple of irises before computing the 72 associated local Hamming distances.

We analyzed how the local Hamming distances are linked to these quality values. More precisely, we divided the couples of sub-images into 6 categories according to the value of their quality and calculated the mean local normalized Hamming distance for each category. The results are presented in Figure 4.6.

It appears that, for the intra-class comparisons, the mean local Hamming distance between the couples of sub-images increases as their quality decreases. On the other hand, for the inter-class comparisons the mean of the local Hamming distances stays stable except for the lowest quality values ( $Q(w) < 0.4$ ) where it increases slightly. So the higher the quality of the sub-images according to our measure, the better the local Hamming distance can discriminate intra-class couples from inter-class couples, which implies that the local recognition performance is improved. This shows that it is relevant to use our quality measure with an iris recognition system based on Daugman’s feature extraction and matching technique, for it enables to know which sub-images in the normalized iris images will be the most discriminating for matching.

Furthermore, relating Figure 4.6 to Figure 4.3. that presents the mean local qualities for different categories of sub-images, we can conclude that both the amount of texture and the presence of artifacts in sub-images have an impact on

$\min(Q(w_A), Q(w_B))$	$\text{mean}(\text{locHD}_{\text{intra}}(w_A, w_B))$	$\text{mean}(\text{locHD}_{\text{inter}}(w_A, w_B))$
[0.8; 1]	0.24	0.45
[0.6; 0.8]	0.25	0.45
[0.4; 0.6]	0.28	0.45
[0.2; 0.4]	0.30	0.46
[0.1; 0.2]	0.32	0.47
[0; 0.1]	0.35	0.46

Figure 4.6: Mean local Hamming distance for intra-class and inter-class comparisons of sub-images, for different quality values of these sub-images.

recognition performance. However, the impact of the second factor is larger than that of the first one. This follows from Figure 4.7, which presents the difference  $\delta = \text{mean}(\text{locHD}_{\text{inter}}) - \text{mean}(\text{locHD}_{\text{intra}})$  for different values of the local quality. The slope of the curve is always positive, but it is steeper for the lowest quality values ( $Q(w) < 0.2$ ). According to Figure 4.3., these low quality values generally correspond to sub-images containing artifacts. This demonstrates that the impact of occlusions on recognition performance is higher than the impact of the amount of texture, even if both factors do influence performance.

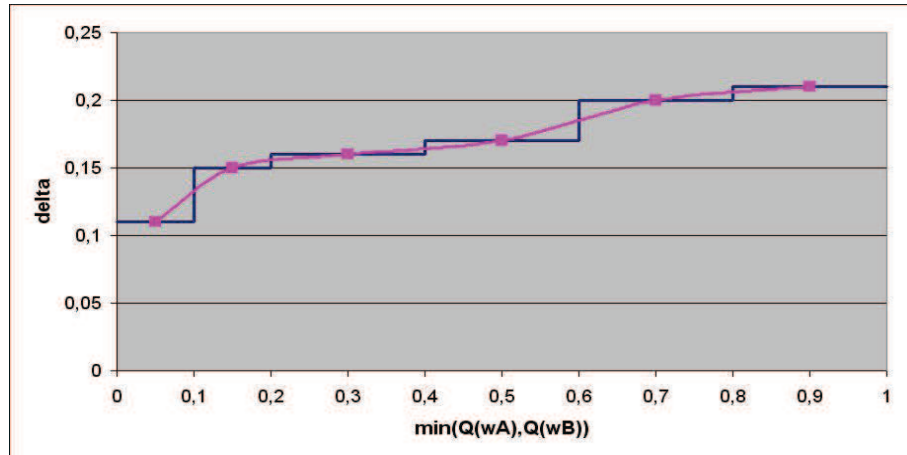


Figure 4.7: The difference  $\delta = \text{mean}(\text{locHD}_{\text{inter}}) - \text{mean}(\text{locHD}_{\text{intra}})$  as a function of the minimum quality values

Moreover, as we have already mentioned, the Hamming distance computed to compare two normalized iris images globally is directly linked to the local Hamming distances. Indeed if the sub-images used to compute the local Hamming distances do not overlap and cover the entire normalized image, then the

global Hamming distance is the mean of the local Hamming distances. So the more discriminating the local Hamming distances are, the more discriminating is the global Hamming distance and the better is the global recognition performance. Since we have demonstrated that our local quality measure can detect the sub-images that will give the most discriminating local Hamming distances, we can use this measure to include only such sub-images in the computation of the global Hamming distance. This will be presented in Chapter 5.

### Link with bit consistency

It is interesting to relate this result to the concept of bit consistency, defined by Hollingsworth et al. [35]. A consistent bit is stable in iris codes resulting from different images of the same iris. The intra-class local Hamming distance ( $locHD_{intra}$ ) defined here above measures the stability of bits in sub-images. Since we have demonstrated that this  $locHD_{intra}$  is linked to our local quality measure, we can deduce that our quality measure is related to bit consistency: a sub-image of high quality according to our measure will with a high probability generate consistent bits with Daugman's feature extraction technique. This proves once again that our quality measure can contribute to improving the performance of a Daugman technique based algorithm. Indeed, it is a way of detecting consistent bits in a normalized iris image, and it has been demonstrated in [35] that the knowledge of the position of consistent bits can improve recognition performance. This is achieved either by masking inconsistent bits, by weighting the bits according to their consistency or by fusing the Hamming distance with a score measuring the overlap in the position of consistent bits in the two images to compare.

The advantage of detecting the consistent bits in a normalized iris image with our quality measure is that no reference images of the same iris are required to learn the position of the consistent bits. Indeed, we can train the GMM used for our quality measure on highly textured and free from occlusion sub-images that come from any pre-processed normalized iris image. Once the GMM is trained we can simply compute the local quality on any test sub-image by applying formula 4.8. Moreover, our quality measure can detect an artifact located at an unusual position since its computation does not rely on the position of the sub-image in the normalized image.

To the contrary, the training method presented in [35], requires several registered images of the same iris and learns the positions that are the most likely to generate inconsistent bits. They consider  $n$  iris codes in the same class to determine for each bit the number of times,  $m_1$ , that it is equal to 1 and the number of times,  $m_0$ , that it is equal to 0. The stability of the bit is then measured by:

$$s = \frac{|m_1 - m_0|}{m_1 + m_0} \quad (4.9)$$

Since this method enables to determine where it is the most likely to find inconsistent bits, one can miss inconsistent bits when using it, if these bits result from an artifact at an unusual position.

**CONFIDENTIAL**

### 4.3.3 Global pertinence of our quality measure

We have demonstrated that the local quality measure defined in Section 4.1 can detect the sub-regions of the normalized iris images that are the most discriminating for matching with Daugman’s technique. We now show that our quality measure can also be used to define different global quality metrics for a pair of normalized iris images and that the images for which our metrics are the highest are those for which Daugman’s recognition algorithm performs best.

#### Definition of our global quality metrics

In the IREX II evaluation, fourteen global quality metrics have been defined and evaluated. It was shown that the metric with the highest impact on recognition performance is the usable iris area. We will compare our global quality measure to the usable iris area metric that can also be measured by the occlusion rate.

When comparing two normalized iris images, the quality of both images can have an impact on recognition performance. Often researchers have computed a global image quality for each image and then fused the global qualities of the two images to compare. For instance, they keep the minimum of the two values or the mean value. However, if you consider the occlusion rates of two images  $OcR_A$  and  $OcR_B$  then these two rates do not give us any information on the overlapping of the occluded zones in images A and B. And yet, it is easy to understand that for set values of  $OcR_A$  and  $OcR_B$ , the recognition accuracy should be higher if the position of the occluded zones is the same in the two images than if these positions are complementary. Indeed in the second case, there is less iris texture left to compare.

For this reason, we have chosen to define a global quality measure that is directly linked to a pair of normalized iris images (A,B), instead of computing a global quality measure for each image and fusing the two global measures. To do so, we have exploited the local quality measure defined in Section 4.1. We have chosen to combine the local qualities of images (A,B), before computing the global quality measure. This way our global quality measure will take into account the overlapping of the good or bad quality regions in images A and B.

More precisely, we have started by computing for each normalized iris image A and B the local qualities of the sub-images centered around  $M$  application points. This gives us  $M$  couples of local quality values  $(Q(w_i^A), Q(w_i^B))$  where  $i \in [1; M]$ . Then, we keep the minimum value for each couple of qualities in the same way as we have done in Section 4.3.2. In result, we obtain  $M$  fused local quality values  $Q^{fus}(w_i) = \min(Q(w_i^A), Q(w_i^B))$  for  $i \in [1; M]$ .

We have chosen two approaches for defining a global quality measure from these  $M$  fused local quality values. In the first approach we compute the mean of these  $M$  values, according to formula 4.10:

$$GlobQ_1 = \text{mean}_{i \in [1; M]}(Q^{fus}(w_i)) \tag{4.10}$$

In result,  $GlobQ_1$  will measure the amount of texture and artifact in the pair of normalized images to compare, when taking into account the worst sub-image out of the two  $(w_i^A, w_i^B)$  considered for each of the  $M$  application points.

In the second approach, we wish to measure the proportion of the image pair that contain no artifact. Since we have shown in Section 4.1, that sub-images

containing artifacts are those that have the lowest local quality measure, we can consider that the artifacts are located in the sub-images for which the local quality is beneath a given threshold that needs to be determined. We therefore consider the three following global quality measures. They are computed after sorting the  $M$  fused local quality measures and associating a rank to each quality  $rank(Q^{fus}(w_i))$  so that the rank associated to the highest fused local quality is equal to 1.

- $GlobQ_2 = \min_{i \in [1;M]}(\text{rank}(Q^{fus}(w_i) \leq 0.2))$  that represents the rank from which the fused local qualities are smaller than 0.2
- $GlobQ_3 = \min_{i \in [1;M]}(\text{rank}(Q^{fus}(w_i) \leq 0.1))$  that represents the rank from which the fused local qualities are smaller than 0.1
- $GlobQ_4 = \min_{i \in [1;M]}(\text{rank}(Q^{fus}(w_i) \leq 0.05))$  that represents the rank from which the fused local qualities are smaller than 0.05

They measure the number of regions in the image pair that have a local quality higher than 0.2, 0.1 or 0.05. These values were chosen empirically based on the statistics presented in Section 4.1. They show that the mean values of the local qualities of sub-images belonging to the categories eye-lid occlusion, eye-lash occlusion and specular reflection are respectively 0.07, 0.14 and 0.05. This gives us a vague indication on how to choose our local quality threshold. Since they are all included in the range  $[0, 0.2]$ , we have tested three different values for the threshold in this range of values.

### Comparison of global quality measures: Experiments and results

We have used global quality metrics to eliminate the image couples of worst quality from the matchings done by OSIRIS. If our metrics are pertinent to be used with this recognition algorithm, then the image couples of worst image quality according to these metrics should be the ones for which the recognition algorithm performs worst. Eliminating these images should therefore improve recognition performance. Moreover, for a given percentage of eliminated images, the most pertinent metric is the metric that leads to the best recognition performance when used for eliminating images. Note that the percentage of eliminated images can also be called Failure to Match rate (FTM) according to the definition given in Chapter 2.

We have elaborated a list of 6 million inter-class and 60 000 intra-class image comparisons using images from the ND-IRIS-0405 database. For a start, we have eliminated 10% of the comparisons of worst quality according to either  $GlobQ_1$ ,  $GlobQ_2$ ,  $GlobQ_3$  or  $GlobQ_4$ . The resulting recognition performances are presented with ROC curves in Figure 4.8.

For an  $FTM = 0.1$ , the best performance is achieved with  $GlobQ_3$ , then with  $GlobQ_1$ ,  $GlobQ_2$  and  $GlobQ_4$ . Eliminating the worst images according to  $GlobQ_3$  brings a 45% improvement of the FRR for  $FAR=10^{-4}$  compared to the FRR obtained without eliminating any image couples.

$GlobQ_3$  can be seen as a measure of the usable iris area, for it measures the number of sub-images out of the total  $M$  sub-images in the normalized sub-image that have a local quality higher than 0.1, and we have seen in Section

**CONFIDENTIAL**



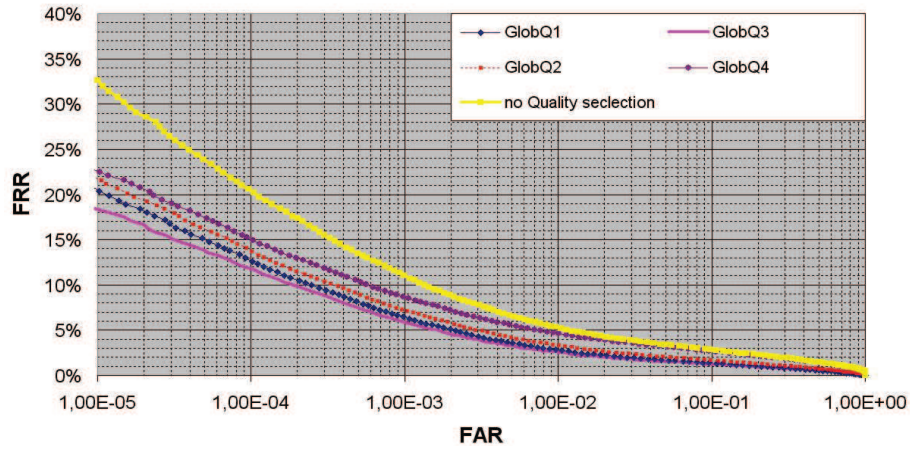


Figure 4.8: ROC curves obtained when eliminating the 10% worst image couples according to different GMM-based quality measures on the ND-IRIS-0405 database.

4.2 that such images are usually occluded images.  $GlobQ_2$  and  $GlobQ_4$  are very similar, but the quality threshold they use is different. This shows that the best local quality threshold, out of the three tested, for eliminating most of the sub-images containing an artifact, without eliminating too many, is  $Q_{th} = 0.1$ .

Since  $GlobQ_3$  can be seen as a measure of the usable iris area, we have compared it to the Occlusion Rate ( $OcR$ ) often used in the literature that is computed with the mask resulting from the segmentation of the iris. To compute it we have used the segmentation mask given by Lefèvre’s segmentation module described in Chapter 3. This Occlusion Rate is the ratio of the number of masked bits on the number of total bits in the binary code. It is inversely proportional to the usable iris area, so we have eliminated the images with the highest Occlusion Rates i.e., images with the lowest usable iris area. This comparison is presented in Figure 4.9.

With an  $FTM = 0.1$  the performance achieved with  $GlobQ_3$  is very close to the one achieved with the Occlusion Rate. As a matter of fact, the two rates seem to be complementary: for an  $FAR > 4 \times 10^{-4}$  the best performance is achieved by eliminating the image couples according to  $OcR$  and for an  $FAR \leq 10^{-4}$  it is achieved by eliminating the couples according to  $GlobQ_3$ . The difference in performance between these two measures of the usable iris areas arises because they use two different criteria to define occlusions: one uses the value of our local GMM-based quality measure and the other one uses a segmentation mask computed according to Lefèvre’s technique.

We have performed the same experiment on the CASIA-IrisV3-Lamp database. More specifically, we have considered 7 million inter-class comparisons and 45000 intra-class comparisons and eliminated 10% of the comparisons of worst quality according to  $GlobQ_1$ ,  $GlobQ_3$  or  $OcR$ . These results are presented in Figure 4.10.

Now the eliminations according to  $GlobQ_1$  and  $GlobQ_3$  lead to very similar performance: for  $FAR=10^{-4}$ , there is a 24% reduction of the FRR compared to

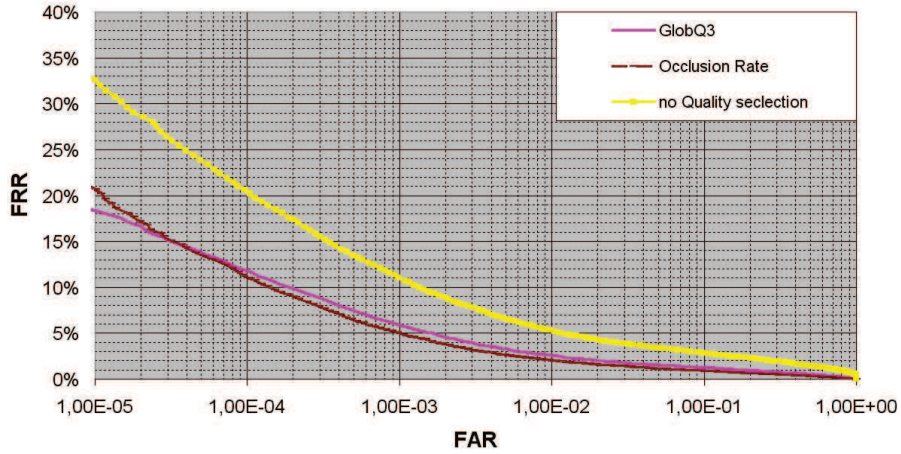


Figure 4.9: ROC curves obtained when eliminating the 10% worst image couples according to a GMM-based or a segmentation-based quality measures on the ND-IRIS-0405 database.

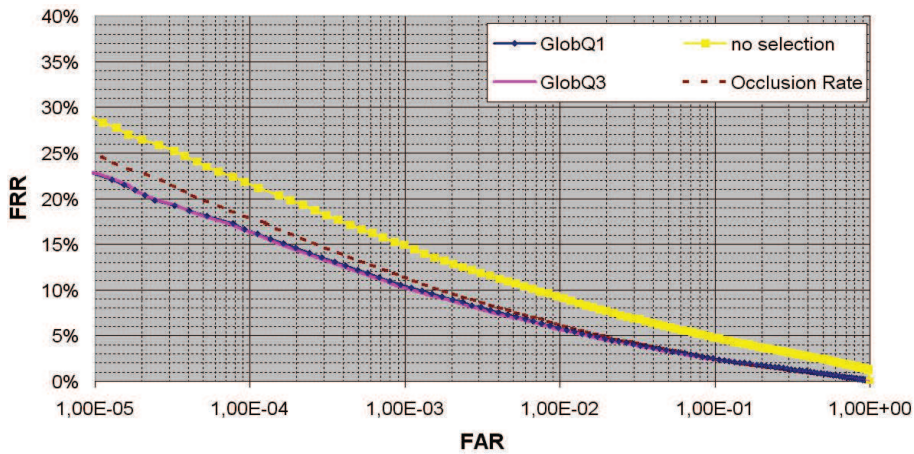


Figure 4.10: ROC curves obtained when eliminating the 10% worst image couples according to different quality measures on the CASIA-IrisV3-Lamp database.

**CONFIDENTIAL**



the case where no images are eliminated. On the other hand, eliminating couples according to *OcR* leads to a smaller performance improvement: only 14 %. This is because the segmentation module performs less well on this database. Indeed it is sensitive to acquisition conditions, so its performance can change from one database to another if they were acquired in different conditions. *GlobQ3* seems to me more robust to those changes. This comes from the fact that *GlobQ3* is based on our local, GMM-based quality measure that is measured on pre-processed normalized iris images as explained in Section 4.1. Consequently, it is robust to illumination variations.

In a second experiment, we have chosen to keep only the 10% image couples with the highest quality from the ND-IRIS-0405 database. This selection is done by using either *GlobQ1*, *GlobQ3* or the Occlusion Rate. Figure 4.11 presents the corresponding ROC curves. It shows that the best performance is now achieved with the selection according to *GlobQ1*. The selections according to *GlobQ3* and the Occlusion Rate give similar performances, but they are lower than the one of *GlobQ1*. This can be explained by the fact that for these images, the Occlusion Rate is very low (< 10%). In consequence variations of the amount of texture in the non-occluded regions has a bigger impact on performance than small variations of the Occlusion Rate. *GlobQ1* gives an indication of the amount of texture as well as the amount of occlusions, while *GlobQ3* and the Occlusion Rate only quantify the amount of occlusion.

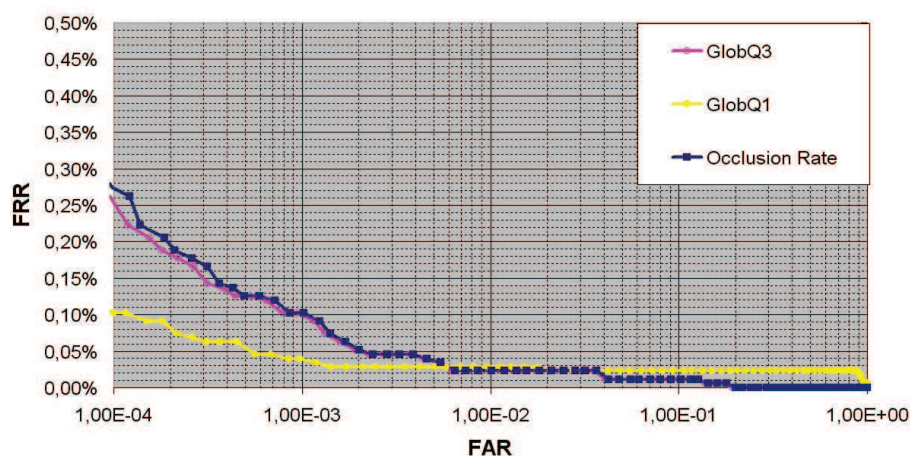


Figure 4.11: ROC curves obtained when selecting the 10% best image couples according to different quality measures on the ND-IRIS-0405 database.

#### 4.4 Pertinence of our quality measure for a Krichen type matching

In this work, we chose to consider a second algorithm for feature extraction and matching, in addition to the one proposed by Daugman: the one described by Krichen et al. in [48]. We chose this algorithm for it uses more information for recognition. This information is extracted from the iris by applying Gabor

filters. However, the resulting Gabor phases are not simply compared bit to bit after phase quantization. Instead, they are compared by performing local cross-correlations, so the entire Gabor phase sub-images are exploited. In consequence, this algorithm may perform better on degraded images than OSIRIS in accordance to what has been shown in [48]. It could therefore be used in addition to OSIRIS to improve performance on low quality images.

The aim of this section is to analyze if it is pertinent to use the quality measure defined in Section 4.1 with a Krichen based recognition algorithm. First we will describe the feature extraction and matching technique that we have applied in line with the description given in [48]. Then we will see if it is relevant to exploit our quality measure with this algorithm.

Throughout this section, the normalized iris images we consider have been obtained with Lefèvre's segmentation algorithm and the normalization module from OSIRIS presented in Chapter 3. They have also been pre-processed as explained in Section 4.1.

#### 4.4.1 Description of the feature extraction and matching technique used

Like Daugman, Krichen exploits 2D Gabor filters to extract information from normalized iris images. However, he applies these filters around all the points in the normalized images, instead of applying them around a limited number of application points. Moreover, the Gabor phases resulting from the application of these filters are not quantized. Consequently, the iris is represented by  $N_f$  real-valued phase images, where  $N_f$  is the number of Gabor filters that are applied. So comparing a couple (A,B) of normalized iris images comes to comparing  $N_f$  couples of phase images  $(PhImA_r, PhImB_r)_{r \in [1; N_f]}$ .

$M$  points of interest are positioned in each phase image, so that the location of these points is the same in all images. To compare two phase images  $PhImA_r$  and  $PhImB_r$ , Krichen considers the sub-images centered around these  $M$  points and computes the normalized cross-correlation between the  $M$  couples of sub-images. The size of the  $M$  sub-images in  $PhImA_r$  is greater than the one from the  $M$  sub-images in  $PhImB_r$ . This way we allow translations between the normalized iris images that can be caused by rotations of the iris from one image to another.

The normalized cross-correlation at pixel  $(u, v)$  between two sub-images  $w_A$  and  $w_B$ , in  $PhImA_r$  and  $PhImB_r$  respectively, is given by formula 4.11.  $w_A$  is centered at an application point number  $i$  located at pixel  $(x, y)$  and  $w_B$  is centered at the pixel  $(u, v)$  in the sub-image  $w_A$ .

$$C_i(u, v) = C_{(x,y)}(u, v) = \frac{\sum_{x,y} [w_A(x, y) - \overline{w_{A_{u,v}}}] [w_B(x - u, y - v) - \overline{w_B}]}{\sum_{x,y} [w_A(x, y) - \overline{w_{A_{u,v}}}]^2 \sum_{x,y} [w_B(x - u, y - v) - \overline{w_B}]^2} \quad (4.11)$$

In formula 4.11,  $\overline{w_B}$  represents the mean of the gray-level values in the sub-image  $w_B$  and  $\overline{w_{A_{u,v}}}$  represents the mean of the gray-level values in the portion of the sub-image  $w_A$  that is centered around pixel  $(u, v)$  and has the same size as sub-image  $w_B$ .

**CONFIDENTIAL**

The correlation is normalized by the energies of the sub-images  $w_A$  and  $w_B$ . This energy is defined as the sum of the squared gray-level values of the sub-images. Consequently, the maximum value of  $C_i(u, v)$  is 1. This value can be reached when  $w_A$  and  $w_B$  have been extracted from two iris images of the same iris. This normalization is meant to eliminate the dependence of the cross-correlation to the illumination variations.

From each cross-correlation we can define the Peak to Slob Ratio (PSR) as well as the Peak Position (PP).

$$PSR((w_A)_i, (w_B)_i) = \frac{\max_{(u,v)}(C_i(u, v)) - \text{mean}_{u,v}(C_i(u, v))}{\text{std}_{(u,v)}(C_i(u, v))} \quad (4.12)$$

The similarity score (SS) between two phase images is computed with formula 4.13.

$$SS_r = \frac{\text{mean}_{i \in [1;M]}(PSR((w_A)_i^{(r)}, (w_B)_i^{(r)}))}{\text{std}_{i \in [1;M]}(PP((w_A)_i^{(r)}, (w_B)_i^{(r)}))} \quad (4.13)$$

Finally, all phase image similarity scores are fused to obtain one single score that represents the similarity between the iris images A and B.

Figure 4.12 presents the distributions of the similarity scores ( $\times 1000$ ) of intra-class and inter-class comparisons of the ND-IRIS-0405 database.

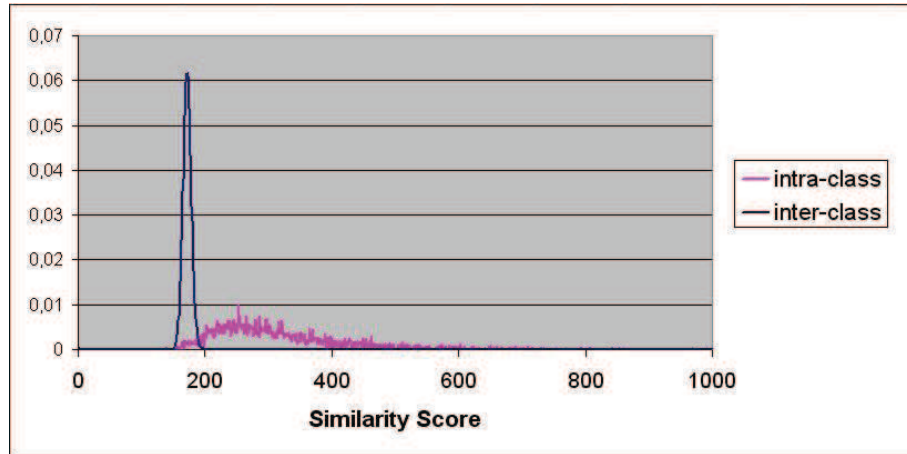


Figure 4.12: Distribution of the similarity scores computed with Krichen's algorithm for inter-class and intra-class comparisons from the ND-IRIS-0405 database.

#### 4.4.2 Global pertinence of our quality metric

We wish to verify that the performance of the Krichen based iris recognition system described here above is correlated to the image quality measure defined in Section 4.1. To this end, we associate a global quality to each normalized iris image couple to compare as in Section 4.3. This global quality measure is

computed by exploiting our local quality measure and was defined in Section 4.3. It is either  $GlobQ_1$  given by formula 4.14 or  $GlobQ_3$  given by formula 4.15 where  $Q^{fus}(w_i)$  is defined in formula 4.16.

$$GlobQ_1 = \text{mean}_{i \in [1;M]} (Q^{fus}(w_i)) \quad (4.14)$$

$$GlobQ_3 = \min_{i \in [1;M]} (\text{rank}(Q^{fus}(w_i) \leq 0.1)) \quad (4.15)$$

$$Q^{fus}(w_i) = \min_{i \in [1;M]} (Q(w_i^A), Q(w_i^B)) \quad (4.16)$$

As in section 4.3, we have eliminated from a pre-established matching list the 10% worst couples of images to compare according to one of our quality measures. Then, we have compared the performance of Krichen's algorithm on the resulting matching lists. First, we have worked with a matching list containing images from the ND-IRIS-0405 database leading to approximately 300 000 inter-class and 30 000 intra-class comparisons. The corresponding results are presented in Figure 4.13. Then, we have worked with a second matching list containing images from the CASIA-IrisV3-Lamp database, also leading to 300 000 inter-class and 30 000 intra-class comparisons. The latter results are shown in Figure 4.14.

