

Biométrie multimodale basée sur l'iris et le visage Nefissa Khiari Hili

▶ To cite this version:

Nefissa Khiari Hili. Biométrie multimodale basée sur l'iris et le visage. Traitement du signal et de l'image [eess.SP]. Université Paris-Saclay; Université de Tunis El Manar, 2016. Français. NNT : 2016SACLE014 . tel-01762973

HAL Id: tel-01762973 https://hal.science/tel-01762973

Submitted on 10 Apr 2018

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.



NNT : AAAASACLYXXX





THESE DE DOCTORAT de L'UNIVERSITE DE TUNIS EL MANAR et de L'UNIVERSITE PARIS-SACLAY préparée à l'Université d'Evry Val d'Essonne

ÉCOLE DOCTORALE N°580 Sciences et Technologies de l'Information et de la Communication ÉCOLE NATIONALE d'INGÉNIEURS de TUNIS

Spécialité de doctorat : Traitement du signal et des images

Par

Mme Nefissa KHIARI-HILI

Biométrie multimodale basée sur l'iris et le visage

Thèse présentée et soutenue à Evry, le 24 Mai 2016

Composition du Jury :

M. Beauseroy Pierre, Professeur, UTTRapporteurM. Farah Imed Riadh, Professeur, ISAMMRapporteurM. Rosenberger Christophe, Professeur des universités, ENSICAENExaminateurMme Lelandais-Bonadé Sylvie, Professeur des universités, UEVEDirectrice de thèseM. Hamrouni Kamel, Professeur, ENITDirecteur de thèseM. Montagne Christophe, Maître de conférences, UEVECo-encadrant



Résumé

Cette thèse vise à apporter une contribution dans le domaine de la biométrie par l'iris (l'une des modalités les plus précises et difficiles à pirater) et le visage (l'une des modalités les moins intrusives et les moins coûteuses).

A travers ce travail, nous abordons plusieurs aspects importants de la biométrie mono et multimodale. Nous commençons par dresser un état de l'art sur la biométrie monomodale par l'iris et par le visage et sur la multimodalité iris/visage, avant de proposer plusieurs approches personnelles de reconnaissance d'individus pour chacune des deux modalités. Nous abordons, en particulier, la reconnaissance faciale par des approches classiques reposant sur des combinaisons d'algorithmes et des approches bioinspirées émulant le mécanisme de la vision humaine. Nous démontrons l'intérêt des approches bio-inspirées par rapport aux approches classiques à travers deux méthodes. La première exploite les résultats issus de travaux neuroscientifiques indiquant l'importance des régions et des échelles de décomposition utiles à l'identification d'un visage. La deuxième consiste à appliquer une méthode de codage par ordre de classement dans la phase de prétraitement pour renforcer le contenu informatif des images de visage.

Nous retenons la meilleure approche de chacune des modalités de l'iris et du visage pour concevoir deux méthodes biométriques multimodales. A travers ces méthodes, nous évaluons différentes stratégies classiques de fusion multimodale au niveau des scores. Nous proposons ensuite une nouvelle règle de fusion de scores basée sur un facteur de qualité dépendant du taux d'occultation des iris. Puis, nous mettons en avant l'intérêt de l'aspect double échantillons de l'iris dans une approche multimodale.

L'ensemble des méthodes proposées sont évaluées sur la base multimodale réelle IV² capturée dans des environnements variables voire dégradés et en suivant un protocole bien précis fourni dans le cadre de la campagne d'évaluation IV². Grâce à une étude comparative avec les algorithmes participants à la campagne IV², nous prouvons la compétitivité de nos algorithmes qui arrivent dans plusieurs cas à se positionner en tête de liste.

Abstract

This thesis aims to make a contribution in the field of biometrics based on iris (one of the most accurate and hard to hack biometrics) in conjunction with face (one of the cheapest and less intrusive biometrics).

Through this work, we discuss several important aspects of unimodal and multimodal biometrics. After an overview on unimodal and multimodal biometrics based on iris and face, we propose several personal approaches of biometric authentication using each single trait. Particularly, we address facial recognition first with conventional approaches based on combined algorithms, then with bio-inspired approaches emulating the human vision mechanism. We demonstrate the interest of bio-inspired approaches over conventional approaches through two proposed methods. The first one exploits the results of neuroscientific work indicating the relevant regions and scales in a face identification task. The second consists in applying a rank order coding method at the preprocessing step so as to enhance the information content of face images.

We keep the best unimodal approach of iris and face recognition to design two multimodal biometric methods. Through these methods, we evaluate different classic strategies of multimodal score-level fusion. Afterwards, we propose a new score-level fusion rule based on a quality metric according to irises occultation rates. Then, we point out the interest of the double-sample iris aspect in a multimodal approach.

All the proposed methods are evaluated on the real multimodal IV^2 database captured under variable to degraded environments, and following a specific protocol provided as part of the IV^2 evaluation campaign. After a comparative study with the participant algorithms in the IV^2 campaign, we prove the competitiveness of our algorithms witch outperform most of the participant ones in the IV^2 campaign in many experiments.

A mes parents bien aimés Souheyla et Mohamed A mon cher époux Riadh A mes adorables petits bouts de choux Adam et Anas A toute ma famille et amis proches...

Remerciements

La route a été longue, parfois semée d'embuches, mais nous y sommes arrivés. Je tiens à remercier tous ceux qui m'ont aidé à la réalisation de ce travail.

J'aimerais avant tout témoigner toute ma gratitude à mes deux directeurs de thèse qui m'ont chaleureusement accueilli dans leurs laboratoires, et ont accepté de diriger ma thèse : Madame Sylvie LELANDAIS et Monsieur Kamel HAMROUNI.

Je tiens à remercier Sylvie pour son investissement dans tous les aspects de mon travail et de ma rédaction, des fois malgré la distance qui nous sépare. Sylvie n'a pas hésité à mettre la main à la pâte quand il le fallait pour m'aider à avancer. Elle m'a soutenue à chaque instant de cette thèse et m'a gardée à un moment où je ne me serais pas gardée moi-même. Si j'ai pu conclure mes travaux de thèse, un grand mérite lui revient. Sylvie ne sera sans doute pas éternellement ma chef, mais je la considérerai à jamais comme faisant partie de ma famille.

Je tiens également à remercier Si Kamel qui a su me faire confiance dès ma deuxième année de Master et qui a toujours été d'excellents conseils ; je lui dois de m'avoir fait découvrir l'univers de la biométrie. Il a su faire preuve d'une pédagogie et d'une gentillesse remarquables. Qu'il trouve ici l'expression de ma gratitude pour son encadrement scientifique et pour toute l'attention qu'il a apportée au déroulement de mes travaux.

Je remercie également Christophe MONTAGNE qui a co-encadré ma thèse et qui a répondu présent à chaque fois où j'en avais besoin. Grâce à ses remarques pertinentes et ses encouragements sans cesse renouvelés, j'ai pu avancer en toute assurance.

Je remercie vivement les membres de jury d'avoir accepté de juger mon travail de thèse. Je suis très reconnaissante à Monsieur Imed Riadh FARAH et Monsieur Pierre BEAUSEROY de s'être penchés avec rigueur et grand intérêt sur ce rapport et de m'avoir fait l'honneur d'en être les rapporteurs. Mes plus sincères remerciements vont aussi à Monsieur Christophe ROSENBERGER qui a examiné ce travail de thèse et a participé au jury.

Merci également aux invités de mes comités de thèse, Madame Bernadette DORIZZI et Monsieur Jean-Yves DIDIER qui ont su, par leurs critiques constructives et leurs conseils, faire avancer ma réflexion.

Je tiens également à remercier toutes les personnes que j'ai pu rencontrer et avec lesquelles j'ai pu échanger ; particulièrement Justin PLANTIER notre collaborateur de l'IRBA pour son aide précieuse dans la partie des approches bio-inspirées et avec qui j'ai vécu l'expérience de rédiger mon premier article de revue, Nadia FEDDAOUI qui m'a aidé dans l'implémentation du module de la reconnaissance par l'iris, Anis CHAARI qui a consacré de son temps pour m'expliquer les détails de la campagne IV² et Ayoub EL GHANAOUI l'étudiant en stage que j'ai eu le plaisir d'encadrer.

Ce travail de thèse a été réalisé dans le cadre d'une co-tutelle entre le laboratoire LR-SITI de l'Ecole Nationale d'Ingénieurs de Tunis (Université de Tunis El Manar – UTM) et le laboratoire IBISC (Université d'Evry Val d'Essonne – UEVE). Je remercie leurs directeurs ainsi que tous les chercheurs de ces laboratoires pour l'accueil et les conditions de travail privilégiées qui m'y ont été offertes. Je n'oublie pas le personnel administratif, en particulier Madame Florence HAMON, Madame Carole TROUSSIER et Madame Marité QUINTIN de l'UEVE, ainsi que Madame Rakia CHALOUATI et Monsieur Salem DRIDI de l'ENIT qui n'ont ménagé aucun effort pour me faciliter les démarches administratives tout au long de cette thèse.

Ce travail n'aurait pu être mené à bien sans l'aide de différents financeurs que sont : l'UEVE, l'UTM, l'Agence Universitaire de la Francophonie ainsi que Campus France.

Mes remerciements vont aussi à ma famille et mes amis qui m'ont permis de ne jamais dévier de mon objectif final. Merci à Khaled mon grand frère et à Fatma, Imène et Khouloud mes sœurs de cœur pour leur affection maintes fois renouvelée. Merci à la famille AHMED pour son hospitalité chaleureuse.

Enfin, je remercie mon cher époux pour son soutien quotidien et sa compréhension à l'égard de mes engagements de thèse. Notre couple a grandi en même temps que mon projet scientifique, le premier servant de socle solide au second.

Ces remerciements ne peuvent s'achever, sans une pensée pour mes parents qui ont tout sacrifié pour leurs enfants, n'épargnant ni santé ni efforts ni argent :

Mon père Mohamed qui a su être là dans tous les moments, les bons comme les difficiles ; son soutien constant est pour moi le pilier fondateur de ce que je suis et de ce que je fais, son ambition illimitée m'a toujours inspirée.

Ma mère Souheyla qui a su m'écouter et pris le temps d'entendre mes histoires de biométrie bien qu'étant littéraire ; son dévouement, sa droiture et sa pédagogie dans l'enseignement ont toujours été un modèle pour moi. Merci maman d'avoir consacré ces deux dernières années à la garde de mes enfants pour que je puisse mener à terme ma thèse. J'en profite pour saluer toutes les nounous qui ont su prendre soin de mes enfants en Tunisie comme en France, me permettant ainsi de travailler avec toute assurance : merci Joular, Marwa, Tata Latifa, Imène, Nadia et Chéma.

Pour clore ce préambule, je souhaite exprimer ma profonde reconnaissance à tous ceux qui ont participé de près ou de loin à l'aboutissement de cette thèse.

Table des matières

Table des figuresxi					
Liste d	es tables		xvii		
Introdu	uction géné	érale	1		
Chapit	re 1 : Conc	epts biométriques et base de données	8		
Ir	ntroduction.		9		
1.	1.1 Modes opératoires d'un système biométrique				
1.	1.2 Evaluation d'un système biométrique				
1.	1.3 Biométrie et multimodalité				
	1.3.1	Taxonomie des Systèmes biométriques multimodaux	13		
	1.3.2	Architectures d'un système multimodal	16		
	1.3.3	Niveaux de fusion multimodale	17		
	1.3.4	f'usion au niveau des scores	20		
_	1.3.5	Normalisation des scores			
1.	.4 Bases d	le données et protocoles de test			
	1.4.1	Conception des bases multimodales	27		
	1.4.2	La base multimodale IV^2			
C	1.4.3	Les évaluations realisées sur la base l'v ²			
C	onclusion				
Chapit	re 2 : Biom	étries par l'iris et par le visage : Approches classiques			
Ir	ntroduction.				
2.	.1 Biomét	rie par l'iris			
	2.1.1	Etat de l'art			
	2.1.2	Approche par les Pyramides Orientables (PO)	54		
	2.1.3	Méthode de référence (Log-Gabor)	61		
	2.1.4	Approche de reconnaissance par l'iris retenue	73		
2.	.2 Biomét	rie par le visage	74		
	2.2.1	Etat de l'art	74		
	2.2.2	Application de la Pyramide Orientable en biométrie faciale	et 81		
	2.2.3	Méthodes de réduction de l'espace caractéristique			
	2.2.4	Proposition de méthodes combinées	94		
	2.2.5	Comparaison à la campagne IV ² et méthode de reconnaissan	ıce		
	fa	aciale classique retenue	109		
С	onclusion		111		
Chapit	re 3 : Appro	oches bio-inspirées appliquées en biométrie faciale	113		
Ir	ntroduction.		113		
3.	.1 Context	te neuro-scientifique	114		
	3.1.1 La vision, de la rétine au cortex 11				
	3.1.2	Approches modélisant le mécanisme de la vision utilisées	en		
	r	econnaissance faciale	121		
	3.1.3	Approches bio-inspirées retenues	129		

3.	2 Re	connaissance bio-inspirée par LBP/DoG1	.29
	3	2.2.1 Concepts théoriques de la méthode LBP/DoG1	.30
	3	2.2 Méthode LBP/DoG	.33
	3	2.2.3 Méthode LBP/DoG pondérée selon les travaux de Gosselin et	26
Schyns			
3.	.3 Ke	connaissance factale bio-inspire par Reconstruction Asynchrone – RA	.41
	3 3	5.3.1 Concepts theoriques de la KA	.41
	0	faciale	48
3.	.4 Ap	ports de la RA sur les meilleures méthodes proposées et choix de la	
	mé	thode unimodale finale du visage1	56
С	onclusi	on1	57
Chanit	re 4 •]	Riométrie multimodale basée sur l'iris et le visage · Approches et	
ez	xpérin	nentation	.60
Ir	- ntroduc	tion	61
4.	1 Et:	at de l'art des méthodes multimodales par fusion de scores basées sur	
	l'ir	is et le visage1	61
4.	2 Sti	ructures des bases multimodales IV ² expérimentées1	64
4.	.3 Mé	éthode bimodale par l'iris droit et le visage1	68
	4	.3.1. Normalisation des scores monomodaux1	.69
	4	.3.2. Fusion de scores par des méthodes usuelles de l'état de l'art 1	.70
	4	.3.3. Approche de fusion par somme pondérée basée sur le taux	-
	1	a occultation de la comparaison des iris – ICORws	.74
	г	paramètres de fusion – BApf 1	.86
4.	4 Mé	éthode multimodale par l'iris droit, le visage et l'iris gauche	.89
	4	.4.1. Fusion multimodale par les règles usuelles de l'état de l'art	.89
	4	.4.2. Comparaison à la méthode biomodale par l'iris droit et le visage 1	93
	4	.4.3. Comparaison à la campagne IV ² 1	.94
4.	.5 Im	pact du rejet des comparaisons selon le taux d'occultation des iris1	.97
C	onclusi	on1	.99
Conclu	sion g	énérale et perspectives2	01
Référei	nces bi	ibliographiques2	05
Liste d	es nuh	lications	18
			10
ANNEX	KEA:	Intervalles de confiance pour l'estimation des erreurs d'un système biométrique2	19
ANNEX	KEB:	Performances de la méthode multimodale par l'iris droit, le visage et l'iris gauche	21
ANNEX	KEC:	Description des algorithmes multimodaux de la campagne IV ² 2	24

Table des figures

1. 1 Evolutions du marché biométrique mondial par technologies. A droite : parts de
marché en 2009, à gauche : prévisions des parts de marché en 20174
1. 2 Processus de fonctionnement d'un système biométrique
1. 3 Systèmes en mode d'identification en groupe ouvert et en groupe fermé10
1. 4 Système en mode vérification10
1. 5 Illustration du FAR et du FRR (a) à travers leur variation en fonction du seuil de
décision et (b) à travers les distributions des authentiques et des imposteurs12
1. 6 Courbes DET : FRR en fonction du FAR, (à gauche) en échelle linéaire et (à droite)
en échelle logarithmique13
1. 7 Courbe ROC : TAR en fonction du FAR
1. 8 Les multiples sources d'information dans un système biométrique multimodal14
1. 9 Architectures d'un système multimodal (a) en série, (b) hiérarchique, (c) parallèle
(adapté à partir de [Fierrez 2006])17
1. 10 Les différents niveaux de fusion dans un système biométrique multimodal (adapté
depuis [Ross 2006])19
1. 11 Fonction de normalisation des scores par double sigmoïde, $t = 200$, $r_1 = 20$ et $r_2 =$
3024
1. 12 Exemple illustrant les étapes de calcul des paramètres d'un estimateur Tanh 25
1. 13 Exemple de normalisation de scores par Tanh26
$\textbf{1. 14} \ \text{Cabine d'acquisition IV}^{2}\text{, capteurs dédiés et illustration des données acquises.} \dots \dots 31$
1. 15 Echantillons de la base d'iris IV ² (a) image d'iris de bonne qualité, (b-d) images
d'iris dégradées par (b) le flou, (c) les lunettes et (d) les occultations32
1. 16 Evolution de la qualité des images de la base IV^2 et ICE (allant de 0 à 1 en partant
de la qualité dégradée vers la bonne qualité), d'après [Petrovska-Delacretaz 2008]
1. 17 Echantillons d'images de visage 2D de la base IV ² , incluant des variabilités
d'expression, d'éclairage, de qualité et de session
2. 1 Anatomie de l'iris : (a) photographie en gros plan [Manvelyan 2015], (b) coupe
transversale, (c) secteur frontal ; (b) et (c) sont extraits de [Wildes 1997], page 3, Fig.2 38
2. 2 Différentes étapes du processus de reconnaissance de l'iris
2. 3 Différence de qualité entre des images du même iris capturé (a) dans le visible, (b)
dans le proche infrarouge
2. 4 Résultats du traitement de l'iris par la méthode de Daugman. Les contours en blanc
montrent les résultats de la localisation de l'anneau de l'iris et des paupières. L'iris code
en haut à gauche est le résultat de l'encodage de phase des coefficients d'ondelettes de
Gabor complexes 2D ayant servi à la caractérisation du modèle de l'iris. Figure extraite
de [Daugman 2004]
2. 5 Résultats du traitement de l'iris par la méthode de Wildes : (a) image initiale ; (b)
image segmentée par la transformée de Hough ; (c) représentation multi-échelle par les
filtres Laplaciens. Images extraites de [Wildes 1997], Figures 5, 6 et 742
2. 6 Résultats du traitement de l'iris par la méthode de Boles et Boashash : (a) image
initiale ; (b) Exemple d'une signature d'iris (signal situé sur un cercle virtuel) de l'iris de

l'image (a); (c) Trois niveaux de résolution de la transformée en ondelettes de la signature (b) ; (d) Représentation des passages par zéro de (c). Images extraites de [Boles **2.8** Illustration idéalisée de la décomposition spectrale par une pyramide orientable à 3 **2. 9** Décomposition d'une image d'iris par une PO à 3 échelles (s = 1..3) et à (n+1) = 4orientations (0, π/4, π/2 et 3π/4)......51 2. 10 Fonction de transfert d'un filtre de Gabor de symétrie paire avant une large bande 2. 11 Comparaison entre (a) la fonction Gabor et (b) la fonction Log-Gabor vues selon l'échelle linéaire (à gauche) et l'échelle logarithmique (à droite). Figure extraite de [Field 2. 12 (a) Processus de linéarisation selon le modèle « rubber sheet » de Daugman avec décalage exagéré entre les centres de l'iris et de la pupille pour des raisons d'illustration. 2.14 (a) image de CASIA V1.0 (masquée), (b) image originale de CASIA-IrisV3-Interval....57 2. 15 (a) Distributions des autorisés (rose) et des imposteurs (bleu) : (b) Courbe DET de la méthode de vérification d'iris par PO sur la base CASIA V1......60 2. 16 (a) Distributions des autorisés (rose) et des imposteurs (bleu) ; (b) Courbe DET de la méthode de vérification d'iris de Masek comparée à la méthode proposée sur la base 2. 17 Etapes de normalisation et de génération d'un masque final (e) à partir d'une image initiale d'iris (a) redimensionnée à la taille de son cercle externe ; (b) étant une reproduction de (a) avec les pixels de bruit marquées en noir, (c) l'iris déroulé et (d) le 2. 20 Distributions inter et intra-classes des scores de la méthode de référence Log-2. 21 Courbes DET associées à la méthode de référence Log-Gabor de reconnaissance de 2. 23 EER et nombre de comparaisons de l'algorithme de référence Log-Gabor sur les 2. 24 EER et nombre de comparaisons de l'algorithme de référence Log-Gabor en fonction du taux d'occultation des iris de la base IV²......69 2. 25 Distributions inter et intra-classes des scores de la méthode Log-Gabor correspondant aux seuils 100% et 35% d'occultation des iris......70 2. 26 Courbes DET associées à la méthode Log-Gabor correspondant aux seuils 100% et 35% d'occultation des iris : (à gauche) échelle linéaire et (à droite) échelle logarithmique.....70 2. 27 Exemple de taux d'occultation d'une comparaison entre deux iris (haut) et illustration de leur masque union (bas)......71 2. 28 EER et nombre de comparaisons de l'algorithme de référence Log-Gabor en fonction du taux d'occultation des comparaisons entre les iris de la base IV².72 2. 29 Etapes du prétraitement dans une chaîne de vérification biométrique de visage : (a) image initiale, (b) détection de la posistion du visage et des yeux, (c) image pivotée, (d) image redimensionnée, (e) région d'intérêt (visage) normalisée......75

2. 30 Catégorisation des méthodes principales de reconnaissance faciale
2. 31 Décomposition d'une image de visage par une PO à 3 échelles (s = 13) et à $(n+1)$ =
4 orientations $(3\pi/4, \pi/2, \pi/4 \text{ et } 0)$
2. 32 Courbes DET de l'algorithme de la PO appliqué aux quatre expériences de la compagne \mathbb{N}^2
2 33 Los sing promiors <i>Figurfages</i> coloulés à partir de l'onsemble d'appropriesage W^2
2. 35 Les cinq premiers <i>Eigenfaces</i> calculés à partir de l'ensemble d'apprentissage IV ²
2. 34 Les cinq premiers <i>Fisnerfaces</i> calcules à partir de l'ensemble d'apprentissage 1v ² 86
2. 35 Principe d'Adaboost
2. 36 Schéma fonctionnel de l'algorithme d'Adaboost
2. 37 Evolution du taux d'énergie résiduelle en fonction de la variance de l'espace de
réduction de l'ensemble d'apprentissage IV ² 90
2. 38 Evolution du EER en fonction de la variance de l'espace de projection pour les
quatre expériences de la base IV ² 90
 2. 39 Performance d'Adaboost en termes d'EER en fonction du nombre d'itérations du modèle d'apprentissage, pour l'expérience 1 d'IV²
2. 40 Génération du vecteur caractéristique final d'une image initiale par l'algorithme de
la PCA/PO (à droite) et les 4 configurations testées des bandes passe-bas d'une PO à 3 échelles et 6 orientations (à gauche). Les bandes pleines sont gardées dans toutes les
2 41 Import de la variation (a) du nombre d'arientations – nombre d'échelles firé à 2
2. 41 Impact de la variation (a) du nombre d'orientations – nombre d'échelles – nombre d'échelles – nombre d'échelles – $\frac{1}{2}$
2 42 Cánáration des vectours constánisticues d'une image initials per l'algorithme de la
2. 42 Generation des vecteurs caracteristiques d'une image initiale par l'algorithme de la $I D A/PO$, avec une PO à 3 échelles et 6 orientations solon le configuration n°2 (bandes en
LDAFFO, avec une FO a 5 echenes et 6 orientations seion la computation il 2 (bandes en
2 13 Application de l'algorithme Adabast/PO sur la PO aptière (3 échelles 6
orientations, Config.2) pour la caractérisation et la classification d'un couple d'images candidates à comparer
2. 44 Performance de l'algorithme Adaboost/PO entière en EER en fonction du nombre
d'itérations dans le modèle Adaboost pour l'expérience 1 d'IV ²
2. 45 Application de l'algorithme Adaboost/PO sur chaque sous bande d'une PO à 3
échelles et 6 orientations (Config.2) pour la caractérisation et la classification d'un
couple d'images candidates à comparer
9.1 Observice since the main second s
3. I Chemin visuel dans le cerveau numain.
3. 2 Reponses des cellules ganglionnaires a centre "ON" (a gauche) et a centre "OFF" (a
droite) a un stimulus lumineux. Le jaune represente la zone du champ recepteur
3. 3 (a) Cellules nerveuses de la retine et (b) modalités de reponses en modes centre UN
et centre OFF a des stimuli d'illumination centrale et (c) peripherique (extrait de
[Dubuc 2013])
3. 4 Aires visuelles et lobes cervicaux
3. 5 Schema de la répartition des différentes fonctions visuelles, intégratives et
cognitives du cortex cérébral humain, à partir de [Birbaumer 2006]118
3. 6 Modèle du système neuronal distribué pour la perception de visage extrait de
[Haxby 2010]
3. 7 Illustration de la prosopagnosie (d'après [Wawrzyniak 2009])120
3. 8 Illusion d'optique Arbre aux dix visages
3. 9 Approximation de l'opérateur Laplacien o ² G (ligne continue) par une DOG (ligne en
pointillés). Les deux profils sont très similaires (extrait de [Marr 1980])124