We can see that, eliminating the 10% worst matches according to  $GlobQ_1$  or  $GlobQ_3$  improves recognition performance and that the improvement is practically the same for the two metrics, whichever database is tested. More precisely, at an FAR= $10^{-4}$  there is a 30% relative decrease of the FRR: it changes from FRR=8% to FRR=5% on the ND-IRIS-0405 database and from FRR=10% to FRR=7% on the CASIA-IrisV3-Lamp database.

This demonstrates that there is a correlation between the performance of Krichen's algorithm and both of the global quality metrics we have defined:  $GlobQ_1$  and  $GlobQ_3$ .

## 4.5 Conclusion

In this Chapter we defined a local, GMM-based quality measure and demonstrated that it can be used to detect sub-images that contain artifacts as well as to quantify the amount of texture in non-occluded regions of the iris. We have also proven that this local quality measure is correlated to the local recognition performance of Daugman's algorithm.

Moreover, we defined different global quality measures from our local measure and showed that these global measures are correlated to the recognition performance of both Daugman's and Krichen's recognition algorithms. The highest correlation was obtained with  $GlobQ_1$  and  $GlobQ_3$ .  $GlobQ_1$  represents the mean of local quality values, so it measures the usable iris area as well as the amount of texture in this area, for a pair of iris images. On the other hand,  $GlobQ_3$  counts the number of regions for which the local quality value is higher than 0.1, so it only measures the usable iris area. Consequently,  $GlobQ_1$  contains more information than  $GlobQ_3$  and we will therefore work with  $GlobQ_1$  in the rest of this work.

**CONFIDENTIAL**

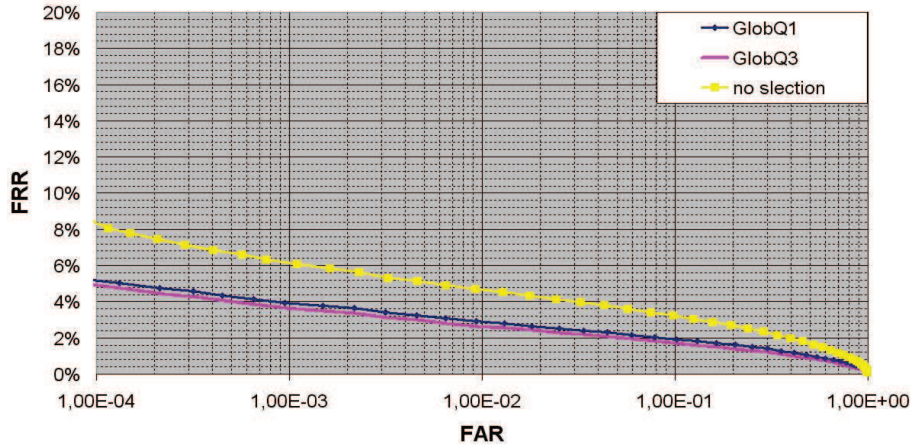


Figure 4.13: ROC curves obtained when selecting the 90% best image couples according to different quality measures on the ND-IRIS-0405 database and using Krichen's algorithm for feature extraction and matching.

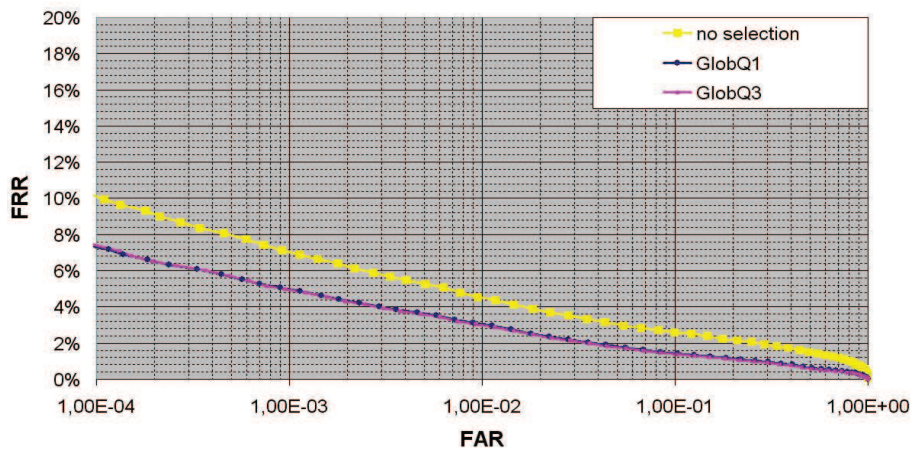


Figure 4.14: ROC curves obtained when selecting the 90% best image couples according to different quality measures on the CASIA-IrisV3-Lamp database and using Krichen's algorithm for feature extraction and matching.



## Chapter 5

# Integration of our quality measure in OSIRIS

The local quality measure defined in Chapter 4 is related to the likelihood that a given sub-image from the normalized iris image is free from occlusion and highly textured. The values of this measure range from 0 to 1. We showed that the lowest values are given to sub-images that contain artifacts and the highest values are given to sub-images that are free from occlusion and highly textured. Sub-images that do not contain any artifact but that are poorly textured are allocated intermediate values.

In Chapter 4 we also demonstrated that the higher the quality of sub-images according to our measure, the more discriminating are these sub-images when used with a Daugman-based iris recognition algorithm such as OSIRIS.

The aim of this Chapter is to present a novel strategy for selecting the regions of normalized iris images used for feature extraction and matching that exploits our local quality measure. We have seen in Chapter 3 that authors have defined local quality metrics and integrated them in Daugman's matching algorithm. In the first section of this chapter, we will integrate our own quality measure in OSIRIS using the methods described in literature, and we will present the drawbacks of these strategies. Then we will describe the method we propose and present the experiments and results that validate our method. Following this, we will analyze these results in some detail. This analysis will lead us to exploiting the global quality measure we have defined in Section 4.3.3 conjointly with our local quality measure.

## 5.1 Drawbacks of the state-of-the-art techniques

### 5.1.1 Creation of a binary quality mask

One way of integrating our local quality measure in OSIRIS would be to use it to mask regions of the iris that have a bad quality, similarly to what is done by Krichen in [46] or by Li et al. in [39]. In this case, we would experimentally define a threshold for the quality value, and sub-images that have a quality below the threshold would be masked. So a binary mask would be associated to each iris code. To evaluate the dissimilarity between two iris images A and B, we

would then compute the Hamming distance according to formula 5.1, in which  $codeA$  and  $codeB$  are the binary codes resulting from the feature extraction of the normalized iris images A and B, and  $maskA$  and  $maskB$  are the binary masks associated to them.  $\otimes$  represents the binary operator XOR and  $\cap$  the operator AND.

$$HD_{mask} = \frac{\|(codeA \otimes codeB) \cap maskA \cap maskB\|}{\|maskA \cap maskB\|} \quad (5.1)$$

However by doing so, we would transform our real-valued quality measure into a binary one by separating abruptly iris sub-images into two categories: good quality and bad quality ones. For instance we could set the threshold so that sub-images containing artifacts are considered of bad quality, and therefore masked, and all other sub-images unmasked. This would lead to losing the information our quality measure gives on the amount of texture in sub-images that are free from occlusion. Since we have demonstrated in Chapter 4 that this information is correlated to the matching performance, we have chosen not to apply such a binary strategy.

### 5.1.2 Weighted quality mask

#### Description of the method

Another way of exploiting local quality in a Daugman based iris recognition system was described in the literature. It consists in incorporating a weighted map in the computation of the Hamming distance as was done by Chen et al. in [19], as well as Krichen in [46] and Dong et al. in [25].

As explained in Chapter 3, Chen et al. define a wavelet transform-based local quality metric that measures the energy in a given concentric band of the iris. He then computes the weighted Hamming  $HD_w$  distance between iris codes  $codeA$  and  $codeB$  of length  $L$  according to formula 5.2.

$$HD_w = \frac{1}{L} \frac{\sum_{i=1}^L \sqrt{E_{g(i)}^A \times E_{g(i)}^B} \times (codeA_i \otimes codeB_i)}{\sum_{i=1}^L \sqrt{E_{g(i)}^A \times E_{g(i)}^B}} \quad (5.2)$$

In 5.2,  $g(i)$  is the index of the band that contains the  $i$ -th bit of the iris code. The symbols  $E_{g(i)}^A$  and  $E_{g(i)}^B$  are the associated local quality measures of the  $g(i)$ -th band in iris images A and B, respectively.

Chen's weighting scheme is such that regions with high quality in both A and B contribute more to the Hamming distance than regions with poor quality.

Krichen and Dong et al. also apply a similar weighting scheme with their own local quality metrics: a GMM-based quality measure and a bit consistency measure, respectively. These measures have been presented in Section 3.2.2. Note that they do not combine the local qualities of images A and B with a multiplication like Chen et al. Instead Krichen considers the minimum of the two values and Dong et al. directly compute local qualities that are associated to a pair of iris codes.

This strategy chosen by Chen, Krichen and Dong, has the advantage of keeping the real value of local quality measures instead of binarizing them.

**CONFIDENTIAL**

We have applied it by following the process after-specified on each couple of normalized iris images (A,B) to compare:

- feature extraction around  $M$  application points uniformly distributed in each one of the normalized iris images A and B
- evaluation of the quality of the  $M$  sub-images centered at the  $M$  application points in image A and image B
- computation of the weighted Hamming distance

The first step is done by considering  $M$  application points  $pt_i$   $i \in [1; M]$  identically located in normalized iris images A and B, and performing feature extraction around each one of these points. More precisely, we perform convolution with two dimensional Gabor filters and the filters' phase responses is quantized in a pair of bits.

In the second step, we pre-process the images A and B as explained in Chapter 4. Then we apply our probabilistic model to each one of the  $M$  sub-images that are centered at the  $M$  application points. As a result we have  $M$  quality values associated to image A,  $Q(wA_i)$  with  $i \in [1; M]$ , as well as  $M$  quality values associated to image B,  $Q(wB_i)$  with  $i \in [1; M]$ .

Finally, we compute the weighted Hamming distance. To do this we associate a quality measure to each bit of the iris code. More precisely, each pair of bits comes from the application of a Gabor filter around one of the  $M$  application points  $pt_i$ . So we associate to that pair of bits the value resulting from the computation of the local quality measure on the sub-image centered at that particular application point  $pt_i$ . We then calculate the weighted Hamming distance, according to the following formula 5.3:

$$HD_{weight} = \frac{\sum_{i=1}^M \|(codeA_i \otimes codeB_i)\| \cdot \min(Q(wA_i), Q(wB_i))}{\sum_{i=1}^M L_i \cdot \min(Q(wA_i), Q(wB_i))} \quad (5.3)$$

where  $codeA_i$  and  $codeB_i$  are the binary codes associated to the sub-images  $wA_i$  and  $wB_i$  centered at point  $pt_i$ .  $L_i$  is the length of these two binary codes and  $Q(wA_i)$  and  $Q(wB_i)$  are the local quality of the sub-images  $wA_i$  and  $wB_i$ .

We have chosen to weight the comparison of each bit by the minimum value between  $Q(wA_i)$  and  $Q(wB_i)$ , because we have demonstrated in Chapter 4 that this value is correlated to the local performance of the recognition. Other authors have chosen different ways to combine local qualities. For instance, we have seen that Chen et al. have chosen to multiply them. However there is no consensus on which combination method is best for integrating local quality measures in the computation of the Hamming distance.

## Experiments and results

We have compared the feature extraction and matching protocol described here above to the one used in OSIRIS-V2, the reference system described in Chapter 3.

Note that in both of the protocols, the same 2D Gabor filters are applied around the same  $M$  application points. The binary codes associated to each

**CONFIDENTIAL**



normalized iris image are therefore identical. The OSIRIS-V2 protocol and ours differ at the matching step. In OSIRIS-V2, two binary codes are compared by computing the Hamming distance according to formula 5.4 or 5.5 whether we wish to take into account the masks generated at the segmentation stage or not. In formula 5.4,  $L$  is the length of each code. In formula 5.5,  $maskA$  and  $maskB$  are the binary masks that are associated to the iris images A and B and that have been obtained with a segmentation module.

$$HD_{withoutMask} = \frac{\|(codeA \otimes codeB)\|}{L} \quad (5.4)$$

$$HD_{withMask} = \frac{\|(codeA \otimes codeB) \cap maskA \cap maskB\|}{\|maskA \cap maskB\|} \quad (5.5)$$

On the other hand, when following the quality weighting strategy, we integrate the real-valued quality metrics associated to each bit in the computation of the weighted Hamming distance according to formula 5.3.

This experiment has been carried out with images from the ICE 2005 database. The circles modeling the inner and outer boundaries of the iris have been determined manually on these images and were given as input to OSIRIS' normalization module. The segmentation masks have been obtained with the active contour algorithm of OSIRIS-V2, initialized with the circles determined manually.

We have worked with 2953 images and performed 13 262 intra-class matches and 1 069 352 inter-class matches. To evaluate performance of each algorithm we have plotted its ROC curve, presenting the False Rejection Rate (FRR) as a function of the False Acceptance Rate (FAR) for these matches. This is presented in Figure 5.1.

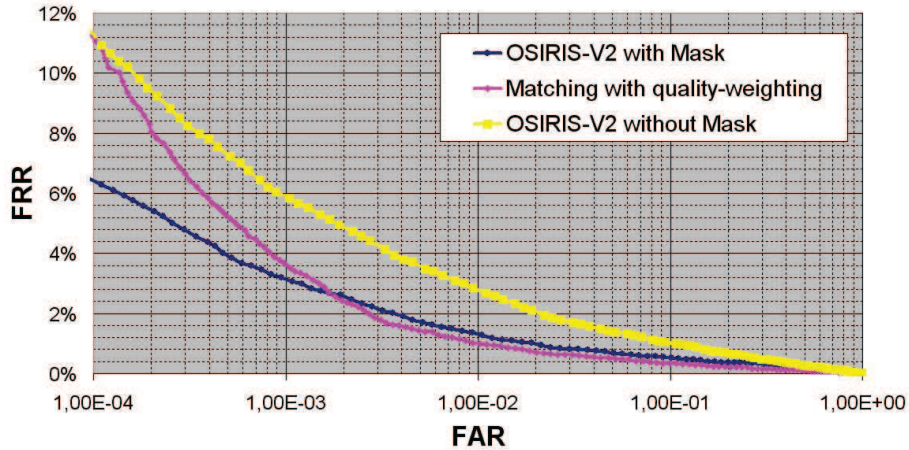


Figure 5.1: ROC curves of 3 iris recognition algorithms based on the same implementation of Daugman's approach but using different formulas to compute the Hamming distance on the ICE 2005 database.

As we can see, at low FAR ( $< 2 \times 10^{-3}$ ) the performance of the quality weighting technique is very close to the one of OSIRIS-V2 without mask. For

**CONFIDENTIAL**

this range of FAR, the OSIRIS-V2 algorithm with Mask performs considerably better than the other two algorithms. For higher FARs ( $> 2 \times 10^{-3}$ ), the performance of the quality weighting technique and OSIRIS-V2 with mask are very close. Consequently, we can conclude that the best algorithm, out of the three tested, is OSIRIS-V2 with mask whatever the value of the targeted FAR.

This is not the result we expected. Indeed, the quality weighting technique enables to exploit our real-valued quality measure. This measure brings more information than a binary mask for it detects occlusions and quantifies the amount of texture in non-occluded regions of normalized iris images, while the binary mask only detects occlusions. So we would have expected the quality weighting technique to perform better than OSIRIS-V2 with mask.

To understand this result, we have analyzed the intra-class matches that had the highest dissimilarity score with the quality weighting technique as well as the inter-class matches with the lowest ones, since these matches are the ones that lead to errors. We have noticed that they correspond to cases where the global quality of the images (according to the metric  $GlobQ_1$  defined in Chapter 4) is very low, for both inter-class and intra-class matches.

This observation has led us to test a new strategy for recognition in which we apply only the quality weighting technique to images that have a sufficient global quality and we apply OSIRIS-V2 with mask to the other images. More specifically, we evaluate  $GlobQ_1$  for each couple of images to compare and :

- if  $GlobQ_1 < 0.1$  we apply OSIRIS-V2 with mask to the couple
- if  $GlobQ_1 \geq 0.1$  we apply the weighting quality technique to the couple

Figure 5.2 presents the comparative performance of this strategy and the three we have already tested (OSIRIS-V2 without mask, OSIRIS-V2 with mask and the quality weighting technique). The proposed strategy performs best. Indeed for an FAR =  $10^{-4}$ , the FRR = 4.7% for this strategy, and FRR = 6.3% for OSIRIS-V2 with mask.

This shows that the quality weighting technique improves performance for images that verify  $GlobQ_1 \geq 0.1$ , but degrades performance for the other images, in comparison to OSIRIS-V2 with Mask. An explanation for this is that a large proportion of bits are associated to very low quality values for the image couples for which  $GlobQ_1 < 0.1$ , so a very low weight is attributed to a large number of bits at the computation of the Hamming distance and the information that they bring is barely taken into consideration for matching. In consequence, too little information is taken into account at the computation of the Hamming distance for this distance to be significant to discriminate authentications from impostors.

In other words the amount of information available for matching with the quality weighting technique is directly linked to the local quality values and can therefore be indirectly associated to the global quality of the normalized iris images: the lower the global quality, the more chances there are that a large number of bits are attributed very low weights and that there is too little information available for the matching to be reliable.

Consequently, we can conclude that reducing the impact of the low quality regions at the matching stage improves recognition performance as long as there

**CONFIDENTIAL**

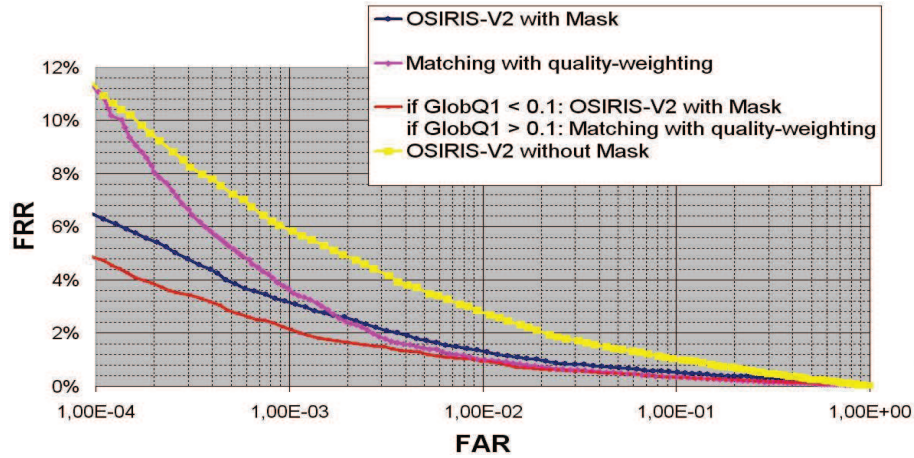


Figure 5.2: ROC curves of 4 iris recognition algorithms based on the same implementation of Daugman's approach but using different formulas to compute the Hamming distance on the ICE 2005 database.

are not too many low quality regions. In other terms it is important to keep enough significant information for the matching.

This point is in line with what Daugman demonstrates in [23]. He shows that when the amount of information used for matching is not controlled, there can be significant variations of this amount from one matching to another, which can lead to a biased inter-class distribution and therefore degrade recognition performance. More precisely, it is important to choose aptly the amount of information used for matching. Too much information would be unnecessary and would increase the chances of including noisy information. On the other hand, too little information would increase the false acceptance rate.

The latter protocol we propose yields a better recognition performance than the one of OSIRIS-V2. Still, we do not fully control the amount of information used for matching, since the quantity of information used for matching is only indirectly linked to the global quality. It would be wiser to quantify directly the amount of information available, which is what we have done in the next section.

## 5.2 Proposed technique

### 5.2.1 Description of the method

In response to the drawbacks of these techniques, we have chosen to integrate our quality measure into Daugman's algorithm in a new and original way that gives us full control over the amount of information used for matching. More specifically, we have decided to use our quality measure to select the regions of the normalized iris images that will be used for feature extraction while choosing wisely the number of regions to select.

The algorithm we propose compares a gallery and a probe iris image that have previously been segmented, normalized and pre-processed as described in

**CONFIDENTIAL**

Chapter 3 (segmentation and normalization) and Chapter 4 (pre-processing). It can be cut into the six steps that are presented in Figure 5.3 followed by bitwise matching.

In the first step we consider the same  $M$  points  $pt_i$  in the gallery and in the probe normalized pre-processed iris images. These points are positioned on a uniform grid in the images.

The second step consists in computing our GMM-based quality measure for the sub-images centered at each of these  $M$  points in the gallery and in the probe image. As a result we have  $M$  quality values for the gallery image and  $M$  quality values for the probe image.

Recall that our final goal is to match the gallery and the probe iris in order to decide whether or not they belong to the same person. Accordingly, when extracting features and matching them, we are interested in both the quality of the probe image and the one of the gallery. Therefore, the third step of our algorithm is to fuse the quality values of the gallery and probe images. It is quite obvious that when matching the features corresponding to a gallery and a probe sub-image, a single sub-image containing an artifact is sufficient to increase the chances of false rejection drastically. So, as explained in Chapter 4, we have considered that the most relevant way to fuse the quality measures is to select the minimum quality value between the gallery and the probe. This is done for the  $M$  couples of reference and test sub-images. As a result, we have  $M$  fused quality measures associated to the  $M$  (gallery, probe) couples of sub-images.

In the fourth step, we sort the  $M$  fused quality measures in decreasing order. Each quality measure being associated to one of the  $M$  points, this sorting enables us to associate a rank to each one of these points.

In the fifth step, we select the points that have the  $N$  lowest ranks, so they correspond to the sub-images with the  $N$  highest quality values.

These points will be used for feature extraction in the sixth step. More precisely, a convolution with Gabor filters is done around them. Each filter's phase response is then quantized in a pair of bits, so the information from the iris image is represented by a binary code, the length of which is proportional to  $N$ .

The parameter  $N$  that is directly linked to the amount of information used for the matching is very important. Setting it too high or too low would reduce recognition performance. Consequently, we have conducted extensive experiments for determining a good range of values for  $N$ . This will be explained in detail in the next section.

The bitwise matching that follows the sixth step of our algorithm is the one that was originally proposed by Daugman [21] and that computes the Hamming distances between the binary codes associated to each iris. To take into account the variations of rotation of the iris from one image to another we perform all these steps for different translations of the probe normalized image along the angular coordinate. The minimum computed Hamming distance is assumed to correspond to the correct alignment of the gallery and probe images.

### 5.2.2 Experiments and results

We consider iris images that have been segmented, normalized and pre-processed. To demonstrate that the improvement in recognition performance achieved with

**CONFIDENTIAL**

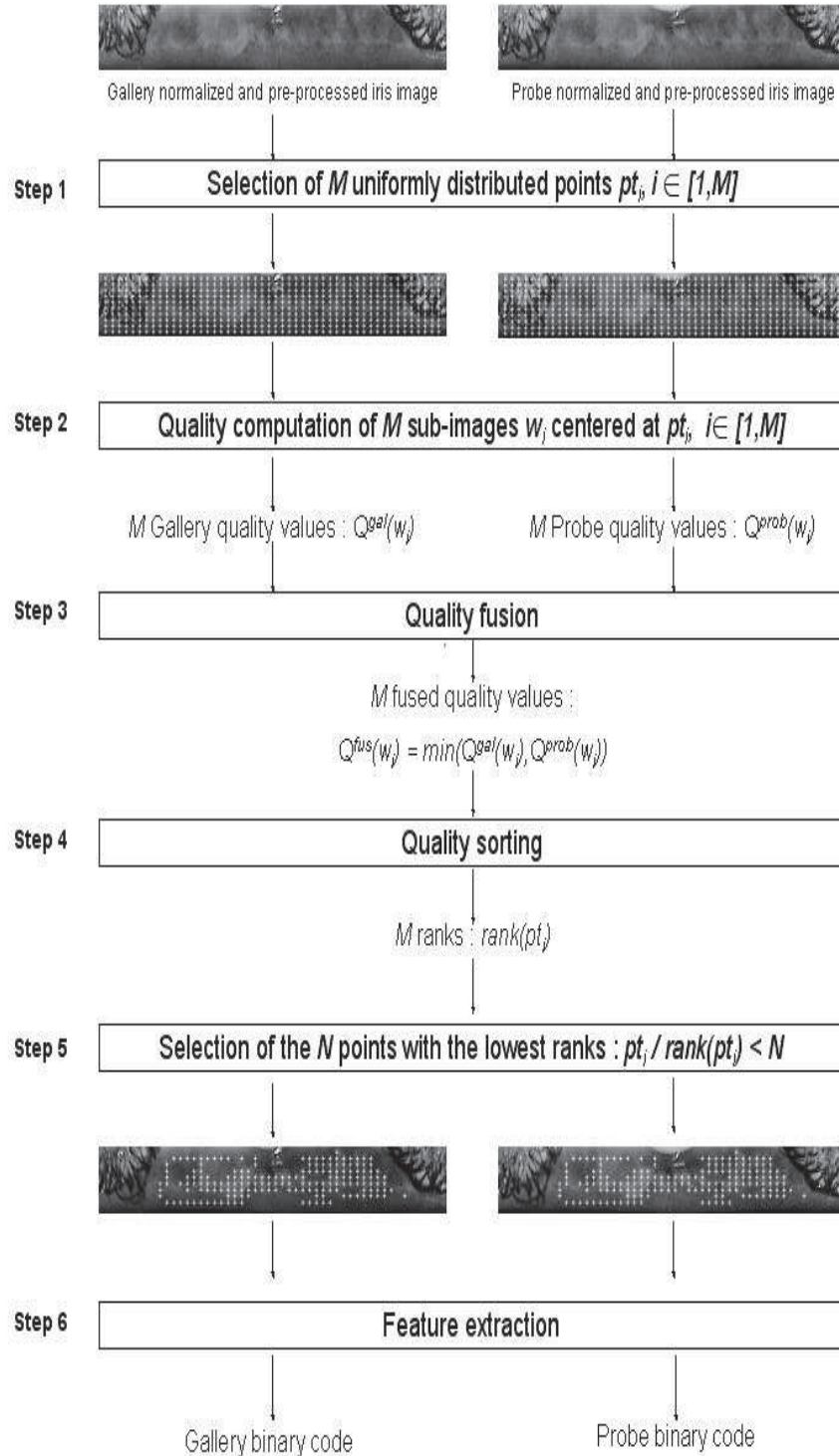


Figure 5.3: Global scheme of the method we propose for pertinent feature extraction using our quality measure.

**CONFIDENTIAL**



our method is not linked to a specific segmentation technique, we have worked with images that have been segmented and normalized with different techniques.

In the first set of experiments we carried out, we use images that have been segmented with a manual initialization. More specifically, the circles that model the iris' inner and outer boundaries have been determined manually and OSIRIS' segmentation module has been used to associate a binary segmentation mask to each iris image. In the second group of experiments, we have worked with images that have been segmented with Lefèvre's algorithm.

For all images considered and whatever technique is used for segmentation, this segmentation is followed by OSIRIS' normalization module to generate the normalized iris images and the associated normalized segmentation masks as explained in Chapter 3. Finally, all images have been pre-processed according to what has been described in Chapter 4 to bring the mean of their gray-levels to a constant value.

### Experiments based on a manual segmentation

For a start, we have compared our feature extraction algorithm to the one in OSIRIS-V2 in which the features of the normalized iris images are extracted around  $M$  points that are uniformly distributed. To the contrary, in the feature extraction algorithm that we propose, we select the  $N$  (out of  $M$ ) best quality sub-images and extract the features around the  $N$  points located at the centers of these sub-images. In OSIRIS-V2, it is possible to take into account the segmentation mask associated to the normalized iris images, which improves recognition performance. However in the algorithm we propose, it is unnecessary to use this mask, since we choose the application points used for feature extraction as a function of the local quality. In both cases, the feature extraction is carried out by convolution with the same two-dimensional Gabor filters.

We have compared our feature extraction algorithm to the one in OSIRIS-V2, with and without using the segmentation masks. More precisely, since both feature extraction techniques lead to a binary code representation of the iris, we have followed the different feature extractions by a bitwise matching step in which we calculate the Hamming distance between their binary codes. We have then compared the performance of the different iris recognition methods thus obtained when taking as input the same normalized iris images.

Since the segmentation module of OSIRIS-V2 has been optimized on images from ICE-2005, it performs best on ND-IRIS-0405 (a superset of ICE-2005 and ICE-2006). We wish to compare our feature extraction algorithm to the one in OSIRIS-V2 when the latter performs best, so we have used images from ND-IRIS-0405 for this comparison. The images we have used are 2136 images from the ND-IRIS database, 6 images for each of the 249 subjects, selected randomly.

To evaluate performance of each algorithm we have plotted its ROC curve, presenting the False Rejection Rate (FRR) as a function of the False Acceptance Rate (FAR). Figure 5.4 presents the results of the ROC curves of OSIRIS-V2 with and without using the segmentation masks and of our algorithm that does not use the segmentation masks.

As we can see, the lowest performance is obtained when using OSIRIS-V2, without segmentation mask. In that case  $FRR=16\%$  for  $FAR=10^{-4}$ . Adding this mask improves recognition performance considerably, since at an  $FAR=10^{-4}$

**CONFIDENTIAL**

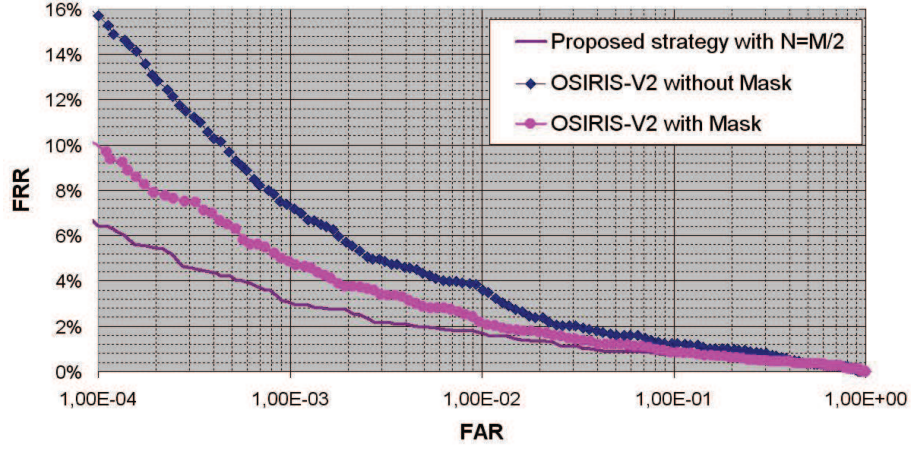


Figure 5.4: ROC curves of 3 iris recognition algorithms based on the same implementation of Daugman's approach but using different strategies for point selection for feature extraction on the ND-IRIS database.

the FRR=10%, so the FRR is 36% lower with mask than without mask. However using our feature extraction algorithm in which we select the  $N (= M/2)$  best quality sub-images for the two dimensional Gabor filters convolution gives even better performances, especially at low FAR. For an FAR= $10^{-4}$  the FRR=6.5%, so it is 59% lower than OSIRIS-V2 without mask and 36% lower than OSIRIS-V2 with mask.

Recall that the choice of the parameter  $N$  is very important because it determines the length of the binary code and the quantity of information that is used for matching. In the first experiment, we chose an arbitrary value for  $N$ , equal to  $M/2$ . Intuitively, it is easy to understand that, for a set value of  $M$ , the value of  $N$  can be optimized in order to compromise between too much information for matching (that would be unnecessary and increase the chances of including noisy information) and too few information that would increase the false acceptance rate. So we compared the performance of our algorithm on the ND-IRIS database for different values for  $N$  and a fixed value for  $M$ . In other words, we changed the proportion of the initial normalized iris image that is kept for feature extraction when applying our quality-based region selection strategy. The performances which we obtained are shown in Figure 5.5.

We can see that, in this experiment with images from the ND-IRIS database, the best value for  $N$  when chosen in  $\{M/6, M/3, M/2, 2M/3, 5M/6\}$  is  $M/3$ . For this value of  $N$ , the FRR at an FAR= $10^{-4}$  is 48% lower than with OSIRIS-V2 with mask.

In order to generalize this result, we have performed the same experiment on all the images from the CASIA-IrisV3-Interval and the CASIA-IrisV3-Twins databases. Figure 5.6 shows the performance of our algorithm on the CASIA-IrisV3-Interval database for different values of  $N$  and a set value of  $M$ , and Figure 5.7 shows the same on the CASIA-IrisV3-Twins database.

**CONFIDENTIAL**

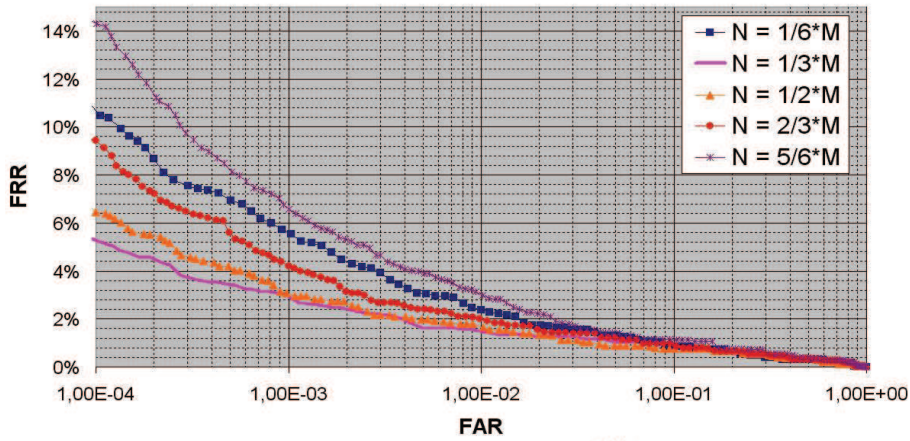


Figure 5.5: ROC curves for different values of  $N$  (number of selected points for feature extraction) and a set value of  $M$  (initial number of points) for the ND-IRIS database.

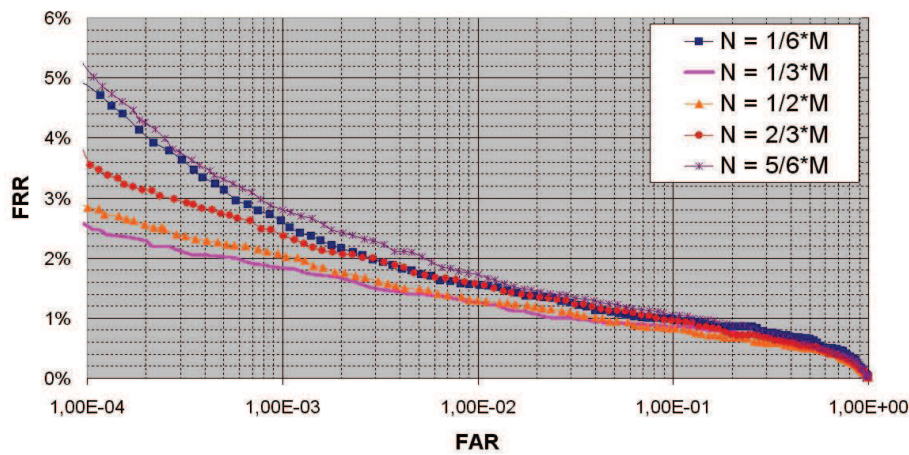


Figure 5.6: ROC curves for different values of  $N$  (number of selected points for feature extraction) and a set value of  $M$  (initial number of points) for the CASIA-IrisV3-Interval database.

CONFIDENTIAL



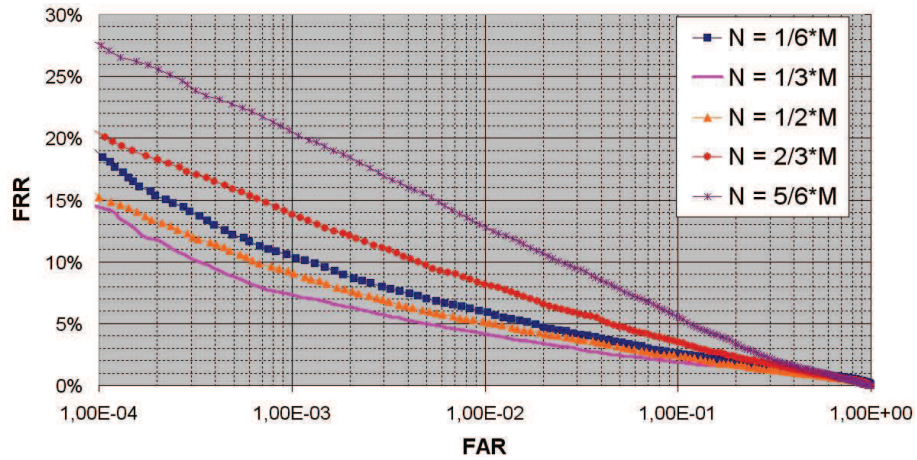


Figure 5.7: ROC curves for different values of  $N$  (number of selected points for feature extraction) and a set value of  $M$  (initial number of points) for the CASIA-IrisV3-Twins database.

On both of these databases the best of the tested values for  $N$  is equal to  $M/3$ , so it is the same as for the ND-IRIS databases.

Figure 5.8 presents, for the three different databases, the False Rejection Rate at a False Acceptance Rate equal to  $10^{-4}$  for different values of  $N$  and shows once again that the lowest values of the FRR are obtained for  $N = M/3$ .

Database \ N	N				
	$1/6 * M$	$1/3 * M$	$1/2 * M$	$2/3 * M$	$5/6 * M$
ND-IRIS-0405	10.5%	5.2%	6.4%	9.4%	14.3%
Casia-IrisV3-Interval	4.9%	2.5%	2.8%	3.6%	5.2%
Casia-IrisV3-Twins	18.8%	14.2%	15.2%	20.1%	27.5%

Figure 5.8: False Rejection Rate (for  $FAR = 10^{-4}$ ) for different databases and for different values of  $N$  (number of selected points for feature extraction) given in function of  $M$  (initial number of points).

Since we have obtained the same result in experiments on three databases that are different in terms of subjects and image quality, we can assume that we would obtain a similar result on other databases.

**CONFIDENTIAL**

### Experiments based on Lefèvre’s segmentation

In order to verify that the results obtained are not linked to a specific type of segmentation, we have repeated the experiments conducted here above with images segmented with Lefèvre’s algorithm.

First, we compare the recognition performance achieved when selecting the regions used for feature extraction and matching according to our method to the performance achieved when exploiting all the unmasked regions in the normalized iris images as is done in OSIRIS-V2. This is presented in Figure 5.9. It shows that the segmentation mask improves recognition performance. For an  $FAR > 2 \times 10^{-3}$  it performs better than our method. However, our method enables to achieve better performance for an  $FAR < 2 \times 10^{-3}$  which usually corresponds to the targeted FAR. This result will be explained in Section 5.3.2 when we will analyze the impact of our method on the intra-class and inter-class score distributions.

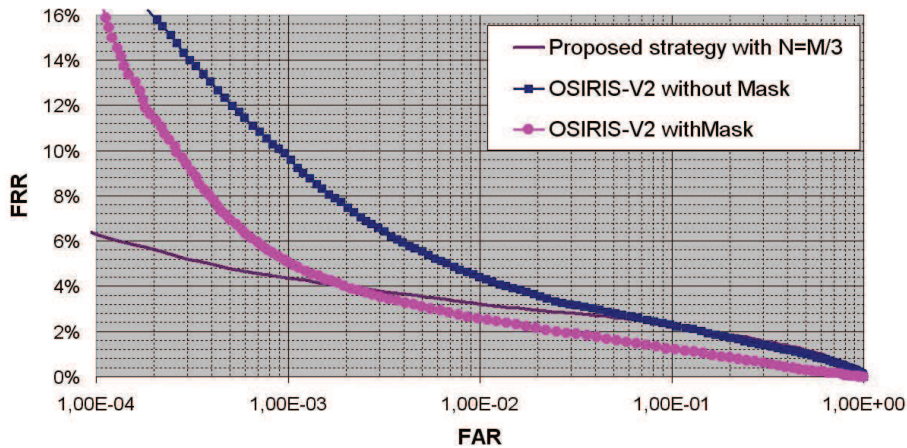


Figure 5.9: ROC curves of 3 iris recognition algorithms based on the same implementation of Daugman’s approach but using different strategies for feature extraction’s point selection, on images from the ND-IRIS database obtained with Lefèvre’s segmentation.

Then, we verified that the best value for  $N$  for a set value of  $M$  is the same with the images resulting from Lefèvre’s segmentation than with the images resulting from a manual segmentation. To do this, we have applied our selection strategy for different values of  $N$  and a set value of  $M$ . This is presented in Figure 5.10. We can see that once, again the best performance is achieved for  $N = M/3$ .

It is interesting to note that when choosing the application points for feature extraction wisely, the best performance is obtained when keeping only 1/3 of the initial application points, whatever database is used and whatever segmentation technique has been applied on the iris images. When more points are selected, the chances of performing feature extraction in bad quality regions rises, which increases the false rejection rate. To the contrary, when keeping fewer points, the amount of information available for the matching decreases too much, which

**CONFIDENTIAL**

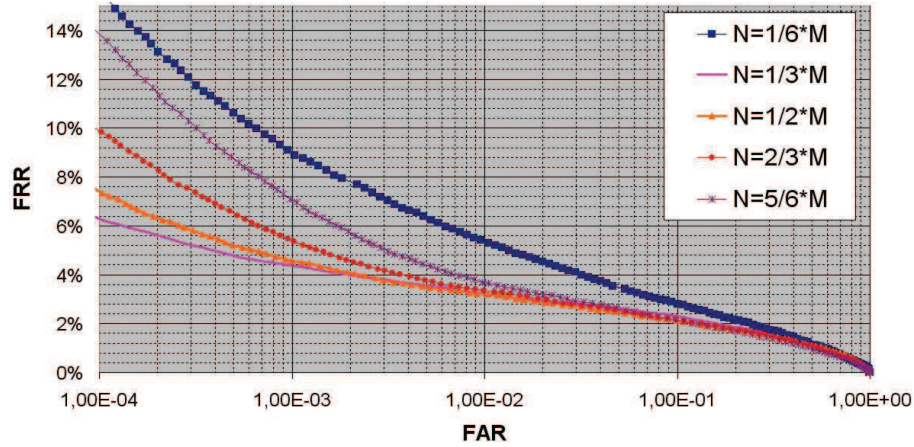


Figure 5.10: ROC curves for different values of  $N$  (number of selected points for feature extraction) and a set value of  $M$  (initial number of points) for images from the ND-IRIS database obtained with Lefèvre's segmentation.

tends to increase the False Acceptance Rate.

For the moment we have tested only 5 possible values for  $N$  when  $M$  is fixed. Consequently, it is important to note that the best value obtained for  $N$  should be seen as an indication of the range of values that are best for  $N$  and not as a specific optimal value.

### 5.3 Understanding of the best value of $N$

In the previous section we have seen that selecting the regions of normalized iris images used for feature extraction and matching according to our local quality measure improves the recognition performance. Moreover, for a set number  $M$  of initial regions, we have demonstrated that there is a number  $N_{best}$  of regions to select that leads to the best performance and that this number  $N_{best}$  is the same for three different databases: ND-IRIS-0405, CASIA-IrisV3-Interval and CASIA-IrisV3-Twins.

The aim of this section is to understand why there is a value for  $N$  that is best. First we will demonstrate that the important parameter to optimize performance is effectively the ratio  $N/M$  that represents the proportion of each initial normalized iris image that is exploited for feature extraction and matching. Then we will analyze the impact of this parameter on the distributions of inter-class and intra-class comparisons, when the iris boundaries were determined manually or with Lefèvre's algorithm.

#### 5.3.1 Importance of the ratio $N/M$

The first important point to understand is the impact of the value of  $M$  on the performance of OSIRIS-V2. The iris images exploited in this section are 2136 normalized iris images from the ND-IRIS-0405 database (6 images for each of the 246 subjects). They were obtained by determining manually the circles that

**CONFIDENTIAL**

model best the irises inner and outer boundaries and then normalizing the iris with OSIRIS' normalization module, as explained in Chapter 3.

We have tested four values of  $M$ : 250, 500, 750 and 1000. For each tested value, the  $M$  application points are uniformly distributed in each normalized iris image. Figure 5.11 presents the comparative performance of OSIRIS-V2 on images from the ND-IRIS database for different values of  $M$ . It turns out that the performance is the same for  $M = 500, 750$  and  $1000$ , but it is less good for  $M = 250$ . This shows us that increasing the amount of information extracted from each normalized iris image by raising the value of  $M$  improves recognition performance until a certain threshold ( $M_{th} = 500$ ) from which the additional information becomes redundant.

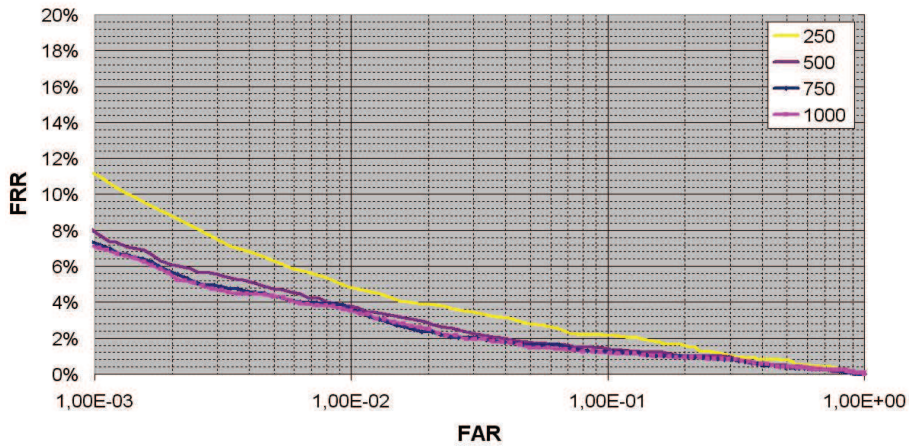


Figure 5.11: ROC curves for different values of  $M$  (number of uniformly distributed application points) for the ND-IRIS-0405 database.

This being established, we examined if the best value for  $N$  depends on  $M$ . To this end, we carried out the same experiment as in Section 5.2.2 on images from the ND-IRIS-0405 database, for different values of  $M$ . More precisely, we have tested our quality-based region selection strategy with different values of  $N$ , for each value of  $M$  considered. Once again, the values tested for  $M$  are: 250, 500, 750 and 1000. For each value of  $M$ , we have tested five values for  $N$ :  $M/6$ ,  $M/3$ ,  $M/2$ ,  $2M/3$  and  $5M/6$ . Table 5.1 presents the different values tested for  $N$  for each value of  $M$ .

Table 5.2 presents the value of  $N$  with which the best recognition performance are achieved for different values of  $M$ , as well as the ratio  $N_{best}/M$ . We can notice that the value of this ratio is approximately the same for all values of  $M$ . This proves that the important parameter is the proportion of the initial image that is exploited for feature extraction and matching and not the absolute number of exploited regions.

**CONFIDENTIAL**

Table 5.1: Values tested for  $N$  for each value of  $M$ 

$M$	Values tested for $N$
250	42, 83, 125, 167, 208
500	83, 167, 250, 333, 417
750	125, 250, 375, 500, 625
1000	167, 333, 500, 667, 833

Table 5.2: Values of  $N$  corresponding to the best performance

$M$	$N_{best}$	$N_{best}/M$
250	83	0.35
500	167	0.35
750	250	0.33
1000	333	0.35

**CONFIDENTIAL**



### 5.3.2 The impact of $N/M$ on score distributions

To understand why there is a value of the ratio  $N/M$  that maximizes performance, we have analyzed the impact of this ratio on the score distributions of inter-class and intra-class comparisons. For this experiment, we have set the value of  $M$  to 750.

#### Images obtained with manual segmentation

First, we have worked with images that were obtained with a manual segmentation. Figure 5.12 presents the mean of the normalized inter-class and intra-class distributions ( $mean_{inter}$  and  $mean_{intra}$ ) for different values of  $N$ .

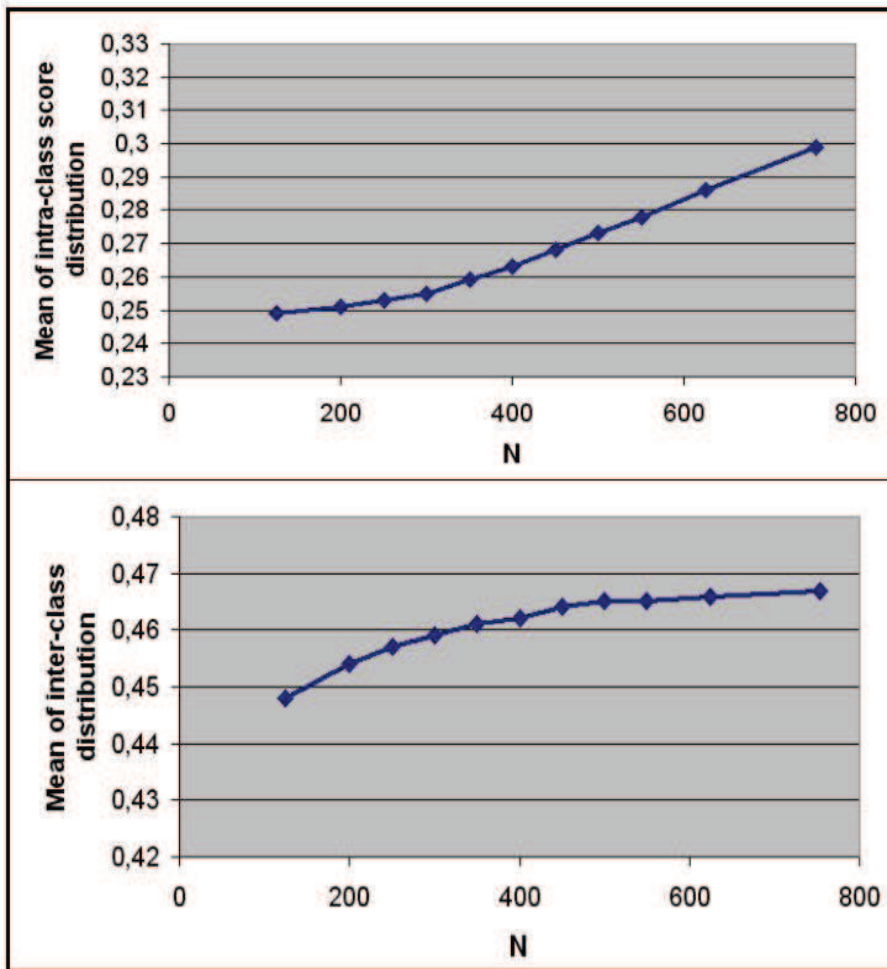


Figure 5.12: Evolution of the means of the normalized inter-class and intra-class score distributions with  $N$ .

We can see that decreasing the value of  $N$  leads to a translation of both the intra-class and inter-class distribution towards the lower Hamming distances. Since, a lower Hamming distance indicates that the similarity between the two

**CONFIDENTIAL**

iris images compared is greater, the best recognition performance is achieved when  $mean_{intra}$  is low and  $mean_{inter}$  is high. So reducing  $N$  improves the intra-class distribution but deteriorates the inter-class distribution. However, the speed at which the  $mean_{inter}$  and  $mean_{intra}$  decrease is not the same.  $mean_{inter}$  decreases slowly at first, then the decline becomes faster. To the contrary,  $mean_{intra}$  starts of by decreasing quickly, then the decline slows down.

Our interpretation of this phenomenon is the following. Since the  $M$  application points are sorted according to their local quality, the regions of worst quality are eliminated first. So for the highest values of  $N$ , the regions that are eliminated are mostly occluded regions. When  $N$  continues to decrease and all occluded regions have been eliminated, then regions that are free from occlusion are eliminated, starting with the ones that are the less textured. We have seen in Chapter 4 that the impact of occlusions on recognition performance is much higher than the impact of poor texture. This explains why  $mean_{intra}$  decreases fast at first, for the eliminated regions are mainly occluded, then it decreases more slowly, when all occluded sub-images have been eliminated. Consequently, the value of  $N$  at which the decrease of  $mean_{intra}$  slows down should depend on the global quality of the image couple to compare.

On the other hand, the  $mean_{inter}$  does not depend on the quality of the sub-images used for feature extraction and matching, as shown in Chapter 4. So eliminating bad quality regions does not increase its value. However, reducing the amount of information used leads to matching less discriminative binary codes and therefore biases the inter-class distribution. This phenomenon is emphasized more when the amount of information is small.

The optimal value of  $N$  will compromise between improving the intra-class distribution without deteriorating the inter-class distribution too much. Since the intra-class distribution is improved rapidly at first and more slowly afterwards, the optimal value of  $N$  corresponds to the point where the slope of  $mean_{intra} = f(N)$  changes. We can see on Figure 5.12, that this happens approximately for  $N = 300$ . Since,  $M = 750$ ,  $N/M=0.4$ . This explains why the best value of  $N/M$  out of the five tested, was  $1/3$ .

### Images obtained with Lefèvre's segmentation

The results presented in the previous paragraph do not explain why Daugman's feautre extraction and matching algorithm performs less well with our region selection strategy than with Lefèvre's segmentation mask when the  $FAR < 2 \times 10^{-3}$ .

To understand this result, we compared the normalized score distributions achieved with these two methods. This is presented in Figure 5.13. We can see that the intra-class distribution is shifted towards the lower Hamming distances with our region selection strategy, compared to the method using Lefèvre's segmentation mask. However, one part of the distribution, corresponding to Hamming distances higher than 0.4, is not shifted.

We selected these intra-class comparisons and observed the associated normalized iris images. These turned out to be images for which the segmentation had completely failed. Figure 5.14 shows an example of such an image. It appears logical that selecting the regions in the normalized iris image for feature extraction and matching will not make up for an incongruous normalization.

**CONFIDENTIAL**

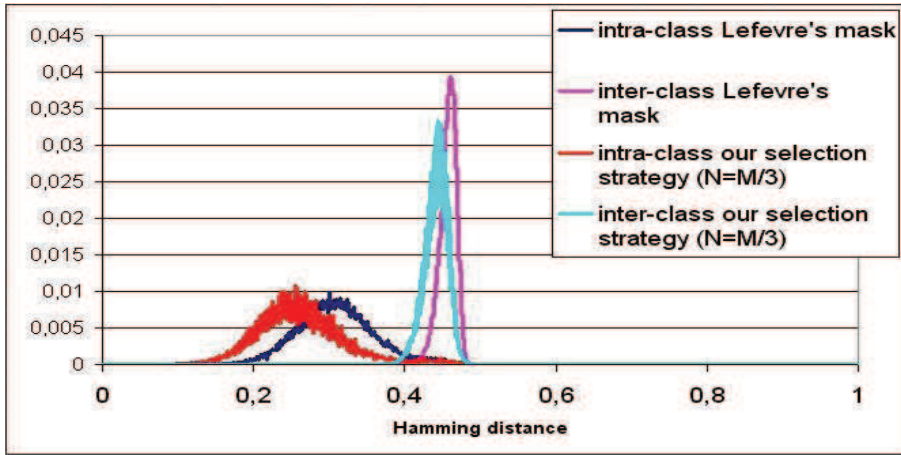


Figure 5.13: Normalized score distributions obtained when applying Daugman's feature extraction and matching algorithm with our region selection strategy or with Lefèvre's segmentation mask.

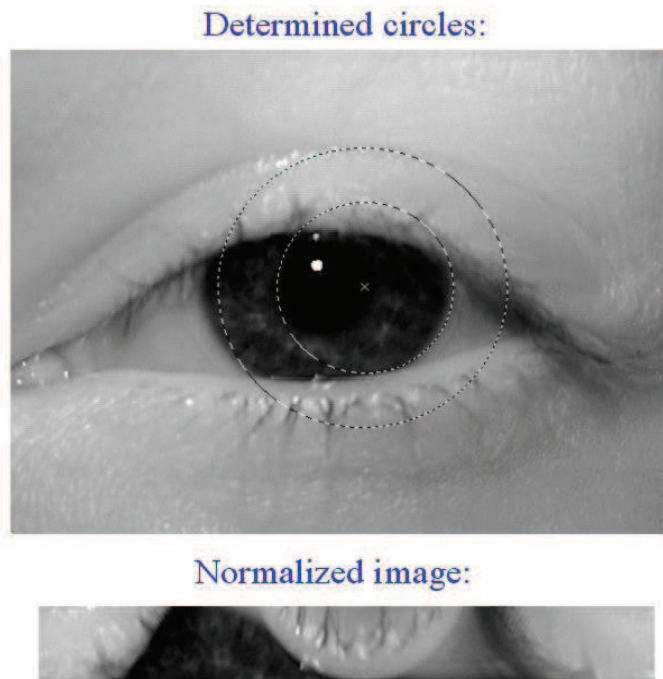


Figure 5.14: Example of an image for which the normalization failed

**CONFIDENTIAL**



This is why our method does not reduce the Hamming distances resulting from the comparison including one of these images.