3. 10 Comparaison des estimations des réponses de cellules X à des enregistrements électrophysiologiques. La première ligne illustre l'estimation S de la réponse par le filtre	
nour un hord isolé une barre mince de largeur w (où w est la largeur de la région	
d'excitation centrale du champ récepteur), et une barre large de largeur 2.5*w. La	
deuxième ligne correspond à la partie positive de S et la quatrième ligne à l'opposé de sa	
partie négative. Les troisième et cinquième lignes correspondent aux réponses	
physiologiques (extrait de [Marr 1979]).	124
3. 11 Méthode des Bubbles : (a) image initiale, (b) images filtrées selon les différentes	
bandes de fréquences (échelles), (c) représentation des <i>bubbles</i> pour chaqur échelle	
constituant l'échantillonnage aléatoire, (d) parties du visage révélées pour chaque	
échelle et (e) visage résultant de la sommation des informations issues de chaque échelle	
(un stimulus) (figure extraite de [Gosselin 2001])	126
3. 12 (a) Régions significatives pour la tâche de reconnaissance du visage et (b)	
histogramme indiquant la proportion des pixels significatifs dans les quatre premières	
échelles. Figure extraite de [Schyns 2002]	127
3. 13 Trois exemples de voisinages pour le calcul de motifs LBP avec $(P; R)$ valant (8;	
1), (12 ; 2,5) et (16 ; 4) en partant de la gauche vers la droite	130
3. 14 Exemple de calcul d'un LBP de base sur un voisinage (<i>P</i> =8 ; <i>R</i> =1)	131
3. 15 Réponse fréquentielle d'une Différence de Gaussienne – DoG	132
3. 16 Exemple de représentation de quatre bandes de fréquence contenant chacune une	
octave d'une décomposition par DoG	133
3. 17 Exemple de décomposition par filtre DoG: (a) image originale, (b) images	
décomposées sur quatre échelles, (c) images LBPU2 de (b) subdivisées en blocs, (d)	
vecteur caractéristique final issu de la concaténation des histogrammes des blocs des	
images de (c).	134
3. 18 Influence du nombre d'échelles de la DoG sur la méthode LBP/DoG (LBP, bloc	105
$(16 \times 16), a1 = 0, b)$	135
3. 19 Selection des regions d'interet sur l'image du visage selon les échelles de	197
2 20 Examples de configurations nouv l'étiquetage des POI dens l'image de visage	197
2. 21 Etanos de la méthode L PD/DeC his inenjuée : (a) image aviginale. (b) décomposition	197
De G sur quatro áchallas (a) sólaction das BOIs (d) application I BPU ² at subdivision on	
blocs (e) calcul des histogrammes par blocs (f) affectation des coefficients de	
pondération $w_{\rm e}$ et $w_{\rm r}$ (g) concaténation des histogrammes pondérés (h) distance x ² entre	
histogrammes H1 et H2.	139
3. 22 Image originale et ses 7 décompositions <i>fine-to-coarse</i> après filtrage par une	
fonction DoG (échelle a_m allant de 0,5 à 32), m étant le niveau de décomposition	142
3. 23 Schéma fonctionnel de la reconstruction asynchrone basée sur le codage par ordre	
de classement et comparaison avec la reconstruction <i>coarse-to-fine</i>	145
3. 24 Exemple comparatif entre reconstruction asynchrone (à droite) et reconstruction	
coarse-to-fine (à gauche) : (a) images reconstruites, (b) images des différences entre les	
images reconstruites et l'image originale	147
3. 25 Evolution de l'EER en fonction des niveaux de reconstruction coarse-to-fine (vert)	
et asynchrone (rouge) pour (a) des images de taille (64×64) et (b) des images de taille	
(128×128) Erreur ! Signet non déf	ini.
3. 26 Evolution des critères adaptatifs – (a) écart-type, (b) entropie et (c) contours perdus	
- en fonction des niveaux de reconstructions <i>coarse-to-fine</i> (bleu) and asynchrone (rouge)	
sur une base de 2725 images	151

3. 27 Représentation en niveaux de gris de la carte des indices locaux $IndC_{sort}$ d'une image reconstruite de manière asynchrone. 154
3. 28 Distribution des échelles de décomposition (axe vertical) sur les huit réponses maximales (axe horizontal) de la reconstruction asynchrone. Les statistiques sont données pour tous les pixels des 2725 images de la base de données (2725×128×128
 155 3. 29 Exemple d'image de la base IV² reconstruite de manière asynchrone en utilisant le 1^{er} et le 4^{ème} rang de reconstruction avec les images différence correspondantes
 4. 1 Distributions du nombre d'images par sujets dans les listes de références originales des modalités de l'iris (à gauche) et du visage (à droite) de la base IV²
/ Visage
 4. 3 Distributions originales des scores authentiques et imposteurs des méthodes monomodales de l'iris droit (à gauche) et du visage (à droite)
 4. 5 Performances de la règle de fusion par produit associée aux normalisations par Tanh (à gauche) et zScores (à droite) sur l'approche bimodale Iris_D/Visage_Exp1. En haut sont montrées les distributions normalisées des scores des approches monomodales correspondantes, au milieu : les courbes DET (écholles linéaire et logarithmique) et en
 bas : les distributions des scores authentiques et imposteurs de l'approche bimodale
4. 7 Schéma fonctionnel du module de fusion en ligne et hors ligne des scores d'une m éthode d'authentification multimodale par l'iris et le visage (règle de fusion : ws + ICORws)
4. 8 Courbes DET en échelle linéaire représentant la performance de l'approche bimodale Iris_D/Visage pour les quatre expériences IV ² en fonction des associations des méthodes de Normalisation/Fusion et comparaison avec les approches monomodales correspondantes
4. 9 Courbes DET en échelle logarithmique représentant la performance de l'approche bimodale Iris_D/Visage pour les quatre expériences IV ² en fonction des associations des méthodes de Normalisation/Fusion et comparaison avec les approches monomodales correspondantes
4. 10 Influence de la variation du <i>Seuil</i> d'COR sur la proportion des comparaisons affectées par la pondération en ligne
4. 11 Influence de la variation du Seuil de l'association Tanh/ICORws sur les distributions des scores de l'approche bimodale Iris_D/Visage pour les quatre expériences IV ²
4. 12 Performances de la règle de fusion par pondération en ligne ICORws proposée, associée à la normalisation Tanh sur la méthode bimodale Iris_D/Visage pour les expériences 1 et 2. Pour chaque expérience sont montrées en haut : les distributions normalisées des scores des méthodes monomodales correspondantes, au milieu : les

courbes DET (échelles linéaire et logarithmique), en bas : les distributions des scores authentiques et imposteurs de la méthode himodale				
4 13 Performances de la règle de fusion nar nondération en ligne et hors ligne proposée				
[ws+ICORws] associáe à la normalisation zScores sur la máthoda himodale Iris. D/Visage				
nour les expériences 3 et 4 Pour chaque expérience sont montrées en haut : les				
distributions normalisées des scores des méthodes monomodales correspondantes au				
milieu : les courbes DET (échelles linéaire et logarithmique) en bas : les distributions				
des scores authentiques et imposteurs de la méthode himodale				
4 14 Courbes DET en échelles linéaire (haut) et logarithmique (has) représentant				
l'amélioration de la performance d'associations de méthodes de normalisation/fusion par				
l'agrandissoment de la BAnf avec l'annroche himodale Ivis D/Visage Evn1				
4 15 Couvhos DET (écholle logerithmique) représentant le performance de l'enprese				
4. 15 Courbes DET (echene logarithinique) representant la performance de l'approche				
Normalization/Fusion at comparaison avec los approchas monomodalos correspondentes 102				
16 Distribution des secons suthantiques et imposteurs des associations de méthodes				
4. 16 Distribution des scores authentiques et imposteurs des associations de methodes				
Luia DAliaara/Luia C aun l'américanas 4				
152 A 17 Comparison do la conference 4.				
4. 17 Comparaison de la performance de l'approche multimodale Iris_D/visage/Iris_G a				
celle de l'approche bimodale Iris_D/Visage pour les quatre expériences IV ² 193				
4. 18 Comparaison de la performance de l'approche proposee par fusion de l'iris droit et W_{2}^{2} de chaine de chaine de l'iris droit et W_{2}^{2} de chaine de c				
gauche et du visage a ceux des participants à la campagne IV ² par fusion du visage 2D,				
du visage 3D et de l'iris droit (extraits de [Colheau 2007])				
4. 19 Influence du rejet des comparaisons suivant le taux d'occultation des iris sur les				
performances des approches multimodales proposees a deux (a gauche) et a trois				
echantilions biometriques (a droite).				
4. 20 Influence du rejet des comparaisons suivant le taux d'occultation des iris sur l'EER				
et sur la composition des bases de test des approches multimodales proposées à deux (à				
gauche) et a trois echantillons biometriques (a droite)				
B. 1 Performances de l'association Tanh/Somme sur la méthode multimodale				
Iris_D/Visage/Iris_G pour l'expérience 1				
B. 2 Performances de l'association Tanh/Somme sur la méthode multimodale				
Iris_D/Visage/Iris_G pour l'expérience 2				
B. 3 Performances de l'association Tanh/Somme sur la méthode multimodale				
Iris_D/Visage/Iris_G pour l'expérience 3				
B. 4 Performances de l'association Tanh/Somme sur la méthode multimodale				
Iris_D/Visage/Iris_G pour l'expérience 4				

Liste des tables

2. 20 Influence de l'inertie cumulée sur l'algorithme de la PCA/PO (Exp1)97
2. 21 Influence de configuration des sous-bandes de la PO sur l'algorithme PCA/PO98
2. 22 Comparaison de la PCA/PO à ses méthodes individuelles
2. 23 Influence de l'orientation de la bande dans l'application de la LDA sur une PO à 6
orientations et 3 echelles dans les 4 experiences IV ²
2. 24 Influence de configuration des sous-bandes de la FO sur l'algorithme LDA/FO
2. 26 Comparaison d'Adaboost/PO à ses méthodes individuelles
2. 27 Exemples de la répartition des caractéristiques sélectionnées par Adaboost sur
chaque bande de la PO avec et sans pondération du nombre de caractéristiques 108
2. 28 Comparaison entre les différentes configurations de l'algorithme Adaboost/PO
entière. 3 b à 3 e : Adaboost sur chaque bande de la PO. Description : nombre de
caractéristiques ; ratio = pondération du nombre de caractéristiques par bande ; s =
fusion des scores par somme ; ws = fusion des scores par somme pondérée109
2. 29 Résultats comparatifs entre les algorithmes combinés proposés (bleu) avec les
algorithmes individuels correspondents (vert) et ceux de la campagne d'evaluation IV^2
9.1 Influence de la meleur de l'échelle initiale de la DeC sur la méthode LDD/DeC (LDD
3. I influence de la valeur de l'échelle initiale de la DoG sur la méthode LBF/DoG (LBF, $NBO = 4$, bloc (16×16))
3. 2 Influence du motif LBP sur la méthode LBP/DoG (NBO = 5, $a_1 = 0.5$, bloc (16×16)) 135
3. 3 Influence de la taille des blocs sur la méthode LBP/DoG (LBP ^{U2} , NBO = 5, $a_1 = 0,5$)135
3. 4 Comparaison de la Méthode LBP/DoG avec la méthode LBP135
3. 5 Poids des régions et des échelles dans la méthode LBP/DoG bio-inspirée
3. 6 Influence de l'approche bio-inspirée (sélection ROIs + pondération) sur la méthode LBP/DoG
3. 7 Résultats comparatifs entre l'algorithme LBP/DoG bio-inspiré proposé (vert) et ceux de la compagne d'évaluation IV ² (noir)
3. 8 Impact en EER du prétraitement sur la performance de la PCA pour les images de taille (64×64) et (128×128). Les échelles variant respectivement de 0.5 à 32 et de 0.5 à
64. Les meilleurs résultats sont surlignés en jaune et les seconds meilleurs en bleu
3. 9 Impact du prétraitement adaptatif sur la performance de la PCA (en EER)152
3. 10 Reconstruction Asynchrone versus Reconstruction <i>coarse-to-fine</i> : histogrammes de
la répartition des images sur les rangs de reconstruction nécessaires pour atteindre les
3. 11 Comparaison des meilleures méthodes de reconnaissance de visage proposées
4. 1 Caractáristiques de la base d'apprentissage chimérique des paramètres de fusion et
des listes de comparaisons d'apprentissage associées
4. 2 Caractéristiques des bases de test réelles expérimentées et des listes de
comparaisons de test correspondantes167
4. 3 Association des méthodes de normalisation aux méthodes de fusion de scores usuelles dans l'approche bimodale Iris_D / Visage_Exp1171
4. 4 Influence de l'association des méthodes de normalisation aux méthodes de fusion de
scores sur les performances de l'approche bimodale Iris_D / Visage pour les quatre
experiences $1v^2$ (EEK en %, intervalle de conflance de conflance a 90% et FKK au point FAR = 1%)
4. 5 Caractéristiques des bases d'apprentissages des paramètres de fusion et des listes
de comparaisons d'apprentissage associées

4. 6 Performances des associations de méthodes de normalisation/fusion suite à	
l'agrandissement de la BApf avec l'approche bimodale Iris_D/Visage Exp11	87
4. 7 Comparaison de la performance des associations de méthodes de normalisation/fusion de scores sur l'approche multimodale Iris droit / Visage / Iris gauche pour les quatre expériences IV^2 (EER en %, intervalle de confiance à 90% et	
FRRs au points FAR = 1% et 0,1%)	91
4. 8 Comparaison entre les caractéristiques des bases de données expérimentées en fusion multimodale des iris droit et gauche et du visage 2D avec celles de la campagne IV ² en fusion multimodale des visages 2D et 3D et de l'iris droit19	.95
A. 1 Table de la loi normale centrée réduite – Table Z2	20
 C. 1 Description des algorithmes monomodaux participant dans la conception des algorithmes multimodaux de la campagne d'évaluation IV²	24
campagne d'évaluation IV ²	24
C. 3 Résultats expérimentaux des algorithmes monomodaux de la modalité du visage 2D ayant participé dans la conception des algorithmes multimodaux de la campagne d'évaluation IV ²	25
C. 4 Description et performances des algorithmes multimodaux de la campagne IV ² pour l'expérience multi-session Test 03a	25
C. 5 Description et performances des algorithmes multimodaux de la campagne IV ² pour l'expérience multi-session Test 03b	25
C. 6 Description et performances des algorithmes multimodaux de la campagne IV ² pour l'expérience mono-session Test 04a	26
C. 7 Description et performances des algorithmes multimodaux de la campagne IV ² pour l'expérience mono-session Test 04b	26

Introduction générale

Historique

L'utilisation de la biométrie non automatisée comme moyen d'authentification remonte au début de la civilisation humaine. Dès 300 ans avant JC, les potiers assyriens utilisaient l'empreinte du pouce comme forme précoce d'identité de marque de leurs marchandises. Un autre exemple des débuts de la biométrie est apparu en Chine classique à travers l'utilisation de signatures manuscrites gravées sur des tampons par les empereurs de Chine pour signer leurs actes officiels. La première instance de système biométrique juridique d'authentification est apparue au cours de la dynastie des Tang (618-906 AD) où les empreintes digitales ont été utilisées pour signer des contrats.

Le développement de systèmes biométriques contemporains peut être considéré comme une conséquence de l'effort d'experts en criminologie et d'organismes judiciaires pour identifier et classer les criminels à la fin du XIXe et au début du XXe siècles. En 1882, Alphonse Bertillon, un criminologue français, a introduit un système de mensurations corporelles appelé anthropométrie pour identifier les criminels. Ce système d'identification a été adopté dans toute l'Europe, puis aux États-Unis, et utilisé jusqu'en 1970 où il a été détrôné par l'utilisation des empreintes digitales. Jusque dans les années 1980, les experts policiers devaient recouper manuellement des milliers de fiches réparties dans différents fichiers régionaux. Ce n'est qu'en 1987 que le procédé fut automatisé par le logiciel Morpho System de SAGEM à travers la création du Fichier Automatisé des Empreintes Digitales (FAED) géré par la police scientifique.

De nos jours, grâce à la puissance de calcul des ordinateurs et aux possibilités de stockage de données, associées à des programmes informatiques complexes, la biométrie ne se limite plus aux empreintes digitales et son utilisation n'est plus réduite à un cadre d'application de loi. Aujourd'hui, une politique de renforcement de l'utilisation des technologies biométriques est engagée par les gouvernements, mais aussi par le secteur privé afin de lutter contre le terrorisme et la fraude, étant donné les enjeux sécuritaires et économiques colossaux qui s'y rattachent.

Dans ce qui suit, nous allons énumérer les principaux secteurs d'application de la biométrie puis nous allons préciser quelques définitions et propriétés biométriques. Ensuite, nous allons présenter les motivations derrière notre choix de la multi-biométrie par l'iris et le visage dans ce travail de thèse et nous finirons par décrire l'organisation de ce manuscrit.

Domaines d'application

Bien qu'il existe de nombreuses applications potentielles de la biométrie, les principales peuvent être divisées en quatre catégories [Lease 2009]. La première est celle des systèmes de sécurité ou **systèmes d'accès logique**. Leur rôle est de surveiller, restreindre ou autoriser l'accès aux données ou aux informations. Dans ce cas, la biométrie remplace ou complémente les codes pin, les mots de passe et les jetons. Le volume et le chiffre d'affaires de l'industrie financière et banquière (notamment à travers le commerce électronique), en plus de la valeur des données personnelles sensibles véhiculées et/ou sauvegardées dans les réseaux et les ordinateurs, font de la biométrie pour sécuriser l'accès logique une industrie beaucoup plus déployée que celle de la sécurité physique.

La deuxième catégorie d'applications de la biométrie concerne les systèmes d'accès aux installations ou **systèmes d'accès physique**. Leur rôle consiste à surveiller, restreindre ou autoriser le mouvement d'une personne dans ou hors une zone spécifique. Dans ce cas, la biométrie remplace ou complète les clés et les cartes d'accès, permettant aux utilisateurs autorisés l'accès aux zones sécurisées. Les systèmes d'accès physique sont souvent déployés dans les principaux périmètres de l'infrastructure publique, tels que les aéroports, les lieux publics très fréquentés (musées, ...) et les installations frontalières, afin de surveiller et de limiter les mouvements de personnes non autorisées ou suspectes. En plus de l'entrée à des espaces sécurisés, les systèmes d'accès physique appliqués dans un cadre commercial constituent un outil d'aide à la gestion de ressources humaines à travers les systèmes de contrôle d'assiduité des employé en combinant la vérification d'accès à un endroit avec l'enregistrement du moment où l'authentification a été produite.

La troisième catégorie concerne les systèmes biométriques qui assurent l'**unicité des individus**. Ces systèmes se concentrent généralement sur la prévention de la double inscription dans les programmes ou les applications. Leur utilisation principale se produit dans le secteur public notamment dans les programmes d'attribution d'aides sociales ou dans les systèmes de vote.

Enfin, les **applications gouvernementales** constituent la dernière et la plus importante catégorie d'application biométrique. Elles couvrent principalement les domaines militaires et de sécurité intérieure ainsi que le domaine judiciaire (criminologie, contrôle des établissements pénitentiaires). Elles couvrent également les secteurs civils tels que les services gouvernementaux, l'éducation, le transport et le domaine de la santé.

Définitions et propriétés

Une définition exacte de la biométrie est donnée par Jain et al. la désignant comme la science d'établir automatiquement l'identité d'un individu sur la base de ses caractéristiques physiologiques ou comportementales [Jain 2007]. Ainsi, la biométrie vérifie l'identité d'un individu par ce qu'il est et non par ce qu'il possède (clé, carte d'accès, ...) ni par ce qu'il sait (mot de passe), ce qui la rend moins vulnérable, que les moyens classiques d'identification, aux tentatives de contrefaçon.

En théorie, la plupart des traits physiologiques ou comportementaux humains peuvent être utilisés en tant que modalités biométriques. Toutefois, pour tenir dans un système biométrique, potentiellement précis, pratique et rentable, le trait/caractéristique utilisé doit également satisfaire à une série d'exigences proposées dans [Jain 1999] :

- *l'universalité* : toute personne doit posséder le trait biométrique,
- *l'unicité* : une probabilité quasi nulle que deux personnes soient les mêmes selon le trait caractéristique biométrique,
- la permanence : la stabilité du trait biométrique dans le temps,
- *la mesurabilité* : la quantification du trait biométrique d'une manière pratique,
- *la performance* : la précision et la vitesse de la reconnaissance à travers la caractéristique biométrique,
- l'acceptabilité : l'accord du public pour la mesure de la caractéristique,
- *la non-circonvention* : le degré de facilité/difficulté avec laquelle le système peut être trompé en falsifiant la caractéristique biométrique.

Les modalités adoptées dans les systèmes biométriques possèdent ces propriétés mais à des degrés différents. Un compromis est alors réalisé lors du choix de la modalité en fonction des besoins de l'application biométrique.

De nombreuses modalités biométriques ont été proposées et sont utilisées dans des applications variées. Les modalités physiologiques se basent sur des caractéristiques morphologiques ou biologiques et comprennent le visage, l'oreille, l'iris, la rétine, les empreintes digitale et palmaire, la géométrie de la main, le réseau veineux, l'ADN, l'odeur de la peau, ... Les modalités comportementales utilisent un trait personnel du comportement tel que la voix, la dynamique de la signature, la démarche, la dynamique de frappe au clavier ou encore le mouvement des lèvres. Certains de ces éléments biométriques ont un long historique et peuvent être considérés comme des technologies matures, tandis que d'autres sont encore de jeunes arènes de recherche.

Acuity Market Intelligence (Acuity MI), entreprise mondiale de conseil en stratégies des technologies émergeantes dotée d'une expérience avérée à anticiper avec précision l'évolution du marché biométrique, fournit des données sur l'évolution des parts de marché de diverses technologies biométriques [Acuity 2009]. Selon les données publiées par Acuity MI en 2009 et ses prévisions pour l'année 2017 (cf. **Figure 1. 1**), l'empreinte digitale connaitra une baisse au profit de l'iris, le visage, la voix, la signature et les veines.

En mettant à part le système gouvernemental AFIS (Automated Fingerprint Identification System), il est prévu que la modalité de l'empreinte digitale détienne toujours la première place (25,4%). L'iris montera en seconde position (18,8%) à la place du visage qui sera au troisième rang (14,3%). Viendront ensuite les modalités utilisant les veines (5,8%) et la voix (5,5%). Il est attendu que la signature persiste sur le marché biométrique avec une légère augmentation (1,57%) contrairement à la modalité basée sur la main qui perdra sa part du marché (0,0%).



Figure 1. 1 Evolutions du marché biométrique mondial par technologies. A droite : parts de marché en 2009, à gauche : prévisions des parts de marché en 2017.

Limites de la monomodalité et solution multimodale

Bien que les applications biométriques à petite et à moyenne échelle (quelques centaines d'utilisateurs) puissent toujours utiliser les systèmes à modalité unique (appelés systèmes monomodaux ou unimodaux), ces derniers souffrent de plusieurs limites qui entravent leur usage lorsqu'il s'agit de construire un système de reconnaissance de haute précision pour des applications à grande échelle [Jain 2004]. Parmi ces limites, citons :

- Le bruit dans les données acquises : Des exemples de données bruitées sont une empreinte digitale avec une cicatrice ou une voix altérée par un rhume. Mais les données bruitées sont généralement dues à des capteurs défectueux ou mal entretenus (par exemple, l'accumulation de la poussière sur un capteur d'empreinte digitale, un mauvais focus de caméra entraînant du flou) ou à des conditions ambiantes défavorables (par exemple, un mauvais éclairage). Les données biométriques bruitées risquent alors d'être mal appariées avec les modèles de la base de données (cf. **Figure 1. 1**) aboutissant à un rejet incorrect d'un utilisateur autorisé.
- Les variations intraclasses : Les données biométriques acquises d'un individu lors de l'authentification peuvent être très différentes des données à partir desquelles a été généré le modèle de cet individu lors de la phase d'enrôlement. Cette variation peut être causée par une mauvaise interaction entre l'utilisateur et le capteur (par exemple en changeant de pose) ou lorsque les caractéristiques du capteur sont modifiées (par exemple, en changeant de capteur), ou par une variation des conditions de l'environnement ambiant (par exemple, un changement d'éclairage) ou même par une variation de la constitution psychologique de l'individu (induisant par exemple un changement d'expression, de ton, de dynamique, que ce soit pour la démarche, la signature ou la frappe au clavier).
- Le manque d'individualité : Alors qu'un trait biométrique est sensé varier considérablement d'un individu à l'autre, il peut y avoir de grandes similitudes interclasses par rapport à ce trait. Par exemple, les vrais jumeaux ont une

apparence faciale quasi-identique. Cette limitation réduit la capacité de discrimination du système biométrique qui présente plus de risques d'accepter des personnes qui ne devraient pas l'être.

- La non universalité : L'utilisation d'un système biométrique unimodal suppose que chaque individu de la population ciblée possède la modalité en question. Hors, cette hypothèse n'est pas toujours vérifiée, et un sous-ensemble de la population risque d'être exclu par le système monomodal si aucune autre alternative ne lui est proposé. Par exemple, une personne muette ne peut pas être enrôlée dans un système de reconnaissance de la voix, ni une personne présentant une paralysie de la main dans un système de reconnaissance de la signature. Dans ces cas, les personnes ne possèdent pas la biométrie et des erreurs d'enrôlement (« Failure To Enroll » ou FTE) sont générées par le système. Un autre problème de non universalité est de posséder le trait biométrique mais pas d'une manière exploitable par le système biométrique. Par exemple, une personne souffrant de cataracte oculaire ne fournira pas d'images d'iris d'assez bonne qualité pour une reconnaissance automatique, ou une personne souffrant d'une sécheresse extrême des doigts ne fournira pas d'images d'empreintes digitales exploitables. Dans ce cas, le système générera une erreur de capture (« Failure To Capture » ou FTC).
- La sensibilité aux attaques : Il est vrai que les traits biométriques sont beaucoup plus difficiles à contrefaire que les moyens classiques d'identification tels que les mots de passe et les cartes d'accès. Cependant, les impostures existent en particulier dans les modalités comportementales telles que la signature et la voix. Les traits physiologiques sont également sensibles aux attaques d'usurpation. En particulier, les empreintes digitales peuvent être reproduites à l'aide de silicone et le visage par une photographie.

Plusieurs limites rencontrées avec les systèmes biométriques unimodaux sont surmontées en ayant recours à la multimodalité. Les systèmes multimodaux sont basés sur l'utilisation simultanée de plusieurs traits biométriques (par exemple les empreintes digitales et le visage). Ces systèmes sont en mesure de satisfaire aux exigences de performance rigoureuses imposées par diverses applications. Ils apportent une solution au problème de la non universalité car, à partir de plusieurs traits, ils assurent une couverture beaucoup plus large de la population. Les systèmes biométrique multimodaux sont également plus fiables en raison de la présence de multiples pièces de preuve d'identité. Ainsi, même si deux individus différents se trouvent confondus par rapport à un trait biométrique (problème de non-individualité), ils seront différenciés par rapport à l'autre. Aussi, les différents traits biométriques ne sont pas affectés de la même manière par le bruit dans les données acquises ni par les variations intraclasses. Par exemple, la qualité de l'image d'un visage est affectée par un changement d'éclairage mais pas celle d'une empreinte digitale. De même, une cicatrice sur un doigt affecte la reconnaissance d'une empreinte digitale mais pas la reconnaissance par les veines. En outre, les systèmes multimodaux offrent de nombreuses possibilités de contrer les tentatives d'usurpation. En premier lieu, il est plus difficile pour un intrus d'usurper à la fois plusieurs traits biométriques d'un utilisateur légitime. Ensuite, les systèmes multimodaux permettent d'établir un type d'authentification question/réponse pour s'assurer qu'un utilisateur est bien présent au moment de l'acquisition des données, par exemple, en demandant à l'utilisateur de présenter un sous-ensemble aléatoire de caractéristiques biométriques (telles que l'index droit et le majeur droit, dans cet ordre).

Contexte et motivations

Ma première expérience dans le domaine de la biométrie remonte au sujet de mon mémoire de mastère avec la reconnaissance de l'iris [Khiari 2008]. Cette expérience m'a permis de m'initier à ce domaine passionnant et de comprendre les contraintes à surmonter lors de la conception d'un système biométrique. A l'issue de cette expérience, la multimodalité constituait une alternative à fort potentiel permettant de compléter le travail réalisé avec la modalité de l'iris. Une collaboration est alors née entre l'équipe « Image et formes » dans mon laboratoire de recherche d'origine LR-SITI (ancienne UR-TSIRF) à l'ENIT en Tunisie et l'équipe SIMOB (ancienne TADIB) dans mon laboratoire d'accueil IBISC à l'université d'Evry Val d'Essonne en France. Cette collaboration a été créée dans le but de valoriser les compétences acquises par les deux laboratoires dans les domaines de la biométrie de l'iris (LR-SITI) d'une part et de la biométrie du visage (IBISC) d'autre part. Le choix a porté sur la fusion de l'iris et du visage pour plusieurs raisons. D'abord, la modalité de l'iris est considérée parmi les plus fiables. Sa stabilité au cours du temps, son unicité même entre vrais jumeaux et sa signature de petite taille constituent d'autres atouts majeurs. De plus, l'iris est difficile à pirater grâce à un test de la réactivité de la pupille (dilatation/contraction) par rapport à la quantité de lumière. Cependant, l'iris reste moyennement mesurable car il s'agit d'extraire une petite zone qui risque d'être occultée (paupières, cils, spots lumineux, ...). Aussi, bien que la capture de l'image d'un iris se fasse à distance et sans contact, elle requiert la coopération de l'utilisateur pour garantir un minimum de qualité (garder la tête droite, les yeux ouverts, ne pas bouger, regarder la caméra, ...). Ces contraintes d'acquisition et ce manque d'acceptabilité se trouvent réduits voire annulés dans la modalité du visage qui est dotée d'un caractère non intrusif offrant jusqu'à la possibilité de réaliser un travail à la volée. De plus, grâce à son faible coût de déploiement, le visage est une modalité privilégiée pour faire partie d'un système multimodal. Sa performance sensible aux changements de l'environnement ambiant est corrigée par l'efficacité de la modalité de l'iris. Ainsi, la complémentarité entre les propriétés biométriques de l'iris et du visage a guidé notre choix pour ces deux modalités.

Organisation du manuscrit et contributions

Le manuscrit de thèse est organisé comme suit :

Le **Chapitre 1** sert à situer le contexte de ce travail de thèse. Il introduit, d'abord, les critères d'évaluation des systèmes biométriques puis il présente les multiples aspects de la multimodalité et les différents niveaux de fusion possibles, en particulier celle des scores. Ce chapitre décrit également la base de données et le protocole IV^2 avec lesquels sont réalisés les tests de toutes les méthodes proposées.

Les deux chapitres suivants sont consacrés à la proposition de méthodes biométriques monomodales par l'iris et le visage et le dernier chapitre est dédié à l'introduction de méthodes multimodales.

Le **Chapitre 2** est constitué de deux grandes parties : Dans la première partie, suite à une étude bibliographique des méthodes de reconnaissance de l'iris, nous réunissons plusieurs modules de référence de l'état de l'art pour construire une méthode biométrique par l'iris. Nous proposons deux critères de qualité, qui constituent une petite première contribution, pour améliorer les performances de la reconnaissance. La deuxième partie du **Chapitre 2** est dédiée à la biométrie faciale par le biais d'approches classiques. Cette partie comprend trois contributions de cette thèse à travers la proposition de trois méthodes d'authentification qui sont basées sur la combinaison d'algorithmes avec plusieurs niveaux de fusion et qui sont évaluées par rapport à la campagne IV² sur des expériences comprenant des variations d'éclairage, de qualité et de session d'enregistrement.

Dans le **Chapitre 3**, où nous proposons deux contributions majeures de cette thèse, nous abordons le problème de la reconnaissance faciale par des approches bio-inspirées. Suite à la description du mécanisme de la perception du visage dans le cerveau humain, nous en présentons quelques modélisations par des algorithmes fondamentaux utilisés en biométrie faciale. Puis, nous élaborons deux approches bio-inspirées s'appuyant sur certaines de ces modalisations. La première approche est une méthode d'authentification de visage qui utilise les résultats de travaux neuroscientifiques indiquant l'importance des régions et des échelles de décomposition utiles à l'identification d'un visage. La deuxième approche bio-inspirée consiste à appliquer une méthode de codage par ordre de classement dans la phase de prétraitement pour renforcer le contenu informatif des images de visage.

Les meilleures approches monomodales par l'iris et par le visage proposées dans les **Chapitres 2 et 3** sont associées dans le **Chapitre 4** pour aboutir à l'objectif initial de ce sujet de thèse : la conception d'une méthode multimodale. Suite à la description de la conception de la base multimodale et du protocole de test, nous élaborons, en fait, deux approches multimodales en utilisant une série de techniques classiques de normalisation et de fusion de scores. La première approche multimodale associe les échantillons de l'iris droit et du visage. Cette approche comprend la principale contribution de ce chapitre à travers la proposition d'une nouvelle règle de fusion basée sur un critère de qualité proposé dans le **Chapitre 2**. La deuxième approche multimodale repose sur une reconnaissance par double iris et par le visage. Les deux approches multimodales proposées sont testées sur des environnements contrôlés et non contrôlés d'acquisition de visage de la base multimodale réelle IV² et plusieurs aspects de la multimodalité sont discutés.

Enfin, nous dresserons une conclusion générale sur ces travaux et nous en présenterons de nombreuses perspectives futures.

Chapitre 1

Concepts biométriques et base de données

Intro	oducti	ion	9
1.1	Mode	es opératoires d'un système biométrique	9
1.2	Eval	uation d'un système biométrique	11
1.3	Bion	nétrie et multimodalité	13
	1.3.1	Taxonomie des Systèmes biométriques multimodaux	
	1.3.2	Architectures d'un système multimodal	
	1.3.3	Niveaux de fusion multimodale	
	1.3.4	Fusion au niveau des scores	
	1.3.5	Normalisation des scores	22
1.4	Base	es de données et protocoles de test	26
	1.4.1	Conception des bases multimodales	
		a) Bases chimériques et bases réelles	
		b) Architecture d'une base multimodale	
	1.4.2	La base multimodale IV ²	
	1.4.3	Les évaluations réalisées sur la base IV ²	
		a) Base d'iris et campagne d'évaluation IV ²	
		b) Base de visage 2D et campagne d'évaluation IV ²	
		c) Protocole IV ² d'évaluation multimodale	

Introduction

Un système biométrique est essentiellement un système de reconnaissance de formes qui fonctionne en quatre étapes. La première est celle de l'acquisition des données biométriques de l'utilisateur. Ensuite, vient l'étape d'extraction des caractéristiques à partir des données acquises, éventuellement précédée d'une phase de prétraitement. La troisième étape se fait à travers la comparaison des caractéristiques extraites contre le modèle figurant dans la base de données et ce en vue de générer des mesures de similarité. En dernier lieu, une étape de décision sert à conclure sur l'identité de l'utilisateur (cf. **Figure 1. 2**). Le point de départ pour le système biométrique est la phase d'enrôlement. Dans cette phase, les données biométriques d'un utilisateur sont initialement collectées et traitées dans un modèle : forme sous laquelle elles sont ensuite stockées pour une utilisation permanente. Les modèles ne sont pas des données brutes ou des images numérisées d'un échantillon biométrique, mais ils sont une représentation mathématique de caractéristiques distinctives extraites par le système biométrique [Jain 2007].



Figure 1. 2 Processus de fonctionnement d'un système biométrique

Les systèmes biométriques peuvent être conçus selon différents modes de fonctionnement qui seront spécifiés dans le paragraphe suivant. Nous décrirons, ensuite, les outils d'évaluation d'un système d'authentification biométrique. Puis, nous aborderons les multiples aspects de la multimodalité en mettant l'accent sur la fusion au niveau des scores. A la fin du Chapitre, nous décrirons la base de données IV² sur laquelle nous avons testé l'ensemble des systèmes biométriques proposés dans cette thèse, ainsi que le protocole IV² ayant servi aux expériences monomodales et multimodales.

1.1 Modes opératoires d'un système biométrique

Il existe deux modes possibles pour un système biométrique : le mode d'identification qui répond à la question « **Qui suis-je** ? » et le mode de vérification (ou d'authentification) qui répond à la question « **Suis-je la personne que je déclare** être ? » [Jain 2007].



Figure 1. 3 Systèmes en mode d'identification en groupe ouvert et en groupe fermé.



Figure 1. 4 Système en mode vérification.

Dans le mode d'identification, aussi appelée reconnaissance « *one-to-many* », la comparaison se fait entre le modèle biométrique du candidat et les modèles de tous les utilisateurs stockés dans la base d'enrôlement. Deux contextes sont envisageables. Le premier est de travailler dans le cadre d'un groupe fermé, c'est-à-dire que l'on est sûr que le candidat appartient à la base de données des autorisés et le problème revient à déterminer laquelle des identités correspond au mieux à ce candidat. Généralement, c'est ce contexte de groupe fermé qui est adopté dans les travaux de recherche. Le deuxième contexte est de travailler dans le cadre d'un groupe ouvert, c'est-à-dire qu'il se peut que le candidat soit un imposteur qu'il faut rejeter. Ainsi, le système peut générer deux décisions : le rejet ou l'acceptation avec la détermination de l'identité du candidat (cf. **Figure 1. 3**).

Dans le mode de vérification, appelée aussi authentification, il s'agit toujours d'un contexte de groupe ouvert, c'est-à-dire que l'on n'est absolument pas sûr que l'identité du candidat est réellement connue par le système. En pratique, le candidat revendique l'identité d'un des individus enregistrés dans la base de données. La comparaison se fait alors uniquement entre le modèle biométrique du candidat et les modèles de l'individu annoncé. Ainsi, il s'agit d'une reconnaissance « *one-to-one* » (cf. **Figure 1. 4**). L'ensemble des systèmes biométriques proposés dans le cadre de cette thèse relèvent du mode de vérification.

1.2 Evaluation d'un système biométrique

Malgré les nombreux avantages des systèmes d'authentification biométriques, leur mise en œuvre comporte un certain nombre de risques. Même le système biométrique le plus précis n'est pas parfait et des erreurs seront générées.

Dans cette thèse, nous allons utiliser des indicateurs de performance qui comparent différents algorithmes, en particulier, dans le cadre d'une authentification d'identité. Nous n'allons pas nous intéresser aux autres indicateurs de performance tels que la complexité de calcul et les ressources requises en mémoire. Les méthodes proposées dans ce travail de thèse ont été testées à travers une implémentation basique sur un ordinateur doté d'un processeur Intel (R) Core (TM) I5, avec une horloge de 2,67 GHz, une RAM de 4Go, et un système d'exploitation Windows7 64 bits.

Un système d'authentification biométrique peut produire deux décisions possibles : l'acceptation ou le rejet d'un candidat. Deux erreurs sont alors susceptibles d'arriver :

- les fausses acceptations : lorsque le système accepte des imposteurs en les considérant à tord comme étant des personnes autorisées,
- les faux rejets : lorsque le système rejette des personnes authentiques en les considérant à tord comme étant des imposteurs.

La mesure de ces taux d'erreurs se fait alors selon les équations (1. 1) et (1. 2) en fonction du seuil de décision τ . Le **Taux de Fausse Acceptation** ou **FAR** (*False Acceptance Rate*) est calculé à partir du rapport entre le nombre de fausses acceptations (c'est-à-dire le nombre d'imposteurs mal-classés) sur le nombre total d'accès imposteurs dans la base. De même, le **Taux de Faux Rejet** ou **FRR** (*False Reject Rate*) est calculé en divisant le nombre de faux rejets (nombre d'authentiques mal-classés) sur le nombre total d'accès authentiques (selon le protocole suivi dans ce travail, un accès imposteur est une comparaison entre deux images provenant de deux individus différents, alors qu'un accès authentique est une comparaison entre images du même individu. cf. § 1.4).

$$FAR(\tau) = \frac{FA(\tau)}{N_{Imp}}$$
(1.1) $FRR(\tau) = \frac{FR(\tau)}{N_{Aut}}$ (1.2)

L'évolution de ces taux d'erreurs en fonction du seuil de décision peut être visualisée à travers les distributions de scores Authentiques et Imposteurs comme représenté sur la Figure 1. 5-a, ou alors en traçant les valeurs de ces erreurs pour chaque seuil, comme illustré dans la Figure 1. 5-b.



Figure 1. 5 Illustration du FAR et du FRR (a) à travers leur variation en fonction du seuil de décision et (b) à travers les distributions des authentiques et des imposteurs.

Idéalement, le choix du seuil de décision devrait correspondre à la valeur de t qui minimise à la fois le FAR et le FRR. Or, ces deux types d'erreurs ont des monotonies contraires puisque FAR est croissant, alors que FRR est décroissant. Le choix du seuil devra alors faire l'objet d'un compromis entre sécurité et confort, qui dépendra des besoins de l'application. Certaines applications exigent un taux de fausse acceptation FAR très faible (voire quasi nul, par exemple, lors de l'accès à des documents confidentiels dans un contexte militaire). D'autres applications ne tolèrent pas un taux de faux rejet FRR élevé (tel que l'accès à un téléphone portable). Dans les systèmes ayant des types d'erreurs qui nécessitent un compromis, la performance d'un système biométrique est évaluée en plusieurs points de fonctionnement ou encore à travers des courbes de performance.

Un point de fonctionnement particulier est l'**EER** (*Equal Error Rate*) ou Taux d'égales erreurs qui est atteint lorsque FAR et FRR sont égaux. Ce point est souvent utilisé en raison de son caractère neutre et indépendant du type de l'application.

Un autre point de fonctionnement est le **FRR en une valeur fixée du FAR**. Si par exemple, le FAR est fixé à 1% ou à 0,1%, le système requiert un taux d'imposture relativement faible (1/100 ou 1/1000) et sa performance est donnée par le FRR en cette valeur de FAR.

Par ailleurs, la courbe de performance la plus utilisée dans la littérature est la **courbe DET** (Detection Error Tradeoff) qui trace le FRR en fonction du FAR pour les différents seuils de décision τ [Martin 1997]. Cette courbe est d'autant meilleure qu'elle est proche de l'origine du repère. L'EER peut y être facilement détecté par l'intersection avec la première bissectrice y = x. La courbe DET, généralement représentée en échelle logarithmique (cf. **Figure 1. 6**), est différente de la **courbe ROC** (Receiver Operating Characteristic), une autre courbe de performance couramment utilisée, qui trace le taux de rejet correct ou *True Accept Rate* (TAR = 1-FRR) en fonction du FAR et qui est représentée en échelle linéaire (cf. **Figure 1. 7**). Une courbe ROC est d'autant meilleure qu'elle se rapproche de l'angle supérieur gauche (FAR proche de zéro et TAR proche de 1).



Figure 1. 6 Courbes DET : FRR en fonction du FAR, (à gauche) en échelle linéaire et (à droite) en échelle logarithmique.



Figure 1. 7 Courbe ROC : TAR en fonction du FAR.

Au cours des évaluations des méthodes monomodales proposées, nous allons fournir les EERs et les courbes DET suivies d'interprétations et de discussions. Dans l'évaluation des méthodes multimodales proposées, nous fournirons en plus les taux de FRR aux deux points où le FAR est égal à 1% et 0,1%. Les taux d'erreurs seront accompagnés d'intervalles de confiance à 90% permettant d'évaluer leur signification statistique en termes de précision (cf. **Annexe A**).

1.3 Biométrie et multimodalité

Cette partie sera dédiée à l'étude de plusieurs aspects de la multimodalité. D'abord, nous exposerons les multiples sources d'information des systèmes multimodaux. Ensuite, nous spécifierons les architectures possibles d'un système multimodal. Puis, nous présenterons les différents niveaux de fusion multimodale en particulier la fusion de scores et la phase de normalisation qu'elle requiert.

1.3.1 Taxonomie des Systèmes biométriques multimodaux

Le principe de la multimodalité au sens large est de combiner plusieurs sources d'information à partir de systèmes monomodaux [Jain 2004]. Selon les sources d'information combinées, cinq scénarios de multimodalité sont envisageables (cf. **Figure 1.8**):

- *Systèmes multi-capteurs* : lorsque différents capteurs sont utilisés pour acquérir le même trait biométrique, par exemple, une caméra digitale, une webcam et une caméra thermique pour la capture d'images de visage.
- *Systèmes multi-échantillons :* lorsque plusieurs échantillons du même trait/modalité sont associés, par exemple, les images des iris gauche et droit ou les empreintes digitales de l'index et du majeur.
- Systèmes multi-instances : lorsque plusieurs instances du même trait, du même échantillon biométrique et ayant été acquises par un même capteur sont associées, par exemple, plusieurs images de visage prises avec des changements de pose (frontale ou de profil), d'expression ou d'éclairage [Allano 2009].
- *Systèmes multi-algorithmes* : lorsque le traitement d'une image se fait par la combinaison de plusieurs algorithmes que ce soit dans le module d'extraction de caractéristiques ou celui de comparaison.
- *Systèmes multi-biométries* : ou systèmes multi-traits ou encore systèmes multimodaux au sens strict du terme, lorsque plusieurs modalités/traits biométriques sont combinés, par exemple, la voix et la signature ou le visage et l'iris.



Caméra digitale et webcam (IV²)

 $Figure \ 1.\ 8 \ Les \ multiples \ sources \ d'information \ dans \ un \ système \ biométrique \ multimodal.$

A partir de ces scénarios, on peut imaginer toute sorte de combinaison. Une comparaison de l'intérêt des scénarios multimodaux a été réalisée dans [Tissé 2003] suivant trois critères (cf. **Table 1. 1**) : le gain en performances, le coût matériel du système, et enfin sa consommation en temps d'acquisition ou d'exécution du traitement numérique.

Table 1. 1 Comparaison des scénarios multimodaux suivant les critères de l'amélioration de la performance, du coût matériel et de la consommation du temps. +/- désigne le niveau de satisfaction (adapté à partir de [Tissé 2003]).

	multi-capteurs	multi- échantillons	multi- instances	multi-algorithmes	multi- biométries
	(One)	10	B		
Performance en FRR et FAR	+	++	-	+	+++
Coût matériel	-	++	++	+	
Consommation en temps	-	-	+	+ +	+

A vrai dire, aucun scénario ne peut seul répondre à toutes les exigences d'un système biométrique. D'après la comparaison de la **Table 1. 1**, ces exigences sont relativement moins satisfaites par les scénarios *multi-capteurs* et *multi-instances* par rapport aux autres (ayant 4 + ou plus).

Le scénario *multi-échantillons* souffre également du rallongement du temps d'acquisition dû à la capture de plusieurs échantillons du même trait biométrique à partir d'un seul capteur. Toutefois, l'utilisation d'un unique capteur est attrayante en termes de coût matériel et fait du scénario multi-échantillons un bon compromis pour augmenter les performances du système de manière significative sans nécessiter de coût supplémentaire.

Le scénario *multi-algorithmes* améliore les performances du système mais de manière moins significative que le scénario *multi-échantillons*. Ceci est dû au fait que la performance de chaque algorithme combiné dépend étroitement de la qualité d'un unique exemple initialement acquis, d'une même biométrie et à travers un même capteur. Le coût d'un tel système est davantage lié aux unités de calcul (plate-forme matérielle) qu'au capteur.