On the other hand, our method induces a slight deterioration of the inter-class distribution: it is shifted towards the lower Hamming distances. Consequently, for the highest FARs, corresponding to the highest Hamming distance thresholds, our method performs less well than the one based on Lefèvre’s segmentation mask.

## 5.4 Combination of global and local information

### 5.4.1 Motivation

In Section 5.2, the proportion of the normalized iris image exploited for feature extraction and matching was the same for all (gallery, probe) couples of images to compare. In this section, we wish to demonstrate that the best value of this proportion, represented by the ratio  $N/M$ , can change for each (gallery, probe) couple depending on the couple’s quality.

It appears obvious that the less occluded an image is, the more regions of the image can be used for matching without taking into account occlusions. Therefore the higher the image quality is in terms of occlusion, the higher the best value of the ratio  $N/M$  should be. Moreover, non-occluded regions that contain very few texture will not necessarily bring discriminating information for matching. So the amount of texture can also influence the best value for  $N/M$ .

To qualify a couple of (probe, gallery) images globally in terms of occlusion and amount of texture we have exploited the global quality measure  $GlobQ_1$  described in Chapter 4. Recall that it is defined as in formula 5.6.

$$GlobQ_1 = \text{mean}_{i \in [1;M]}(\min(Q^{gal}(w_i), Q^{probe}(w_i))) \quad (5.6)$$

$GlobQ_1$  is directly linked to the local measure used for feature selection. We therefore expect the ratio  $N/M$  to depend on  $GlobQ_1$ .

### 5.4.2 Experiments and results

In this section we have exploited normalized iris images from the ND-IRIS-0405 and the CASIA-IrisV3-Lamp databases. These images were obtained with Lefèvre’s segmentation module followed by the normalization module of OSIRIS. Both of these modules have been described in Chapter 3. They have also been pre-processed to bring the mean of their gray-levels to a constant value as explained in Chapter 4.

We have applied the recognition process described in section 5.2 on these normalized iris images. More specifically, in order to demonstrate the link between the global quality defined in formula 5.6 and the best value of  $N/M$  for a given pair of images, we have generated matching lists for these databases and then divided them into five categories according to the value of the global quality associated to each (gallery, probe) couple. Then we have applied our recognition process on each category for different values of  $N$  and compared the achieved performances.

**CONFIDENTIAL**

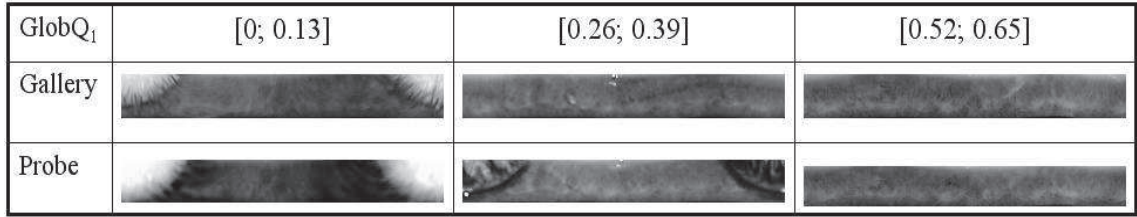


Figure 5.15: Examples of (gallery, probe) normalized iris image couples for different global qualities.

The values of the global quality measure that define the boundaries of each of the five categories were determined by analyzing the range of the global quality on the entire list, and dividing the resulting quality interval into five equal-sized intervals. Table 5.3 presents the percentage of the full matching list in each category for the ND-IRIS-0405 and CASIA-IrisV3-Lamp databases. We have considered a total of 6 million inter-class and 60 000 intra-class matches from the ND-IRIS-0405 database and 5 million inter-class and 20 000 intra-class matches from the CASIA-IrisV3 database.

Figure 5.15 gives examples of couples of (gallery, probe) normalized iris images belonging to categories 1, 3 and 5.

Table 5.3: Percentage of the full matching list for each category of  $GlobQ_1$

Categories	1	2	3	4	5
$GlobQ_1$	[0; 0.13]	[0.13; 0.26]	[0.26; 0.39]	[0.39; 0.52]	[0.52; 0.65]
ND-IRIS-0405	11%	16%	34%	33%	6%
CASIA-IrisV3-Lamp	13%	21%	39%	24%	3%

We have started by applying our recognition process to the first category of matches ( $0 < GlobQ_1 < 0.13$ ) for different values of  $N$ . The values tested for  $N/M$  were 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9. Figure 5.16 presents the ROC curves, i.e the False Rejection rate (FRR) as a function of the False Acceptance rate (FAR), obtained for 4 values of  $N/M$  on this category of images of the CASIA-IrisV3-Lamp database. We have chosen to plot the curves for only 4 values to avoid overloading the graph. We have kept the value for which the best performance is achieved and chosen a pace equal to 0.2.

It shows that the best performance is achieved for  $N/M = 0.4$ . We have carried out the same experiment on the four other categories of matches. Table 5.4 presents the values of  $N/M$  that lead to the best performance for each category for the CASIA-IrisV3-Lamp and the ND-IRIS-0405 databases.

It appears that the best value of  $N/M$  is the same for the two first categories and increases with  $GlobQ_1$  for the three other ones. This can be explained by the fact that a higher value of  $GlobQ_1$  implies that the quality of the images is better in terms of occlusion and amount of texture, so taking into account more regions for feature extraction and matching adds useful information for

**CONFIDENTIAL**

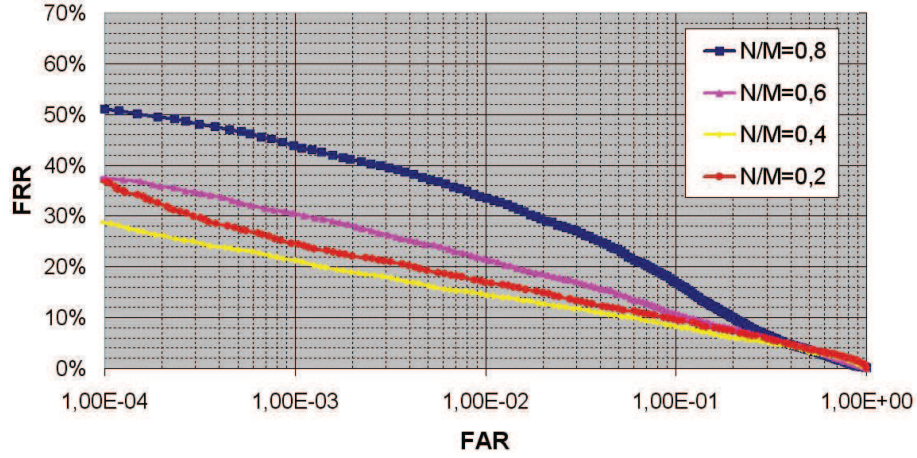


Figure 5.16: ROC curves for 4 different values of  $N/M$  on matches from the CASIA-IrisV3-Lamp database for which  $0 < GlobQ_1 < 0.13$ .

Table 5.4: Best values for  $N/M$  for different databases and different values of  $GlobQ_1$  when using Daugman's algorithm

$GlobQ_1$	[0; 0.13]	[0.13; 0.26]	[0.26; 0.39]	[0.39; 0.52]	[0.52; 0.65]
ND-IRIS-0405	0.4	0.4	0.5	0.6	0.7
CASIA-IrisV3-Lamp	0.4	0.4	0.5	0.6	0.7

recognition.

It is interesting to note, that the optimal value of  $N/M$  is the same for categories 1 and 2, which means that when  $GlobQ_1$  goes beneath a certain threshold,  $N/M$  stays stable and does not decrease anymore. One reason for this is that this value of  $N/M$  is the value underneath which too few information is taken into account for the matching to be significant. Consequently, under a certain value of  $GlobQ_1$ , taking into account regions of low quality has a less negative impact on recognition performance than taking into account too few regions.

Table 5.4 shows that the best value of  $N/M$  for  $(0.39 < GlobQ_1 < 0.52)$  is 0.6. On the other hand, Figure 5.16 shows that using  $N/M = 0.6$  for  $0 < GlobQ_1 < 0.13$ , instead of  $N/M = 0.4$  leads to a 30% increase of the FRR for an  $FAR=10^{-4}$ . This demonstrates that setting  $N/M$  appropriately with the knowledge of  $GlobQ_1$  can considerably improve performance compared to the use of a different value of  $N/M$ .

Moreover, for a given range of  $GlobQ_1$ , the best value for  $N/M$  is the same whether the exploited database is ND-IRIS-0405 or CASIA-IrisV3-Lamp.

**CONFIDENTIAL**

### Link with the best value for $N/M$ obtained in Section 5.2

In section 5.2 we had applied our region selection strategy for feature extraction using the same value of  $N/M$  for all couples of images considered in the ND-IRIS-0405 database. We had tested five values for  $N/M \in \{1/6; 1/3; 1/2; 2/3; 5/6\}$  and demonstrated the best performance were achieved for  $N/M = 1/3$ .

In section 5.4 we have divided the pairs of images to compare into five categories according to their global quality values and have determined the best value of  $N/M$  for each category. This time the values tested for  $N/M$  were  $2/10, 3/10, 4/10, 5/10, 6/10, 7/10, 8/10$  and  $9/10$ . Table 5.5 gives the percentage of images in each category as well as the optimal value of  $N/M$  for the ND-IRIS-0405 database and for different values tested in  $E_{1/6} = \{\frac{1}{6}; \frac{1}{3}; \frac{1}{2}; \frac{2}{3}; \frac{5}{6}\}$  and  $E_{1/10} = \{\frac{2}{10}; \frac{3}{10}; \frac{4}{10}; \frac{5}{10}; \frac{6}{10}; \frac{6}{10}; \frac{8}{10}; \frac{9}{10}\}$ .

Table 5.5: Proportion of the full ND-IRIS-0405 database in each range of  $GlobQ_1$  and best value of the ratio  $N/M$  for each quality range

$GlobQ_1$	[0; 0.13]	[0.13; 0.26]	[0.26; 0.39]	[0.39; 0.52]	[0.52; 0.65]
Percentage of the total database	11%	16%	34%	33%	6%
$(N/M)_{best} \in E_{1/10}$	4/10	4/10	5/10	6/10	7/10
$(N/M)_{best} \in E_{1/6}$	1/3	1/3	1/2	2/3	2/3

Given this information, it is hard to understand why the best value for  $N/M$  on the entire database is  $1/3$ . For this to become apparent, we need to consider the absolute and relative performance improvement achieved by using the optimum value of  $N/M$  instead of another value. This is presented in Table 5.6

Table 5.6: Absolute and relative improvement of the FRR for  $FAR = 10^{-4}$  for different global quality ranges

$GlobQ_1$	[0; 0.13]	[0.13; 0.26]	[0.26; 0.39]	[0.39; 0.52]	[0.52; 0.65]
FRR for $(N/M)_{best}$	49%	15%	4%	1%	0%
FRR for $N/M=1$	70%	43%	17%	2%	0.2%
Absolute performance improvement	21%	28%	13%	1%	0.2%
Relative performance improvement	30%	65%	76%	50%	100%

It appears that the absolute performance variation is much more important for the categories of lower quality. So these categories will weigh much more than the others on the performance of the entire database. This explains

**CONFIDENTIAL**

why the best value on the entire database is 4/10 when the values tested are  $\{2/10, 3/10, 4/10, 5/10, 6/10, 7/10, 8/10, 9/10\}$  and 1/3 when the values tested are  $\{1/6; 1/3; 1/2; 2/3; 5/6\}$ .

Either way, the experiments we conducted only give an indication on the range of the best values for  $N/M$  and do not allow to determine a precise value of this ratio that would be optimal. Note that given our confidence intervals it would not be possible to determine such a value precisely.

## 5.5 Conclusion

In this chapter, we integrated our local quality metric, described in Chapter 4, into the feature extraction and matching algorithms proposed by Daugman [21].

First we included it in the computation of the Hamming distance by defining a weighted Hamming distance similarly to what has been done in literature. We showed that this strategy has its drawbacks: it does not allow any control on the amount of information that is considered for matching. Consequently, it performs badly on images of low quality for which a high proportion of bits will be weighted by low local quality values. For such images, only few bits contribute significantly to the computation of the Hamming distance.

Then we proposed our own method for integrating our quality metric into Daugman's algorithm. It consists in sorting the regions of normalized iris images according to their local quality and selecting the  $N$  regions with the highest qualities for feature extraction and matching. We demonstrated that the recognition performance achieved with this method depends highly on the value of  $N$ . Moreover, there is a value of  $N$  for which the achieved performance is the highest. We showed that this value was the same for images coming from different databases and obtained with different segmentation techniques.

We conducted further experiments to understand this result. They showed that the important parameter was not  $N$ , but  $N/M$  i.e., the proportion of the normalized image that is selected for feature extraction and matching. The best value of  $N/M$  compromises between improving the intra-class distribution without deteriorating the inter-class distribution too much. Furthermore, the improvement of the intra-class distribution depends on the quality of the iris images considered. The lower is this quality, the lower  $N/M$  needs to be set for the intra-class distribution to be significantly improved. However, there is a threshold value of  $N/M$  under which the improvement of the intra-class distribution does not compensate for the deterioration of the inter-class distribution. Consequently, the best value of  $N/M$  is never below this threshold.

Following this, we showed that the optimal value of  $N/M$  for each pair of images to compare, depends on their global quality, that can be measured with the metric  $GlobQ_1$  we defined in Chapter 4. Setting  $N/M$  inaccurately will have a higher impact on low quality images than on high quality images. Consequently, it is mainly the low quality images that impact the best value of  $N/M$  in an entire database, which explains why the best value of this ratio was the same for the different databases we worked with.

**CONFIDENTIAL**

## Chapter 6

# Integration of our quality measure in Krichen's algorithm

We have seen in the previous chapter that exploiting the local quality measure described in Chapter 4 with an iris recognition system based on Daugman's technique can improve recognition accuracy. The quality measure is used to select the regions of the normalized iris images that are used for feature extraction and matching. This way, the primarily selected regions are the ones that are free from artifacts and highly textured. Moreover we have shown that it is important to control the proportion of the normalized images that is selected. More precisely, there is a value of this proportion for which the recognition performance is best and it depends on the global quality of each normalized iris image couple to compare.

The objective of this chapter is to study if we can exploit our local quality measure in the same way with an iris recognition system based on Krichen's technique. We will evaluate the improvement in recognition performance that this can achieve and verify that we can find a proportion of the normalized iris images to select for feature extraction and matching that optimizes performance. It will be interesting to examine if the best value for this proportion is the same for a Daugman-based and a Krichen-based algorithm.

The images exploited in this Chapter were segmented with Lefèvre's algorithm and normalized with OSIRIS' normalization module as explained in Chapter 3. They were also pre-processed to bring the mean of their gray-level values to a constant value as describes in Chapter 4.

## 6.1 Integration of the region selection routine

### 6.1.1 Description of the proposed technique

As in Chapter 5, we consider iris images that have previously been segmented, normalized and pre-processed according to the procedures described in Chapter 3 (segmentation and normalization) and Chapter 4 (pre-processing). We have applied the process illustrated in Figure 6.1 to compare a probe image to a

gallery image. Step 1 to 5 are identical to the first steps of the process described in Chapter 5 but steps 6 and 7 comply with the feature extraction and matching technique proposed by Krichen in [48]. More precisely:

- **Step 1:**  $M$  uniformly distributed points are chosen in each of the gallery and probe images.  $M$  has been set to a constant value.
- **Step 2:** the local qualities of the sub-images centered at these  $M$  points is computed.
- **Step 3:** the  $M$  local qualities from the gallery image are fused with the  $M$  local qualities from the probe image: the minimum value is selected.
- **Step 4:** the fused local qualities are sorted decreasingly in order to allocate a rank to the associated points
- **Step 5:** the  $N$  points with the lowest ranks are selected. They correspond to the sub-images of highest quality.
- **Step 6:** convolution with  $N_f$  2D-Gabor filters is performed for the  $N$  sub-images centered around the  $N$  selected points. This is done both for the gallery and the probe images
- **Step 7:** normalized cross-correlations between each of the  $N_f \times N$  couples of sub-images are computed, and the similarity score is determined

As in Chapter 5, we use our quality measure to select the  $N$  points related to the sub-images of highest quality and perform convolution with 2D Gabor filters around these points. However instead of quantizing the Gabor phases we perform local normalized cross-correlation between the phase sub-images. Each cross-correlation leads to a Peak to Slob Ratio (PSR) value and a Peak Position, as explained in Section 4.4.1. The similarity score is computed following formula 6.1

$$SS = \text{mean}_{k \in [1; N_f]} \left( \frac{\text{mean}_{i \in [1; N]}(PSR(w_i))}{\text{std}_{i \in [1; N]}(PP(w_i))} \right) \quad (6.1)$$

### 6.1.2 Experiments and results

In this section, we will show that the process described in Section 6.1.1 for a set value of  $M$ :

- performs better than Krichen's algorithm alone
- achieves best performance for a given value of  $N = N_{best}$
- the value determined for  $N_{best}$  is the same for the process based on Krichen's matching as for the process based on Daugman's matching

To this end, we applied the process described in Section 6.1.1, to iris image couples from the ND-IRIS-0405 database for a fixed value of  $M$  and different values of  $N \in \{M/6; M/3; M/2; 2M/3; 5M/6\}$ . We also applied the algorithm Krichen describes in [48] to the same images. The main difference of the algorithms is that Krichen's uses all of the  $M$  sub-images for feature extraction and matching,

**CONFIDENTIAL**



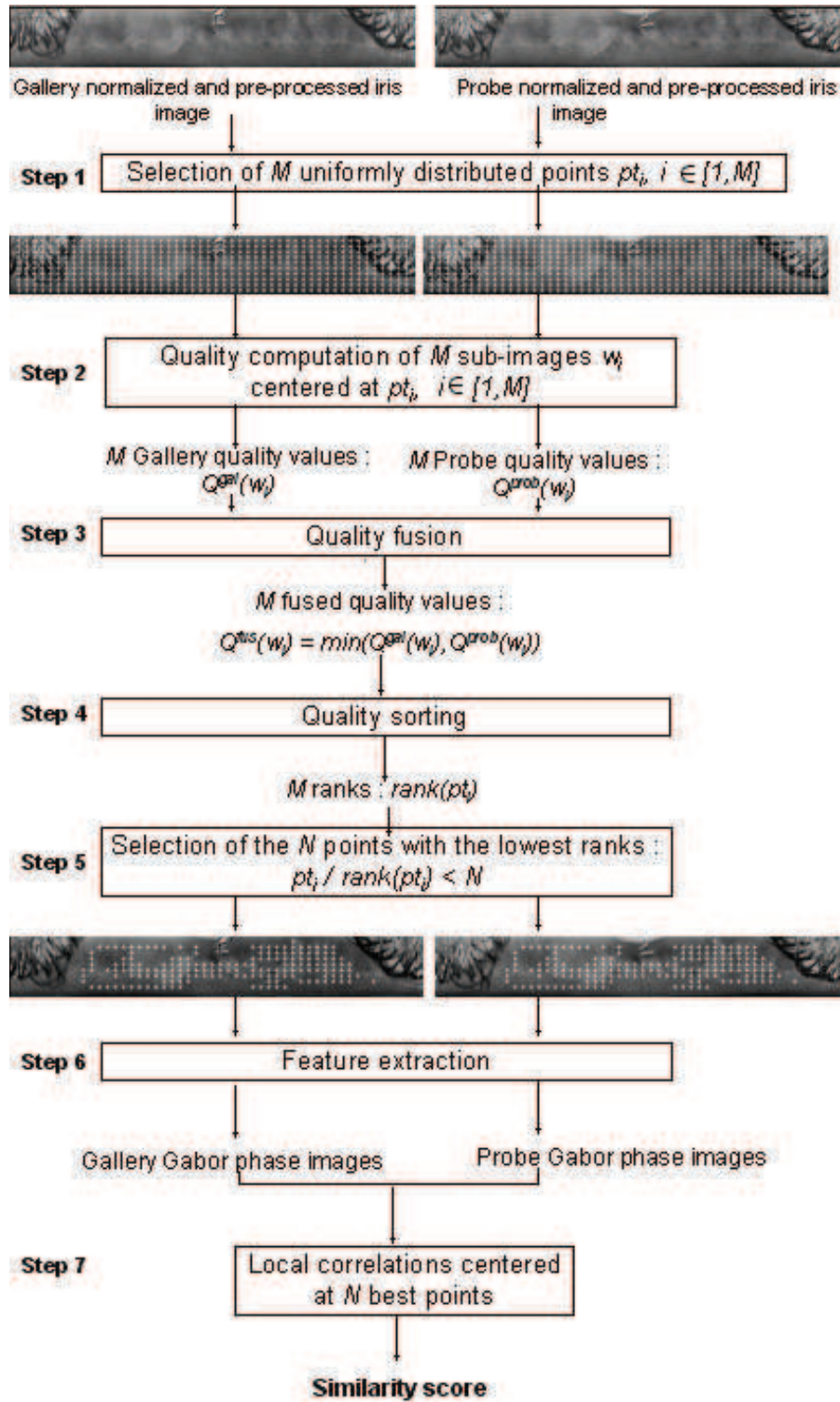


Figure 6.1: Global scheme of the method we propose for pertinent feature extraction and matching according to Krichen's technique using our quality measure.

**CONFIDENTIAL**



whereas our algorithm selects the  $N$  best sub-images according to our quality measure. Figure 6.2 presents the comparative performance obtained for 300 000 inter-class and 10 000 intra-class matches selected randomly in matching lists generated with the entire ND-IRIS-0405 database.

We can see that the best performance is achieved for  $N = M/3$ . For this value of  $N$  our process leads to an FRR=4.7% for FAR= $10^{-3}$  while Krichen's algorithm gives an FRR=6.7% at the same FAR. Consequently, performance is improved by 30%.

We have carried out the same experiment with images from the CASIA-IrisV3-Lamp database. Figure 6.3 presents the ROC curves obtained with 300 000 inter-class and 10 000 intra-class comparisons selected randomly from the matching lists of all possible comparisons for the CASIA-IrisV3-Lamp database.

For this database, the best value for  $N/M$  out of the five values tested is  $1/3$ , which confirms the result we have obtained with the ND-IRIS-0405 database. For this value of  $N/M$ , FRR=4.4% for an FAR= $10^{-3}$ , while FRR=7.5% for the same FAR with Krichen's algorithm.

These two experiments lead us to conclude that selecting the best sub-images in normalized iris images for feature extraction in matching improves recognition performance whether the technique used is Daugman's or Krichen's. In both cases the selected proportion of the normalized images that optimizes performance is equal for both algorithms.

## 6.2 Combination of global and local information

### 6.2.1 Motivation

We have demonstrated in Chapter 5 that the optimal proportion of the normalized iris images to keep for feature extraction and matching with Daugman's technique depends on the global quality associated to the considered image couple. This result is obtained because the suppression of regions implies a trade-off between the improvement of the intra-class distribution caused by the elimination of bad quality regions and the deterioration of the inter-class distribution caused by the reduction of discriminant information available for matching. Moreover, the performance improvement caused by the elimination of regions of the normalized images depends on the quality of the suppressed regions: the lower the quality of the eliminated sub-images, the higher is the improvement in performance. So the higher the global quality of the images, the less regions need to be suppressed in order to achieve optimal performance. This explains why the proportion of the image to keep depends on the global quality of the image pair.

We expect this phenomenon to be independent of the technique used for feature extraction and matching. Consequently, we can imagine that the proportion of the normalized iris images to keep for feature extraction and matching with Krichen's technique will also depend on the global quality of the image to compare. However, we cannot assume a priori that the function relating this proportion to the global quality will be the same as when the feature extraction and matching is done according to Daugman's technique. Indeed, the speed

**CONFIDENTIAL**

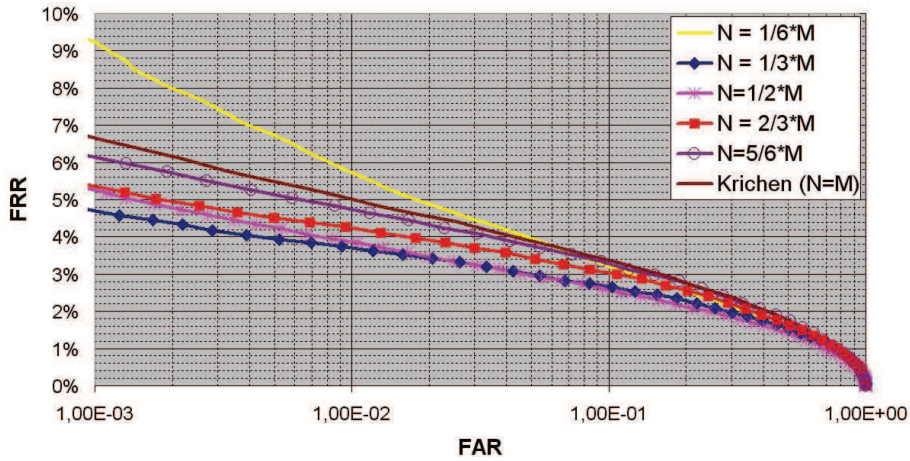


Figure 6.2: ROC curves for different values of  $N$  (number of points selected for feature extraction) and a set value of  $M$  (initial number of points) for the ND-IRIS-0405 database.

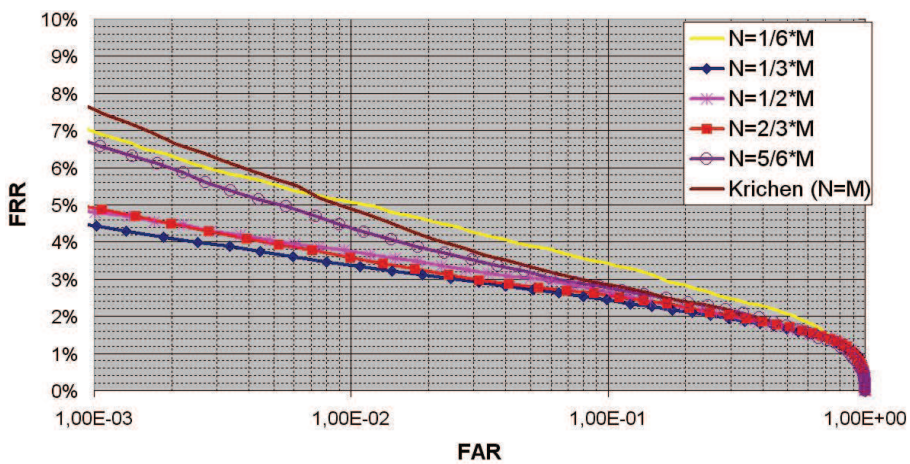


Figure 6.3: ROC curves for different values of  $N$  (number of points selected for feature extraction) for the CASIA-IrisV3-Lamp database.

**CONFIDENTIAL**

of degradation of the inter-class score distribution with the suppression of the exploited regions is likely to depend on the matching technique used.

The aim of this section is to verify that the proportion of the normalized iris images to keep for feature extraction and matching with Krichen's technique will depend on the global quality of the image to compare. Moreover, we wish to compare the function that relates this proportion to the global quality with Krichen's and with Daugman's techniques.

## 6.2.2 Experiments and results

Similarly to what has been done in Section 5.4 we have generated matching lists with images from the ND-IRIS-0405 and CASIA-IrisV3-Lamp databases and divided them into five categories according to the global quality associated to each (gallery, probe) couple. These created matching lists are the same as the ones in Section 5.4: 6 million inter-class and 60 000 intra-class comparisons from the ND-IRIS-0405 database, and 5 million inter-class and 20 000 intra-class comparisons from CASIA-IrisV3-Lamp. Once these matching lists were divided into five categories according to their global quality, we have randomly selected 120 000 inter-class and 5000 intra-class comparisons for each category. Then we have applied our recognition process to these comparisons for different values of  $N$  and compared the achieved performance.