Le scénario de la *multi-biométries* est potentiellement celui qui arrive à atteindre les meilleures performances en termes de FAR et de FRR. Cependant, ce scénario présente des inconvénients liés à la facilité et au coût d'utilisation. En effet, en utilisant plusieurs modalités biométriques, le temps d'acquisition et le temps de traitement risquent de se rallonger. Si l'on suppose que la capture des multiples biométries et le traitement de l'information peuvent s'effectuer de manière parallèle, alors ce scénario répond au critère de la consommation en temps. Par contre, l'utilisation de plusieurs capteurs implique forcément une augmentation du coût matériel. C'est pour cette raison qu'en pratique, les systèmes multimodaux comprenant plus de deux modalités ne sont quasiment jamais utilisés.
Maintenant, si l'on considère d'autres critères liés aux problèmes de la nonuniversalité et aux tentatives d'usurpation d'identité, le scénario de la *multi-biométries* est le seul qui soit capable d'y apporter des solutions. En effet, les quatre premiers scénarios combinent des informations issues toutes d'un unique trait biométrique, ce qui ne leur permet pas de résoudre efficacement tous les problèmes rencontrés par les systèmes monomodaux bien qu'ils arrivent à en améliorer les performances. Finalement, le choix d'un scénario, ou du mixage de plusieurs d'entre eux, est une affaire de compromis entre les différents critères et les besoins de l'application.

L'ensemble des méthodes que nous allons présenter dans ce travail de thèse couvrent la totalité des scénarios multimodaux envisageables. En effet, les méthodes de reconnaissance faciale qui seront proposées reposent toutes sur des *combinaisons d'algorithmes*. De plus, la base d'images sur laquelle nous travaillons présente des variabilités d'éclairage, d'expression et de qualité. Les deux premières variabilités nous situent dans le contexte d'un scénario *multi-instances*, alors qu'avec la dernière variabilité nous travaillons avec un système *multi-capteurs* (caméra digitale et webcam). Les deux scénarios restants seront couverts à travers les deux systèmes multimodaux que nous allons présenter dans le dernier chapitre. Ces systèmes se basent sur la *multibiométries* de l'iris et du visage et l'un d'entre eux combine les *deux échantillons* droit et gauche de l'iris.

1.3.2 Architectures d'un système multimodal

Un système biométrique multimodal peut être conçu selon trois architectures [Jain 2004] [Fierrez 2006] : l'architecture en série (ou en cascade), l'architecture en parallèle et l'architecture hiérarchique (ou en arbre). Une représentation graphique des trois catégories est donnée dans la **Figure 1.9**.

Dans *l'architecture en série* (**Figure 1. 9-a**), les systèmes individuels sont invoqués en séquence. Certains d'entre eux peuvent n'être utilisés que lorsqu'une condition éventuelle se produit à la sortie des systèmes invoqués précédemment, ce qui permet d'aboutir à une décision sans forcément faire intervenir la totalité de ces systèmes. Cette architecture peut être utilisée comme schéma d'indexation pour réduire le nombre d'identités possibles avant l'utilisation de la donnée suivante. Elle permet également d'augmenter l'efficacité en faisant appel d'abord aux systèmes peu coûteux et moins précis et ensuite aux systèmes onéreux mais plus précis.

Dans *l'architecture hiérarchique* (**Figure 1. 9-b**), les différents systèmes individuels sont combinés en une structure arborescente. Cette architecture est considérée comme étant la plus flexible et permet de faire face aux problèmes de données manquantes ou de mauvaise qualité souvent rencontrées dans les systèmes biométriques.

Dans *l'architecture en parallèle* (**Figure 1. 9-c**), l'information provenant des différents systèmes est utilisée simultanément pour réaliser la tâche de la reconnaissance. L'utilisation de toutes les informations biométriques est alors requise pour produire une décision, ce qui est susceptible d'apporter plus d'amélioration que

dans le cas d'une *architecture en série*. Ces avantages ont fait que la plupart des méthodes proposées dans la littérature appartiennent à cette catégorie d'architecture, ce qui est également le cas des méthodes proposées dans ce travail de thèse.

Le choix de l'architecture d'un système dépend des besoins de l'application. Dans les applications conviviales et à faible risque, *l'architecture en série* est privilégiée pour ses avantages de coût en temps et en matériel par rapport à *l'architecture en parallèle* nécessitant l'acquisition et le traitement d'un grand nombre de données biométriques. En revanche, dans les applications où la sécurité est d'une importance primordiale (par exemple, dans les applications militaires), l'utilisation d'une *architecture en parallèle* est plus appropriée. L'architecture arborescente est privilégiée dans les applications présentant un plus grand risque d'acquisition de données biométriques de mauvaise qualité (cas d'un système de contrôle d'assiduité à multiples échantillons par les empreintes digitales dans un milieu hospitalier). Elle est également utilisée dans les applications dans lesquelles l'une ou l'autre des modalités risque d'être manquante (cas d'un système de contrôle d'assiduité à multiples échantillons par les applications dans lesquelles l'une ou l'autre des modalités risque d'être manquante (cas d'un système de contrôle d'assiduité à multiple échantillons par les empreintes digitales dans une usine de bois).



Figure 1. 9 Architectures d'un système multimodal (a) en série, (b) hiérarchique, (c) parallèle (adapté à partir de [Fierrez 2006]).

1.3.3 Niveaux de fusion multimodale

Un système biométrique est constitué principalement de 4 modules : un module d'acquisition des données, un module d'extraction de caractéristiques, un module de mise en correspondance ou de *Matching* (appelé aussi, par abus de langage, module de « classification ») et un module de décision. La fusion multimodale peut s'opérer à la sortie de n'importe lequel de ces modules donnant lieu à quatre niveaux de fusion (cf. Figure 1. 10) qui peuvent être groupés en deux grandes familles : fusion avant la mise en correspondance, et fusion après la mise en correspondance [Jain 2005].

La fusion pré-classification : se fait avant le module de mise en correspondance, au niveau du capteur ou au niveau des caractéristiques.

- Fusion au niveau des capteurs : il s'agit de fusionner des données brutes provenant du ou des capteurs. Ce type de fusion ne peut se faire qu'entre différentes instances de la même biométrie et nécessite d'avoir des données compatibles, par exemple, la combinaison de plusieurs images de visage en différentes poses pour former un modèle 3D.
- Fusion au niveau des caractéristiques: il s'agit de combiner des vecteurs caractéristiques provenant de différents capteurs ou obtenus suite à l'application d'algorithmes différents sur une même donnée biométrique. Lorsque les vecteurs caractéristiques sont homogènes, par exemple plusieurs impressions d'empreinte digitale d'un doigt, une somme pondérée peut être appliquée pour obtenir un seul vecteur caractéristique. Si les vecteurs caractéristiques sont hétérogènes, par exemple s'ils sont issus de différents algorithmes de caractérisation ou de différents traits biométriques, et s'ils sont compatibles, ce qui n'est pas toujours le cas comme pour les minuties et les coefficients d'analyse en composantes principales d'une empreinte digitale, ils peuvent être concaténés en un unique vecteur caractéristique. La fusion au niveau des caractéristiques offre certes une information plus riche mais, néanmoins, elle est difficile à réaliser dans la pratique pour les raisons suivantes :
 - i) La dépendance entre les espaces caractéristiques à fusionner qui est généralement inconnue et qui nécessite d'éliminer les caractéristiques fortement corrélés.
 - ii) La concaténation des vecteurs caractéristiques risque de conduire à un espace caractéristique excessivement grand.
 - iii) Les vecteurs caractéristiques générés par les systèmes biométriques commerciaux sont la propriété des vendeurs qui, la plupart du temps n'y donnent pas accès.

L'une des approches classiques de reconnaissance de visage que nous allons proposer dans le chapitre suivant opère une fusion au niveau des caractéristiques.

Au vu des multiples contraintes imposées par la fusion pré-classification, la plupart des chercheurs s'orientent vers l'étude de schémas de **fusion post-classification** qui se font suite au module de mise en correspondance, au niveau des scores ou au niveau des décisions.

- *Fusion au niveau des décisions* : il s'agit de combiner les décisions des systèmes biométriques qui donnent chacun une réponse (accepté : 1 ou rejeté : 0, dans le cas de la vérification) selon l'entrée qui leur est présentée. Ce niveau de fusion peut s'opérer par des règles simples telles que le ET, le OU et le vote à la majorité, ainsi que par des règles plus complexes par vote pondéré ou par classification dans

l'espace des décisions [Jain 2004]. La fusion au niveau des décisions a l'avantage d'être simple. Par contre, l'information dont elle fait usage est très limitée (0 ou 1).

Fusion au niveau des scores : il s'agit de combiner les scores issus de chacun des modules de comparaison individuels. Les scores ont l'avantage d'être indépendants des systèmes biométriques et donc beaucoup plus faciles d'accès que les caractéristiques. Les scores contiennent l'information la plus riche à propos du modèle d'entrée, après les vecteurs caractéristiques. De plus, ils permettent de surmonter les contraintes de compatibilité et de grande dimension de l'espace caractéristique rencontrées avec la fusion au niveau des caractéristiques. Ainsi, la fusion au niveau des scores donne le meilleur compromis entre la richesse de l'information et la facilité d'implémentation, ce qui fait d'elle l'approche la plus répandue dans les systèmes biométriques multimodaux. D'ailleurs, c'est l'approche que nous retenons dans toutes les méthodes classiques de reconnaissance du visage proposées ainsi que dans les méthodes multimodales par l'iris et le visage. Ce niveau de fusion sera présenté plus en détails dans § 1.3.4.



Figure 1. 10 Les différents niveaux de fusion dans un système biométrique multimodal (adapté depuis [Ross 2006]).

1.3.4 Fusion au niveau des scores

Dans le contexte d'un système de vérification, deux approches sont envisageables pour fusionner les scores issus des différents comparateurs ou *matchers* individuels : l'approche par classification et l'approche par combinaison de scores.

Dans **l'approche par classification**, un vecteur caractéristique formé par les scores des *matchers* individuels est attribué à l'une des classes « accepté » (appelé aussi autorisé / authentique / client) ou « rejeté » (imposteur). Généralement, le classifieur utilisé est capable de prendre une telle décision indépendamment de la manière avec laquelle le vecteur caractéristique des scores a été généré, ce qui permet de fusionner des scores non homogènes en sortie des différents *matchers* (distance ou similarité, différents intervalles de scores, ...) sans aucun traitement préalable avant de les passer au classifieur.

Plusieurs classifieurs ont été utilisés pour consolider les scores des matchers individuels et arriver à une décision finale. Nous pouvons citer, comme exemple, le travail de Wang et al. qui ont appliqué une LDA (Linear Discriminant Analysis) et un réseau de neurones **RBF** (Radial Basis Function) pour classifier un vecteur caractéristique 2D construit à partir de scores issus de deux modules de reconnaissance par le visage et par l'iris [Wang Y. 2003]. Ross et Jain ont proposé une fusion de scores par arbre décisionnel et par LDA dans un système multimodal par le visage, l'empreinte digitale et la géométrie de la main [Ross 2003]. Dinerstein et al. ont aussi fusionné les scores issus de trois modules de reconnaissance par le visage, par la géométrie de la main et par l'ADN moyennant un classifieur multi-SVM (Support Vector Machine) à noyaux RBF [Dinerstein 2007]. Un autre classifieur k-NN (k-Nearest Neighbor) a été utilisé pour affecter des poids statiques à des scores venant de trois modules différents de reconnaissance du visage par lumière visible et infrarouge afin de résister à la variation d'éclairage et aux occlusions [Annis-Fathima 2014]. Kumar et Devi ont fusionné les scores d'un système d'identification par l'iris et l'empreinte digitale via une méthode de regroupement (*clustering*) appelée k_means couplée à un classifieur SVM [Kumar 2014]. Fakhar et al. ont fusionné des scores de visage et d'empreintes digitales à l'aide d'un clasifieur flou basé sur l'intégrale supérieure et les mesures floues de Sugeno [Fakhar 2015].

D'un autre côté, **l'approche par combinaison** combine les scores issus des *matchers* individuels en formant un unique score qui sera utilisé par la suite pour prendre la décision finale. De ce fait, les scores des différents *matchers* doivent d'abord subir une normalisation pour être transformés dans un domaine commun assurant une cohérence des données lors de la combinaison. Kittler et al. ont développé un formalisme mathématique pour consolider la combinaison de plusieurs classifieurs en utilisant des schémas évidents tels que la règle somme, la règle produit, la règle maximum, la règle minimum entre autres [Kittler 1998]. Pour pouvoir utiliser ces schémas, les scores doivent, d'abord, être transformés en des probabilités *a posteriori* conformément à un utilisateur autorisé ou imposteur. Kittler et al. considèrent le problème de classer un modèle d'entrée Z en l'une des m classes possibles (où m = 2 pour un système de

vérification) d'après l'information fournie par R matchers différents. Soit \vec{x}_i le vecteur caractéristique (construit à partir du modèle d'entrée Z) présenté au $i^{\text{ème}}$ matcher. Soient $P(\omega_j | \vec{x}_i)$ les sorties des matchers individuels, c'est-à-dire les probabilités a posteriori du modèle Z appartenant à la classe ω_j sachant le vecteur caractéristique \vec{x}_i . Soit $c \in \{1, 2, ..., m\}$ le numéro de la classe ω_c à laquelle le modèle d'entrée Z est finalement assigné ; cpeut être alors déterminé en utilisant les règles suivantes [Jain 2005] :

La règle Produit : Cette règle est basée sur l'hypothèse d'une indépendance statistique des représentations $\vec{x}_1, \vec{x}_2, ..., \vec{x}_R$; ce qui est généralement vérifié dans le cas de multimodalité utilisant différents traits (iris, visage, empreinte digitale, ...). c est alors donné par l'équation (1.3) :

$$c = \arg\max_{j} \prod_{i=1}^{R} P(\omega_{j} | \vec{x}_{i})$$
(1.3)

La règle Somme : En plus de l'hypothèse d'indépendance statistique des représentations, la règle somme suppose que les probabilités *a posteriori* calculées à partir des *matchers* individuels ne dévient pas beaucoup par rapport aux probabilités *a priori*. Cette règle s'avère plus efficace que la règle produit dans le cas d'un haut niveau de bruit menant à une ambiguïté dans le problème de classification. Le modèle Z est assigné à la classe c par la règle somme selon l'équation (1. 4) :

$$c = \arg\max_{j} \sum_{i=1}^{R} P(\omega_{j} | \vec{x}_{i})$$
(1.4)

La règle Maximum : La moyenne des probabilités *a posteriori* est approximée par la valeur maximale. La classe *c* du modèle d'entrée *Z* est alors donnée par l'équation (1. 5):

$$c = \arg\max_{j} \left(\max_{i} P(\omega_{j} | \vec{x}_{i}) \right)$$
(1.5)

La règle Minimum : Cette règle approxime la classe c du modèle Z selon l'équation (1. 6) par le minimum des probabilités *a posteriori*.

$$c = \arg\max_{i} \left(\min_{i} P(\omega_{i} | \vec{x}_{i}) \right)$$
(1.6)

Ceci étant dit, l'hypothèse d'indépendance statistique des représentations $\vec{x}_1, \vec{x}_2, ..., \vec{x}_R$ peut ne pas être vérifiée dans un système biométrique multimodal utilisant différentes représentations caractéristiques et différents algorithmes de classification sur *le même trait biométrique* [Prabhakar 2002]. C'est pourquoi, d'autres alternatives ont été proposées. Certaines s'appuient sur une **estimation non paramétrique de densité** et utilisent un test de ratio de vraisemblance pour prendre la décision finale [Prabhakar 2002] [Murakami 2015]. D'autres utilisent des **statistiques Bayésiennes** prenant en compte la précision estimée des *matchers* individuels lors du processus de fusion [Bigun 1997].

La combinaison de scores peut également se calculer par une **somme pondérée** des scores issus des *matchers* individuels. Certains travaux assignent des poids statiques aux scores des *matchers* individuels [Wang Y. 2003]. D'autres proposent d'utiliser des

poids spécifiques à chaque utilisateur pour les différentes modalités [Jain 2002]. D'autres encore présentent une manière adaptative de combiner les scores par somme pondérée via un processus d'apprentissage de poids hors ligne se basant sur les EERs des *matchers* individuels et des poids en ligne attribuant une plus grande importance aux scores situés aux premiers rangs de l'identification [Ben Soltana 2010].

Il est important de noter que Ross et Jain ont montré que l'approche par combinaison de scores s'avère plus efficace que certaines approches par classification, telles que les arbres de décision et l'annalyse linéaire discriminante [Ross 2003]. Cependant, ils affirment qu'il n'existe pas de schema unique, que ce soit par classification ou par combinaison de scores, qui soit le meilleur en toutes circonstances. Aussi, lorsque les sorties des *matchers* individuels sont de simples scores, sans mesure de la qualité de l'échantillon, il est plus judicieux de combiner directement les scores en utilisant la méthode appropriée plutôt que de les convertir en probabilités [Ross 2003] [Jain 2005]. Ce type d'approche par combinaison sera adopté dans les méthodes monomodales par le visage et dans la fusion multimodale basée sur l'iris et le visage.

1.3.5 Normalisation des scores

Le choix d'une approche de fusion par combinaison de scores implique une étape préalable de normalisation avant la fusion proprement dite. Cette étape s'avère indispensable pour les trois raisons suivantes :

- les sorties des modules de correspondance ou *matchers* individuels peuvent être non homogènes (distances/similarités),
- les scores risquent de se situer dans des intervalles différents,
- les distributions statistiques des sorties de chaque *matcher* peuvent être différentes.

Ceci explique la nécessité de transformer les scores des *matchers* individuels dans un domaine commun avant de les combiner, par un changement des paramètres de position (moyenne) et d'échelle (écart-type) : la normalisation. Pour avoir une bonne normalisation, les estimateurs des paramètres de position et d'échelle doivent être à la fois robustes et efficaces. La robustesse se traduit par une insensibilité par rapport aux valeurs aberrantes ou « *outliers* » tandis que l'efficacité exprime la proximité de l'estimateur obtenu par rapport à l'estimateur optimal lorsque la distribution des données est connue. Une présentation détaillée des concepts de robustesse et d'efficacité des procédures statistiques se trouve dans [Huber 1981].

Parmi les techniques de normalisation les plus utilisées [Jain 2005], nous pouvons citer les suivantes :

Min-Max : C'est la technique la plus simple et la plus adaptée lorsque les valeurs minimales et maximales des *matchers* individuels sont connues. Il suffit alors de les translater, respectivement, vers 0 et 1. Cependant, même si les scores des *matchers* ne sont pas bornés, les bornes peuvent être estimées par un jeu de scores d'entraînement et appliquées par la suite. La normalisation min-max est définie pour un score donné s_k selon l'équation (1. 7), où min et max désignent les scores minimal et maximal du matcher k.

$$s_k' = \frac{s_k - \min}{\max - \min} \tag{1.7}$$

Dans le cas de l'estimation des bornes *min* et *max*, cette normalisation n'est pas robuste aux valeurs aberrantes qui sont situées en dehors de l'intervalle estimé par le jeu de scores d'entraînement.

Decimal-scaling : cette méthode peut être utilisée dans le cas où les scores des *matchers* individuels évoluent dans des intervalles à rapport logarithmique. Par exemple, les scores de l'un sont dans l'intervalle [0, 1] et les scores des autres dans l'intervalle [0, 1000]. La normalisation peut alors être appliquée selon l'équation (1. 8). Les défauts de cette approche sont le manque de robustesse et l'hypothèse de variation des différents *matchers* selon un facteur logarithmique.

$$s_k' = \frac{s_k}{10^{\log_{10}(max)}}$$
(1.8)

zScores : C'est la technique la plus courante (équation (1. 9)). Elle se calcule à partir de la moyenne arithmétique et de l'écart-type d'une distribution donnée. Elle a pour effet de centrer les scores des différents *matchers* autour de zéro. Il est attendu que ce schéma soit bon si on dispose de connaissances *a priori* sur le score moyen et les variations des scores du *matcher*. Dans le cas contraire, une estimation préalable de ces grandeurs est nécessaire à partir d'un ensemble de scores donné. Cependant, la normalisation zScores est sensible aux valeurs aberrantes, ce qui la rend non robuste. Enfin, il est vrai que cette technique est optimale s'il s'agit d'une distribution gaussienne car la moyenne et l'écart-type ne sont des estimateurs optimaux pour les paramètres de position et d'échelle seulement que si la distribution est gaussienne. Dans le cas d'une distribution arbitraire, ce ne sont plus des estimateurs optimaux mais ils restent raisonnablement bons selon [Jain 2005].

$$s_k' = \frac{s_k - \mu}{\sigma} \tag{1.9}$$

Médiane et écart absolu médian MAD (*Medium Absolute Deviation*) : La normalisation par la médiane et le MAD se fait selon l'équation (1. 10) et s'avère robuste car insensible aux valeurs aberrantes. Toutefois, contrairement à la moyenne et l'écart-type, la médiane et le MAD constituent des estimateurs faibles de la position et de l'échelle lorsque la distribution des scores est non gaussienne. C'est pourquoi cette technique est d'une faible efficacité, ne préserve pas la distribution d'entrée en sortie et ne transforme pas les scores normalisés des différents *matchers* en un intervalle commun.

$$s'_{k} = \frac{s_{k} - median}{MAD} \qquad \text{avec } MAD = median(|s_{k} - median|) \qquad (1. 10)$$

Fonction double sigmoïde : La normalisation double sigmoïde est donnée par l'équation (1. 11) où t désigne le point de référence, r_1 et r_2 les bornes entre lesquelles la fonction double sigmoïde est linéaire.

$$s_{k}' = \begin{cases} \frac{1}{1 + \exp\left(-2((s_{k} - t/r_{1}))\right)} & \text{si } s_{k} < t\\ \frac{1}{1 + \exp\left(-2((s_{k} - t/r_{2}))\right)} & \text{sinon} \end{cases}$$
(1.11)

Cette technique garantit la transformation des scores dans un intervalle commun entre 0 et 1 mais demande un réglage prudent des paramètres t, r_1 et r_2 pour avoir une bonne efficacité. Les scores qui se situent dans l'intervalle $[t-r_1, t+r_2]$ correspondent à la zone de recouvrement des distributions authentiques et imposteurs et subissent une transformation linéaire, alors que les scores en dehors de cette marge sont transformés d'une façon non linéaire. Un exemple de fonction de normalisation par double sigmoïde est présenté dans la **Figure 1. 11**, où t = 200, $r_1 = 20$ et $r_2 = 30$.



Figure 1. 11 Fonction de normalisation des scores par double sigmoïde, t = 200, $r_1 = 20$ et $r_2 = 30$.

Estimateurs Tanh (Tangente hyperbolique) : Ils ont été proposés par Hampel et al. et sont connus pour leur robustesse et leur efficacité [Hampel 1986]. Pour une distribution initiale des scores de développement située dans l'intervalle [0, 1] (cf. Figure 1. 12-a), une transformation préalable des scores, qui est proche de celle du min-max, est requise (cf. équation (1. 12)) pour translater les scores dans l'intervalle [-1, 1] (cf. Figure 1. 12-b).

$$st_k = \frac{2s_k - (max + min)}{max - min}$$
(1.12)

La première étape réside dans l'application de la fonction d'influence de Hampel définie par l'équation (1. 13) et dont la courbe est tracée sur la Figure 1. 12-c, sur les scores de développement transformés dans l'intervalle [-1, 1]. Cette fonction réduit l'influence des points situés aux extrémités de la distribution (identifiés par a, b et c) durant l'estimation des paramètres de position et d'échelle. Ce qui rend cette méthode insensible aux valeurs aberrantes.





Figure 1.12 Exemple illustrant les étapes de calcul des paramètres d'un estimateur Tanh.

Cependant, il y a un compromis à faire lors du choix des paramètres a, b et c. En effet, si l'influence d'un grand nombre de points aux extrémités de la distribution est réduite, l'estimateur est plus robuste mais moins efficace. Inversement, si plusieurs points aux extrémités influencent l'estimation, l'estimateur est plus efficace mais moins robuste. C'est pourquoi les paramètres a, b et c de l'estimateur de Hampel sont généralement choisis tels que 70% des scores translatés dans l'intervalle [-1, 1] soient inclus dans [m-a, m+a], 85% des scores inclus dans [m-b, m+b] et 95% des scores inclus dans [m-c, m+c], avec m le score médian. Selon [Jain 2005], le calcul des paramètres de la technique de normalisation Tanh conduit à de meilleurs résultats dans la phase de test, s'il se fait à partir de la distribution authentique plutôt que la distribution totale de l'ensemble de développement.

On obtient une nouvelle distribution de scores authentiques de Hampel (de développement), à partir de laquelle on calcule les paramètres : moyenne μ_{AH} et écart-

type σ_{AH} , comme illustré dans la **Figure 1. 12-d**. Ces paramètres serviront au calcul final des scores normalisés par les estimateurs Tanh de Hampel, dans la phase de test, selon l'équation (1. 14). La **Figure 1. 13** montre un exemple de normalisation des scores de test par l'estimateur Tanh de la **Figure 1. 12**. Les scores obtenus sont centrés sur 0,5.

$$s_{k}^{'} = \frac{1}{2} \left\{ tanh\left(0.01 \left(\frac{s_{k} - \mu_{AH}}{\sigma_{AH}} \right) \right) + 1 \right\}$$
Scores de test initiaux, $s_{k} \begin{bmatrix} 80 & 250 \end{bmatrix}$
Scores de test normalisés par Tanh, $s_{k}^{'} \begin{bmatrix} 80 \end{bmatrix}$



Figure 1. 13 Exemple de normalisation de scores par Tanh.

La **Table 1. 2** résume les caractéristiques de chacune des techniques de normalisation sus-présentées, en termes de robustesse et d'efficacité. Dans les fusions multimodales que nous allons présenter dans cette thèse, nous utilisons les techniques min-max pour sa simplicité, zScores pour son efficacité et Tanh à la fois pour son efficacité et sa robustesse.

Technique de normalisation	Robustesse	Efficacité
Min-Max	Non	N/A
Decimal scaling	Non	N/A
Z-Score	Non	Elevée (optimale pour des distributions gaussiennes)
Médiane et MAD	Oui	Modérée
Double sigmoïde	Oui	Elevée
Estimateurs Tanh	Oui	Elevée

Table 1. 2 Récapitulatif des techniques de normalisation de score.

1.4 Bases de données et protocoles de test

Dans les recherches en biométrie, il est crucial d'avoir un ensemble de données assez grand sur lequel l'efficacité des algorithmes proposés peut être évaluée. Certaines bases de données sont mises à la disposition des chercheurs. Cependant, la plupart d'entre elles n'offrent pas suffisamment de données que ce soit en matière de nombre ou de variabilité. C'est dans ce sens que la base de données IV² a été conçue.

Dans cette partie, nous allons d'abord donner quelques notions à propos des bases de données multimodales qui permettront de situer la base IV^2 parmi les bases multimodales existantes. Ensuite, nous détaillerons le contenu de la base IV^2 ainsi que les protocoles d'évaluation de la campagne IV^2 pour les modalités de l'iris et du visage ainsi que pour les tests multimodaux.

1.4.1 Conception des bases multimodales

Lorsqu'il s'agit de travailler en biométrie multimodale, plusieurs questions se posent d'abord au sujet de la base de données sur laquelle la performance de la méthode multimodale proposée sera testée. Cette partie apporte des éléments de réponses sur le type de la base de données expérimentée ainsi que sur son architecture.

a) Bases chimériques et bases réelles

Dans le domaine de la biométrie multimodale, le manque de bases de données réelles constitue un vrai problème. Le terme « réel » traduit le fait que les modalités biométriques proviennent de la même personne. Une autre alternative qui est devenue une pratique courante dans les expériences biométriques multimodales est la création de bases de données chimériques provenant de différentes bases de données [Dorizzi 2006]. Cette opération se fait par assignation arbitraire des échantillons d'une personne X d'une modalité aux échantillons d'une personne Y d'une modalité différente, en vue de créer des identités « virtuelles ».

i) Choix du type de la base

L'utilisation de bases de données chimériques est basée sur trois principaux arguments :

1- le manque de bases de données multimodales réelles publiques et la taille réduite des bases disponibles,

2- l'hypothèse de l'indépendance entre les modalités à fusionner,

3- un problème éthique lié à la vie privée : dans les bases multimodales réelles, le rassemblement de plusieurs traits biométriques d'une même personne au même endroit présente un risque plus élevé d'usurpation d'identité.

Bien que le problème lié à la vie privée se trouve en effet résolu avec les utilisateurs virtuels, des études réalisées au sein le l'école polytechnique fédérale de la Lausanne [Poh 2005] [Poh 2006] ont remis en question l'hypothèse de l'indépendance entre les modalités biométriques. Ces études ont permis de conclure que la performance mesurée sur une base de données chimérique n'est pas un bon prédicateur de la performance mesurée sur une base de données réelle [Poh 2005]. Par ailleurs, les travaux suivants des mêmes auteurs [Poh 2006] ont montré que l'utilisation de bases de données chimérique multimodal par comparaison à l'apprentissage de cette même règle de fusion sur une base de données réelle ; à condition que les tests soient réalisés uniquement sur une base de données réelle. L'avantage réside alors dans la possibilité de générer, de manière artificielle, des bases d'apprentissage de taille plus élevée et de palier ainsi le manque de données multimodales d'entraînement.

Des études postérieures conduites par [Scheidat 2007] visant à tester l'indépendance entre deux modalités ont prouvé qu'en général les tests sur une base chimérique conduisaient à une minoration de la performance d'un système biométrique comparés aux tests sur une base réelle. Dans leurs travaux, Scheidat et Vielhauer ont également montré que l'utilisation de données chimériques pour mesurer la performance n'était pas toujours justifiée et que l'hypothèse d'indépendance entre les modalités était naturellement dépendante des modalités utilisées. De plus, des algorithmes de comparaisons seraient plus sévèrement affectés par les dépendances qui pourraient exister entre les modalités.

Les deux études réalisées par Poh et Bengio et par Scheidat et Vielhauer ont abouti à la conclusion qu'il était plus approprié d'utiliser une base de test réelle pour évaluer la performance d'un système multimodal. Les travaux de Poh et Bengio ont démontré de plus que l'utilisation d'une base de données d'apprentissage chimérique n'améliore ni ne dégrade les performances. Sur cette base, nous avons opté pour l'utilisation d'une base d'entraînement chimérique pour apprendre les paramètres de fusion, et d'une base de test réelle pour mesurer la performance des méthodes multimodales proposées.

ii) Bases multimodales réelles existantes

De nos jours, il existe diverses bases de données multimodales réelles. La **Table 1.3** présente les principales caractéristiques de certaines d'entre elles. Parmi les bases comprenant à la fois les modalités de l'iris et du visage, la base Biosecure est la plus complète en termes de nombre de sujets. Cependant, elle n'est pas gratuite. Tel est aussi le cas des bases BioSecur-ID, BioSec et WVU/CLARKSON.

A notre connaissance, les seules bases de données multimodales réelles gratuites combinant les modalités de l'iris et du visage sont les bases MBGC (*Multiple Biometric Grand Challenge*) et IV^2 (Identification par l'Iris et le Visage via la Vidéo). Le laboratoire IBISC étant le coordinateur du projet IV^2 , nous avons eu l'opportunité d'avoir à disposition les protocoles de test ainsi que les rapports techniques de la campagne d'évaluation IV^2 . C'est pourquoi, nous avons choisi de travailler avec la base et le protocole IV^2 .

	Modalités	Nb. sujets	Projet	Gratuite	Autres détails
BT-DAVID [Chibelushi 1999]	Visage, voix	124	-	-	5 sessions, vidéos visage face/profil + enregistrements audio
XM2VTS [Messer 1999]	Visage 2D + 3D, voix	295	EU ACTS programme	Frais de production	Extension M2VTS, images visage haute qualité, fichiers audio, séquences vidéo, modèles 3D, 4 sessions 3 scénarios de qualité
[Bailly-Baillière 2003]	Visage, voix	208	BANCA	Non	12 sessions, images, séquences vidéo, fichiers audio
MCYT [Ortega-Garcia 2003]	Empreinte digitale, signature écrite	75 et 100	-	Sous- ensemble empreinte digitale	2 versions : signatures hors ligne et en ligne
BIOMET [Garcia-Salicetti 2003]	Visage, voix, empreinte digitale, forme de la main, signature écrite	130, 106, 91 selon session	Multi-modal Biometric Identity Verification	Frais de production	3 sessions
MyIdea [Dumas 2005]	Visage, voix, empreintes digitale et palmaire, signature, écriture, géométrie de la main	104	IM2 & BioSecure	Frais de production	Visage parlant, multi- capteurs, multi-scénarios
FRGC [Phillips 2005] [NIST FRGC 2006]	Visage 2D + 3D	741	FRGC challenge	Oui	Multisessions, env. contrôlé / non contrôlé (éclairage), variation expression
BioSec [Fierrez 2007]	Empreinte digitale, iris, voix, visage	250	Projet integré FP6 EU BioSec	Non	4 sessions
WVU/CLARKSON [Crihalmeanu 2007]	Iris, visage, voix, empreintes digitale et palmaire, géométrie de la main	Entre 112 et 274 selon modalité	JAMBDC	Non	Images et sequences vidéo
IV² [Petrovska-Delacrétaz 2008]	Visage 2D + 3D, iris	315, 77,19 selon session	IV ² , programme Techno Vison	Oui	3 sessions, env. contrôlé / non contrôlé (éclairage, qualité), Capteur iris portable
MBGC [Phillips 2009] [NIST MBGC 2010]	Visage, iris	-	MBGC challenge	Oui	Visage : éclairage visible contrôlé/non contrôlé Iris droit et gauche + NIR vidéo, capteur Iris On the Move
BioSecur-ID [Fierrez 2010]	Visage, empreintes digitale et palmaire, géométrie de la main, voix, iris, signature écrite, écriture, dynamique de frappe au clavier	400	BiosecurID	Non	4 sessions
BioSecure [Ortega-Garcia 2010]	Signature, empreinte digitale, main, visage, iris	1000, ~700, ~700 selon la session	Biosecure Network of Excellence	Non	Extension BioSec, 2 sessions multi-scénarios (3), multi-environnements
MOBIO [McCool 2012]	Visage, voix	160	MOBIO	Oui	capteurs mobiles 6 sessions

Table 1.3	Exemples de	bases	biométriques	multimodales	réelles.

b) Architecture d'une base multimodale

Lorsqu'il s'agit de travailler avec plusieurs modalités, il faut que la base de données multimodale soit structurée par utilisateur. Selon [Morizet 2009], au moins deux architectures sont possibles :

L'Architecture I comprend une seule base de données contenant un dossier (et son identifiant) par personne. Chaque dossier rassemble les données biométriques d'un utilisateur, provenant de différentes modalités biométriques. Cette architecture présente l'avantage d'être pratique à mettre en œuvre. Toutefois, le rassemblement de la totalité des traits biométriques d'une même personne au même endroit risque de rendre le système plus vulnérable aux pannes et aux tentatives d'usurpation d'identité.

L'Achitecture II comprend plusieurs bases de données (une base par modalité biométrique) contenant un dossier (et son identifiant) par personne. Par contre, chaque dossier ne contient que les informations biométriques de la modalité correspondante à la base de données à laquelle il appartient. Dans ce cas, les dossiers contenant les informations biométriques relatives à une même personne doivent avoir le même identifiant. Cette architecture est beaucoup plus lourde à gérer que la première mais permet de résoudre le problème éthique lié à la vie privée en évitant de concentrer toutes les données biométriques au même endroit ; par exemple en stockant les informations biométriques des utilisateurs sur différents serveurs. Elle permet également d'être plus efficace contre les dysfonctionnements qui peuvent mettre en panne l'exploitation d'une modalité biométrique, en offrant la possibilité de toujours faire fonctionner au moins l'autre modalité. La base multimodale IV^2 a été structurée selon l'architecture II.

1.4.2 La base multimodale IV²

La base de données publique IV^2 provient de 315 sujets. Elle a été conçue dans le but d'offrir une panoplie de situations de tests confrontant les algorithmes à des variabilités très critiques pour la performance des systèmes biométriques, telles que la pose, l'expression, l'éclairage, la qualité et la session. La base IV^2 permet d'effectuer des expériences monomodales mais aussi multimodales utilisant des données acquises sur le visage par différents capteurs. Un package d'évaluation ainsi que deux ensembles de données disjoints ont été constitués à partir de la totalité des données acquises, l'un pour la phase de développement (52 sujets) et l'autre pour la phase d'évaluation [Petrovska-Delacretaz 2008].

Les données présentes dans la base IV^2 et relatives au visage ou à des parties du visage sont : des séquences audio-vidéo 2D de visages parlants issues d'une caméra digitale à haute définition et d'une webcam (basse qualité d'image), des données stéréoscopiques acquises avec deux paires de caméras synchronisées, des données de visage 3D acquises par un scanner laser et des images d'iris acquises par une caméra infrarouge portable. Cette base de données a été prise par plusieurs opérateurs (Télécom SudParis, Laboratoire IBISC, Ecole Centrale de Lyon – LIRIS, Institut EURECOM) et à

des endroits différents. A chaque endroit, une cabine d'acquisition sous la forme d'un « photomaton » biométrique a été montée comme illustré dans la **Figure 1.14** [Lelandais 2007].



Figure 1. 14 Cabine d'acquisition IV², capteurs dédiés et illustration des données acquises.

1.4.3 Les évaluations réalisées sur la base IV^2

Une campagne d'évaluation a été lancée comme étape finale du projet IV^2 . Cette campagne concerne la reconnaissance de l'iris, la reconnaissance du visage 2D et 3D et la reconnaissance du visage parlant. Les algorithmes participant à cette campagne d'évaluation ont été ensuite fusionnés au niveau des scores pour faire l'objet d'une comparaison entre algorithmes de reconnaissance multimodaux.

L'évaluation d'un système biométrique monomodal nécessite d'avoir une base de données partitionnée en deux sous-ensembles disjoints, l'un pour le réglage des paramètres du système et éventuellement le choix du seuil de décision et l'autre pour le test des performances. Un système biométrique multimodal nécessite d'avoir une base de données répartie en trois sous-ensembles disjoints : en plus de l'ensemble de développement et celui de test, le troisième ensemble sert à l'apprentissage des paramètres de la fusion.

A partir de l'ensemble de test, des comparaisons sont établies entre un sousensemble de sujets enrôlés et un sous-ensemble de sujets candidats (ou de test) pour produire des scores de similarité. Une manière courante de travailler est de générer toutes les combinaisons possibles de comparaisons entre les images d'enrôlement et de test. Ceci conduit à un nombre pléthorique de comparaisons interclasses (entre deux images provenant de sujets différents) et un nombre limité de comparaisons intraclasses (entre deux images provenant du même sujet). Ce cas de figure est loin de refléter les conditions réelles d'application des systèmes biométriques où, en général, les accès clients (comparaisons intraclasses) sont beaucoup plus nombreux que les accès imposteurs (comparaisons interclasses). Le protocole IV^2 procède plutôt par un tirage aléatoire qui garantit à peu près le même nombre de comparaisons intraclasses et interclasses. Cette stratégie est plus proche de la réalité et permet d'optimiser le nombre de comparaisons à réaliser pour évaluer correctement la distribution des scores obtenue. Les comparaisons entre les images d'enrôlement et de test sont fournies sous forme de listes de comparaisons deux à deux, c'est-à-dire qu'une seule image d'enrôlement est comparée à une seule image de test. En se basant sur les résultats des comparaisons, le protocole évalue les performances du système en termes de FAR, de FRR et d'EER [Colineau 2007].

a) Base d'iris et campagne d'évaluation IV²

La base d'iris IV^2 contient 8141 images capturées à partir de 296 sujets. Elle présente plusieurs types de dégradations, comme montré dans la **Figure 1.15**.



Figure 1. 15 Echantillons de la base d'iris IV^2 (a) image d'iris de bonne qualité, (b-d) images d'iris dégradées par (b) le flou, (c) les lunettes et (d) les occultations.

La difficulté de travailler sur la base IV^2 provient non seulement de la nature des dégradations qui y sont présentes mais aussi du nombre d'images affectées par ces dégradations. Pour mieux illustrer les défis contenus dans la base d'iris IV^2 , nous allons la comparer à la base ICE2005 [Phillips 2008] [NIST ICE 2011], une base d'iris référencée [Daugman 2007] [Proença 2010], incluant elle aussi plusieurs dégradations de la qualité. Krichen et al. ont défini une mesure de qualité probabiliste globale basée sur les Modèles à Mixture de Gaussiennes – GMM – pour évaluer la qualité des images d'iris [Krichen 2007]. La **Figure 1. 16** montre les distributions de cette mesure de qualité pour des images de la base IV^2 et ICE. D'après les courbes de la **Figure 1. 16**, la base IV^2 ont une mesure de qualité inférieure à 0,5, 20% des images ont une mesure de qualité inférieure

à 0,6 et 50% des images ont une mesure de qualité inférieure à 0,7 contre, respectivement, seulement 2%, 8% et 23% de la base ICE. C'est pourquoi, la base IV² présente un bon choix pour l'évaluation de la performance de systèmes face à des images d'iris dégradés.



Figure 1. 16 Evolution de la qualité des images de la base IV² et ICE (allant de 0 à 1 en partant de la qualité dégradée vers la bonne qualité), d'après [**Petrovska-Delacretaz 2008**].

La collecte des iris de la base IV^2 s'est produite en deux sessions. La première session comprend les images d'enrôlement avec 4 images par sujet et la deuxième session comporte les images de test avec 10 images par sujet. Le port de lunettes de vue a été autorisé seulement dans la seconde session (cf. **Table 1. 4**). Un ensemble disjoint constitué d'iris capturés à partir de 53 sujets (5 images par sujet) a également été mis à disposition. Dans l'évaluation de la modalité de l'iris, deux jeux de test (ou expériences) ont été réalisés : un pour l'échantillon de l'iris droit et l'autre pour celui de gauche. Un descriptif de ces expériences est montré dans la **Table 1. 4** [Lelandais 2007] [Petrovska-Delacretaz 2008].

Table 1. 4 Descriptif des expériences de la modalité de l'iris dans la campagne IV² (N.= nombre et comp. = comparaisons).

Technique de normalisation	Iris droit	Iris gauche	
N. Images enrôlement	1152	1176	
N. Images test	2874	2939	
N. comp. intraclasses	11496	11520	
N. comp. interclasses	11756	11760	

b) Base de visage 2D et campagne d'évaluation IV²

La base d'images de visage 2D est issue de 315 sujets ayant participé à une première session dont 77 ont participé à une seconde session et 19 ont participé à une troisième session. Une partie de cette base a été capturée dans un environnement contrôlé et comprend une variation limitée de l'expression du visage. L'autre partie de la base a été acquise dans un environnement non contrôlé comprenant plusieurs variabilités notamment celles de l'éclairage, de la qualité et de la session (cf. **Figure 1. 17**). Un autre ensemble disjoint de 156 images, issus de 52 sujets (3 images par sujet) a également été fourni et constitue la base de développement. Les images de visage 2D mises à disposition sont accompagnées de métadonnées spécifiant la position des centres des yeux. Cette annotation a été réalisée manuellement sur les images de développement et ensuite, de manière automatique pour les images d'évaluation grâce aux algorithmes fournis par la société URATEK, partenaire du projet IV².



Figure 1. 17 Echantillons d'images de visage 2D de la base IV², incluant des variabilités d'expression, d'éclairage, de qualité et de session.

Dans l'évaluation de la modalité du visage 2D, la stratégie de tester une seule variabilité à la fois a été adoptée (expression, éclairage, qualité ou session) dans le but d'évaluer le degré du défi apporté par chacune d'entre elles sur les performances des systèmes biométriques. Un ensemble de plus de 15000 images réparties en 4 sousensembles a été utilisé. Une description des images de test selon les expériences correspondantes est donnée dans la **Table 1.5** [Mellakh 2009-b].

Expériences	Exp1	Exp2	Exp3	Exp4
Session	Mono	Mono	Mono	Multi
Qualité	Bonne	Bonne	Mauvaise	Bonne
V. expression	Limitée	Non	Non	Non
V. éclairage	Non	Oui	Non	Non
N. sujets	289	287	207	93
N. img/sujet enrôlement	3	3	3	4
N. img/sujet test	3	3	3	5
N. comp. intraclasses	2595	2503	1774	1796
N. comp. interclasses	2454	2364	1713	1796

Table 1. 5 Description des expériences de visages 2D (V.= variation, N.= nombre, img = image et comp. = comparaisons).

Au total, la société Thalès Research and Technology – TRT, partenaire du projet IV^2 , a reçu en tant qu'évaluateur la participation de cinq algorithmes de reconnaissance de visage 2D lors de la campagne d'évaluation IV^2 . Les résultats de ces algorithmes serviront de référence pour la comparaison avec les méthodes de reconnaissance de visage élaborées dans le cadre de ce travail de thèse.

c) Protocole IV² d'évaluation multimodale

A l'issue de la campagne d'évaluation IV², une fusion multimodale au niveau des scores a été opérée en utilisant les méthodes participantes. La fusion a été réalisée pour trois types de sources d'information :

- une fusion multi-capteurs : entre la meilleure méthode de reconnaissance de visage 2D et celle de visage 3D,
- une fusion multi-algorithmes : pour la modalité du visage 2D entre l'algorithme de la PCA et l'algorithme LDA-Gabor,
- une fusion multi-biométries : en effectuant trois combinaisons de traits. D'abord l'iris et le visage 2D, ensuite l'iris et le visage 3D et enfin, l'iris, le visage 2D et le visage 3D.

Le protocole d'évaluation et le partitionnement de la base de données multimodale n'ont pas été donnés en détails dans le rapport technique de la campagne d'avaluation [Colineau 2007]. Toutefois, il a été précisé dans le cas de la fusion multimodale de l'iris, du visage 2D et du visage 3D que deux séries de test ont été établies dont l'une est en mono-session et l'autre en multi-session. Les caractéristiques de ces expériences sont décrites dans la **Table 1. 6**.

Etant donné le manque d'information sur le protocole d'évaluation multimodal IV² et la non disponibilité des listes de comparaisons multimodales, nous avons créé notre propre protocole et nos propres listes de comparaison pour tester les approches multimodales proposées dans le **Chapitre 4**. Néanmoins, nous allons établir, à titre indicatif, une comparaison entre les méthodes multimodales proposées et celles d'IV², étant donné qu'elles ont été évaluées sur la même base de données.

	Expérience	e mono-session	Expérience multi-session		
	Bapf Base d'évaluation		Bapf	Base d'évaluation	
N. total de sujets retenus/modalité	88	176	-	(≤77)	
N. total de comparaisons/modalité	712	1831	3690	1470	

Table 1. 6 Descriptif des expériences multimodales IV² de fusion des modalités de l'iris, du visage2D et du visage 3D (N.= nombre, BApf = Base d'apprentissage des paramètres de la fusion).

Conclusion

Après avoir introduit les modes de fonctionnement et les outils d'évaluation d'un système biométrique, nous avons vu, dans ce premier chapitre, que la conception d'un système biométrique multimodal dépend de plusieurs facteurs tels que les sources d'informations fusionnées, l'architecture et le niveau de fusion. Nous avons porté une attention particulière à la fusion au niveau des scores notamment par les règles de combinaison dont nous allons faire usage dans les méthodes proposées.

Ensuite, nous avons décrit les types et les architectures possibles d'une base de données multimodale. Suite à un état de l'art sur les bases de données multimodales réelles existantes, nous avons justifié notre choix pour la base IV^2 . Le contenu de cette base a été détaillé pour les traits de l'iris et du visage 2D. Les protocoles d'évaluation entrepris dans le cadre de la campagne IV^2 ont également été présentés pour chacune des deux modalités et pour la fusion multimodale.

Dans les chapitres suivants, nous allons présenter les méthodes d'authentification proposées qui couvrent plusieurs aspects de la multimodalité. Ces méthodes seront évaluées avec le protocole IV² et comparées à celles de la campagne IV².