Figure 6.4 represents the performance achieved on the first category of iris images ( $0 \leq GlobQ_1 \leq 0.13$ ) for four values of the ratio  $N/M$ . It shows that on this category the optimal value for  $N/M$  is 0.4.

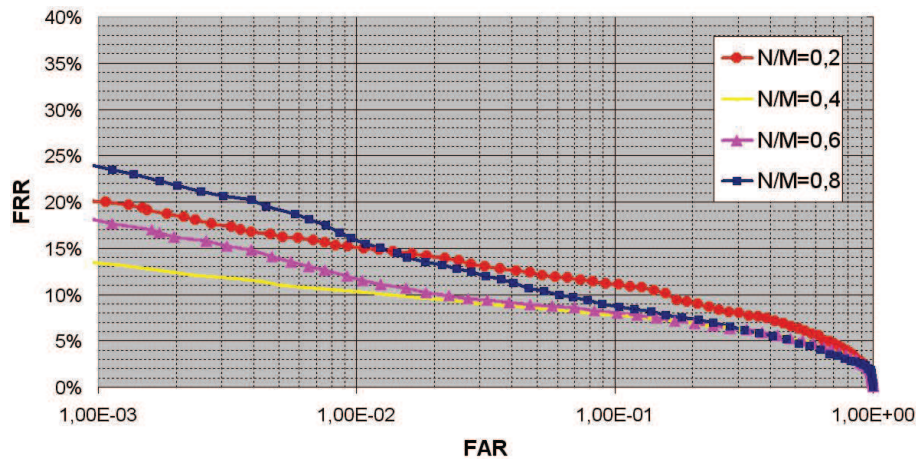


Figure 6.4: ROC curves for different values of  $N/M$  for matches from the CASIA-IrisV3-Lamp database for which  $0 \leq GlobQ_1 \leq 0.13$ .

Table 6.1 presents the values of  $N/M$  that lead to the best performance for each category for the CASIA-IrisV3-Lamp and the ND-IRIS-0405 database.

We can see that the best value for the ratio  $N/M$  increases when the global quality of the images improves. This confirms that the best value of  $N/M$  depends on the global quality of the compared images as we had expected.

Furthermore it is interesting to notice that for each category of  $GlobQ_1$ , the optimal value of  $N/M$  when using Krichen's feature extraction and matching

**CONFIDENTIAL**

Table 6.1: *Best values for  $N/M$  for different databases and different values of  $GlobQ_1$  when using Krichen's algorithm*

$GlobQ_1$	[0; 0.13]	(0.13; 0.26]	(0.26; 0.39]	(0.39; 0.52]	(0.52; 0.65]
<b>ND-IRIS-0405</b>	0.4	0.4	0.5	0.6	0.7
<b>CASIA-IrisV3-Lamp</b>	0.4	0.4	0.5	0.6	0.7

technique is the same as the one we had obtained in Chapter 5 with Daugman's technique (*cf* Table 5.4). More precisely, the optimum value of  $N/M$  stays constant for the first two categories and increases linearly for the next categories. This means that the trade-off between the improvement of the intra-class score distribution and the deterioration of the inter-class distribution caused by the reduction of amount of information taken into account for matching is optimum for the same proportion of suppressed information, whether the technique used for feature extraction and matching is Daugman's or Krichen's.

We saw in Section 5.3.2 that the best proportion to keep in the normalized iris images corresponds to a change of the speed at which the intra-class scores are shifted towards the lower Hamming distances. The value of  $N/M$  at which this change occurs depends on the global quality of each image pair. Changing the feature extraction and matching technique may have an impact on the absolute translation speeds of the inter-class and intra-class score distributions. However the experiments conducted in this chapter demonstrate that the speed transition of the intra-class distribution, for images belonging to a given category of global qualities, occurs for the same value of  $N/M$ . More surprisingly, the threshold value of  $N/M$  underneath which the improvement of the intra-class distribution does not compensate for the deterioration of the inter-class distribution is the same for the two algorithms considered. We can suppose that this result is linked to the fact that the same Gabor filters are used in the extraction with Daugman's algorithm as with Krichen's algorithm.

### 6.3 Conclusion

We demonstrated in this chapter that the results obtained in Chapter 6 with Daugman's feature extraction and matching technique can also be achieved with Krichen's technique. More specifically, we showed that applying our quality-based region selection strategy with Krichen's algorithm enables to improve recognition performance. Moreover, there is a specific value of the proportion of the normalized iris image to keep in order to optimize performance. This value is the same whether the applied feature extraction and matching technique is Daugman's or Krichen's. Finally, we showed that the optimal proportion to keep in the normalized images is a function of the global quality of the image couple and that this function stays the the same with Krichen's and Daugman's technique.



## Chapter 7

# Adapting the technique used for matching to image quality

In the previous chapters we used two algorithms for comparing a couple of (gallery, probe) normalized iris images. The first one was proposed by Daugman [21]. It applies 2D Gabor filters around application points in the normalized iris images and quantizes the resulting Gabor phases so that each iris is represented by a binary code. Two irises are then compared with the computation of the Hamming distance between the binary codes. The second one was proposed by Krichen et al. [48]. It applies 2D Gabor filters on the normalized iris images in the same way as Daugman's algorithm; however the real values of the Gabor phases are kept for matching. More precisely, local cross-correlations are performed between the phase images leading to the computation of a similarity score.

In Chapters 5 and 6, we proposed a novel way of selecting the regions on which the Gabor filters are applied in the normalized iris images, that can be used with both of these two algorithms. It takes into account the quality of sub-images of the normalized images to position the application points in the regions that are free from artifacts and the most highly textured. Performance is largely improved when the proportion of the normalized images used for feature extraction is set correctly whether this selection strategy is applied along with Daugman's algorithm or with Krichen's. In both cases, the optimal value of this proportion depends on the global quality of the image couple to compare.

The first aim of this Chapter is to compare the performance of Daugman's algorithm to the one of Krichen's when these two algorithms use our region selection strategy. This comparison will be carried out on different categories of images built according to their quality. We will use different quality measures. The first one, used in Section 7.1, measures the amount of artifact contained in the normalized iris images as well as the amount of texture in the non-occluded regions of the image. The second one, used in Section 7.2 detects iris images that are off-angle.

Then we will present in Section 7.3 a complete iris recognition strategy that combines Krichen's and Daugman's algorithms as well as the local and global



use of our GMM-based quality measure. This strategy also exploits the off-angle score mentioned here above.

## 7.1 Quality according to our GMM-based measure

### 7.1.1 Matching performance per quality category

In this section we consider lists of images to compare from the ND-IRIS-0405 and the CASIA-IrisV3-Lamp databases. More precisely, we consider 6 million inter-class and 60 000 intra-class comparisons from the ND-IRIS-0405 database and 5 million inter-class and 20 000 intra-class comparisons from the CASIA-IrisV3-Lamp database. The images we work with have been previously segmented according to Lefèvre’s segmentation algorithm and normalized with OSIRIS, as explained in Chapter 3. Moreover, the normalized images have been pre-processed as described in Chapter 4.

For each (gallery, probe) couple of images to match, we compute the global quality measure  $GlobQ_1$  defined in Chapter 4 by formula 7.1.

$$GlobQ_1 = \text{mean}_{i \in [1; M]}(\min(Q^{gal}(w_i), Q^{probe}(w_i))) \quad (7.1)$$

where  $Q^{gal}(w_i)$  and  $Q^{probe}(w_i)$  are the local qualities of the sub-images in the gallery and in the probe images that are centered around the application point  $pt_i$  with  $i \in [1; M]$ . These local qualities are computed using a previously trained GMM, as explained in Chapter 4. We have shown in that chapter that their values are the highest when the sub-images are free from occlusion and highly textured and the lowest when the sub-images contain artifacts. Intermediate values are given to sub-images that are free from occlusion but poorly textured. So  $GlobQ_1$  gives us an indication of the amount of occlusion in the images to compare and on the amount of texture contained in non-occluded regions.

We divide the image lists into five categories according to the values of  $GlobQ_1$ , in the same way as in Section 6.4.2. For each category of images, we select randomly 120 000 inter-class and 5000 intra-class comparisons from each database. Then we compare the performance of Daugman’s and Krichen’s algorithms when they are used with the local-quality-based region selection strategy for feature extraction as in Chapters 5 and 6, for each category of images. We use the results from Tables 5.4 and 6.1, to set the value of  $N/M$  for each category as the value that maximizes performance. The comparative performance of the Daugman and Krichen based processes is presented in Table 7.1 and Table 7.2 for the ND-IRIS-0405 database and the CASIA-IrisV3-Lamp database respectively.

These tables show that the algorithm based on Krichen’s feature extraction and matching technique performs better than the one based on Daugman’s technique on the categories of lower global quality. However, the performance is equivalent on the two categories of best global quality.

**CONFIDENTIAL**

Table 7.1: *FRR for FAR=10<sup>-3</sup> on the ND-IRIS-0405 database for different values of GlobQ<sub>1</sub> and N/M when using the recognition process based on Daugman's and Krichen's techniques*

<i>GlobQ<sub>1</sub></i>	[0; 0.13]	[0.13; 0.26]	[0.26; 0.39]	[0.39; 0.52]	[0.52; 0.65]
<i>N/M</i>	0.4	0.4	0.5	0.6	0.7
<b>Daugman</b>	33%	11%	3%	1%	0%
<b>Krichen</b>	19%	6%	2%	1%	0%

Table 7.2: *FRR for FAR=10<sup>-3</sup> on the CASIA-IrisV3-Lamp database for different values of GlobQ<sub>1</sub> and N/M when using the recognition process based on Daugman's and Krichen's techniques*

<i>GlobQ<sub>1</sub></i>	[0; 0.13]	[0.13; 0.26]	[0.26; 0.39]	[0.39; 0.52]	[0.52; 0.65]
<i>N/M</i>	0.4	0.4	0.5	0.6	0.7
<b>Daugman</b>	17%	11%	3%	1%	0%
<b>Krichen</b>	13%	7%	2%	1%	0%

### 7.1.2 Result interpretation

We have observed that the algorithm we apply that uses Krichen's feature extraction and matching technique performs better than the one that uses Daugman's technique, on image couples for which  $GlobQ_1 \in [0; 0.39]$ . Given the definition of  $GlobQ_1$  in formula 7.1, this means that Krichen's technique is more efficient on images that are highly occluded or poorly textured. Let's try to understand this result.

Krichen and Daugman's algorithm use the same 2D Gabor filters for feature extraction. However, Daugman only convolves the filters around the selected application points, while Krichen convolves the filters around every pixel from the sub-images centered around each application point. As a result Daugman obtains  $N_f \times N_{pt}$  Gabor phases while Krichen obtains  $N_f \times N_{pt}$  Gabor phase sub-images. Consequently, Krichen extracts more information than Daugman. Furthermore, Daugman quantizes the extracted information to obtain a binary code representing each iris while Krichen keeps the real-valued phase images.

Following this, they have chosen different strategies for matching. Daugman compares the binary codes by computing the Hamming distance allowing only global translations along the angular coordinates and Krichen performs local cross-correlations between the phase sub-images and computes a similarity code that takes into account both the value and the position of each correlation peak.

In consequence, one can expect that Krichen's algorithm should be more robust to local distortions as well as to a reduced amount of texture available. This is because Krichen's matching allows local translations along the radial and angular directions while Daugman only allows global angular translations. Krichen therefore handles better local distortions. Moreover, the feature vectors that Krichen exploits contain more information than Daugman's, so they are less sensitive to the reduction of the amount of texture. In addition to exploiting local information on the iris features by extracting the Gabor phases, it utilizes information in the entire sub-images by performing cross-correlation between them. When the iris is poorly textured, this global information can be very

**CONFIDENTIAL**



useful for comparing iris images.

Both the local distortions and the amount of texture are linked to our global quality measure:

- local distortions mostly occur in the presence of occlusions which are detected by our quality measure
- the amount of texture is also directly measured by our quality measure

We can therefore understand why Krichen’s algorithm performs better than Daugman’s on couple of images that have a low global quality according to our measure.

Moreover, it is important to note that the algorithm based on Krichen’s technique is more than 10 times slower than Daugman’s. The best compromise between speed and precision would therefore be to apply Krichen’s technique only on the image couples for which it will perform best. In practice, this can be done by utilizing our global quality measure, since the benefit of Krichen’s algorithm towards Daugman’s depends on  $GlobQ_1$ . This idea will be developed in Section 7.3.

## 7.2 Off-angle quality

In this section we consider a different quality criteria to categorize iris images. We attempt to divide the ND-IRIS-0405 database in two: off-angle and non off-angle iris images. We use the original iris images to categorize each image. Once the image is categorized, we have segmented, normalized and preprocessed these images with the same techniques as in Section 7.1.

### 7.2.1 Off-angle quality measure

#### Description of the measure

To detect off-angle images we have exploited the positions of specular reflections in the iris image similarly to what has been done by Li et al. in [53]. More specifically, we measure the distance between the main specular reflection and the center of the pupil that has been determined at the segmentation step. If this distance is higher than the pupil radius, then chances are high that the considered iris image is off-angle.

To detect the principal specular reflection, we have considered that it could be characterized by the following elements:

- a high brightness
- a significant magnitude gradient at the contour line
- an elliptical shape

We have exploited a region map elaborated by Lefèvre in his thesis. For each region:

- we apply a boundary detection

**CONFIDENTIAL**

- we use a two dimensional principal component analysis (PCA) on the contour line. The two obtained eigenvalues  $\lambda_1$  and  $\lambda_2$  give an indication on the distribution of the points from the contour line. If  $\lambda_1 \approx \lambda_2$  and the two eigenvalues are large enough then the spot has a circular shape.
- We compute the mean intensity of the pixels inside the contour line.

This allows us to position the principal specular reflection.

Following this, we can compute the distance  $d$  between the pupil center and the principal specular reflection. To decide if an iris image is off-angle or not we compute the score  $s$  given by formula 7.2, where  $r$  is the pupil radius and  $k$  is a constant value chosen for normalizing  $s$  so that it is comprised between 0 and 1.

$$s = k \frac{d}{r} \quad (7.2)$$

The closer the score  $s$  is to 0 the less chances there are that the considered image is off-angle.

### Performance

To evaluate the pertinence of the off-angle detection score, we have manually labeled 100 off-angle and 1000 non-off-angle images and calculated the score defined in formula 7.2. We compare these scores  $s$  to a threshold value  $th$  that can vary from 0 to 1.

- If  $s \geq th$  then we decide that the image is off-angle
- If  $s < th$  then we decide that the image is not off-angle

For each value of the threshold  $th$  we can compute:

- the false positive rate that represents the ratio of the number of images that are decided to be off-angle, when in fact they are not, on the total number of non-off-angle images.
- the false negative rate that represents the ratio of the number of images that are decided to be non off-angle, when in fact they are off-angle, on the total number of off-angle images.

Figure 7.1 represents the corresponding ROC curve depicting the False Negative Rate (FNR) as a function of the False Positive Rate (FPR).

It shows that the Equal Error Rate (EER) is given by 6% which means that there is a value of the threshold  $th$  for which the FPR and the FNR are both equal to 6%. This error rate is still pretty high, however as we will see in the next sub-section, we do not need the off-angle/non-off-angle classification to be more precise than this.

### 7.2.2 Comparative matching performance on off-angle images

We wish to compare the performance of Daugman's and Krichen's algorithms when using our region selection strategy for two categories of iris image comparisons:

**CONFIDENTIAL**

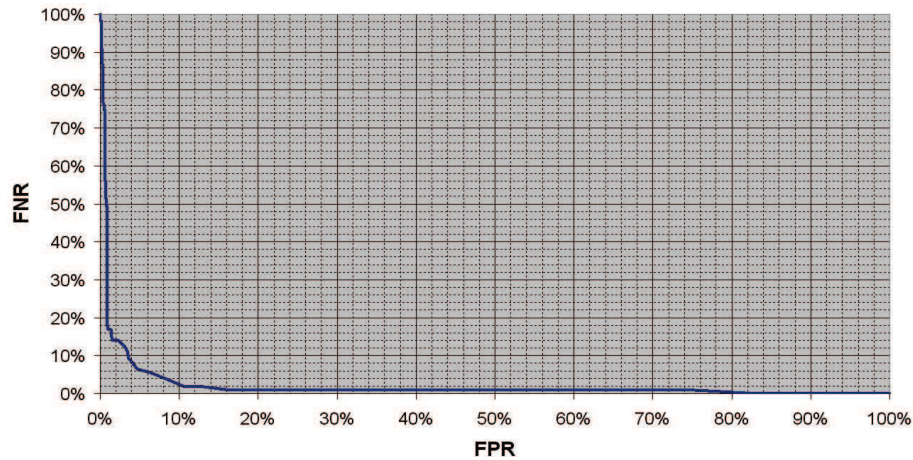


Figure 7.1: ROC curve for the detection of off-angle images using the score defined in formula 7.2.

- the comparisons for which at least one of the two images is off-angle
- the comparisons of two non-off-angle images.

To this end we need to split the ND-IRIS-0405 database in two categories: the off-angle and the non off-angle images. The score defined in section 7.2.1 gives us a good indication of the off-angle characteristic of an image, as shown by our previous experiment. We will therefore use it to split the database into these two categories, after setting a threshold value for this score.

We wish to minimize the False Positive Rate so that the chances that a non-off-angle image is sorted into the off-angle category are very low. To choose the value of the threshold  $th$  that separates the off-angle scores from the non-off-angle scores we use the same curve as the one in Figure 7.1 zoomed around the lowest values of the FPR. This way we can observe the curve's behavior more precisely and choose the threshold at a point where the curve presents a rupture. This is presented in Figure 7.2.

Given this curve, we choose to work with the threshold  $th$  for which  $FPR=1,5\%$  given a  $FNR=15\%$ . This being set, we have computed the off-angle score for all images of the ND-IRIS-0405 database and compared each score to the threshold to categorize the corresponding image. Then we built a set of intra-class and inter-class matching lists containing matches for which at least one image is off-angle. Finally, we have compared the performance of Daugman's and Krichen's algorithms when using our region selection strategy with  $N/M = 0.4$  on these matching lists. Figure 7.3 presents the comparative performance achieved.

It appears that both algorithms perform badly on off-angle images. However the algorithm exploiting Krichen's feature extraction and matching method performs significantly better than the one that uses Daugman's method.

The bad performance of both algorithms on off-angle images can firstly be explained by the fact that the number of aberrant segmentations is much higher in the category of off-angle images than in the category of non-off-angle images. Recall that we have shown in Section 5.3.2 that our local selection strategy is not

**CONFIDENTIAL**

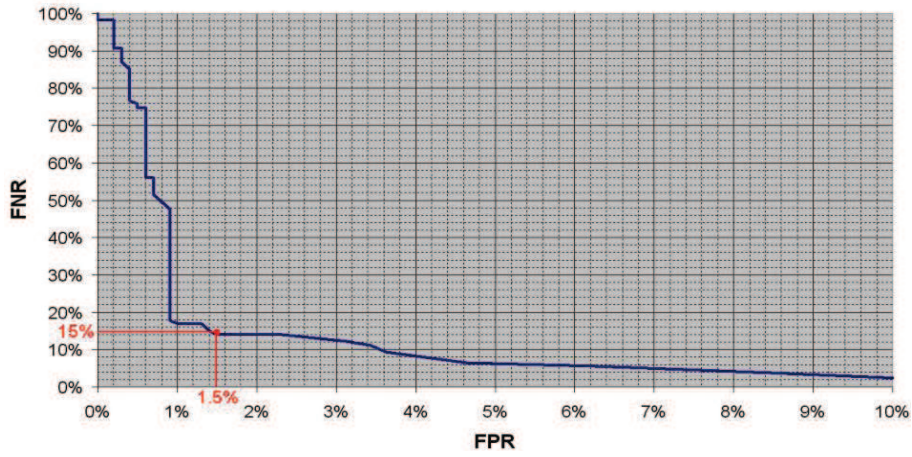


Figure 7.2: Zoomed ROC curve for the detection of off-angle images using the score defined in formula 7.2.

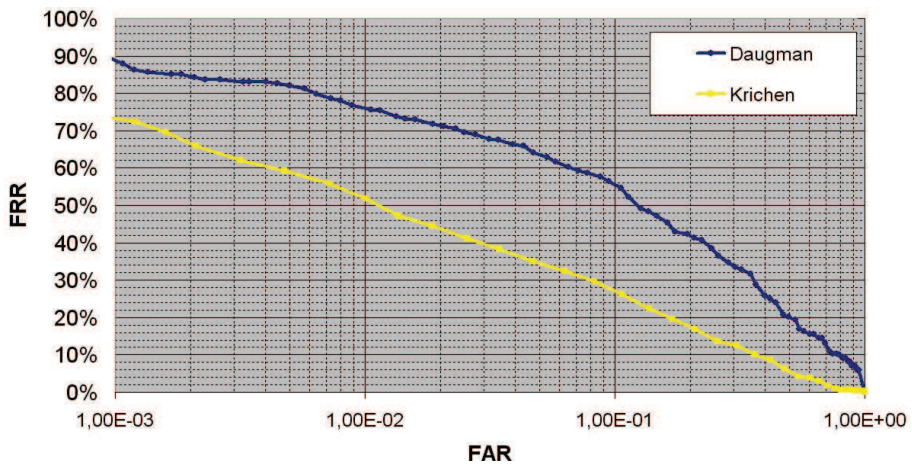


Figure 7.3: ROC curves for iris recognition algorithms using Krichen’s or Daugman’s feature extraction and matching technique and the quality based region selection strategy with  $N/M=0.4$  on off-angle images.



able to compensate for such big segmentation errors. Secondly, off-angle images for which the segmentation is as accurate as possible present strong distortions in the normalized image. This is because the segmentation module we use models these boundaries of the iris by circles. In the case of off-angle images these boundaries are not circular but elliptic so that they are necessarily inaccurately modeled. This generates strong distortions in the normalized images.

Figure 7.4 presents an example of an off-angle and a non-off-angle iris image of the same subject and the associated normalized images in the case where the segmentation is pretty accurate. Note that the normalized image corresponding to the off-angle image contains distortions compared to the other one.

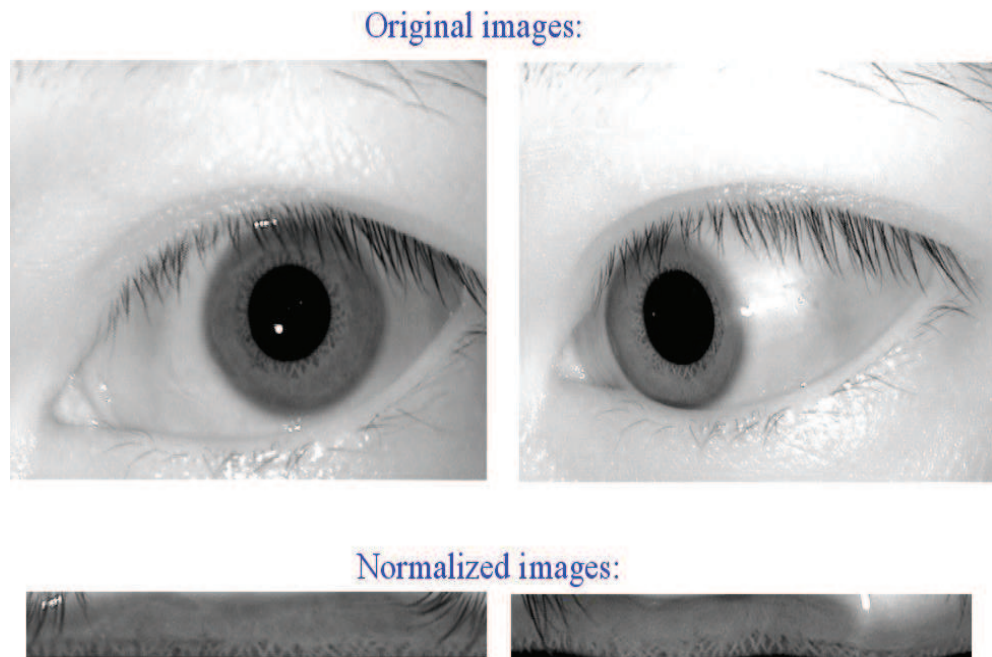


Figure 7.4: Example of an off-angle and a non-off-angle iris image of the same subject and the associated normalized images

Since we have shown in Figure 7.3 that the algorithm based on Krichen's method performs better on the off-angle category than the algorithm based on Daugman's method and that the images in the off-angle category contain higher distortions, we can conclude once more that the algorithm based on Krichen's method is more robust to distortions.

### 7.3 Final strategy proposed for iris recognition

We have demonstrated that using the local GMM-based quality measure proposed in Chapter 4 to select the regions of the normalized iris images that are used for feature extraction and matching, improves recognition performance, whether the extraction and matching techniques used are Daugman's or Krichen's. Furthermore, we have shown that the global quality measure defined in Chapter

4 can be used to set the proportion of the normalized iris images that should be exploited, so as to maximize performance for each image couple to compare. This global quality measure can also be used to examine which technique, out of Daugman's or Krichen's, is the best adapted to each couple in terms of recognition accuracy and speed. A score that measures the chances that a given image is off-angle can also be used to this end.

Based on these results, we propose in this chapter a complete iris recognition strategy that combines Krichen's and Daugman's algorithms as well as the local and global use of our GMM-based quality measure. This strategy also exploits the off-angle score we have defined in Section 7.2.

### 7.3.1 Method description

Our strategy follows the segmentation and the normalization of the iris images presented in Chapter 3, as well as the pre-processing of the normalized iris images described in Chapter 4, to bring the image's mean intensity to a constant value. Its objective is to compare two normalized and pre-processed iris images: a gallery and a probe image. Note that we have exploited the images before normalization to compute the off-angle score  $Q_{off-angle}$  associated to each image as described in Section 7.2. Our strategy can be cut into 7 steps as illustrated in Figure 7.5.

- **Step 1:**  $M$  points are selected in each one of the two images to compare. These points are uniformly distributed in each image and are identically localized in the gallery and in the probe image.
- **Step 2:** For each iris image (gallery or probe), we consider the  $M$  sub-images  $w_i$  centered around the  $M$  points  $pt_i$  selected in Step 1 and compute the associated local quality measures according to formula 7.3.

$$Q(w) = \exp\left(-\frac{1}{ab} \sum_{i=1}^{ab} |\log(P(x_i|\lambda)) - \overline{LL}_{train}|\right) \quad (7.3)$$

In this formula  $ab$  is the size of the sub-image  $w_i$ ,  $x_i$  is the input vector of our GMM associated to a pixel  $i$ ,  $P(x_i|\lambda)$  is the likelihood given by our GMM to this input vector and  $\overline{LL}_{train}$  is the mean log-likelihood on the training set.

As a result, we obtain  $M$  gallery local quality values  $Q^{gal}(w_i)$  and  $M$  probe local quality values  $Q^{probe}(w_i)$ .

- **Step 3:** For each value of  $i \in [1; M]$ , we fuse the local gallery quality value  $Q^{gal}(w_i)$  with the local probe quality value  $Q^{probe}(w_i)$ . The fusion operator selects the minimum of the two values. This gives us  $M$  fused local quality values:  $Q^{fus}(w_i) = \min(Q^{gal}(w_i), Q^{probe}(w_i))$ .
- **Step 4:** This step combines

- \* the sorting of the  $M$  fused local qualities. The lowest rank is given to the highest quality value. Since each local quality value can be associated to a point  $pt_i$ , this sorting attributes a rank to each point:  $rank(pt_i)$ .