Chapitre 2

Biométries par l'iris et par le visage : Approches classiques

Intro	oducti	ion	38
2.1	Bion	nétrie par l'iris	38
	2.1.1	Etat de l'art	39
		 a) Processus de reconnaissance par l'iris b) Etude bibliographique des principales méthodes de reconnaissance par l'iris 	39
		c) Concepts théoriques des méthodes expérimentées	47
	2.1.2	Approche par les Pyramides Orientables (PO)	54
		a) Description de la méthodeb) Evaluation sur la base CASIA V1	54 56
	2.1.3	Méthode de référence (Log-Gabor)	61
		 a) Description de la méthode b) Evaluation par le protocole IV² c) Impact de facteurs de qualité basés sur l'occultation des iris d) Comparaison avec des méthodes existantes 	61 64 66 72
	2.1.4	Approche de reconnaissance par l'iris retenue	73
2.2	Bion	nétrie par le visage	74
	2.2.1	Etat de l'art	74
		 a) Les approches locales b) Les approches globales c) Les approches hybrides d) Approche adoptée 	75 76 79 81
	2.2.2	Application de la Pyramide Orientable en biométrie faciale et limites	81
	2.2.3	Méthodes de réduction de l'espace caractéristique	84
		 a) Concepts théoriques des méthodes de réduction b) Résultats des méthodes de réduction c) Limite des méthodes de réduction 	84 89 93
	2.2.4	Proposition de méthodes combinées	94
		a) Couplage de la PO à des méthodes projectivesb) Couplage de la PO à une méthode adaptative : Adaboost/PO	95 102
	2.2.5	Comparaison à la campagne IV^2 et méthode de reconnaissance faciale classique retenue	109
Con	clusio	n	111

Introduction

Les deux modalités devant intervenir dans la conception de notre approche biométrique multimodale étant l'iris et le visage, nous consacrons ce chapitre à l'étude de chacune des méthodes monomodales correspondantes.

Pour chacune des deux modalités, une étude succincte de l'état de l'art sera d'abord présentée. Ensuite, les méthodes expérimentées seront décrites et leurs résultats seront discutés. A l'issue de ce chapitre, nous retiendrons deux approches monomodales, l'une par l'iris et l'autre par le visage.

2.1 Biométrie par l'iris

Ces dernières années ont connu un riche développement de la recherche dans le domaine de la biométrie par l'iris. Avant d'analyser l'intérêt de l'iris en tant que caractéristique biométrique, nous allons commencer par détailler l'anatomie de cet organe (**Figure 2.1**).



Figure 2. 1 Anatomie de l'iris : (a) photographie en gros plan [Manvelyan 2015], (b) coupe transversale, (c) secteur frontal ; (b) et (c) sont extraits de [Wildes 1997], page 3, Fig.2.

L'iris est un organe visible et coloré de l'œil situé entre la sclère (blanc de l'œil) et la pupille. L'étymologie de son nom, iris, qui, en grec ancien veut dire arc-en-ciel, vient du fait que les iris soient de couleurs très variées. Sa principale fonction est de moduler le diamètre de la pupille afin de contrôler la quantité de lumière qui parvient à la rétine.

Si l'on considère sa coupe transversale (**Figure 2. 1-b**), l'iris est constitué de 5 couches superposées : la face antérieure, le stroma, le sphincter irien, le muscle dilatateur irien et la face postérieure. Vu de face (**Figure 2. 1-a, c**), c'est un anneau divisé en deux principales zones : la zone ciliaire et la zone pupillaire qui sont séparées

par la collerette. La texture de l'iris est très riche en sillons (tubes fins dont le diamètre est inférieur à celui d'un cheveu), cryptes (creux) et naevi (crêtes). Cette texture contient une quantité d'information particulièrement importante, comparable à celle contenue dans l'ADN.

La biométrie par l'iris a la réputation d'être parmi les plus sûres. Selon les estimations de Daugman, la texture de l'iris possède une densité d'information très élevée, avec une entropie d'environ 3.2 bits/mm². Cela implique que la probabilité que deux iris coïncident par hasard à plus de 75% est inférieure à un sur un billiard c'est-àdire 10⁻¹⁵ [Daugman 2001]. Cette probabilité infime reste inchangée même entre iris provenant de vrais jumeaux et entre iris gauche et droit d'une même personne.

En complément de cette singularité, la texture de l'iris a l'avantage d'être stable au cours du temps. En effet, contrairement à sa couleur qui peut varier au fil des ans et à cause de certaines maladies, la structure de l'iris reste pratiquement fixe durant la vie d'un individu.

La technologie par l'iris présente d'autres avantages. En effet, elle est non intrusive car l'image est acquise à distance et sans contact. De plus, elle possède une signature de petite taille qui est difficile à pirater grâce à une vérification de la réactivité de la pupille (dilatation/contraction) face à une lumière d'intensité variable qui peut être conduite en parallèle avec l'acquisition.

Tous ces éléments nous ont incités à intégrer la modalité de l'iris dans notre méthode biométrique bimodale. Nous allons par la suite présenter l'état de l'art de la biométrie par l'iris. Ensuite, nous allons détailler les deux méthodes expérimentées dans cette thèse. Il est, par ailleurs, à noter que la première méthode a été développée dans le cadre d'un travail de master [Khiari 2008] et que la seconde a été développée à partir de méthodes de références comme une seconde alternative. Deux facteurs de qualité basés sur l'occultation de l'iris seront proposés à la fin de cette partie dans le but d'améliorer les taux de reconnaissance obtenus.

2.1.1 Etat de l'art

Avant de présenter une sélection de méthodes référencées dans la littérature, il est utile de rappeler les différentes étapes du processus de reconnaissance par l'iris.

a) Processus de reconnaissance par l'iris

Cinq étapes principales constituent la chaîne de reconnaissance par l'iris, (cf. Figure 2. 2):

- i) acquisition de l'image de l'iris,
- ii) segmentation de la région de l'iris,
- iii) normalisation des dimensions des données,
- iv) analyse et représentation de la texture de l'iris,
- v) comparaison des représentations de l'iris ou matching.



Figure 2. 2 Différentes étapes du processus de reconnaissance de l'iris.

Acquisition : C'est une étape critique et difficile en raison de la petite taille de l'iris et de sa couleur relativement foncée (spécialement pour les asiatiques et les arabes). De plus, les images acquises souffrent souvent de problèmes de non qualité tels que la non visibilité de la totalité de l'iris, la variabilité de l'éclairage, les reflets, ... Bien que possible avec une caméra monochrome, l'acquisition s'effectue généralement dans le domaine du proche infrarouge (non visible, donc non éblouissant pour l'utilisateur) pour minimiser la réflexion spéculaire des sources lumineuses environnantes sur la cornée (cf. **Figure 2. 3**).



Figure 2. 3 Différence de qualité entre des images du même iris capturé (a) dans le visible, (b) dans le proche infrarouge.

Segmentation : Le but de cette étape est d'isoler la région de l'iris. Cette région peut être approximée comme étant comprise entre deux cercles, généralement non concentriques, l'un désigne la frontière entre l'iris et la sclère, et l'autre, interne au premier, désigne la frontière entre l'iris et la pupille. La plupart du temps, seule une partie de l'iris est exploitable à des fins de caractérisation, l'autre partie étant couverte de bruits tels que les cils, les paupières, les réflexions des spots lumineux, ... Une technique de segmentation est donc requise pour localiser la région circulaire de l'iris tout en éliminant le bruit.

Normalisation : Les éventuelles variations de l'éclairage et de la distance œilcaméra engendrent des dimensions d'iris différentes (pour une même personne) et des déformations élastiques de la texture de l'iris qui peuvent fausser complètement les résultats de comparaison, d'où la nécessité de compenser ces variations et de respecter l'invariance par translation, par rotation et par échelle.

Caractérisation : L'iris est composé de plusieurs petits blocs irréguliers distribués aléatoirement autour de la pupille tels que les tâches, les rayures, les sillons, ... L'objectif de cette étape est de transformer les données physiques caractérisant la texture de l'iris, en données quantifiées. La représentation obtenue est appelée signature biométrique. Elle doit coder l'information la plus discriminante et servira par la suite dans l'étape du *matching*.

Matching : Cette étape consiste à choisir une métrique de comparaison qui génère une mesure de similarité ou de dissimilarité entre deux signatures d'iris et qui permet de conclure si elles proviennent du même iris ou non. Ces mesures sont classées en deux ensembles de valeurs : l'un provenant de comparaisons intra-classes, obtenu lors de la mise en correspondance entre deux signatures du même iris ; et l'autre provenant de comparaisons inter-classes, c'est-à-dire à partir de deux iris différents.

b) Etude bibliographique des principales méthodes de reconnaissance par l'iris

L'idée d'utiliser l'iris comme biométrie remonte à l'année **1936** où l'ophtalmologiste américain Frank Burch l'annonça officiellement lors d'une conférence pour l' « American Academy of Ophthalmology ». **En 1987**, deux ophtalmologistes américains Leonard Folm et Aran Safir essayent de transposer l'idée de Burch en développant un modèle conceptuel non implémenté d'un système de reconnaissance de l'iris qu'ils ont breveté dans [Flom 1987].

Une collaboration est ensuite née entre Flom et Safir et John Daugman, alors professeur à l'Université Harvard (et maintenant à l'Université de Cambridge) qui est venu les aider en mettant au point une approche mathématique pour l'analyse des motifs aléatoires de l'iris. Cette collaboration a abouti à un prototype fonctionnel breveté **en 1994** [Daugman 1994]. En outre, il est à noter que la plupart des systèmes biométriques par l'iris existants sur le marché sont basés sur l'approche de Daugman du fait que les brevets de Flom et Safir (modèle conceptuel) et de Daugman (prototype basé sur le modèle conceptuel de Flom et Safir) ont été détenus pendant une bonne période par la même compagnie.

Le système introduit par Daugman est décrit dans [Daugman 1993] [Daugman 1994] [Daugman 2001] [Daugman 2006] dans lesquels il propose un opérateur intégrodifférentiel pour détecter les deux cercles de l'iris dans l'espace paramétrique. Ensuite, il normalise la région annulaire de l'iris en un modèle rectangulaire linéarisé de taille fixe. Dans son système, Daugman analyse la texture de l'iris par des ondelettes complexes de Gabor 2D non orthogonales et génère des coefficients complexes dont la phase de chacun est quantifiée, selon un codage 4 quadrants, en un code à 2 bits. L'iris est ainsi représenté par une signature de code binaire compacte de taille 256 octets appelée « *iris code* » (cf. **Figure 2. 4**). Dans la phase de *matching*, Daugman propose une distance de Hamming normalisée qui mesure la portion de bits différents entre deux iris codes localisés dans des zones non occultées simultanément dans les deux images.



Figure 2. 4 Résultats du traitement de l'iris par la méthode de Daugman. Les contours en blanc montrent les résultats de la localisation de l'anneau de l'iris et des paupières. L'iris code en haut à gauche est le résultat de l'encodage de phase des coefficients d'ondelettes de Gabor complexes 2D ayant servi à la caractérisation du modèle de l'iris. Figure extraite de [Daugman 2004].

En 1997, Wildes présente un système biométrique par l'iris développé dans les laboratoires Sarnoff au New Jersey qui est complètement différent de celui de Daugman et qui repose sur une segmentation de l'iris par transformée de Hough. La normalistation de la région de l'iris se fait par recalage et la caractérisation de la texture de l'iris se base sur une représentation pyramidale multirésolution par des filtres Laplaciens (cf. Figure 2. 5). A la fin, la classification se fait par une mesure de corrélation normalisée [Wildes 1997].



Figure 2. 5 Résultats du traitement de l'iris par la méthode de Wildes : (a) image initiale ; (b) image segmentée par la transformée de Hough ; (c) représentation multi-échelle par les filtres Laplaciens. Images extraites de [Wildes 1997], Figures 5, 6 et 7.

En 1998, Boles et Boashash présentent un algorithme dont l'analyse de la texture se fait à travers une transformée unidimensionnelle en ondelettes orthogonales avec huit échelles. Le signal transformé est également unidimensionnel et est acquis suite à une normalisation par cercles virtuels (concentriques) au moment de la comparaison entre deux iris, de façon à avoir un diamètre constant avec une image de référence et assurant que les mêmes points de la texture de l'iris soient échantillonnés indépendamment de la taille de l'iris dans l'image. Les cercles virtuels donnent lieu à n « signatures d'iris » sur chaque échelle. L'encodage est ensuite réalisé pour chacune de ces n signatures sur trois des échelles les plus basses (4^{ème}, 5^{ème} et 6^{ème}) et ce, par la détection des passages par zéro (*zero-crossings*) de la dérivée seconde de la transformée par ondelettes lissée (cf. **Figure 2.** 6). La comparaison est assurée à travers des fonctions de dissimilitude telles que la différence d'énergies et le coefficient normalisé de corrélation croisée entre deux représentations de passages par zéro [Boles 1998].



Figure 2. 6 Résultats du traitement de l'iris par la méthode de Boles et Boashash : (a) image initiale ; (b) Exemple d'une signature d'iris (signal situé sur un cercle virtuel) de l'iris de l'image (a) ; (c) Trois niveaux de résolution de la transformée en ondelettes de la signature (b) ; (d) Représentation des passages par zéro de (c). Images extraites de [Boles 1998].

Depuis, les études sur la biométrie par l'iris ont connu un essor grandissant. Chaque étude apporte des contributions dans une ou plusieurs étapes du processus de la reconnaissance. Certaines se focalisent sur l'acquisition de l'image de l'iris, d'autres sur la segmentation de la région d'intérêt. D'autres contributions couvrent le codage de la texture et le *matching* ou encore l'amélioration de la performance dans des conditions variées.

L'état de l'art que nous présentons dans ce qui suit traite principalement les travaux apportant des contributions pour le codage de la texture et le *matching*.

En 2000, Zhu et al. comparent deux méthodes d'analyse de la texture de l'iris : les ondelettes réelles de Gabor non orthogonales à 6 niveaux et 4 orientations et l'analyse par ondelettes orthogonales de Daubechies d'ordre 4. Ensuite, la moyenne et l'écart-type sont extraits à partir des sorties des filtrages utilisées comme signatures. A la fin, la comparaison est assurée par la Distance Euclidienne Pondérée (Weighted Euclidean Distance, WED). Les résultats de ces expériences ont été nettement en faveur des ondelettes de Gabor non orthogonales [Zhu 2000].

En 2001, Lim et al. proposent une méthode se basant sur une transformée en ondelettes de Haar à 4 niveaux pour la caractérisation de la texture de l'iris. Une

signature à valeurs réelles est obtenue en combinant les coefficients des détails diagonaux à l'échelle la plus basse avec les intensités moyennes des coefficients des détails diagonaux des 3 niveaux supérieurs. Chaque coefficient de la signature obtenue est ensuite codé avec seulement un bit selon son signe, donnant lieu à des signatures binaires de taille réduite. La mise en correspondance entre deux signatures binaires se fait par un réseau de neurones. Une comparaison avec les ondelettes de Gabor a été réalisée et les résultats étaient légèrement en faveur des ondelettes de Haar [Lim 2001].

En 2002, Tissé et al. proposent un système dans lequel l'analyse de la texture se fait par calcul de l'image analytique en sommant l'image originale avec sa transformée de Hilbert 2D. Ensuite, un codage similaire à celui de Daugman est réalisé mais, au lieu de la phase locale, Tissé et al. codent la phase instantanée et la fréquence émergente de la Transformée de Hilbert 2D. A la fin, le *matching* entre les deux iris codes se fait par la distance de Hamming sur des modèles prédéfinis permettant d'omettre les calculs sur les régions du haut et du bas de l'iris où les occultations par les cils et les paupières sont les plus fréquentes [Tissé 2002].

En 2003, Masek propose d'extraire les caractéristiques de l'iris par des filtres Log-Gabor, introduits par Field [Field 1987] comme une meilleure alternative que les filtres de Gabor pour l'analyse de la texture (voir § 2.1.1-c) pour plus de détails). Il est à noter que le code *open source* fourni par Masek [Masek 2003] a fait de sa méthode une méthode de référence à laquelle plusieurs études récentes se comparent.

En 2004, Son et al. essayent plusieurs combinaisons de méthodes basées sur les ondelettes de Daubechies, l'Analyse en Composantes Principales (Principle Component Analysis – PCA) et l'Analyse Linéaire Discriminante (Linear Discriminant Analysis – LDA) avec sa variante directe (Direct-LDA ou DLDA), pour analyser la texture de l'iris. Pour la classification, deux techniques sont utilisées, à savoir : les Machines à Vecteurs de Support (Support Vector Machines – SVM) et les k plus proches voisins (k-Nearest Neighbours – kNN). La méthode qu'ils retiennent analyse la texture de l'iris par les ondelettes de Daubechies 2D, puis le vecteur caractéristique obtenu est réduit par une DLDA et le *matching* se fait par les SVM [Son 2004].

En 2005, Gu et al. proposent une méthode basée sur le calcul de la dimension fractale sur la transformée en pyramide orientable des images d'iris afin de représenter ses irrégularités. La prise de décision est réalisée par des SVM non symétriques [Gu 2005].

En 2006, Thoonsaengngam et al. présentent un système qui exploite les variations d'intensité locales des textures visibles de l'iris telles que les cryptes et les nævi (cf. Figure 2. 1). Les textures sont extraites en utilisant une égalisation d'histogramme locale et une technique de seuillage par quotient. Cette dernière technique qu'ils ont conçu consiste à binariser les images d'iris à l'aide d'un seuil déterminé de manière adaptative, de telle sorte que le rapport entre le premier plan et l'arrière plan de chaque image d'iris soit préservé (c'est-à-dire, soit le même pour toutes les images). En fixant ce rapport, appelé « ratio de décision », les variations d'éclairage à travers les images sont compensées donnant lieu à des textures d'iris informatives similaires à des blobs (régions

ayant des propriétés constantes ou qui varient dans un intervalle de valeurs prescrit). Le *matching* est réalisé en maximisant le nombre de pixels alignés au premier plan à travers des rotations et des translations respectivement de +/- 10 degrés et +/- 10 pixels [Thoonsaengngam 2006].

En 2007, Yu et al. appliquent un filtrage de Gabor 2D multicannaux sur l'iris linéarisé. Sur chaque bloc filtré de l'image, ils extraient les valeurs maximales qu'ils appellent « points clés ». Ensuite, ils cherchent le barycentre de ces points clés sur chaque canal et calculent les distances entre le barycentre et tous les points clés de chaque bloc filtré. L'ensemble de ces distances, appelées « distances relatives », constituent le vecteur caractéristique le l'iris qui est comparé à ceux des autres iris par une distance Euclidienne [Yu 2007].

En 2008, Roy et Bhattacharya réduisent le temps de comparaison entre les iris en explorant deux méthodes de sélection de caractéristiques : les algorithmes génétiques et les SVM à élimination récursive de caractéristiques (SVM with Recusive Feature Elimination – SVM-RFE) [Roy 2008].

En 2009, Krichen et al. explorent l'utilisation d'une approche par corrélation de phase normalisée dans la phase de *matching* comme alternative au code iris binaire standard. Ils comparent leurs résultats aux algorithmes OSIRIS [Biosecure 2009] et Masek [Masek 2003] sur les bases ICE 2005 et CAISA-Biosecure [Krichen 2009].

En 2010, Rathgeb et Uhl développent une approche qui utilise une égalisation d'histogramme adaptative à contraste limité et qui trace les variations de l'intensité des pixels le long des lignes de l'image normalisée (cercles concentriques de la région de l'iris) qu'ils appellent « chemins de pixels » (pixel paths). Ils appliquent leur approche sur la base CASIA v3 [Rathgeb 2010].

Des études bibliographiques plus complètes autour des principaux travaux biométriques par l'iris réalisés jusqu'en 2010 sont décrites dans [Daugman 2007] [Bowyer 2008] [Bowyer 2013]. Des exemples de recherches conduites pendant les quatre dernières années sont présentées dans ce qui suit.

En 2011, Hollingsworth et al. exploitent la position des bits fragiles (un bit est labélisé fragile si sa valeur change à travers les iris codes créés à partir de différentes images d'un même iris) coïncidant entre deux modèles d'iris, au lieu de les ignorer complètement, afin d'améliorer la précision des comparaisons. Ils trouvent que la fusion des scores issus de la distance par bits fragiles avec ceux issus de la distance de Hamming donnent de meilleurs résultats que la distance de Hamming seule [Hollingsworth 2011].

En 2012, Li et al. proposent une méthode similaire au SIFT (Scale Invariant Feature Transform) [Lowe 2004], de par ses propriétés d'insensibilité aux variations d'éclairage et au bruit, qu'ils baptisent Histogramme de Phase de Co-occurrence Pondérée (Weighted Phase Co-occurrence Histogram - WPCH). Le WCPH implique un calcul d'histogramme d'orientation du gradient tout comme les SIFT. Cependant, à leur différence, il considère à la fois l'information issue de la phase d'angle et l'information spatiale. De plus, le WPCH inclut une fonction de pondération permettant à un angle de phase d'avoir des contributions dans plusieurs classes adjacentes. Dans l'étape du *matching*, Li et al. introduisent une distance de Bhattacharyya [Bhattacharyya 1943] entre deux WCPHs [Li 2012].

En 2014, Chen et al. subdivisent l'iris normalisé en plusieurs bandes pour comparer l'importance relative des différentes régions de l'iris. L'extraction des caractéristiques de chaque bande se fait par un filtrage de Gabor Multicanaux 2D puis par une matrice de co-occurrence de niveaux de gris (Gray Level Cooccurrence Matrix – GLCM). Une analyse de l'information texturale est ensuite produite via des paramètres locaux d'évaluation de la qualité qui sont pondérés par une technique d'optimisation par essaims particulaires (Particle Swarm Optimization – PSO) permettant également d'attribuer des coefficients pondérés aux différentes bandes de l'iris. La classification est assurée par les SVM et les kNN. Chen et al. trouvent que les poids des bandes du milieu sont supérieurs à ceux des bandes internes et externes dans un iris (donc plus importantes), mais les résultats de reconnaissance de ces bandes internes restent inférieurs à ceux de l'iris complet. Finalement, les scores issus des bandes individuelles sont fusionnés selon leurs poids correspondants et une amélioration sensible est observée par rapport à l'iris entier. Les expériences sont conduites sur des images des bases CASIA v1, CASIA v3 Interval, MMU v1 et JLUBRIRIS v1 [Chen 2014].

La **Table 2. 1** reporte les performances de quelques travaux cités dans l'état de l'art en spécifiant les bases d'iris utilisées ainsi que le nombre d'images associées.

Algorithme	Base (Nb. Images)	CMR	EER	(FAR ; FRR)
[Masek 2003]	CASIA v1 (444)	97,96%	5,8%	(0,02%;0,94%)
[Daugman 2006]	UAE national border-crossing system (632 500)	100%	0%	(0%;0%)
[Yu 2007]	Privée (1016) CASIA (756)	99,74% 97,18%	-	(0,0015%; 0,6%) (0,008%; 10%)
[Roy 2008]	ICE (-) WVU (-)	97,81% 96,23%	-	(- ; -) (- ; -)
[Krichen 2009]	ICE (2093) Casia v2 BioSecure CBS (240) IV ² (4026)	-	0,68% 1% 0,78%	(0,1%;1%) (0,1%;2%) (0,1%;1%)
[Rathgeb 2010]	CASIA V3 (-)	-	entre 1 et 2%	(-;-)
[Li 2012]	UBIRIS v2 (1000)	-	0,23%	(- ; -)
[Chen 2014]	CASIA v1(260) CASIA v3 (240) MMU v1 (179) JLUBRIRIS v1 (729)	- - -	3,23% 4,55% 7,16% 2,98%	(-;-) (-;-) (-;-) (-;-)

Table 2. 1 Résultats reportés de quelques algorithmes de la littérature jusqu'en 2014.

Différentes mesures de performances sont données selon que le mode biométrique testé soit en identification ou en vérification. Pour le mode identification, c'est le Taux de Comparaison Correct (Correct Match Rate – CMR) qui est spécifié ; alors que, pour le mode vérification, deux mesures sont présentées : l'EER (FAR = FRR) et un autre point de fonctionnement (pour lequel FAR et FRR sont différents) selon les besoins en sécurité de l'application.

Suite à cette étude bibliographique, nous pouvons distinguer deux voies principales d'analyse de la texture de l'iris : celle du filtrage fréquentiel – que ce soit par ondelettes ou par filtres orientables – suivie, éventuellement, d'une étape de sélection des caractéristiques ; et celle de l'analyse des intensités locales des pixels (par exemple par LBP, SIFT, ...).

Plusieurs auteurs tels que [Marr 1979] et [Adelson 1985] préconisent plutôt la caractérisation du spectre local pour l'analyse locale d'une image. Cette recommandation se base sur le fait que des études sur la modélisation du système visuel humain suggèrent que le cortex cérébral est sensible à une gamme de fréquences et d'orientations (cf. Chapitre 3). Ceci a motivé notre choix de travailler avec deux méthodes de filtrage fréquentiel : la Pyramide Orientable (PO) et les filtres Log-Gabor.

c) Concepts théoriques des méthodes expérimentées

Au vu de l'état de l'art précédent sur la biométrie par l'iris, nous avons choisi de travailler avec deux méthodes d'analyse de la texture que nous avons jugées pertinentes :

La première est celle de la Pyramide Orientable qui combine la décomposition multiéchelle permettant d'extraire la fréquence de l'information texturale abondante de l'iris, avec les mesures différentielles capables de capturer l'orientation présente dans la structure de l'iris.

La deuxième méthode consiste à utiliser les filtres Log-Gabor dont la performance a été démontrée dans le système Masek [Masek 2003].

Ce qui suit est dédié à la présentation des concepts théoriques des pyramides orientables et des filtres Log-Gabor. Une description des principes de conception des filtres et de leurs contraintes est également fournie.

i) Pyramides Orientables – PO

La pyramide orientable (PO) a été introduite par [Simoncelli 1995]. C'est une décomposition linéaire multi-orientation multi-résolution qui apporte des solutions à plusieurs applications de traitement d'image et, en particulier, à l'analyse de la texture. Les fonctions de base d'une PO sont des opérateurs de type dérivées directionnelles à différentes orientations et échelles.

Le schéma fonctionnel d'une PO [Karasaridis 1996] est donné dans la **Figure 2. 7** aussi bien pour l'analyse que pour la reconstruction. Les points noirs au milieu représentent les images sous-bandes décomposées.



Figure 2.7 La première échelle du schéma fonctionnel d'une pyramide orientable.

Les transformées sont calculées récursivement en utilisant des opérations de convolution et de décimation. Dans la partie analyse, l'image est d'abord décomposée en deux sous-bandes : une composante qui fournit les détails du signal, par un filtre passe-haut H_{θ} , et une composante lisse qui donne l'allure générale du signal, par un filtre passe-bas L_{θ} . La composante passe-bas continue à se décomposer en une collection de (n+1) sous-bandes passe-bandes orientées $B_{\theta}, B_1, \ldots, B_n$ et une sous-bande passe-bas L_I de fréquence inférieure à la première L_{θ} . Ensuite, la dernière sous-bande passe-bas L_I est sous-échantillonnée d'un facteur 2 dans les directions x et y. Ce processus constitue la première échelle de décomposition d'une PO. Répéter la partie encadrée (en bleu) à la sortie du sous-échantillonnage donnera la structure récursive de la pyramide, et donc les échelles suivantes. Dans la partie de synthèse, l'image reconstruite est obtenue en sur-échantillonnant la sous-bande passe-bas d'un facteur 2 et en ajoutant la collection des sous-bandes passe-bandes et la sous-bande passe-haut.

Cette représentation a l'avantage d'être invariante par translation (sans repli de spectre) et par rotation (les sous-bandes sont orientées). Plus important encore, la transformée est auto-inversible; c'est à dire, les mêmes filtres utilisés pour la décomposition sont ensuite utilisés pour la reconstruction. La pyramide a été conçue pour générer n'importe quel nombre k souhaitable de bandes d'orientation. Le premier inconvénient qui s'en suit est que la transformée résultante présente alors un facteur de redondance de 4k/3 (contre 4/3 pour la pyramide Laplacienne). De plus, la conception du filtre est un problème très complexe car il doit obéir à des contraintes concernant l'échelle et l'orientation.

Conception du filtre

La pyramide orientable est mieux définie dans le domaine de Fourier où elle peut être écrite sous forme polaire séparable permettant une représentation indépendante de l'échelle et de l'orientation. La **Figure 2. 8** décrit les réponses fréquentielles des sousbandes d'une PO à 3 échelles et 4 orientations. Les bandes passe-bandes orientées B_{i} (*i* = *o..3)* sont montrées en différentes couleurs et les échelles sont représentées par les cercles concentriques.



Figure 2. 8 Illustration idéalisée de la décomposition spectrale par une pyramide orientable à 3 échelles et 4 orientations (dérivée $3^{\text{ème}}$). Les axes fréquentiels { ω_x , ω_y } ϵ [- π , π].

L'équation (2. 1) donne la formule du $i^{\text{ème}}$ filtre passe bande orienté B_i écrit sous forme polaire séparable.

$$B_i(\vec{w}) = A(\theta - \theta_i)B(\vec{w})$$
(2.1)

avec $A(\theta - \theta_i)$ la composante angulaire et $B(\vec{w})$ la composante radiale dont les expressions sont données par les équations (2. 2) et (2. 3).

$$A(\theta - \theta_i) = [-j\cos(\theta - \theta_i)]^n, \quad \theta = \arg(\vec{w}) = \arctan\left(\frac{w_y}{w_x}\right), \quad \theta_i = \frac{i\pi}{n+1}, \ i\in[0, \ n]$$
(2.2)

$$B(\vec{w}) = \sqrt{\sum_{i=0}^{n} |B_i(\vec{w})|^2}$$
(2.3)

La conception du filtre de la PO doit obéir à un ensemble de contraintes d'orientabilité et d'échelle [Simoncelli 1995] [Karasaridis 1996].

Contraintes d'orientabilité :

Dans la décomposition angulaire, les filtres doivent former une base orientable :

- ils sont des copies les uns des autres par rotation ;
- la synthèse d'un filtre à n'importe quelle orientation θ peut être écrite comme une combinaison linéaire des filtres de la base.

L'exemple le plus simple d'une base orientable est un ensemble de (n+1) dérivées n^{imes} directionnelles (cf. équations (2. 2) et (2. 3)), où (n+1) désigne le nombre d'orientations de la PO c'est-à-dire le nombre de filtres passe-bandes à chaque échelle.

Contraintes d'échelle :

La décomposition radiale doit satisfaire un compromis entre le désir de construire une décomposition récursive (en utilisant un algorithme pyramidal), et le besoin d'éviter l'apparition de l'effet de moiré lors de la reconstruction. Les contraintes suivantes doivent, alors, être respectées :

- Limitation de la bande pour éliminer les erreurs de reconstruction (repli de spectre) dues à l'opération de sous-échantillonnage dans les filtres circulaires symétriques :

$$L_1(\vec{w}) = 0, \quad |\vec{w}| < \frac{\pi}{2}$$
 (2.4)

- La fonction de transfert du système doit être égale à 1 afin d'éviter les distorsions d'amplitude :

$$|L_0(\vec{w})|^2 \left[|L_1(\vec{w})|^2 + \sum_{i=0}^n |B_i(\vec{w})|^2 \right] + |H_0(\vec{w})|^2 = 1$$
(2.5)

- La décomposition doit être récursive :

$$|L_1(\vec{w}/2)|^2 \left[|L_1(\vec{w})|^2 + \sum_{i=0}^n |B_i(\vec{w})|^2 \right] = |L_1(\vec{w}/2)|^2$$
(2. 6)

 L_0 étant le filtre passe-bas initial, H_0 le filtre passe-haut, L_1 le filtre passe-bas de la partie récursive, B_i est le i^{ème} filtre passe-bande orienté et n le nombre total de filtres orientés.

Généralement, on choisit $|L_0(\vec{w})| = |L_1(\vec{w/2})|$ pour que la forme initiale du filtre passe-bas soit la même forme utilisée à l'intérieur de la partie récursive.

La **Figure 2. 9** illustre une décomposition en PO d'une image d'iris segmenté à 3 échelles et 4 orientations (n = 3, dérivée troisième). Sur la figure sont affichées : l'image passe-haut puis, pour chaque échelle, les 4 images passe-bandes orientés et enfin la dernière image passe-bas (le résidu). Il est important de souligner que seule la partie analyse du schéma fonctionnel de la PO est appliquée lors de l'extraction des caractéristiques de la texture de l'iris.



Figure 2. 9 Décomposition d'une image d'iris par une PO à 3 échelles (s = 1..3) et à (n+1) = 4 orientations (0, $\pi/4$, $\pi/2$ et $3\pi/4$).
ii) Filtres Log-Gabor

Les filtres Log-Gabor ont été introduits par Field en 1987 comme une meilleure alternative que les filtres de Gabor pour l'analyse de la texture [Field 1987]. En effet, les filtres de Gabor complexes sont définis, dans le domaine temporel, comme le produit d'une enveloppe gaussienne par une sinusoïde complexe (cf. **équation (2.7)**).

$$G(t) = A \exp\left(\frac{-t^2}{2\sigma^2}\right) \exp(i\omega_0 t + \varphi), \quad avec \ A = \frac{1}{\sigma\sqrt{2\pi}}$$
(2.7)

 ω_0 étant la pulsation unique du filtre ($\omega_0 = 2\pi f_0$, f_0 étant la fréquence centrale du filtre); o l'écart type de l'enveloppe gaussienne permettant de régler la bande passante fréquentielle; et φ la phase de la sinusoïde : la partie réelle et la partie imaginaire de l'ondelette complexe de Gabor sont à la fois utilisées, ce qui donne deux ondelettes : une ondelette paire et une ondelette impaire.

Le problème majeur des filtres de Gabor est qu'ils souffrent d'avoir une composante continue (moyenne du signal dans le domaine fréquentiel) non nulle au niveau du filtre symétrique pair dès que la bande passante dépasse 1 octave [Field 1987]. Il y aura alors un recouvrement des gaussiennes à l'origine dans l'espace fréquentiel, ce qui fait que le filtre passe-bande se transforme en un filtre passe-bas en éliminant certaines hautes fréquences (c.f. **Figure 2. 10**).



Figure 2. 10 Fonction de transfert d'un filtre de Gabor de symétrie paire ayant une large bande passante.

A contrario, le filtre Log-Gabor permet d'avoir une composante continue nulle pour n'importe quelle bande passante en utilisant des fonctions de Gabor qui sont gaussiennes lorsqu'elles sont vues selon une échelle logarithmique. La réponse fréquentielle d'un filtre Log-Gabor est donnée par :

$$G(f) = exp\left(\frac{-(\log{(f/f_0)})^2}{2(\log(\sigma/f_0))^2}\right)$$
(2.8)

où f_0 est la fréquence centrale et σ la bande passante du filtre.

Mise à part cette caractéristique importante des fonctions Log-Gabor qui, par définition, ne possèdent jamais de composante continue (la moyenne des gaussiennes étant nulle du fait qu'elles sont parfaitement centrées sur l'axe logarithmique), la fonction de transfert de la fonction Log-Gabor possède une extrémité étendue aux hautes fréquences comme illustré sur la **Figure 2. 11** [Field 1987].



Figure 2. 11 Comparaison entre (a) la fonction Gabor et (b) la fonction Log-Gabor vues selon l'échelle linéaire (à gauche) et l'échelle logarithmique (à droite). Figure extraite de [Field 1987].

Field suggère que les fonctions de Log-Gabor, possédant des extrémités étendues à l'échelle fréquentielle linéaire et une forme gaussienne à l'échelle logarithmique, devraient être capables d'encoder des images naturelles de manière plus efficace que des fonctions de Gabor ordinaires qui, comme le montre la **Figure 2. 11**, ont tendance à sur-représenter les basses fréquences (filtrage passe-bas) et sous-représenter les hautes fréquences (mauvais filtrage passe-bande).

Conception du filtre

Malheureusement, à cause de la singularité de la fonction log à l'origine, on ne peut pas construire une expression analytique de la fonction Log-Gabor dans le domaine spatial. On en est réduit à concevoir les filtres dans le domaine fréquentiel puis à opérer par une transformée de Fourier inverse [Kovesi 2003]. L'apparence de la fonction Log-Gabor est similaire à celle de Gabor mais sa forme devient plus « nette » au fur et à mesure que la bande passante augmente. Les formes des deux fonctions sont à peu près identiques pour des bandes passantes inférieures à un octave.

Etant donné qu'un banc de filtres Log-Gabor – comme celui de filtres de Gabor, d'ailleurs – ne forme pas une base orthogonale, il n'existe pas d'arrangement idéal ou unique des filtres. La conception des bancs de filtres est alors une tâche qui requiert certaines précautions et est soumise à un nombre de contraintes [Kovesi 2003].

Le but est de produire un banc de filtres permettant une couverture paire de la section du spectre que l'on souhaite représenter. Ceci peut être accompli à travers un recouvrement assez large des fonctions de transfert des filtres individuels de telle sorte que leur somme réalise une couverture paire du spectre. Pour des raisons de complexité de calcul, on aimerait, de plus, que cette couverture soit réalisée avec un nombre minimal de filtres.

Un autre objectif, qui est en conflit avec le premier, est d'assurer que les sorties des filtres individuels soient aussi indépendantes que possible. Pour assurer l'indépendance entre les sorties, les filtres devraient avoir le minimum de recouvrement entre leurs fonctions de transfert.

En conclusion, les fonctions de transfert des filtres doivent avoir le recouvrement minimal nécessaire permettant de réaliser une couverture spectrale paire équitable. De plus amples détails sur la conception du filtre Log-Gabor se trouvent dans [Field 1987].

2.1.2 Approche par les Pyramides Orientables (PO)

Dans ce paragraphe, nous décrivons, d'abord, les différentes étapes du processus de la reconnaissance par la méthode proposée se basant sur les pyramides orientables [Khiari 2008], puis nous exposons les résultats obtenus après son évaluation sur la base CASIA V1.

a) Description de la méthode

La méthode proposée se compose de quatre principales étapes, à savoir : 1) la localisation des cercles de l'iris, 2) la normalisation du modèle de l'iris, 3) l'extraction des caractéristiques et 4) la décision.

Localisation : Nous avons opté pour la méthode de Hough comme moyen de localisation de l'iris. Les bords interne et externe de l'anneau de l'iris peuvent être approximés par deux cercles non concentriques qui sont déterminés en 2 étapes selon la transformée de Hough : (1) trouver les points candidats pouvant appartenir au cercle ; (2) garder le cercle qui rassemble le maximum de votes dans l'espace paramétrique des cercles. Les points candidats sont détectés en appliquant des gradients verticaux et horizontaux pour la détection du cercle interne et en appliquant seulement des gradients verticaux pour la détection du cercle externe et ceci dans le but de minimiser l'effet des cils et des paupières [Masek 2003].

Normalisation : Dans le but de compenser les incohérences de rotation dues aux inclinaisons de têtes au moment de l'acquisition, nous procédons à la rotation de l'iris localisé selon plusieurs angles. La méthode devient alors invariante par rotation. Par ailleurs, pour compenser l'effet de la dilatation de la pupille et les incohérences liées aux variations de tailles (distance œil/caméra) et pour briser l'effet de non concentricité de l'iris et de la pupille, nous avons choisi de linéariser l'anneau circulaire de l'iris par une transformation pseudo-polaire suivant le modèle *rubber sheet* proposé dans [Daugman 1993]. L'anneau de l'iris est retracé sous la forme d'un bloc de texture rectangulaire de taille fixe, suite à une transformation de référence pseudo-polaire. La méthode devient alors invariante à l'échelle. La **Figure 2. 12** montre un modèle rectangulaire d'iris obtenu après le déroulement de l'anneau.



Figure 2.12 (a) Processus de linéarisation selon le modèle « rubber sheet » de Daugman avec décalage exagéré entre les centres de l'iris et de la pupille pour des raisons d'illustration. (b) Modèle linéarisé (64×512) d'un iris de la base CASIA V1.

Comme les modèles des iris normalisés sont encore affectés par les spots lumineux, les cils et les paupières, une simple technique de seuillage a été utilisée pour minimiser l'effet des cils et des spots lumineux. La région affectée par les occlusions des paupières a été localisée dans deux zones angulaires correspondant aux intervalles $[\pi/4; 3\pi/4]$ et $[5\pi/4; 7\pi/4]$, situées du côté externe et ayant une largeur d'un demi-rayon de l'anneau de l'iris. La région d'intérêt (*Region Of Interest* – ROI) est alors obtenue comme montré dans la **Figure 2. 13**.

Extraction des caractéristiques : Afin d'extraire l'information texturale de l'iris, nous avons choisi la pyramide orientable qui est une méthode d'analyse locale multi-résolution multi-orientation. Il est important de noter que seule la partie de l'analyse du schéma fonctionnel de la PO est appliquée lors de l'extraction des caractéristiques de l'iris, c'est-à-dire que l'on ne procède pas à la reconstruction.

Le modèle linéarisé de l'iris obtenu suite à la phase de la normalisation est d'abord subdivisé en blocs de taille fixe $(p \times p)$. La PO est ensuite opérée sur chaque bloc ; puis, l'énergie (cf. **équation (2.9)**) est calculée sur chaque bloc filtré. La concaténation de ces énergies donne lieu à un vecteur caractéristique de taille assez réduite qui est égale au nombre de blocs comme montré sur la **Figure 2.13** (vecteur C_i , $i \in [1, 128]$).

$$E = \sum_{i=1}^{p} \sum_{j=1}^{p} |X(i,j)|$$
(2.9)

Décision : Il s'agit de la dernière phase du processus d'identification. La comparaison d'un vecteur caractéristique x de taille k stocké dans la base d'enrôlement est comparé à celui d'un candidat y moyennant la norme L1 ou distance City-block dont la formule est donnée par l'équation (2. 10).

Norme_L1 =
$$\sum_{i=1}^{k} |x_i - y_i|$$
 (2.10)

La **Figure 2. 13** illustre les différentes étapes du processus de la reconnaissance de l'iris présenté dans cette section.



Figure 2. 13 Processus de reconnaissance de l'iris par PO sur CASIA V1.

b) Evaluation sur la base CASIA V1

Nous avons évalué la performance de notre algorithme sur la base d'images infrarouge CASIA V1 ([CASIA-IrisV1 2010] auparavant [CASIA-V1 2007]). Puis, nous avons comparé la méthode proposée à celle de Masek dans les deux modes identification et vérification [Masek 2003]. Pour ce faire, nous avons mené une étude sur l'influence des différents paramètres de la méthode afin de déterminer les paramètres optimums donnant les taux de reconnaissances les plus élevés.

i) Base d'images CASIA V1

Nous avons utilisé la base d'images CASIA V1 fournie par le National Laboratory of Pattern Recognition (NLPR), Institute of Automation (IA), Chinese Academy of Sciences (CAS). La base comprend 756 images infrarouges issues de 108 yeux de 80 sujets en deux sessions à un mois d'intervalle [CASIA-IrisV1 2010].

Il est à noter que, dans le but de protéger l'architecture de sa caméra (spécialement le schéma de l'illumination proche infrarouge), les régions des pupilles dans toutes les images d'iris de la base CASIA V1 ont été automatiquement détectées et remplacées par une région circulaire d'intensité constante afin de masquer les réflexions spéculaires de la lumière proche infrarouge. Une telle modification rend la détection du contour de l'iris beaucoup plus facile, mais n'a quasiment aucun effet sur les autres phases du processus de reconnaissance, telles que la caractérisation et la classification. Après avoir obtenu la licence sur l'architecture de la caméra CASIA, les images originales non masquées ont été mises à la disposition du public, plus tard, dans la version CASIA-IrisV3-Interval. Un exemple d'une image d'iris masquée et son image originale correspondante est donné sur la **Figure 2. 14** [Tieniu 2005].



Figure 2.14 (a) image de CASIA V1.0 (masquée), (b) image originale de CASIA-IrisV3-Interval.

En mode identification, nous avons utilisé des images issues de 50 yeux, 7 images/œil donnant un total de 760 585 comparaisons. Tandis qu'en mode vérification, l'ensemble utilisé en identification a été gardé pour constituer l'ensemble des « autorisés » (3 images ont servi à l'apprentissage et 4 images au test). Un autre ensemble d'images d'iris « imposteurs » a été constitué à partir de 14 autres yeux, 7 images/œil donnant un total de 873 160 comparaisons (280 875 à l'apprentissage et 592 285 au test).

ii) Résultats expérimentaux

Etant donné que notre travail se focalise sur la caractérisation, nous avons choisi de tester séparément les phases de localisation et d'extraction des caractéristiques. La totalité de la base CASIA V1 est utilisée dans les tests de la phase de localisation mais, seuls les iris bien segmentés serviront dans les phases suivantes.

Localisation : Nous avons adapté la partie du code open-source de Libor Masek, correspondant à la détection des cercles de l'iris par la méthode de Hough, à notre algorithme. Un taux de 88.1% de bonne localisation a été atteint (l'attribution des étiquettes de « bonne » ou de « mauvaise » localisation a été réalisée à l'œil nu). Il est à noter que le cercle interne (celui de la pupille) est toujours bien détecté. Les mauvaises segmentations du cercle externe de l'iris sont dues, la plupart du temps, à une zone visible très réduite de l'iris où à des occlusions trop importantes par les cils.

Mode identification et Paramètres optimums : La méthode dépend de plusieurs paramètres dont l'influence de certains a été étudiée et sera présentée dans ce qui suit. Ces paramètres concernent le type de prétraitement (**Table 2. 2**), les paramètres de normalisation (nombre de rotations de l'iris, déroulement de l'anneau de l'iris, choix de la région d'intérêt, contribution de la région bruitée dans les calculs), les paramètres d'extraction de caractéristiques (**Table 2. 3** à **Table 2.6**) et la métrique de comparaison.

Un grand nombre de tests a été engagé. Dans chacun d'entre eux, nous avons fait varier un seul paramètre en fixant tous les autres. Les **Table 2. 2** à **Table 2. 7** en montrent quelques résultats. Il est important de souligner que les tests ont été d'abord implémentés en mode identification pour déterminer les paramètres optimums, puis ces mêmes paramètres ont été gardés en mode vérification. Les performances seront données en termes de taux d'identification correspondant au rapport entre le nombre de signatures d'iris correctement identifiées et le nombre total de signatures candidates à identifier. En fait, il existe deux manières de considérer une signature comme étant correctement identifiée par le système. La première suppose que le système renvoie une seule identité pour une signature candidate donnée : celle qui correspond à la distance minimale. Dans ce cas, on parle de **taux d'identification au rang 1**. La deuxième manière suppose que le système biométrique renvoie les k identités correspondant au k distances minimales avec la signature candidate. Ainsi, une signature est correctement identifé si son identité figure parmi les k identités retournées par le système. Dans ce cas, on parle de **taux d'identification au rang** k. Il s'agit donc de la formule généralisée du taux d'identification au rang 1 dans laquelle le système est d'autant plus tolérent à l'erreur que le rang k est grand. Dans l'évaluation de la méthode proposée, nous nous positionnons dans le premier cas de figure. C'est pourquoi, les résultats seront donnés en taux d'identification au rang 1.

Trois approches ont été évaluées dans la phase de prétraitement : le filtrage Gaussien, l'égalisation d'histogramme et la combinaison des deux (**Table 2. 2**). Puisque le filtrage Gaussien a donné de meilleurs EERs que l'égalisation d'histogramme, et que la combinaison des deux ne fournit aucun apport, nous avons conduit le restant des tests (**Table 2. 3** à **Table 2. 7**) avec ce type de prétraitement pour déterminer les paramètres optimaux. Mais lorsque nous avons complètement éliminé la phase de prétraitement, les résultats sont passés de 97,45% à 98,47%. La dégradation des résultats par le filtrage Gaussien comparé à l'utilisation d'images non prétraitées peut être expliquée par le fait que le filtrage Gaussien fait perdre de l'information utile à la PO qui est elle même une succession de filtrages.

Dans la phase de caractérisation, nous avons fait varier la taille des blocs filtrés : (64×64), (32×32), (16×16) et (8×8). Les résultats de la **Table 2.3** sont en faveur de la taille (16×16). Nous avons également étudié l'impact des paramètres de la pyramide orientable, à savoir, le nombre d'échelles s (s $\in \{1, 2, 3\}$) et le nombre de rotations (n+1) $((n+1) \in \{0, 2, 4, 6\})$ (Table 2. 4 et Table 2. 5). Il est à noter qu'étant donné que le paramètre étudié dans la Table 2. 4 est le nombre d'échelles, seules les tailles de blocs permettant une analyse multi-résolution sont présentées (c'est-à-dire les tailles (64×64) et (32×32) ; l'application de la PO sur des tailles de (16×16) et de (8×8) se limite à une seule échelle). D'après la Table 2. 4, il est vrai que, pour une taille de bloc donnée, les résultats sont en faveur du plus grand nombre d'échelles. Cependant, l'application de la PO poussée à son échelle maximale (s = 2) sur des blocs de taille (32×32) donne de meilleurs résultats que l'application de la PO sur des blocs de (64×64) et à s = 3 échelles (92,86% contre 59,69%). Nous pouvons alors en déduire que la taille des blocs filtrés a une plus grande influence que celle du nombre d'échelles de la PO sur les taux de reconnaissance. La Table 2. 3 montre également l'importance de la taille des blocs à travers un EER qui passe de 59,18% avec des blocs de taille (64×64) à 97,45% avec des blocs de taille (16×16). Ensuite, nous avons testé l'influence du choix de la mesure statistique calculée sur les blocs filtrés mais aussi la manière de la calculer (sur la PO entière ou sur chaque sous bande de la PO) sur chaque bloc filtré. D'après la Table 2.7, l'énergie décrit mieux les caractéristiques des blocs filtrés que la variance ou l'entropie.