**CONFIDENTIAL**

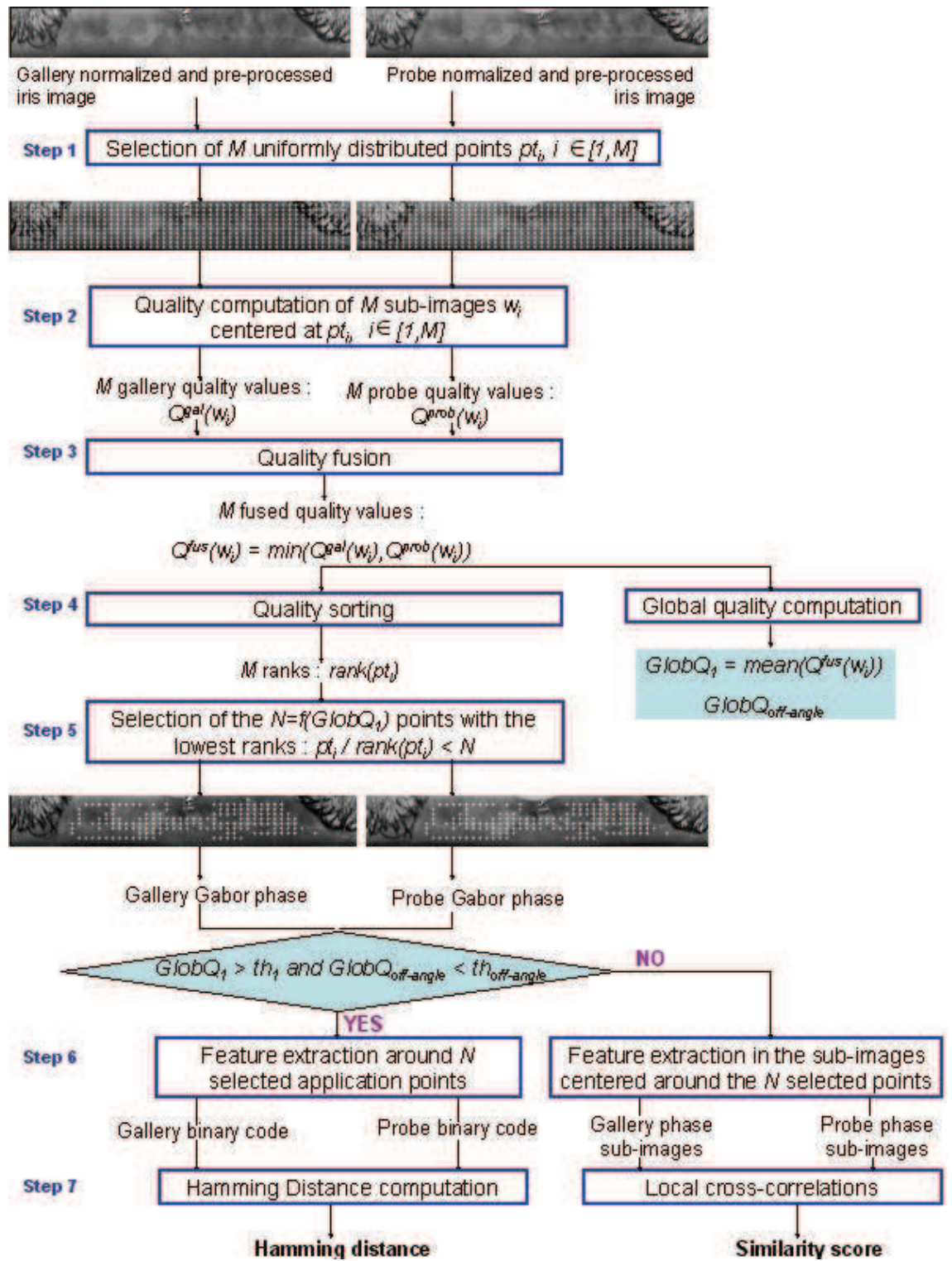


Figure 7.5: Proposed strategy for iris recognition combining Krichen's and Daugman's algorithms as well as local and global quality measures.

CONFIDENTIAL



- \* the computation of the global GMM-based quality score and the global off-angle quality score according to formula 7.4 and 7.5 respectively.

$$GlobQ_1 = \text{mean}_{i \in [1;M]}(\min(Q^{gal}(w_i), Q^{probe}(w_i))) \quad (7.4)$$

$$GlobQ_{off-angle} = \max(Q_{off-angle}(img_{gal}), Q_{off-angle}(img_{prob})) \quad (7.5)$$

Recall that the closer  $Q_{off-angle}(image)$  is to 1, the higher are the chances are that the image is off-angle. Consequently,  $GlobQ_{off-angle}$  increases with the chances that at least one of the images to compare is off-angle.

- **Step 5:** The  $N$  points with the lowest ranks are selected.  $N$  is a function of  $GlobQ_1$ .
- **Step 6:** Depending on the values of  $GlobQ_1$  and  $GlobQ_{off-angle}$ , one out the two following feature extractions is performed
  - \* if  $GlobQ_1 > th_1$  and  $GlobQ_{off-angle} < th_{off-angle}$ , convolutions with  $N_f$  2D Gabor filters are performed around the  $N$  selected points  $pt_i$  in the gallery and in the probe images and the resulting Gabor phases are quantized. So each image is represented by a binary code of length  $2 \times N_f \times N$ .
  - \* if  $GlobQ_1 \leq th_1$  or  $GlobQ_{off-angle} \geq th_{off-angle}$ , convolutions with 2D Gabor filters are performed around all the pixels in the sub-images centered around the  $N$  selected points. So each image is represented by  $N_f \times N$  phase sub-images.
- **Step 7:** Depending on the values of  $GlobQ_1$  and  $GlobQ_{off-angle}$ , one out of the two following matching techniques is performed
  - \* if  $GlobQ_1 > th_1$  and  $GlobQ_{off-angle} < th_{off-angle}$  the Hamming distance between the gallery and the probe binary codes is computed.
  - \* if  $GlobQ_1 \leq th_1$  or  $GlobQ_{off-angle} \geq th_{off-angle}$  cross-correlations are performed between the phase sub-images resulting in the computation of a similarity score.

Consequently, either the value of the Hamming distance or the value of the similarity score will be used to decide if the gallery and probe iris images match or not.

### 7.3.2 Experiments and results

The aim of this section is to test this strategy on iris image comparisons. To this end, we need to set the values of the strategy's parameters namely:

- the function  $f$  that relates  $N$  to  $GlobQ_1$  for a given value of  $M$ .
- the value of the threshold  $th_1$  to which we compare  $GlobQ_1$  for each couple of images

**CONFIDENTIAL**

- the value of the threshold  $th_{off-angle}$  to which we compare  $GlobQ_{off-angle}$  for each couple of images

To do this, we have used the results established in the previous Chapters.

In Chapters 5 and 6, we have found that the optimal value of  $N/M$  for each of the five categories of images generated according to the value of  $GlobQ_1$  did not depend on the database (ND-IRIS-0405 or CASIA-IrisV3-Lamp) or on the technique used for feature extraction and matching (Daugman's or Krichen's). Table 7.3 recalls the best values of  $N/M$  obtained for each category of images.

Table 7.3: Best values for  $N/M$  for different values of  $GlobQ_1$

$GlobQ_1$	[0; 0.13]	[0.13; 0.26]	[0.26; 0.39]	[0.39; 0.52]	[0.52; 0.65]
$N_{best}/M$	0.4	0.4	0.5	0.6	0.7

Consequently, we have chosen to define  $f$  empirically as follows:

- if  $GlobQ_1 \in [0; 0.26]$ ,  $f : GlobQ_1 \mapsto N = 0.4 \times M$
- if  $GlobQ_1 \in [0.26; 0.39]$ ,  $f : GlobQ_1 \mapsto N = 0.5 \times M$
- if  $GlobQ_1 \in [0.39; 0.52]$ ,  $f : GlobQ_1 \mapsto N = 0.6 \times M$
- if  $GlobQ_1 \in [0.52; 0.65]$ ,  $f : GlobQ_1 \mapsto N = 0.7 \times M$

Moreover, we have demonstrated in Section 7.1 (Table 7.1) that using Krichen's technique for feature extraction and matching improves recognition performance on image couples for which  $GlobQ_1 \in [0; 0.39]$ . For the other image couples the accuracy is the same whether feature extraction and matching is done according to Daugman's or Krichen's technique. Since Krichen's algorithm is more than ten times slower than Daugman's, we chose to use it only for the image pairs for which Krichen's algorithm performs better. Consequently, we have set the value of the threshold  $th_1$  to 0.4.

In Section 7.2, we have plotted the ROC curve for the detection of off-angle images using the off-angle score we have defined (Formula 7.2) and chosen  $th_{off-angle}$  so that the FPR=1,5% and the FNR=15%. For our final matching strategy, we choose to keep the same value for  $th_{off-angle}$ .

Figure 7.6 and 7.7 present the ROC curves obtained when applying the iris recognition process presented in Figure 7.5 to comparisons from the ND-IRIS-0405 database. More specifically, 300 000 inter-class and 10 000 intra-class comparisons were performed. Note that the images used for these comparisons are different ones than those used in the previous chapters to set the values of  $f(GlobQ)$ . To plot these ROC curves we have separated the comparisons into two categories:

- bad quality images: images for which  $GlobQ_1 < 0.4$  or  $GlobQ_{off-angle} > th_{off-angle}$
- good quality images: images for which  $GlobQ_1 \geq 0.4$  and  $GlobQ_{off-angle} \leq th_{off-angle}$ .

**CONFIDENTIAL**

We have compared our strategy to two other algorithms:

- **Krichen04**: using the local quality-based region selection strategy with the same value of the  $N/M$  parameter for all images :  $N/M = 0.4$  and Krichen’s feature extraction and matching technique
- **Daugman06**: using the local quality-based region selection strategy with the same value of the  $N/M$  parameter for all images :  $N/M = 0.6$  and Daugman’s feature extraction and matching technique

For each of these two algorithms, the constant value of  $N/M$  for all comparisons has been chosen as the one that maximizes performance for the images for which  $GlobQ_1 < 0.4$  or  $GlobQ_1 \geq 0.4$ . More precisely:

- The algorithm that uses Krichen’s feature extraction and matching technique as well as our region selection strategy performs best on image couples for which  $GlobQ_1 < 0.4$  when  $N/M = 0.4$ .
- The algorithm that uses Daugman’s feature extraction and matching technique as well as our region selection strategy performs best on image couples for which  $GlobQ_1 \geq 0.4$  when  $N/M = 0.6$ .

The figure shows that the best performance is achieved with the method we propose in Figure 7.5, whether the images belong to the category “good quality” or “bad quality”. When they belong to the “bad quality” category, the performance is very close to the one achieved with Krichen04, but it outperforms Krichen04 when they belong to the category “good quality”. On the contrary, the performance of our method is very close to the one of Daugman06 on the “good quality” images but it exceeds Daugman06 performance on the “bad quality” images.

Since we showed in Sections 7.1 and 7.2 that Krichen’s algorithm performs at least as well as Daugman’s for all image categories, we could have chosen to use this algorithm for all image couples, all in adapting  $N/M$  to each couple. However, Krichen’s algorithm is more than ten times slower than Daugman’s and Daugman’s performs as well as Krichen’s in terms of accuracy for the matches for which  $GlobQ_1 \geq 0.4$  (40% of the image pairs considered). Consequently, using only Krichen’s algorithm would have led to multiplying the computation time of the entire matching list by at least 4.6.

## 7.4 Conclusion

In this chapter we compared the performance of Daugman’s and Krichen’s feature extraction and matching technique when combined with our region selection strategy. We showed that their performances in terms of accuracy are the same for image pairs for which  $GlobQ_1 \geq 0.4$ . However, Krichen’s technique outperforms Daugman’s for off-angle images or for images that verify  $GlobQ_1 < 0.4$ . We explained this result by the fact that Krichen’s algorithm is more robust to a decrease of the usable iris area and the richness of its texture (measured by  $GlobQ_1$ ) since it extracts more information from the normalized iris image. In addition, Krichen’s algorithm is more robust to local distortions, very common in off-angle images, since it allows local translations between the normalized

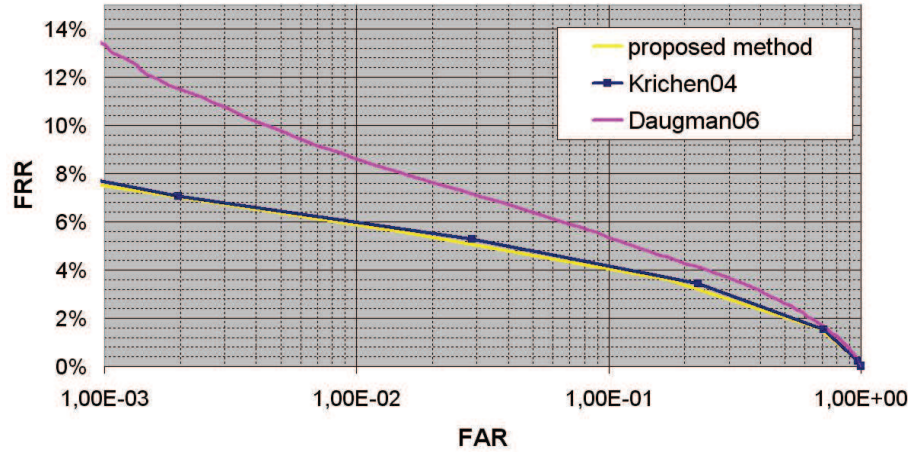


Figure 7.6: Comparative ROC curves for image couples for which  $GlobQ_1 < 0.4$  or  $GlobQ_{off-angle} > th_{off-angle}$ .

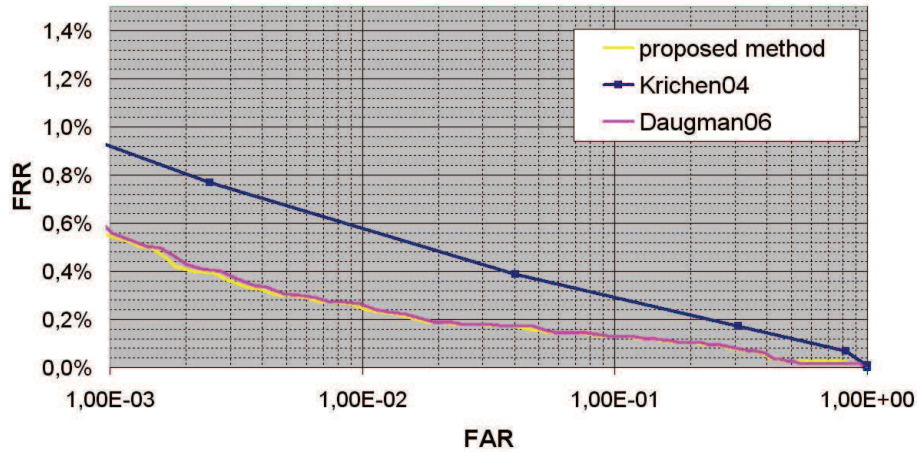


Figure 7.7: Comparative ROC curves for image couples for which  $GlobQ_1 \geq 0.4$  or  $GlobQ_{off-angle} \leq th_{off-angle}$ .

CONFIDENTIAL

images along the radial and angular directions while Daugman's algorithm only allows global angular translations.

Finally, Daugman's algorithm being more than ten times faster than Krichen's, we proposed a comprehensive strategy for feature extraction and matching that uses Krichen's algorithm only for image couples for which we have shown that it will perform significantly better than Daugman's. This way we optimize performance in terms of accuracy and speed. Furthermore, our strategy adapts the proportion of the normalized iris image that is used to the global quality of the considered image pair, which leads to a recognition accuracy improvement.



## Chapter 8

# Conclusions and discussion

In this thesis we have proposed methods to improve the robustness of iris recognition systems to degradations of the quality of the input images. We have worked only on the feature extraction and matching stages of the global system. Our input images are iris images that have previously been segmented and normalized.

The original acquired images can present degradations such as blur, occlusions or an off-angle gaze. These degradations may have repercussions on the segmentation and normalization of the original images and lead to normalized iris images that are affected by distortions or contain non-detected artifacts. Moreover, even if the segmentation and normalization are successful in spite of degradations in the input image, the feature extraction and matching stages of iris recognition systems face other difficulties caused by these degradations. For instance, the amount of useful information available in highly occluded iris images is reduced, even when the occlusions have been correctly located, which makes image comparison less accurate.

In this thesis, we have proposed a strategy to compare normalized iris images that improves the way degradations, such as the presence of undetected artifacts or distortions in normalized iris images, are handled. It therefore presents a certain robustness to segmentation inaccuracy. Our strategy also deals with variations of the amount of information available in the normalized iris images for comparison.

The work we have presented in this thesis, aimed at improving the techniques of comparing two normalized iris images, is original on many levels:

First of all, we do not use any information resulting from segmentation such as a binary mask that would mask occlusions. Instead we use our own local and GMM-based quality measure to detect the sub-images in the normalized iris images that contain artifacts. The Gaussian Mixture Model (GMM) we use has been trained on sub-images that are free from any artifact and highly textured. The quality measure we propose allocates values between 0 and 1. We showed in Chapter 4 that the lowest values are given to sub-images that contain artifacts, the highest ones to sub-images that are artifact free and highly textured, and intermediated values are allocated to sub-images that are artifact free but poorly textured. Consequently this quality measure can be exploited to avoid



taking into account regions containing artifacts.

Secondly, to select the regions in the normalized iris images that are exploited for feature extraction and matching, we privilege the most highly textured regions among the artifact free ones, in addition to avoiding regions containing artifacts. This is also done thanks to our local and GMM-based quality measure, since it allocates higher values to the sub-images that are highly textured among the free-from-artifact ones. We have demonstrated, in Chapter 4, that the higher the quality of sub-images according to our measure, the more discriminatory is the information these sub-images bring for matching. This confirms the purpose of using our quality measure for selecting the regions in the normalized iris images that are exploited for feature extraction and matching.

In practical terms we compute the local qualities for sub-images identically located in the two normalized iris images to compare, and keep the minimum value for each couple of corresponding sub-images. Then, we sort the sub-image couples according to their minimum quality value in order to select the couples with the highest quality for feature extraction and matching. This was presented in Chapters 5 and 6.

The third original point of our strategy is that it controls the amount of information taken into account for matching by choosing the proportion of the normalized iris image selected for feature extraction. We have demonstrated, in Chapter 5 and 6, that there is a value of this proportion that maximizes recognition performance. Indeed, reducing the exploited proportion of each normalized iris image, improves the similarity scores corresponding to intra-class matches and degrades the similarity scores of the inter-class matches at a different speed. As long as the regions that are eliminated have a low quality, the improvement of the intra-class scores will be significant and will highly compensate the deterioration of the inter-class scores, leading to a performance enhancement. But once the quality of the eliminated regions is good enough, the deterioration of the inter-class scores gains the upper-hand and the performance is impaired. From this, it is easily understood that the optimal proportion of the normalized iris image to keep for feature extraction and matching will vary for each image couple to compare and depend on the global image quality.

Fourthly, we have associated a global quality measure to each couple of images to compare. This measure was defined in Chapter 4 and exploits the local and GMM-based quality measure we have defined to evaluate the quality of the image couple in terms of amount of occlusions and amount of texture. It is computed by considering sub-images identically located in the two images to compare and measuring the associated local qualities. We then keep the minimum value for each couple of corresponding sub-images and finally calculate the mean of these values. The advantage of this global quality measure is that it indicates where the good quality regions of the images overlap since it is computed from fused local qualities instead of two global quality values corresponding to each image separately.

We have demonstrated, in Chapters 5 and 6, that there is a relation between the optimal proportion of the normalized images to keep for feature extraction and matching, and the global quality value associated to the image couple, which confirms what we have established in the previous paragraph.

**CONFIDENTIAL**

A fifth point is that we have used two different techniques for feature extraction and matching. The first one was proposed by Daugman [21]. It applies 2D Gabor filters around application points in the normalized iris images and quantizes the resulting Gabor phases so that each iris is represented by a binary code. Two irises are then compared with the computation of the Hamming distance between the binary codes. The second one was proposed by Krichen et al. [48]. It applies 2D Gabor filters on the normalized iris images in the same way as Daugman's algorithm; however the real values of the Gabor phases are kept for matching. More precisely, local cross-correlations are performed between the phase images leading to the computation of a similarity score.

We have used both of these techniques to validate the results presented on our region selection strategy in the normalized iris images. This proves that these results are not just verified for one specific type of feature extraction and matching, but can be generalized.

Sixthly, we have demonstrated in Chapter 7 that the benefit of Krichen's technique compared to Daugman's depends on the global quality of the images to compare. Indeed Krichen's technique is more robust to local distortions in the normalized images as well as a reduction of the amount of texture. These elements can be measured by the global quality measure we have defined. Furthermore, they are very common in off-angle images. Consequently, considering both our global quality measure and an off-angle quality score can indicate us which technique out of Daugman's or Krichen should be used in order to compromise between accuracy and speed for recognition.

Finally we have proposed in Chapter 7 an automatic system for comparing two normalized iris images which exploits the results presented here above. It combines our local and global quality measures, as well as Daugman's and Krichen's feature extraction and matching techniques. It computes the local and global image qualities for each image couple to compare and uses these values to:

- decide what proportion of the images should be used for feature extraction and matching
- select the best regions to exploit for feature extraction and matching
- choose between Krichen's and Daugman's techniques

The work done in this thesis opens up to many perspectives and directions for future work:

First of all, we could incorporate other quality metrics in our iris recognition system that would measure different image quality characteristics than the ones considered in this thesis. Locally, the quality metric we use measures the amount of artifact and of texture in images. Other local quality metrics could be defined to measure distortion levels or entropy, for example. Globally, we exploit two quality measures: one characterizes the images in terms of occlusion and amount of texture, while the other one detects off-angle irises. Many other

**CONFIDENTIAL**

global criteria can be measured in iris images, such as pupil dilation, iris/sclera contrast, resolution, or the presence of interlacing or contact lenses.

Many of these local and global quality metrics could probably be exploited for iris recognition in the same way as we have used our metrics to improve recognition performance. They could be used for selecting the best regions in normalized images, for deciding what proportion of these images should be exploited for feature extraction and matching, and they could be used for choosing between Krichen's and Daugman's algorithms for each couple of iris images to compare.

Moreover we could compare the performance of our quality metrics locally and globally to the ones of these other quality metrics. Our local quality metric has not been elaborated to detect any specific types of artifact. Consequently, it might be less precise than others that would have been designed to this end, but is probably easier to use.

Secondly, there would be many other ways of exploiting local and global quality measures than the three presented here above. Other parameters than the proportion of normalized iris images exploited for feature extraction and matching could be set for each image couple to compare according to global quality criterion. For instance, we could have chosen to adapt the filters used for feature extraction to the quality of each couple. Moreover we could extend the techniques used for feature extraction and matching beyond Daugman's and Krichen's and determine the best technique for a couple of images according to various quality criteria.

A third point to work on would be to increase the interactions between the segmentation/normalizations modules and the feature extraction/matching modules. The segmentation and normalization modules could also use global quality metrics to adapt their parameters to each image. Moreover, the segmentation module could evaluate the accuracy of the performed segmentation and the following modules could exploit this information.

Fourthly, it would be interesting to model mathematically the intra-class and inter-class distributions, as well as the way they evolve, when the amount of information considered for matching changes. This could be done for Daugman's matching technique and Krichen's. This way we could demonstrate what amount of information optimizes recognition performance in theory and compare this result to the one obtained in this thesis.

Finally, the work presented in this document will need to be followed by an industrialization process before being deployed in operational systems. Our feature extraction and matching techniques will need to be tested with the latest versions of Lefevre's segmentation and normalization algorithms. Moreover, the processing time of the entire recognition system will need to be optimized. Work will also have to be done on reducing the storage volume of the data.

**CONFIDENTIAL**

# Appendix A: “Improving normalization circle detection for iris recognition using a variational approach”

[51] T. Lefèvre. Improving normalization circle detection for iris recognition using a variational approach. Technical report, Institut Mines-Telecom, Telecom SudParis, 2012.

# Improving Normalization Circle Detection for Iris Recognition Using a Variational Approach

Thierry Lefevre

thierry.lefevre@it-sudparis.com

## Abstract

*In this paper, we present a new approach for refining the location of the normalization circles in iris recognition, taking benefit of an accurate segmentation stage. This process which relies on Parametric Deformable Models (PDM) is composed of three-steps : first, a Circular Hough Transform (CHT) generates a set of initial circles. Then, a variational optimization drives the circles towards the exact iris borders. Finally, the best circles for normalization are selected based on the results of the optimization.*

*Extensive experiments are conducted on three widely-used and challenging databases (ICE 2005, Casia-Interval and a subset of 7047 images of ND-iris) for which we dispose of a manual segmentation. The results show the interest of our multi step algorithm and an overall performance in the range of the State of the Art.*

## 1. Introduction

Among the various biometric modalities available, iris is considered, since the earliest works of Daugman [2] and Wildes [8], as a very powerful modality for identifying people. In practice, this statement has been validated on high quality images but many degradations, such as blur, occultation by eyelids or eyedrops can severely degrade the performance of such systems, spurring interest for further studies.

In an iris recognition system, the normalization stage is designed to make the texture of the iris independent of the acquisition conditions. The most widespread choice is the "rubber sheet" model introduced by Daugman in [2]. Iris borders are modeled as two non-concentric circles and the iris texture is unwrapped into a rectangle regarding those circles. Modeling iris borders by circles is still the most common choice in the literature as it allows a simple geometric transformation. However, this model does not stand for off-angle images or for images with large dilatation of the pupil. For those reasons, recent works tend to use new contour models such as ellipses [10, ?] or Fourier coefficients [1] to unwrap the iris. In this article, however we will stick to circular normalization circles for simplicity purpose and give some foreseen extensions in the following.

The ongoing research drives the following observations : on one hand it is known that the precision of the normalization circle is critical for the performance of the global system as shown in [5]. On the other hand improving the segmentation stage is the object of many attentions in the current iris literature [10, 1, ?, ?, ?]), but the way those precise segmentations can improve the normalization stage is not clearly settled.

Indeed, a survey of recent literature shows different behaviors: on one hand, in some papers [?, ?] the segmentation is initialized using a simple approximation of the iris borders (circles or ellipses) and the normalization is done using those initial contours without any improvements. On the other hand, some authors [6, ?] take benefit of the accurate contour provided by the segmentation stage in order to improve the precision of the normalization contour, even if they still use a circle or an ellipse to this end.

The present article follows this last tendency. We indeed propose a novel approach allowing to optimize the position of a normalization circle starting from an initial estimation of the circular contours and a precise segmentation. Let us note that the presented approach could be extended to elliptic curves.

Our main originality is to consider the circles as Parametric Deformable Models on which we apply a variational optimization. Note that as variational approaches are highly sensitive to the initialization, we consider several contours during the process to get a better exploration of the circle parameters space.

Due to a lack of space we will not describe our segmentation and extraction stages; our segmentation is based on snakes in the same way as [1]. The extraction and matching stages are also very standard and follow the approach in [2].

This paper is organized as follows: Section 2 presents the different steps of our novel algorithm for finding the normalization circles. Section 3 validates the interest of our three-steps algorithm on some challenging databases (ICE2005, Casia-Interval and ND-iris). Finally, Section 4 concludes this article.

## 2. Accurate Normalization Circles Search

### 2.1. Overall Algorithm

Two historical approaches have been proposed for circle normalization. In [2] Daugman use an integrodifferential operator to find the normalization circles. This criteria is known to be precise but to suffer from a lack of robustness. On the other hand, Wildes [8] propose to apply a Canny edge detector followed by a CHT (Circular Hough Transform). This method is very robust, but the precision is directly affected by the performance of the edge detector. Note that both methods can be very time consuming if no optimization is done.

Some authors have tried to overcome the limitation of those methods. Li et al. in [3] use a CHT to generate a set of circles and select the best one using Daugman's criterion. Tan et al. in [7] use an integrodifferential constellation in order to improve the robustness of Daugman's operator.

Our article follow the ideas of [3] but we find out that adding an optimization stage after the CHT significantly improve the final circle precision. In practice, our novel algorithm corresponds to three consecutive steps :

- The algorithm starts with a CHT to get a set of initial circles close to the expected circle (Figure 1b).
- A variational optimization on the set of initial circles is performed to guide them towards the exact borders of the iris (Section 2.2 ; Figure 1c).
- The algorithm selects the final circle from the set of optimized circles (Figure 1d).

Figure 1 illustrates the three steps of our algorithm. Starting from the segmentation results in Figure 1a (here a Snake algorithm) we generate the set of initial circles displayed in Figure 1b. Using the segmentation results instead of an edge detector on the whole image is important to insure that most of the initial circles are close to the expected circle. Also note that CHT can be sub-sampled to speed up the process as this is an initialization step. Consequently none of the circles of Figure 1b are at the expected location.

Those initial circles of Figure 1b are optimized using our variational scheme presented in detail in Section 2.2 and in this case reach two local minimums displayed in Figure 1c. The importance of taking several circles during the initialization step is outlined at this stage as the circles can reach different local minimum depending of their initial position.

Finally the circle reaching the lower minimum during the optimization is selected as displayed in Figure 1c.

By applying those three steps one after the other, the algorithm benefits on one hand of the robustness of the CHT in getting close to the optimum and, on the other hand, of the preciseness of an energy-based method to build and select the final circles.

### 2.2. Variational optimization

In this work we consider circles as Parametric Deformable Models (PDM) [?] on which we apply a variational optimization. PDM are a type of Active Contours (AC) described by parametric equations and the aim of the method is to express the variational optimization on the parameters of the equation instead of the curve itself.

The advantage of such a method is that the contours have a defined parametric formulation which is convenient to proceed the "rubber sheet" normalization. We focus this article on circles, but it is clear that using other parametric templates such as ellipses is a straightforward extension. PDM are very classical in Computer Vision but are almost absent from iris recognition literature. Only reference being Miyazawa et al. in [4] where a model with 10 parameters is used to find simultaneously inner and outer iris circles and both eyelids.

In this work we adapt classical AC energy for circular contours. In AC models, the objective energy is usually composed of three terms : an edge term, a region term and a regularization term. The edge term is defined in order to promote areas of strong gradient. The region term aims at separating as well as possible regions inside and outside the contour based on their statistical descriptions. The regularization term ensures that the contour keeps a regular shape. As our energy is defined over a fixed shape model (a circle), this term is not pertinent. Thus, the global energy can be written as :

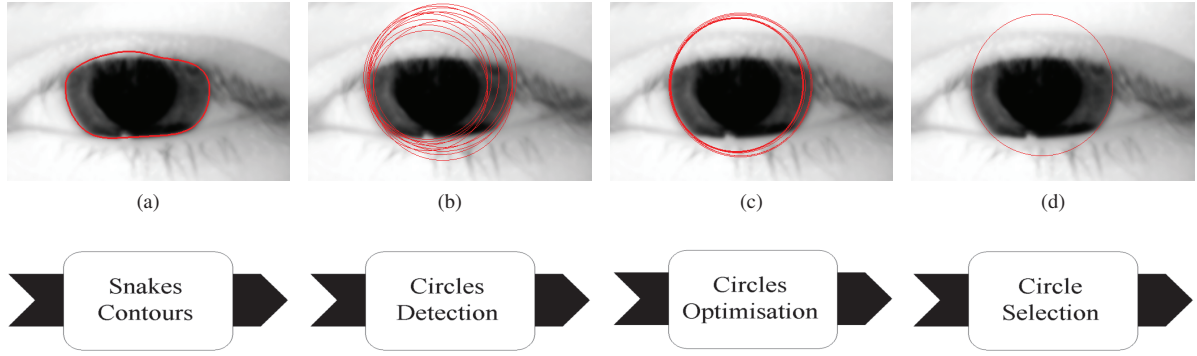


Figure 1: Work flow of our algorithm : Snakes contours from the segmentation stage (a). Set of circle generated using a Circular Hough Transform (b). Optimized circles (c). Best circle selected for Normalization (d).

$$E_g(\mathcal{C}) = E_e(\mathcal{C}) + E_r(\mathcal{C}), \quad (1)$$

for a circle  $\mathcal{C}(\mathbf{x}_c, r_c)$  with center  $\mathbf{x}_c$  and radius  $r_c$  modeled as :

$$\begin{aligned} x &= x_c + r_c \cos(\theta), \\ y &= y_c + r_c \sin(\theta), \end{aligned} \quad (2)$$

where  $\theta \in [0; 2\pi[$ .

A good candidate for  $E_e$  would have been Daugman's operator, unfortunately it is not differentiable as outlined by Tan et al. in [7]. So it is not possible to apply a variational optimization on it. For this reason, we use a standard energy function which keeps the idea that the gradients of the image are important across the border of the circle:

$$E_e(\mathcal{C}) = - \int_0^{2\pi} \left\| \nabla I(\mathbf{x}(\theta)) \cdot \vec{n}_\theta \right\| d\theta, \quad (3)$$

where  $\vec{n}_\theta$  is the unit outward vector at point  $\mathbf{x}(\theta)$ .

The region term we use in equation (1) is classical in the AC literature and is known as region competition:

$$E_r(\mathcal{C}) = - \int_{R_{in}} \log p_{in}(I(\mathbf{x})) d\mathbf{x} - \int_{R_{out}} \log p_{out}(I(\mathbf{x})) d\mathbf{x}, \quad (4)$$

where  $p_{in}$  and  $p_{out}$  are respectively the probabilities of a given intensity to be inside circle  $\mathcal{C}$  or to be outside circle  $\mathcal{C}$  (we note  $R_{in}$  the region inside  $\mathcal{C}$  and we note  $R_{out}$  the region outside  $\mathcal{C}$ ).

Due to space limitation we will not give the derivatives of those equations, but derivatives of equation (3) are obtained by straightforward calculus and details for the calculation of derivatives of equation (4) can be found in [9].

Finally, the associated iterative process is :

$$\begin{aligned} r^{k+1} &= r^k - \frac{\partial}{\partial r} (E_e + E_r), \\ \mathbf{x}_c^{k+1} &= \mathbf{x}_c^k - \frac{\partial}{\partial \mathbf{x}_c} (E_e + E_r). \end{aligned} \quad (5)$$

### 3. Evaluation

We have done the evaluation on three databases for which we have a manual annotation for the circles : ICE 2005, Casia-Interval and a sub-set of 7047 images of ND-Iris.

We have considered the given scenarios to highlight the advantages of our method :

*ref.* The manually annotated circles are used for the normalization.



	ICE 2005		Interval		ND-Iris	
	$10^{-3}$	$10^{-4}$	$10^{-3}$	$10^{-4}$	$10^{-3}$	$10^{-4}$
FAR	2.2	4.4	1.4	3.3	5.0	8.5
<i>ref.</i> (%)	2.1	3.3	5.6	10.2	5.0	8.1
<i>Hough</i> (%)	2.4	3.5	5.4	8.7	4.8	7.6
<i>prop.</i> (%)	1.0	1.8	1.7	2.9	3.6	5.9

Table 1: Evaluation of the different scenarios

*Hough* The circle for normalization is generated using a standard CHT (only 1 circle).

*var.* The circle from *Hough* is optimized using our variational optimization (Section 2.2).

*prop.* The circle is generated using our full algorithm described Section 2 with 10 circles for the initialization.

Table 1 shows the performance of the four scenarios described above.

It is interesting to note that taking several circles during the initialization is an important part of our algorithm because taking only one circle (*var.*) gives roughly the same performance as the CHT (*Hough.*). This can be explained by the fact that taking several circles allow to explore a larger parameter space and avoid being stuck in irrelevant local minimums.

We also observe that our algorithm (*prop.*) reaches better performance than using manual annotations (*ref.*). This tends to confirm that manual annotation are subject to a bias induced by human operator.

Table 2 shows performance of our overall algorithm compared to some references of the literature for a rough benchmarking.

Comparisons with the results of Vasta et al. [?] are interesting because they use a segmentation algorithm close to ours (Geodesic Active Contours) and one of their two matching algorithm (Texture) is also very close to ours. However they do not optimize their normalization circles after their segmentation. Our results are better than theirs on ICE 2005 and lower on Casia-Interval. We assume that this is because our segmentation algorithm has some difficulties to handle the very specific specular reflexions on this database. Direct comparison with the SVM-approach of Vasta et al. is not fair. Indeed, in their system, the authors use in fact a fusion of two classifiers which improves significantly their results. We therefore can expect similar improvement if this approach were used in our system.

In [6] Shah et al. also use a GAC and a matching algorithm close to ours but the way they exploit the segmentation results to generate the normalization circles seems to be a weak point of their proposed algorithm. Consequently using an improved selection of the normalization circles, our algorithm performs better on Casia-Interval despite the segmentation problems described above.

We note that very few research results are evaluated on ND-Iris which is a challenging database. In Table 2 we have anyhow mentioned the performance of the commercial product VeryEye developed by the company Neuro Technology [?]. Those results can be considered as as an optimal solution.

	ICE 2005		Interval		ND-Iris	
	$10^{-3}$	$10^{-4}$	$10^{-3}$	$10^{-4}$	$10^{-3}$	$10^{-4}$
FAR	1.0	1.8	1.7	2.9	3.6	5.9
proposed	—	—	4.5	5.5	—	—
Shah et al. [6]	1.75	2.37	0.55	0.77	—	—
Vasta et al. [?] (texture only)	0.5	0.74	0.2	0.38	—	—
Vasta et al. [?] (with SVM)	0.05	0.12	—	—	1.1	1.8
VeryEye	—	—	—	—	—	—

Table 2: Performance comparison

## 4. Conclusion

We have proposed a novel and efficient way for finding the circles for iris normalization. This process is composed of three steps, whose importance has been assessed thanks to several experiments.

The main perspective we consider for our algorithm is an extension of the variational formulation for ellipses because they can handle off-angle images which is the main problem of our current algorithm. This extension is not straightforward as the formulation of variational equations become more difficult and even more sensitive to local minimums.

## References

- [1] J. Daugman. New Methods in Iris Recognition. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 37(5):1167–1175, September 2007. 1, 2
- [2] J. G. Daugman. High confidence visual recognition of persons by a test of statistical independence. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15:1148–1161, November 1993. 1, 2
- [3] P. Li, X. Liu, L. Xiao, and Q. Song. Robust and accurate iris segmentation in very noisy iris images. *Image and Vision Computing*, 28(2):246 – 253, 2010. Segmentation of Visible Wavelength Iris Images Captured At-a-distance and On-the-move. 2
- [4] K. Miyazawa, K. Ito, T. Aoki, K. Kobayashi, and H. Nakajima. An effective approach for iris recognition using phase-based image matching. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30:1741–1756, October 2008. 2
- [5] H. Proenca. Short communciation: Iris recognition: Analysis of the error rates regarding the accuracy of the segmentation stage. *Image Vision Comput.*, 28:202–206, January 2010. 1
- [6] S. Shah and A. Ross. Iris segmentation using geodesic active contours. *Trans. Info. For. Sec.*, 4:824–836, December 2009. 1, 4
- [7] T. Tan, Z. He, and Z. Sun. Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition. *Image and Vision Computing*, 28(2):223 – 230, 2010. Segmentation of Visible Wavelength Iris Images Captured At-a-distance and On-the-move. 2, 3
- [8] R. Wildes. Iris recognition: an emerging biometric technology. *Proceedings of the IEEE*, September 1997. 1, 2
- [9] S. C. Zhu and A. Yuille. Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 18:884–900, September 1996. 3
- [10] J. Zuo and N. A. Schmid. On a methodology for robust segmentation of nonideal iris images. *Trans. Sys. Man Cyber. Part B*, 40:703–718, June 2010. 1

# Appendix B: “How a local quality measure can help improving iris recognition”

S. Cremer, S. Garcia-Salicetti, B. Dorizzi and N. Lempérière. How a local quality measure can help improving iris recognition. *BIOSIG 2012*, to appear

# How a local quality measure can help improving iris recognition

Sandra Cremer, Sonia Garcia-Salicetti,  
Bernadette Dorizzi

Nadège Lempérière

Département EPH  
Institut Télécom, Télécom SudParis  
9 rue Charles Fourier  
91011 Evry, France  
sandra.cremer@it-sudparis.eu  
sonia.garcia@it-sudparis.eu  
bernadette.dorizzi@it-sudparis.eu

Thales Communications & Security  
20-22 rue Grange Dame Rose  
78141 Vélizy-Villacoublay, France  
nadege.lempriere@thalesgroup.com

**Abstract:** The most common iris recognition systems extract features from the iris after segmentation and normalization steps. In this paper, we propose a new strategy to select the regions of normalized iris images that will be used for feature extraction. It consists in sorting different sub-images of the normalized images according to a GMM-based local quality measure we have elaborated and selecting the  $N$  best sub-images for feature extraction. The proportion of the initial image that is kept for feature extraction has been optimized in order to compromise between minimizing the amount of noise taken into account for feature extraction and maximizing the amount of information available for matching. By proceeding this way, we privilege the regions for which our quality measure gives the highest values, namely regions of the iris that are highly textured and free from occlusion, and minimize the risks of extracting features in occluded regions to which our quality measure gives the lowest values. We also control the amount of information we use for matching by including, if necessary, regions that are given intermediate values by our quality measure and are free from occlusion but barely textured. Experiments were performed on three different databases: ND-IRIS-0405, Casia-IrisV3-Interval and Casia-IrisV3-Twins, and show a significant improvement of recognition performance when using our strategy to select regions for feature extraction instead of using a binary segmentation mask and considering all unmasked regions equally.

## 1 Introduction

Iris possess a very rich pattern that is believed to be different between persons, therefore iris recognition has become one of the most reliable and accurate biometric identification systems available. A detailed survey on iris recognition has been published by Bowyer et al. [1].

The first successful algorithm for iris recognition was proposed by John Daugman [4] and is used in most of today's commercial iris recognition systems. After a

preprocessing of the iris that includes a segmentation and a normalization step, this algorithm uses convolution with two dimensional Gabor filters to extract the texture from the normalized iris image. Each filter's phase response is then quantized in a pair of bits, so the information from the iris image is represented by a binary code. Following this, different images of irises can be compared by an efficient comparison of their binary codes using bitwise operations.

It has been shown that recognition performance is the highest when the feature extraction is done in regions of the normalized iris image that are of good quality, namely well textured parts of the iris that do not contain artifacts such as eyelid or eyelash occlusion, specular reflection and excessive blur [7].

Several methods have been tested in the literature in order to prevent the feature extraction from being performed on noisy regions of the iris. A first strategy is to systematically exclude the regions that are considered the most likely to contain noise. For example, in [4] Daugman chose to exclude two portions of the iris: one at the top that is often occluded by the upper eyelid, and another at the bottom where specular reflection is common. The drawback of such an exclusion strategy is that it doesn't take into account the specificity of each iris image. Even though the excluded regions are those that are the most likely to contain artifacts, these can unpredictably be found in other regions. This is particularly common when the acquisition conditions have been loosened, which is the case of most of the public databases available since the 2005 Iris Challenge Evaluation [11].

In response to this difficulty, it is common to generate a mask that is specific to each iris image and allows a deterministic decision between what is an artifact and what isn't [5, 9]. It was shown in these works that such masks improve considerably the performance of recognition. However, when applying a mask, the unmasked regions will all be considered equally at the extraction step. This can be a problem for two reasons: firstly, masks are rarely perfect, so some parts of the unmasked regions may still contain artifacts. Secondly, some regions are more informative than others because they are highly textured so it would be wise to privilege the feature extraction in those regions.

As an answer to these issues we propose in this paper, a new way of choosing the regions of the normalized iris image that will be used for feature extraction. This technique has already been briefly described in the patent we have submitted [3]. The first step of our algorithm is to use a GMM-based quality measure to estimate the quality of different regions of the normalized iris image, especially to quantify the amount of artifacts in each region as well as the amount of texture. Then we sort these regions depending on their quality and select the  $N$  best regions for the feature extraction, where  $N$  is a set parameter that we have determined experimentally. This way, the priority for the extraction is given to the regions that are free from occlusion and among these, to the ones that are the most highly textured. Additionally, we can choose  $N$  wisely in order to compromise between keeping too many regions, which would mean taking into account regions containing artifacts for the feature extraction step, and too few regions which would give a biased inter-class distribution and decrease recognition performance.

We have chosen to work with three public databases: ND-IRIS-0405 [12], CASIA-IrisV3-Interval [16] and CASIA-IrisV3-Twins [16]. These databases contain a great variety of deteriorations such as occlusions, blur and specular reflection.

This paper is organized as follows. Section 2 presents our GMM-based quality measure and shows how it is correlated to local recognition performance. Section 3 describes our technique for selecting the regions exploited for feature extraction as well as the related performance on the three databases mentioned above. Finally conclusions are given in Section 4.

## 2 GMM-based Iris Quality Measure

### 2.1 Description of the measure

Good quality iris images are the key to high iris recognition performance. For this reason, many authors have proposed quality metrics for iris images and used them to improve system performance. Some have defined global quality metrics, meaning metrics that quantify the quality of an entire iris image. They often measure focus [14, 5, 6], occlusion [10] or pupil dilation [10]. Recently, the National Institute of Standards and Technology (NIST) has done an Iris Quality Calibration and Evaluation (IREX II – IQCE) that is aimed at “defining and quantifying iris image properties that are influential on performance of iris recognition” [13]. Such global quality metrics can be used to screen out poor-quality images before recognition, initiate the acquisition of new data or influence a multimodal biometric fusion process.

Other authors have defined local quality metrics to quantify the quality of different regions in a given iris image. Chen et al. [2] proposed a wavelet transform based quality measure. Krichen et al. [8] as well as Li and Savvides [9] have proposed a local (pixel-level) quality measure relying on a Gaussian Mixture Model. Note that Li and Savvides use two GMMs, which are learned respectively on good quality and low quality images while Krichen et al. use only one GMM learned on the “good” quality texture. Krichen therefore does not need any database of noisy images and can adapt easily to any type of new artifacts.

The quality measure we have exploited in this paper is based on an extension of our Gaussian Mixture Model presented in Krichen’s work in [8]. Our measure enables us to discriminate the noisy iris portions from the good ones as in [8], but also to distinguish highly textured regions from poorly textured ones. This will be shown in section 2.2.

In the same way as in [8], we have used a single GMM to characterize high quality sub-images. However this time we have chosen different sub-images from those used in Krichen’s work to train our model and we have added local observations to the input vector  $x_i$  to enrich the model. The chosen sub-images are not only free from occlusion and well-focused, but also highly textured. In practical terms, we have selected 50 such sub-images of dimension 11x51 from three different databases ND-IRIS-0405, Casia-IrisV3-Interval and Casia-IrisV3-Twins. The input vector  $x_i$  has four components: the

pixel  $i$  grey-level, the local mean, local variance and local contrast measured in a  $5 \times 5$  neighbourhood of pixel  $i$ .

As in [8], we do not explicitly model statistically poor textured or noisy sub-images. Both of these 2 types of sub-images will be characterized by observations that are different from those of our good quality model. This means that the likelihood given by the model will be lower than the one obtained on the high quality images.

The quality measure associated to a sub-image  $w$  will be given by the formula:

$$Q(w) = e^{-\frac{1}{a*b} \sum_{i=1}^{a*b} |\log(p(x_i/\lambda)) - \overline{LL}_{BA}|} = e^{-\frac{1}{a*b} \sum_{i=1}^{a*b} d_{ll_i}}$$

where  $a*b$  is the size of the sub-image  $w$  and  $x_i$  is the input vector of our GMM described here above.  $p(x_i/\lambda)$  is the likelihood given by the GMM  $\lambda$  to the input vector  $x_i$  and  $\overline{LL}_{BA}$  is the mean log-likelihood on the training set. We subtract this mean log-likelihood from the log-likelihood given by the GMM and consider the absolute value of the result. We call this value the log-likelihood distance  $d_{ll}$ . The lower this distance is, the closest the vector  $x_i$  is to the training set. We then calculate the mean of these distances for all the pixels of the sub-image  $w$ . Finally, we use a negative exponential to bring the result  $Q(w)$  back to the likelihood space and therefore obtain a value between 0 and 1. The closest its value will be to 1, the highest are the chances that the sub-image  $w$  is of good quality, namely free from occlusion and highly textured.

## 2.2 Experiments and results

We have analyzed the values given by our quality measure to sub-images belonging to different categories: sub-images containing eyelash occlusion, eyelid occlusion or specular reflection and sub-image that are poorly or highly textured. Figure 1 presents, for each category, one example of a sub-image from this category (localized by a white or black rectangle), the local quality measure  $Q(w)$  associated to this example and the mean of the local qualities of 500 sub-images from this category that have been chosen manually. The images were chosen from the three databases mentioned earlier. We can see that the lowest values are given to occluded sub-images, whether the occlusion comes from eyelashes, eyelids or specular reflections. Regions of the iris that are free from occlusion but very lowly textured are given intermediate values. The highest values are given to highly textured sub-images that are free from occlusion.

We have done experiments to verify that the sub-images of highest quality, according to our measure, are the ones that lead to the best performance. To do this we have considered 900 intra-class and 900 inter-class couples of normalized irises selected randomly in the ND-IRIS-0405, CASIA-IrisV3-Interval and CASIA-IrisV3-Twins databases. We consider 72 sub-images per image and perform Daugman's feature extraction on them, namely convolution with Gabor filters and quantization of the output phase in a binary code. For each of the 1800 couples of irises we calculate 72 local



normalized Hamming distances that are the normalized Hamming distances between the binary codes of the 72 couples of sub-images ( $w_{gal}, w_{prob}$ ). On the other hand, we also have computed the quality of all the sub-images we consider and associated to each couple of sub-images the minimum value of the two qualities.






Category	Presence of artifact	Example of a sub-image	Associated local quality: $Q(w)$	Mean local quality on category: $mean(Q(w))$
Eyelash Occlusion	Yes		0.01	0.14
Eyelid Occlusion	Yes		0.01	0.07
Specular Reflection	Yes		0.19	0.05
Lowly textured	No		0.41	0.55
Highly textured	No		0.89	0.75

Figure 1: Sub-images from different categories and value of the corresponding local quality. The mean local quality value on each category is also provided in the last column.

We have analyzed how the local normalized Hamming distances are linked to these quality values. More precisely, we have divided the couples of sub-images into 6 categories according to the value of their quality and calculated the mean local normalized Hamming distance for each category. The results are presented in Figure 2.

$\min(Q(w_{gal}), Q(w_{prob}))$	$mean(HD_{loc}^{intra}(w_{gal}, w_{prob}))$	$mean(HD_{loc}^{inter}(w_{gal}, w_{prob}))$
[0.8; 1]	0.24	0.45
[0.6; 0.8]	0.25	0.45
[0.4; 0.6]	0.28	0.45
[0.2; 0.4]	0.30	0.46
[0.1; 0.2]	0.32	0.47
[0; 0.1]	0.35	0.46

Figure 2: Mean local normalized Hamming distance for intra-class and inter-class comparisons of sub-images, for different quality values of these sub-images.

As we can see, for the intra-class comparisons, the mean local normalized Hamming distance between the couples of sub-images increases when their quality decreases. On the other hand, for the inter-class comparisons the mean of the local normalized

Hamming distance stays stable except for the lowest quality values ( $Q(w) < 0.4$ ) where it increases slightly. So the higher the quality of the sub-images according to our measure, the higher is the matching accuracy between these sub-images. This proves that it is relevant to use our quality measure to select the sub-images used for feature extraction and matching, following Daugman's approach, which is what we will present in the following section.

### 3 Exploiting our quality measure for pertinent feature extraction

#### 3.1 Description of the method

We have seen in the previous section that the higher is the quality of sub-images in the normalized iris images, the more discriminating are these sub-images for matching according to Daugman's technique. Consequently, we wish to exploit our quality measure to select the regions in the normalized iris images that will be used for feature extraction.

One way of proceeding would be to use our quality measure to mask regions of the iris that have a bad quality, similarly to what is done in [9]. In this case, we would experimentally define a threshold for the quality value and sub-images that have a quality below the threshold would be masked. However, by doing so, we would transform our real-valued quality measure into a binary one by separating abruptly iris sub-images into two categories: good quality and bad quality ones. So we would lose information. Furthermore, we would also lose control on the number of regions we wish to select for feature extraction: some irises would have a big portion of masked regions and others wouldn't have any at all. Daugman has shown in [6] that this can lead to a biased inter-class distribution and therefore degrade recognition performance. More precisely, it is important to choose aptly the amount of information that we use for matching. Too much information would be unnecessary and increase the chances of including noisy information and too few information would increase the false acceptance rate.

For these reasons, we have chosen to couple our quality measure to Daugman's algorithm in a new and original way that gives us full control over the number of regions we use for the feature extraction. Consequently, we control the amount of information used for matching. Our final goal is to match the gallery and the probe iris in order to decide if they belong to the same person or not. The algorithm we propose follows the isolation and normalization steps of the gallery and probe images and precedes the bitwise matching stage. It can be cut into six steps:

- **Step 1:** we choose to consider  $M$  points in each one of the gallery and probe normalized iris images. These points are identically located in the two images and uniformly distributed through each image.
- **Step2:** we compute the local qualities  $Q(w_i)$   $i \in [1; M]$  of the sub-images  $w_i$  centered at each of these  $M$  points in the gallery and in the probe image. As a result we have  $M$  quality values for the gallery image and  $M$  quality values for the probe image.

- **Step 3:** Given that our final goal is to match the gallery and the probe iris in order to decide if they belong to the same person or not, when extracting features and matching them, we are interested in the quality of the probe image as well as the one of the gallery. Therefore the third step of our algorithm is to fuse the quality values of the gallery and probe images. It is quite obvious that when matching the features corresponding to a gallery and a probe sub-image, one of both sub-images containing an artifact is enough for the chances of false rejection to increase drastically. So the most relevant way to fuse the quality measures is to select the minimum value between the gallery and the probe quality value. This is done for the  $M$  couples of gallery and probe sub-images. As a result, we have  $M$  fused quality measures associated to the  $M$  (gallery, probe) couples of sub-images.
- **Step 4:** we sort the  $M$  fused quality measures from the highest value to the lowest one. Each quality measure being associated to one of the  $M$  points, this sorting enables us to associate a rank to each one of these points.
- **Step 5:** we select the points that have the  $N$  lowest ranks, so they correspond to the sub-images with the  $N$  highest quality values.
- **Step 6:** These points will be used for feature extraction in the sixth step. More precisely, a convolution with Gabor filters is done around them. Each filter's phase response is then quantized in a pair of bits, so the information from the iris image is represented by a binary code, of which the length is proportional to  $N$ .

As explained before, the parameter  $N$  that is directly linked to the amount of information used for the matching is very important. Setting it too high or too low would reduce recognition performance. So we have conducted experiments for determining the best value for  $N$ . This will be explained in detail in Section 3.2.

The bitwise matching that follows the sixth step of our algorithm is the one that was originally proposed by Daugman [4] and that computes the normalized Hamming distances between the binary codes associated to each iris. To take into account the variations of rotation of the iris from one image to another we perform all of the six steps described here above as well as the bitwise matching for different translations of the probe normalized image along the angular coordinate. The minimum computed normalized Hamming distance is assumed to correspond to the correct alignment of the gallery and probe images.

### 3.2 Experiments and results

This section presents the experiments we have conducted to test the method that we have presented here above and that enables us to select the regions in the normalized iris image exploited for performing feature extraction.

As explained in Section 3.1, our algorithm follows a segmentation step and a normalization step. These aren't the topic of this paper, therefore the input to all of our experiments are normalized iris images and in some cases the segmentation masks that

are associated to them. In this paper, these images have been obtained by isolating the iris manually by two circles that represent its inner and outer boundaries and then normalized by applying Daugman's rubber-sheet model to generate a rectangular image of pre-set dimensions [4]. The segmentation masks associated to them have been obtained by applying the active contour segmentation of OSIRIS-V2 [15] initialized by the manually-determined circular boundaries of the iris (by using the manual initialization option in OSIRIS-V2). However we also could have used normalized images coming from different isolation and normalization methods as well as segmentation masks coming from different segmentation techniques, the important point being that we use the same normalized images and associated masks when comparing different protocols for feature extraction.

For a start, we have compared our feature extraction algorithm to the one in OSIRIS-V2, our reference system implementing Daugman's approach based on bi-dimensional Gabor filter feature extraction. In OSIRIS-V2 the features of the normalized iris images are extracted around  $M$  points that are uniformly distributed. To the contrary, in the feature extraction algorithm that we propose, we select the  $N$  (out of  $M$ ) best quality sub-images and extract the features around the  $N$  points located at the centers of these sub-images. In OSIRIS-V2, it is possible to take into account the segmentation mask associated to the normalized iris images, which improves recognition performance. However in the algorithm we propose, it is unnecessary to use this mask, since we choose the application points used for feature extraction wisely. In both cases, the feature extraction is done by a convolution with the same two-dimensional Gabor filters.

We have compared our feature extraction algorithm to the one in OSIRIS-V2 with and without using the segmentation masks. More precisely, since both feature extraction techniques lead to a binary code representation of the iris, we have followed the different feature extractions by a bitwise matching step in which we calculate the Hamming distance between their binary codes. We have then compared the performance of the different iris recognition methods thus obtained when taking as input the same normalized iris images.

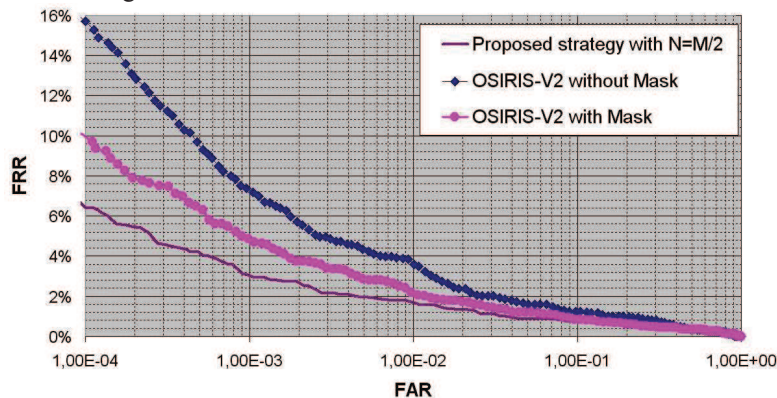


Figure 3: ROC curves of 3 iris recognition algorithms based on the same implementation of Daugman's approach but using different strategies for feature extraction's point selection on the ND-IRIS-0405 database.

Since the segmentation module of OSIRIS-V2 has been optimized on images from ICE-2005, it performs well on ND-IRIS-0405 (a superset of ICE-2005 and ICE-2006). We wish to compare our feature extraction algorithm to the one in OSIRIS-V2 when the latter performs best, so we have used images from ND-IRIS-0405 for this comparison. The images we have used are 2136 images from the ND-IRIS-0405 database, 6 images for each of the 249 subjects, selected randomly. These input images have previously been normalized as explained before.

To evaluate performance of each algorithm, we have plotted its ROC curve, presenting the False Rejection Rate as a function of the False Acceptance Rate. The closest the ROC curve is to the axis, the better are the performance of the algorithm. Figure 3 presents the results of the ROC curves of OSIRIS-V2 with and without using the segmentation masks and of our algorithm that doesn't use the segmentation masks.

As we can see, the lowest performance is obtained when using OSIRIS-V2, without segmentation mask. Adding this mask improves recognition performance considerably, since the FRR at an FAR =  $10^{-4}$  is 36% lower with mask than without mask. However, using our feature extraction algorithm in which we select the  $N (=M/2)$  best quality sub-images for the two dimensional Gabor filters convolution gives an even better performance, especially at low FAR. For an FAR= $10^{-4}$ , the FRR is 59% lower than OSIRIS-V2 without mask and 36% lower than OSIRIS-V2 with mask.

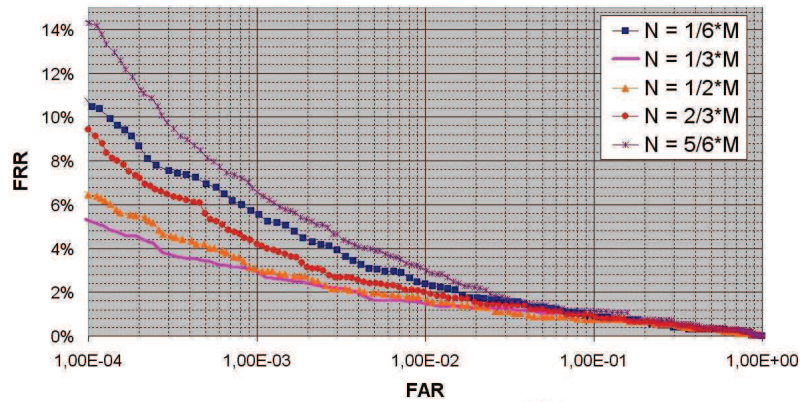


Figure 4: ROC curves for different values of  $N$  (number of selected points for feature extraction) and a set value of  $M$  (number of initial points) for the ND-IRIS-0405 database.

As explained in Section 3.1, the choice of parameter  $N$  is very important because it determines the length of the binary code and the quantity of information that is used for matching. In the first experiment we have pursued, we had chosen an arbitrary value for  $N$ , equal to  $M/2$ . Intuitively, it is easy to understand that, for a set value of  $M$ , the value of  $N$  could be optimized in order to compromise between, on one hand, too much information for matching, that would be unnecessary and increase the chances of including noisy information and, on the other hand, too few information that would increase the false acceptance rate. So we have compared the performance of our



algorithm on the ND-IRIS-0405 database for different values for  $N$ , and a set value of  $M$ . In other words, we have changed the proportion of the initial normalized iris image that is kept for feature extraction when applying our quality-based region selection strategy. The results of this experiment are presented in Figure 4.

We can see in this experiment with images from the ND-IRIS-0405 database, that the optimum value for  $N$  is  $M/3$ . For this value of  $N$ , the FRR at an  $FAR=10^{-4}$  is 48% lower than with OSIRIS-V2 with mask.

In order to generalize this result, we have pursued the same experiment on all the images from the CASIA-IrisV3-Interval and the CASIA-IrisV3-Twins databases. Figure 7 presents, for the three different databases, the False Rejection Rate at a False Acceptance Rate equal to  $10^{-4}$  for different values of  $N$  and a set value of  $M$ . It shows that the performance of our algorithm on the two CASIA-IrisV3 databases for different values of  $N$  with a set value of  $M$  follows the same trend as it does on ND-IRIS-0405 meaning that it is best for  $N = M/3$ .

<b>Database \ N</b>	<b>1/6 * M</b>	<b>1/3 * M</b>	<b>1/2 * M</b>	<b>2/3 * M</b>	<b>5/6 * M</b>
<b>ND-IRIS-0405</b>	10.5%	5.2%	6.4%	9.4%	14.3%
<b>Casia-IrisV3-Interval</b>	4.9%	2.5%	2.8%	3.6%	5.2%
<b>Casia-IrisV3-Twins</b>	18.8%	14.2%	15.2%	20.1%	27.5%

Figure 5: False Rejection Rate (for  $FAR = 10^{-4}$ ) for different databases and for different values of  $N$  (number of selected points for feature extraction) given in function of  $M$  (initial number of points).

So as to verify that the optimal value of  $N$  is a function of  $M$ , we have conducted the experiment described here above for different values of  $M$ . Table 1 presents the values of  $M$  we have considered as well as, for each value of  $M$ : the values tested for  $N$ , the value of  $N$  that optimizes performance ( $N_{opti}$ ) and the value of the ratio  $N_{opti}/M$ . We can see that the value of  $N_{opti}$  changes, but the ratio  $N_{opti}/M$  stays constant and equal to  $1/3$ . This demonstrates that the parameter that has an influence on performance is the proportion of the normalized iris image that is exploited and that this parameter stays the same whatever is the density of the uniformly distributed points considered in the normalized iris images.

It is interesting to see that when choosing the application points for feature extraction wisely, the best performance is obtained when keeping only  $1/3$  of the initial application points. When more points are selected, the chances of performing feature extraction in bad quality regions rises, which increases the false rejection rate. To the contrary, when

keeping fewer points, the amount of information available for the matching decreases too much, which tends to increase the false acceptance rate.

Table1: Values tested for  $N$  (number of selected points for feature extraction), values of  $N$  that optimize performance ( $N_{\text{opti}}$ ) and values of  $N_{\text{opti}}/M$ , for different values of  $M$  (initial number of points).

Values of $M$	Values tested for $N$	$N_{\text{opti}}$	$N_{\text{opti}}/M$
240	40, 80, 120, 160, 200	80	1/3
480	80, 160, 240, 320, 400	160	1/3
720	120, 240, 360, 480, 600	240	1/3
960	160, 320, 480, 640, 800	320	1/3

## 5 Conclusions

In this paper, we have presented a new method for choosing the regions of normalized iris images that will be used for feature extraction. Our method uses the GMM-based local quality measure we elaborated that gives the highest values to sub-images that are highly textured and free from occlusion, the lowest ones to noisy sub-images and intermediate values to regions that are free from artifacts but barely textured. The idea is to select the best sub-images of the iris according to this quality measure for the feature extraction. More precisely, after evaluating the quality of  $M$  sub-images in the normalized gallery and probe iris images that we wish to compare, the quality values of each couple of gallery and probe sub-images are fused and the couples are sorted according to their quality. The  $N$  best couples are then selected for the feature extraction.

We have tested this method for selecting the regions in normalized iris images used for feature extraction on three different databases: ND-IRIS-0405, CASIA-IrisV3-Interval and CASIA-IrisV3-Twins. We have demonstrated that the achieved recognition performance depends highly on the proportion of the initial normalized image that is selected and that there is a value of this proportion that optimizes performance. For this value, the recognition performance was considerably better than the one obtained when using the feature extraction module of OSIRIS-V2, a reference system based on Daugman's recognition algorithm. Our solution, in the same way as the OSIRIS-V2 algorithm when used with a segmentation mask, chooses the regions for the feature extraction specifically for each couple of irises we wish to match, and avoids taking into account regions containing artifacts. It has the additional advantage of exploiting a real-valued quality measure, which allows us to quantify the amount of texture in the artifact-free portions of the iris in order to privilege the highly textured regions for the feature extraction. Furthermore, our method gives us full control on the amount of information used for matching which also tends to increase the performance of our iris recognition algorithm, in opposition to the methods that use segmentation or quality masks.



In this paper, we have chosen to use the same value for  $N$  for all the couples of images we compare. We have demonstrated how to choose  $N$  to achieve optimal performance when following this strategy. In future work, we will consider a different strategy in which we will adapt the value of  $N$  to each couple of images : this way the value of  $N$  will be selected according to the quality of the corresponding images.

## References

- [1] Kevin W. Bowyer, Karen Hollingsworth, Patrick J. Flynn, 2008. Image understanding for iris biometrics: a survey, *Computer Vision and Image Understanding* 110, 281-307
- [2] Yi Chen, Sarat C. Dass, Anil K. Jain, 2005. Localized iris image quality using 2-D wavelets, in: Springer LNCS 3832: International Conference on Biometrics, pp. 373-381
- [3] Sandra Cremer, Nadège Lemperière, Sonia Garcia, Bernadette Dorizzi, Procédé de comparaison d'images d'iris par sélection intelligente de zones texturées, Patent n°10 04942, 2010, Thales
- [4] John Daugman, 1993. High confidence visual recognition of persons by a test of statistical independence, *IEEE Trans. Pattern Anal. Mach. Intell.* 15 (11), 1148-1161
- [5] John Daugman, 2001. Statistical richness of visual phase information : update on recognizing persons by iris patterns, *Int. J. Comput. Vis.* 45(1), 25-38.
- [6] John Daugman, 2007. New methods in iris recognition, *IEEE Trans. Syst. Man Cyber – B37* (5), 1167-1175
- [7] Nathan D. Kalka, Jinyu Zuo, Natalia A. Schmid, Bojan Cukic, 2006. Image quality assessment for iris biometric, in *SPIE 6202: Biometrics Technology for Human Identification III*, pp. 6202:D1-D11
- [8] Emine Krichen, Sonia Garcia-Salicetti, Bernadette Dorizzi, 2007. A new probabilistic iris quality measure for comprehensive noise detection, in: *Biometrics: Theory, Applications, and Systems*
- [9] Yung-hui Li, Marios Savvides, 2009. A pixel-wise, learning-based approach for occlusion estimation of iris images in polar domain, *icassp*, pp.1357-1360, 2009 IEEE International Conference on Acoustics, Speech and Signal Processing
- [10] Pan Lili, Xie Mei, 2005. The Algorithm of Iris Image Preprocessing, *Automatic Identification Advanced Technologies, Fourth IEEE Workshop on Automatic Identification Advanced Technologies*
- [11] P. Jonathon Phillips, Kevin W. Bowyer, Patrick J. Flynn, Xiaomei Liu, and W. Todd Scruggs. "The Iris Challenge Evaluation 2005", IEEE Second International Conference on Biometrics: Theory, Applications and Systems (BTAS 08)
- [12] P. Jonathon Phillips, W. Todd Scruggs, Alice J. O'Toole, Patrick, J. Flynn, Kevin W. Bowyer, Cathy L. Schott, Matthew Sharpe, "FRVT 2006 and ICE 2006 Large-Scale Experimental Results," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, in press. <http://doi.ieeecomputersociety.org/10.1109/TPAMI.2009.59>
- [13] E. Tabassi, P. Grother, W. Salamon, 2011. "IREX II – IQCE. Iris Quality Calibration and Evaluation", NIST Interagency Report 7820, <http://www.nist.gov/itl/iad/ig/irexii.cfm>
- [14] Zhang et al., 1999. Method of measuring the focus of close-up image of eyes. U.S. Patent No. 5,953,440
- [15] E. Krichen, A. Mellakh, S. Garcia-Salicetti, B. Dorizzi, Osiris (open source for iris) reference system. BioSecure Project, <http://www.biosecure.info>, 2008.
- [16] CASIA-IrisV3, <http://www.cbsr.ia.ac.cn/IrisDatabase>

# Bibliography

- [1] <http://iris.nist.gov>.
- [2] <http://iris.nist.gov/irex/>.
- [3] <http://www.nist.gov/itl/iad/ig/mbgc.cfm>.
- [4] Multimedia University iris database. Available from: <http://pesona.mmu.edu.my/~ccteo/>.
- [5] National Institute of Standards and Technology. Iris challenge evaluation data. Available from: <http://iris.nist.gov/ice/>.
- [6] University of bath iris image database. available from: <http://www.bath.ac.uk/elec-eng/research/sipg/irisweb/database.htm>.
- [7] The results of the NICE.II Iris biometrics competition. *Pattern Recognition Letters*, 33(8):965 – 969, 2012.
- [8] Onsy Abdel Alim and Maha Sharkas. Iris recognition using discrete wavelet transform and artificial neural networks. In *Circuits and Systems, 2003 IEEE 46th Midwest Symposium on*, volume 1, pages 337 – 340 Vol. 1, dec. 2003.
- [9] AOptiX. <http://www.aoptix.com>.
- [10] C. Belcher and Yingzi Du. A selective feature information approach for iris image-quality measure. *Information Forensics and Security, IEEE Transactions on*, 3(3):572 –577, sept. 2008.
- [11] A. Bertillon. *La couleur de l'iris*. 1885.
- [12] R. Bodade and S. Talbar. Dynamic iris localisation: A novel approach suitable for fake iris detection. In *Ultra Modern Telecommunications Workshops, 2009. ICUMT '09. International Conference on*, pages 1 –5, oct. 2009.
- [13] W. Boles and B. Boashash. A human identification technique using images of the iris and wavelet transform. *Signal Processing, IEEE Transactions on*, 46(4):1185 –1188, apr 1998.
- [14] Kevin W. Bowyer, Karen Hollingsworth, and Patrick J. Flynn. Image understanding for iris biometrics: A survey. *Computer Vision and Image Understanding*, 110(2):281 – 307, 2008.

- [15] Kevin W. Bowyer, Karen P. Hollingsworth, and Patrick J. Flynn. A survey of iris biometrics research: 2008-2010. *Handbook of iris recognition*.
- [16] K.W. Bowyer, K. Hollingsworth, and P.J. Flynn. Factors that degrade the match distribution in iris biometrics. *Identity in the Information Society*, 2(3):327–343, 2009.
- [17] C. Boyce, A. Ross, M. Monaco, L. Hornak, and Xin Li. Multispectral iris analysis: A preliminary study. In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW '06. Conference on*, page 51, june 2006.
- [18] CASIA-IrisV3. <http://www.cbsr.ia.ac.cn/irisdatabase>.
- [19] Y. Chen, S.C. Dass, and A.K. Jain. Localized iris image quality using 2-D wavelets. *Springer LNCS 3832: International Conference on Biometrics*, pages 373–381, 2005.
- [20] Daugman. The importance of being random: statistical principles of iris recognition. *Pattern Recognition*, 36(2):279 – 291, 2003.
- [21] J. Daugman. High confidence visual recognition of persons by a test of statistical independence. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 15(11):1148 –1161, nov 1993.
- [22] J. Daugman. How iris recognition works. *Circuits and Systems for Video Technology, IEEE Transactions on*, 14(1):21 – 30, jan. 2004.
- [23] J. Daugman. New methods in iris recognition. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 37(5):1167 –1175, oct. 2007.
- [24] Michal Dobes and Libor Machala. Iris database, available from: <http://www.inf.upol.cz/iris/>.
- [25] Wenbo Dong, Zhenan Sun, and Tieniu Tan. Iris matching based on personalized weight map. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(9):1744 –1757, sept. 2011.
- [26] Vivekanand Dorairaj, Natalia A. Schmid, and Fahmy Gamal. Performance evaluation of iris based recognition system implementing PCA and ICA encoding techniques. *SPIE 5779: Biometric Technology for Human Identification II*, 5779:51 – 58, 2005.
- [27] Yingzi Du, Craig Belcher, and Zhi Zhou. Scale invariant gabor descriptor-based noncooperative iris recognition. *EURASIP J. Adv. Signal Process*, 2010:37:1–37:13, February 2010.
- [28] Yingzi Du, Craig Belcher, Zhi Zhou, and Robert Ives. Feature correlation evaluation approach for iris feature quality measure. *Signal Process.*, 90:1176–1187, April 2010.
- [29] Zhang et al. Method for measuring the focus of close-up image of eyes. U.S. Patent No. 5,953,440, 1999.

**CONFIDENTIAL**

- [30] Leonard Flom and Aran Safir. Iris recognition system. U.S. Patent 4 641 349.
- [31] Junying Gan and Yu Liang. Applications of wavelet packets decomposition in iris recognition. In David Zhang and Anil Jain, editors, *Advances in Biometrics*, Lecture Notes in Computer Science. Springer Berlin / Heidelberg.
- [32] Feng Hao, J. Daugman, and P. Zielinski. A fast search algorithm for a large fuzzy database. *Information Forensics and Security, IEEE Transactions on*, 3(2):203–212, june 2008.
- [33] Z. He, Z. Sun, T. Tan, and Z. Wei. Efficient iris spoof detection via boosted local binary patterns, 2009.
- [34] Zhaofeng He, Zhenan Sun, Tieniu Tan, Xianchao Qiu, Cheng Zhong, and Wenbo Dong. Boosting ordinal features for accurate and fast iris recognition. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8, june 2008.
- [35] K. Hollingsworth, K.W. Bowyer, and P.J. Flynn. All iris code bits are not created equal. In *Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on*, pages 1–6, sept. 2007.
- [36] K.P. Hollingsworth, K.W. Bowyer, and P.J. Flynn. Using fragile bit coincidence to improve iris recognition. In *Biometrics: Theory, Applications, and Systems, 2009. BTAS '09. IEEE 3rd International Conference on*, pages 1–6, sept. 2009.
- [37] Huifang Huang and Guangshu Hu. Iris recognition based on adjustable scale wavelet transform. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 7533–7536, jan. 2005.
- [38] Ya-Ping Huang, Si-Wei Luo, and En-Yi Chen. An efficient iris recognition system. In *Machine Learning and Cybernetics, 2002. Proceedings. 2002 International Conference on*, volume 1, pages 450–454 vol.1, 2002.
- [39] Yung hui Li and M. Savvides. A pixel-wise, learning-based approach for occlusion estimation of iris images in polar domain. In *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pages 1357–1360, april 2009.
- [40] Jain Jang, Kang Ryoung Park, Jinho Son, and Yillbyung Lee. A study on multi-unit iris recognition. In *Control, Automation, Robotics and Vision Conference, 2004. ICARCV 2004 8th*, volume 2, pages 1244–1249 Vol. 2, dec. 2004.
- [41] P.A. Johnson, P. Lopez-Meyer, N. Sazonova, F. Hua, and S. Schuckers. Quality in face and iris research ensemble (Q-FIRE). In *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, pages 1–6, sept. 2010.

**CONFIDENTIAL**

- [42] R. Johnston. Can iris patterns be used to identify people? Chemical and Laser Sciences Division Annual Report LA-12331-PR, Los Alamos National Laboratory, June 1992.
- [43] N. D. Kalka, V. Dorairaj, Y. N. Shah, N. A. Schmid, and B. Cukic. Image quality assessment for iris biometric. In *Proceedings of the 24th Annual Meeting of the Gesellschaft für Klassifikation*, pages 445–452. Springer, 2002.
- [44] Byung Jun Kang and Kang Ryoung Park. A study on iris image restoration. In *Proceedings of the 5th international conference on Audio- and Video-Based Biometric Person Authentication, AVBPA'05*, pages 31–40, Berlin, Heidelberg, 2005. Springer-Verlag.
- [45] R. Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 1995.
- [46] E. Krichen. *Reconnaissance des personnes par l'iris en mode dégradé*. PhD thesis, Institut National des Telecommunications d'Evry, 2007.
- [47] E. Krichen, S. Garcia-Salicetti, and B. Dorizzi. A new probabilistic iris quality measure for comprehensive noise detection. In *Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on*, pages 1–6, sept. 2007.
- [48] E. Krichen, S. Garcia-Salicetti, and B. Dorizzi. A new phase-correlation-based iris matching for degraded images. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 39(4):924–934, aug. 2009.
- [49] E. Krichen, A. Mellakh, S. Salicetti, and B. Dorizzi. Osiris (open source for iris) reference system.
- [50] Yooyoung Lee, R.J. Micheals, and P.J. Phillips. Improvements in video-based automated system for iris recognition (VASIR). In *Motion and Video Computing, 2009. WMVC '09. Workshop on*, pages 1–8, dec. 2009.
- [51] T. Lefevre. Improving normalization circle detection for iris recognition using a variational approach. Technical report, Institut Mines-Telecom, Telecom SudParis, 2012.
- [52] T. Lefvre. *Iris segmentation for degraded images*. PhD thesis, Institut Mines-Telecom, Telecom SudParis, 2013.
- [53] Xingguang Li, Zhenan Sun, and Tieniu Tan. Comprehensive assessment of iris image quality. In *Image Processing (ICIP), 2011 18th IEEE International Conference on*, pages 3117–3120, sept. 2011.
- [54] Pan Lili and Xie Mei. The algorithm of iris image preprocessing. In *Automatic Identification Advanced Technologies, 2005. Fourth IEEE Workshop on*, pages 134–138, oct. 2005.
- [55] X. Liu. *Optimizations in iris recognition*. PhD thesis, University of Notre Dame, 2006.

**CONFIDENTIAL**

- [56] Li Ma, Tieniu Tan, Yunhong Wang, and Dexin Zhang. Personal identification based on iris texture analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(12):1519 – 1533, dec. 2003.
- [57] Li Ma, Tieniu Tan, Yunhong Wang, and Dexin Zhang. Efficient iris recognition by characterizing key local variations. *Image Processing, IEEE Transactions on*, 13(6):739 – 750, june 2004.
- [58] Li Ma, Tieniu Tan, Yunhong Wang, and Dexin Zhang. Local intensity variation analysis for iris recognition. *Pattern Recognition*, 37(6):1287 – 1298, 2004.
- [59] L. Masek and P. Kovesi. Matlab source code for a biometric identification system based on iris patterns, 2003. The University of Western Australia, <http://www.csse.uwa.edu.au/pk/studentprojects/libor>.
- [60] J.R. Matey, O. Naroditsky, K. Hanna, R. Kolczynski, D.J. LoIacono, S. Mangru, M. Tinker, T.M. Zappia, and W.Y. Zhao. Iris on the move: Acquisition of images for iris recognition in less constrained environments. *Proceedings of the IEEE*, 94(11):1936 – 1947, nov. 2006.
- [61] K. Miyazawa, K. Ito, T. Aoki, K. Kobayashi, and H. Nakajima. An efficient iris recognition algorithm using phase-based image matching. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, volume 2, pages II – 49–52, sept. 2005.
- [62] D.M. Monro, S. Rakshit, and D. Zhang. DCT-based iris recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(4):586 – 595, april 2007.
- [63] A. Murugan and G. Savithiri. Feature extraction on half iris for personal identification. In *Signal and Image Processing (ICSIP), 2010 International Conference on*, pages 197 – 200, dec. 2010.
- [64] C. Oyster. The human eye structure and function. *Sinauer Associates*, 1999.
- [65] Chul-Hyun Park and Joon-Jae Lee. Extracting and combining multimodal directional iris features. In David Zhang and Anil Jain, editors, *Advances in Biometrics*, Lecture Notes in Computer Science. Springer Berlin / Heidelberg.
- [66] M.B. Pereira and A.C.P. Veiga. Application of genetic algorithms to improve the reliability of an iris recognition system. In *Machine Learning for Signal Processing, 2005 IEEE Workshop on*, pages 159 – 164, sept. 2005.
- [67] M.B. Pereira and A.C.P. Veiga. A method for improving the reliability of an iris recognition system. In *Communications, Computers and signal Processing, 2005. PACRIM. 2005 IEEE Pacific Rim Conference on*, pages 665 – 668, aug. 2005.
- [68] P.J. Phillips, K.W. Bowyer, P.J. Flynn, X. Liu, and W.T. Scruggs. The iris challenge evaluation 2005. In *Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on*, pages 1 – 8, 29 2008–oct. 1 2008.

**CONFIDENTIAL**

- [69] P.J. Phillips, A. Martin, C.L. Wilson, and M. Przybocki. An introduction evaluating biometric systems. *Computer*, 33(2):56–63, feb 2000.
- [70] P.J. Phillips, W.T. Scruggs, A.J. O’Toole, P.J. Flynn, K.W. Bowyer, C.L. Schott, and M. Sharpe. FRVT 2006 and ICE 2006 large-scale experimental results. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(5):831–846, may 2010.
- [71] Hugo Proenca and Luis Alexandre. UBIRIS: A noisy iris image database. available from: <http://iris.di.ubi.pt/>.
- [72] Hugo Proenca and Luis Alexandre. The NICE.I: Noisy iris challenge evaluation - part i. *Biometrics: Theory, Applications, and Systems*, sept 2007.
- [73] D. A. Reynolds and C. Rose. Robust text-independent speaker identification using gaussian mixture speaker models. *IEEE Transactions on Speech and Audio Processing*, 3(1), january 1995.
- [74] S. Ring and K.W. Bowyer. Detection of iris texture distortions by analyzing iris code matching results. In *Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on*, pages 1–6, 29 2008-oct. 1 2008.
- [75] G. Saporta. *Probabilites, analyses de donnees et statistiques*. Editions technip, deuxieme edition.
- [76] S.A.C. Schuckers, N.A. Schmid, A. Abhyankar, V. Dorairaj, C.K. Boyce, and L.A. Hornak. On techniques for angle compensation in nonideal iris recognition. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 37(5):1176–1190, oct. 2007.
- [77] Zhenan Sun, Tieniu Tan, and Yunhong Wan. Robust encoding of local ordinal measures: a general framework of iris recognition. *Proc. BioAW Workshop*, pages 270–282, 2004.
- [78] Zhenan Sun, Yunhong Wang, Tieniu Tan, and Jiali Cui. Improving iris recognition accuracy via cascaded classifiers. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 35(3):435–441, aug. 2005.
- [79] Guillaume Sutra, Sonia Garcia-Salicetti, and Bernadette Dorizzi. The viterbi algorithm at different resolutions for enhanced iris segmentation. *The 5th IAPR International Conference on Biometrics: ICB’12*, 2012.
- [80] E. Tabassi, P. Grother, and W. Salamon. IREX II - IQCE. Iris Quality Calibration and Evaluation. Interagency Report 7820, NIST.
- [81] J. Thornton, M. Savvides, and B.V.K. Kumar. An evaluation of iris pattern representations. In *Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on*, pages 1–6, sept. 2007.
- [82] Christel-Loc Tisse, Lionel Martin, Lionel Torres, and Michel Robert. Person identification technique using human iris recognition. In *Proc. of Vision Interface*, pages 294–299, 2002.

**CONFIDENTIAL**



- 
- [83] M. Vatsa, R. Singh, and A. Noore. Reducing the false rejection rate of iris recognition using textural and topological features. *International Journal of Signal Processing*, 2(2):66–72, 2005.
- [84] S. Venugopalan, U. Prasad, K. Harun, K. Neblett, D. Toomey, J. Heyman, and M. Savvides. Long range iris acquisition system for stationary and mobile subjects. In *Biometrics (IJCB), 2011 International Joint Conference on*, pages 1–8, oct. 2011.
- [85] B.V.K. Vijaya Kumar, Chunyan Xie, and Jason Thornton. Iris verification using correlation filters. In Josef Kittler and Mark Nixon, editors, *Audio- and Video-Based Biometric Person Authentication*, Lecture Notes in Computer Science. Springer Berlin / Heidelberg.
- [86] R.P. Wildes. Iris recognition: an emerging biometric technology. *Proceedings of the IEEE*, 85(9):1348–1363, sep 1997.
- [87] Peng-Fei Zhang, De-Sheng Li, and Qi Wang. A novel iris recognition method based on feature fusion. In *Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on*, volume 6, pages 3661–3665 vol.6, aug. 2004.