Table 2. 2 Taux d'identification au rang 1 en fonction du type de prétraitement (PO à 1 échelle et 6 orientations).

Prétraitement	Aucun	Filtrage Gaussien	Egalisation	Egalisation+ Filtrage Gaussien		
Taux d'identification	98,47%	97,45%	95,92%	97,45%		

Table 2. 3 Taux d'identification au rang 1 en fonction de la taille des blocs (PO à 1 échelle et 6 orientations).

Taille des blocs	(64×64)	(32×32)	(16×16)	(8 × 8)	
Taux d'identification	59,18%	87,24%	97,45%	96,43%	

Table 2. 4 Taux d'identification au rang 1 en fonction du nombre d'échelles de la PO (PO à 6 orientations).

Taille des blocs		(64×64)	(32×32)			
Nbr. échelles	1	2	3	1 2		
Taux d'identification	59,18%	59,69%	59,69%	87,24%	92,86%	

Table 2. 5 Taux d'identification au rang 1 en fonction du nombre d'orientations de la PO (PO à 1 échelle sur des blocs (16×16)).

Nbr. rotations	0	2	4	6	
Taux d'identification	95,41%	92,86%	92,86%	97,45%	

Table 2. 6 Taux d'identification au rang 1 en fonction de la statistique extraite des blocs filtrés.

Mesure statistique	Variance	Entropie	Energie
Taux d'identification	92,35%	95,92%	97,45%

Table 2. 7 Taux d'identification au rang 1 en fonction de la métrique de comparaison (PO à 1 échelle et 6 orientations).

Métrique de comparaison	Euclidienne	Euclidienne City-bloc		Cosinus	Corrélation	Chebychev	
Taux d'identification	96,94%	97,45%	96,94%	89,8%	89,29%	81,63%	

L'algorithme atteint un taux de reconnaissance de 98,47% en mode identification avec les paramètres optimums suivants :

1) Les images d'iris ne subissent aucun prétraitement.

2) La normalisation a été opérée en pivotant l'iris suivant 5 angles (-4°, -2°, 0°, 2° et 4°) en prenant un angle de ROI égal à $\pi/4$ et en remplaçant la zone bruitée par une constante calculée à partir des pixels valides (non affectés par le bruit).

3) Le modèle linéarisé de l'iris de taille (64×512) est subdivisé en blocs de taille (16×16).

4) Une PO à 6 orientations (dérivées 5^{iemes} , n = 5) et une seule échelle (s = 1) est opérée sur chaque bloc (le nombre d'échelles est contraint par la taille réduite du bloc).

5) L'énergie est ensuite calculée sur chaque bloc filtré.

6) La comparaison se fait par la distance L1.

Mode vérification : La **Figure 2. 15-a** illustre les distributions des autorisés (en rose) et des imposteurs (bleu). Les deux distributions présentent une zone de chevauchement, qui constitue une zone à risque, dans laquelle un imposteur peut être reconnu comme autorisé et vice-versa. Etant donné que le choix du seuil de décision dépend de la nature de l'application, nous avons plutôt pris comme points de fonctionnement ceux qui constituent des références quelque soit l'application future de la méthode : le taux d'égales erreurs ou EER c'est-à-dire le seuil pour lequel FAR=FRR et le point de fonctionnement où FAR est égal à 1%. Nous rappelons que les résultats reportés dans la partie vérification sont obtenus en adoptant les paramètres optimums du mode identification donnés préalablement.



Figure 2. 15 (a) Distributions des autorisés (rose) et des imposteurs (bleu) ; (b) Courbe DET de la méthode de vérification d'iris par PO sur la base CASIA V1.

La courbe DET de la **Figure 2. 15-b** illustre l'évolution du FAR en fonction du FRR, l'EER étant de 7,4% avec un FFR de 20,4% au point où le FAR est égal à 1%.

iii) Comparaison avec des méthodes existantes

Etant donné que les travaux de la littérature sur la biométrie par l'iris n'ont pas été réalisés sur des bases d'images communes, et comme les taux n'ont pas été obtenus en suivant les mêmes étapes du processus de reconnaissance, la comparaison entre ces travaux est loin d'être évidente. Cependant, grâce au code open-source fourni par Libor Masek, nous avons pu comparer notre méthode à la sienne en utilisant la même base d'images et le même processus de tests [Masek 2003].



Figure 2. 16 (a) Distributions des autorisés (rose) et des imposteurs (bleu) ; (b) Courbe DET de la méthode de vérification d'iris de Masek comparée à la méthode proposée sur la base CASIA V1.

Dans sa méthode, Masek applique d'abord une transformée de Hough circulaire pour détecter l'iris puis une transformée de Hough linéaire pour détecter les paupières. Après, l'iris est normalisé selon le modèle *rubber sheet* de Daugman avant d'être filtré par un filtre Log-Gabor. L'image filtrée est alors encodée suivant le modèle 4 quadrants proposé par Daugman. Enfin la comparaison entre les codes binaires de l'iris est assurée par la distance de Hamming. La méthode de Masek (cf **Figure 2. 16**) obtient un taux d'EER = 6,0%.

La **Table 2. 8** présente les résultats des deux méthodes en termes de Taux de Comparaison Correct (Correct Match Rate – CMR) en mode identification et en termes d'EER (où FAR = FRR) en mode vérification. La comparaison des résultats sur la même base et avec le même protocole montre qu'en mode vérification, la méthode de Masek est meilleure aux faibles FAR et produit un EER de 6,0% contre 7,4% pour notre méthode. Cependant, la méthode proposée basée sur la pyramide orientable est plus performante aux FAR plus élevés (cf **Figure 2. 16**). Elle surpasse également la méthode de Masek en mode identification en réalisant un CMR de 98,47% contre 97,96%.

Table 2. 8 Comparaison de la méthode proposée en 2008 à la méthode de [Masek 2003].

Alg	gorithme	Base, Nbr. images	CMR	EER	(FAR, FRR)
[M	asek 2003]	CASIA v1, 444	97,96%	6,0%	(1%, 6,8%)
Mé [Kl	ethode proposée hiari 2008]	CASIA v1, 444	98,47%	7,4%	(1%, 20,4%)

Ceci étant dit, il faut rappeler que la méthode par PO a été réalisée dans le cadre du travail de master [Khiari 2008] et que les résultats étaient encore à améliorer. Nous avons, par la suite, essayé d'adapter la méthode à la base IV^2 en vue d'une fusion dans un algorithme bimodal avec celui développé pour le visage. Malheureusement, les résultats ont subi une chute drastique (environ 30% d'EER) sur la base d'iris IV^2 qui est une base beaucoup plus contraignante de par son nombre d'images plus important et sa qualité plus dégradée. Ces faits nous ont poussés à chercher une combinaison de méthodes de référence permettant d'obtenir de meilleurs résultats sur la base IV^2 .

2.1.3 Méthode de référence (Log-Gabor)

Ce paragraphe sera d'abord consacré à la description des différentes étapes de la méthode Log-Gabor de reconnaissance de l'iris puis à l'analyse de ses résultats.

a) Description de la méthode

La méthode de référence proposée se divise principalement en quatre étapes : d'abord la localisation se fait par des opérateurs intégro-différentiels puis la normalisation par le modèle *rubber sheet* de Daugman; ensuite, l'extraction des caractéristiques est réalisé au moyen de filtres Log-Gabor. Un codage quatre quadrants est réalisé pour générer des vecteurs caractéristiques sous forme de codes binaires qui vont être comparés au moyen de la distance de Hamming. **Localisation :** Lors de cette phase, nous avons eu l'opportunité de profiter des paramètres des cercles d'iris et des arcs des paupières fournies avec la base d'iris IV² et dont l'algorithme de détection se base sur les opérateurs intégro-différentiels initialement proposés pour la segmentation par Daugman [Daugman 1993] ; la méthode de Hough avec le code open-source de Masek [Masek 2003] testée sur la base IV² menant à un nombre trop élevé de fausses détections.

Les opérateurs intégro-différentiels sont des détecteurs très efficaces permettant d'estimer les trois paramètres des cercles interne et externe de l'iris, à savoir : le rayon et les coordonnées du centre. Les paramètres des deux cercles doivent être déterminés de manière séparée à cause de leur non-concentricité. L'opérateur agit, en effet, comme un détecteur de contour circulaire qui cherche à maximiser, itérativement, la différence entre la moyenne des gradients de deux cercles de rayons r et r+1 à des échelles successivement plus fines de l'analyse à travers l'espace des trois paramètres jusqu'à atteindre une détection précise. Une fois les contours des cercles interne et externe de l'iris détectés, une approche similaire de détection des bords curvilignes est utilisée afin de localiser les frontières des paupières supérieure et inférieure et ce en utilisant des rayons très grands approximant ainsi des arcs de cercles.

Normalisation : La même technique de linéarisation de l'anneau de l'iris par une transformation pseudo-polaire suivant le modèle « rubber sheet » de Daugman (décrite dans § 2.1.2) a été utilisée dans cette méthode afin de briser la non concentricité des cercles de l'iris et de combler l'effet d'élasticité de la texture de l'iris. Une simple technique de seuillage nous a également permis de minimiser l'effet des cils et des spots lumineux. Les paupières, les cils et les spots lumineux étant détectés, les zones correspondantes sont ensuite marquées dans un masque qui servira à ne pas prendre en considération les pixels associés lors de la phase de décision. Le masque obtenu subit un autre traitement : en le subdivisant en blocs de (4×4), on regarde si le nombre de pixels représentant le bruit est supérieur à un certain seuil déterminé de manière empirique (ici au tiers) à l'intérieur de chaque bloc. Si oui, tout le bloc est éliminé, sinon, tout le bloc est gardé. Ce traitement génère un masque final permettant de préserver une certaine continuité que ce soit dans les zones représentant le bruit ou dans la ROI. La Figure 2. 17 illustre les étapes de génération du masque final à partir d'une image d'iris.



Figure 2. 17 Etapes de normalisation et de génération d'un masque final (e) à partir d'une image initiale d'iris (a) redimensionnée à la taille de son cercle externe ; (b) étant une reproduction de (a) avec les pixels de bruit marquées en noir, (c) l'iris déroulé et (d) le masque d'origine associé.

Extraction des caractéristiques : La caractérisation se fait par un filtrage Log-Gabor 1D du modèle normalisé de l'iris - initialement à deux dimensions - qui est brisé en un ensemble de signaux 1D constitués par les lignes du modèle normalisé. Chaque ligne correspond à un anneau circulaire sur la région de l'iris. La direction angulaire a été choisie plutôt que la radiale (correspondant aux colonnes du modèle normalisé), car c'est elle qui présente le maximum d'indépendance [Masek 2003]. Les signaux 1D sont ensuite convolués avec des ondelettes de Gabor 1D.

La sortie du filtrage est codée selon le procédé de codage de phase 4 quadrants proposé dans [Daugman 1993]. Le choix a porté sur la quantification de la phase Log-Gabor, qui est plus informative, plutôt que son amplitude qui risque de souffrir de problèmes d'illumination [Oppenheim 1981]. En effet chaque phase Log-Gabor sera codée sur 2 bits parmi les codes binaires 00, 01, 10 et 11 selon sa position dans le cercle trigonométrique divisé en quatre quarts comme illustré sur la **Figure 2. 18**. Il est à noter que chaque passage entre un quadrant et un quadrant adjacent entraine un changement d'un seul bit. Ceci permet de limiter les erreurs si la phase calculée est à la frontière entre deux quadrants adjacents.



Figure 2. 18 Principe du codage de phase quatre quadrants.

Le processus d'encodage produit un code binaire, contenant des bits d'information, mais aussi un masque correspondant dans lequel sont marquées les zones corrompues (paupières, cils, spots lumineux) du code binaire associées au modèle normalisé de l'iris. Le nombre total de bits de chaque code/masque sera la résolution angulaire (nombre de lignes) multipliée par la résolution radiale (nombre de colonnes) multipliée par 2 (puisque chaque phase est codée sur 2 bits).

Décision : La distance de Hamming a été choisie puisqu'il était question de comparer des codes binaires. La formule de la distance de Hamming, donnée par l'équation (2.11), inclut des masques où le bruit est marqué, en vue de ne prendre en considération que les bits significatifs dans le calcul de la distance entre les deux modèles normalisés des iris à comparer.

$$HD = \frac{\|(codeA \otimes codeB) \cap maskA \cap maskB\|}{\|maskA \cap maskB\|}$$
(2. 11)

codeA et codeB sont deux codes d'iris, maskA et maskB sont les masques associés et \otimes représente l'opérateur OU exclusif. Littéralement la distance de Hamming calcule le nombre de bits différents et valides pour les deux iris entre le codeA et le codeB.

Afin de remédier aux incohérences liées aux inclinaisons de têtes au moment de l'acquisition, lorsque la distance de Hamming est calculée entre deux *iris codes*, l'un subit des translations de bits successives à droite et à gauche dans la direction horizontale. Chaque translation (de deux bits adjacents) correspond à un angle particulier de rotation de l'image de référence (où l'anneau de l'iris est circulaire). La comparaison de deux modèles d'iris revient donc à comparer un iris code avec un nombre d'iris codes correspondant aux différents angles de rotation. Seule la distance minimale est gardée puisqu'elle représente la meilleure correspondance entre les deux modèles (cf. **Figure 2. 19**).



Figure 2. 19 Processus de translation et de calcul de la distance de Hamming entre deux iris.

b) Evaluation par le protocole IV²

La méthode de référence Log-Gabor a été testée en utilisant la base d'iris et le protocole IV² [Petrovska-Delacretaz 2008] présentés dans le chapitre précédent (cf. § 1.4). Dans cette partie, nous exposons d'abord les résultats de la localisation. Ensuite, nous présentons les paramètres optimaux et les taux de reconnaissance de la méthode en mode vérification. Par ailleurs, l'impact de certains facteurs de qualité basés sur le taux d'occultation des iris sera étudié. Les résultats de cette étude vont nous aider à décider si une image est exploitable à des fins d'authentification. Le cas échéant, elle sera rejetée.

Corpus de données expérimenté :

Au cours de nos expériences en mode vérification, nous avons utilisé un total de 4026 images infrarouges de la base d'iris IV^2 provenant d'environ 120 personnes dont 1152 images ont servi à l'enrôlement (entre 1 et 20 images/œil) et 2874 images au test (entre 1 et 20 images/œil) donnant lieu à un total de 17156 comparaisons (dont 8681 comparaisons intra-classes (cf. § 1.4.3-a) pour de plus amples de détails sur la base d'iris IV^2).

Localisation :

La méthode de localisation basée sur les gradients circulaires de Daugman a conduit à 88 mauvaises segmentations sur un total de 4026 images d'iris, menant à un taux d'erreur de 2,2%, ce qui est un taux très réduit. L'attribution des étiquettes de « bonne » ou de « mauvaise » segmentation a été réalisée à l'œil nu. La plupart des images mal segmentées présentent de fortes dégradations dues à des zones d'occultations importantes ou à une mauvaise qualité de la région d'intérêt. C'est pourquoi elles ont été écartées dans les étapes suivantes du processus de reconnaissance.

Résultats en mode vérification :

La méthode de référence dépend de plusieurs paramètres dont l'influence a été étudiée. Ces paramètres concernent principalement le filtre Log-Gabor, la résolution du modèle rectangulaire et le nombre de décalages.

Les paramètres du filtre Log-Gabor doivent respecter certaines contraintes : les sorties de chaque filtre doivent être indépendantes de telle sorte à obtenir des *iris codes* non corrélés ; autrement, les filtres seraient redondants. Pour un maximum d'indépendance, les bandes-passantes de chaque filtre ne doivent pas présenter de recouvrement dans le domaine fréquentiel. Le nombre de filtres, leurs fréquences centrales, et les paramètres de leur enveloppe gaussienne ont fait l'objet d'une étude empirique menée par Masek à laquelle nous nous sommes référés. Il s'agit d'utiliser un seul filtre Log-Gabor à une échelle, ayant une fréquence centrale correspondant à une longueur d'ondes de 18 pixels et une bande passante d'environ 2 octaves ($\sigma/f = 0.5$). De plus amples détails sur la variation des paramètres du filtre se trouvent dans [Masek 2003]. Par ailleurs, suite à la variation des résolutions radiale et angulaire du modèle rectangulaire de l'iris lors de la phase de linéarisation de l'anneau circulaire, nous avons gardé la meilleure résolution de (32×240) pixels. A la fin, pendant la phase de *matching*, la variation du nombre de translations de 0 à 10 pixels à droite et à gauche conduit au meilleur taux de reconnaissance pour 4 translations. En utilisant les paramètres optimaux précités, la méthode réalise un EER de 3,6% (±0,3) et un FRR de 12% au point où le FAR est à 0,1%.

La Figure 2. 20 illustre les distributions inter et intra-classes. On constate que la zone de recouvrement est relativement faible.



Figure 2. 20 Distributions inter et intra-classes des scores de la méthode de référence Log-Gabor.

La Figure 2. 21 présente les courbes DET (linéaire -a- et logarithmique -b-) associées à la méthode de référence. L'échelle logarithmique permet de visualiser les résultats de manière différente en accentuant les fausses acceptations représentées par des valeurs réduites sur l'échelle linéaire.



Figure 2. 21 Courbes DET associées à la méthode de référence Log-Gabor de reconnaissance de l'iris : (a) échelle linéaire et (b) échelle logarithmique.

c) Impact de facteurs de qualité basés sur l'occultation des iris

Au cours de notre étude, nous avons choisi de tester des facteurs de qualité qui se basent sur le taux d'occultation des iris, donné par l'équation (2. 12). La Figure 2. 22 montre un exemple de masque sur lequel est calculé le taux d'occultation du modèle d'iris associé comme étant le rapport entre le nombre de pixels représentant le bruit (marqués en blanc) et le nombre total de pixels du modèle.

$$Taux_occult = \frac{nb_pixels_marqués}{nb_pixel_tot}$$

$$(2. 12)$$

$$Taux_occult = 15\%$$

Figure 2. 22 Exemple d'un masque d'iris et de son taux d'occultation.

Les pixels représentant le bruit sont les pixels occultés par les paupières, les cils, les spots lumineux ou les réflexions spéculaires ambiantes provenant par exemple de la lumière du soleil, des lunettes, du nez, de la scène...

Selon le rapport officiel IREX II du NIST -National Institute of Standards and Technologies- sur la calibration et l'évaluation de la qualité d'iris -IQCE- [IREX II 2011], l'aire utilisable de l'iris (égale à 1 – le taux d'occultation) a une influence très importante sur la performance du système de reconnaissance. En effet, les algorithmes participant à l'évaluation IQCE montrent qu'un taux d'occultation faible minimise les scores intraclasses et maximise les scores interclasses. Inversement, un taux d'occultation élevé translate la distribution des scores autorisés vers la droite (c'est-à-dire vers celle des imposteurs) et la distribution des scores imposteurs vers la gauche (c'est-à-dire vers celle des autorisés) favorisant, ainsi, un recouvrement plus prononcé des deux distributions. Il est, cependant, à noter que l'effet de l'occlusion sur la distribution des autorisés est plus important que sur celle des imposteurs. Néanmoins, la translation des deux distributions implique que les taux d'erreurs FAR et FRR sont tous les deux affectés par le taux d'occultation. Plus encore, il a été montré par les mêmes expériences de l'évaluation IQCE qu'un rejet de 2% d'images avec un fort taux d'occultation améliore le FRR de l'ordre de 10%. Dans le même but d'étudier l'influence du taux d'occultation sur la performance des algorithmes de reconnaissance par l'iris, la base d'iris a été partitionnée en 3 sous-ensembles : « Best » désigne les 15% d'images présentant le plus faible taux d'occultation, « Worst » désigne les 15% d'images présentant les taux d'occultation les plus forts et « Medium » désigne le reste de la base. Les courbes DET des algorithmes participant à l'évaluation montrent, comme on pourrait s'y attendre, que les meilleures performances sont associées aux images ayant les occlusions les plus faibles et que les pires performances sont associées aux images ayant les occlusions les plus prononcées.

Dans ce qui suit, une étude qui se réfère au partitionnement de la base d'images effectuée par le NIST dans l'évaluation IQCE est réalisée sur la base IV². Ensuite, deux critères de qualité différents basés sur le taux d'occultation des iris sont présentés.

i) Partitionnement de la base IV² selon l'évaluation IQCE du NIST

Nous avons partitionné la base IV^2 comme décrit dans le rapport IREX II du NIST. Les 15% d'images d'iris appartenant à la partition « Best » correspondent à des iris ayant un taux d'occultation entre 0 et 6,6%. Tandis que les 70% d'images d'iris de la partition « Medium » ont un taux d'occultation entre 6,6% et 36,7%. Les 15% d'images restantes appartenant à la partition « Worst » ont des iris occultés à plus de 36,7% (cf. **Table 2. 9**). Il est utile de noter que ce premier test ne prend pas en compte les comparaisons interpartitions. Les résultats sont illustrés dans la **Figure 2. 23** et la **Table 2. 9** et confirment ceux de la compétition IQCE sur d'autres bases telles que la base ICE 2006. Ce qui est à remarquer à partir de la **Table 2. 9**, c'est que ce sont les 15% pires occultations (partition « Worst ») qui détériorent notoirement les résultats avec un EER = 16,4% et un FRR = 95,9% au point de fonctionnement où FAR = 0,001.

L'élimination de cette partition à partir de la liste des images à considérer pourrait être une solution potentielle pour améliorer la performance de la méthode de reconnaissance sans pour autant altérer les statistiques. En effet, cela conduirait au rejet d'un nombre limité de comparaisons de test par rapport au nombre initial de comparaisons qui est égal à 17156. Plus précisément, 4320 comparaisons faisant intervenir des images de la partition « Worst » seraient rejetées, que ce soit des comparaisons intra-partition (Worst-Worst, qui sont au nombre de 828, comme indiqué dans la **Figure 2. 23** et la **Table 2. 9**) ou inter-partitions (Worst-Medium et Worst-Best).

Table 2. 9 Résultats de l'algorithme de référence Log-Gabor sur les partitions de la base IV²selon le taux d'occultation des iris.

Partition IV ²	Intervalle taux_occult (%)	%EER	%FRR (FAR = 0,1%)	N. comp_tot	N. comp_intra
Best	[0;6,6]	0,0 (±0,0)	0,0	857	698
Medium	[6,6;36,7]	1,4 (±0,3)	2,3	9017	4840
Worst	[36,7; 100]	16,4 (±3,4)	95,9	828	631



Figure 2. 23 EER et nombre de comparaisons de l'algorithme de référence Log-Gabor sur les partitions Best, Medium et Worst de la base IV².

ii) Critères de qualité basés sur l'occultation des iris

Dans cette partie, nous décrivons d'abord deux critères de qualité des images d'iris basés sur l'occultation : le taux d'occultation des iris et le taux d'occultation des comparaisons d'iris. Il s'agit, ensuite, de rejeter, à partir de la liste de comparaisons de test, les iris dont le facteur de qualité est supérieur à un certain seuil. Les seuils relatifs aux deux critères présentés sont déterminés de manière empirique. Pour ce faire, nous étudions l'influence des seuils sur la performance de la méthode, mais aussi sur la validité des statistiques suite au rejet des iris/comparaisons ayant des occultations audelà des seuils autorisés.

Taux d'occultation des iris

Au vu des résultats du partitionnement de la base selon l'évaluation IQCE du NIST, nous avons voulu investiguer de plus près l'effet du critère de qualité « Taux d'occultation des iris » défini dans § 2.1.3-c) par l'équation (2.12) et schématisé dans la Figure 2. 22. Pour ce faire, et contrairement au travail réalisé dans l'étude précédente, nous allons dans la suite faire varier de manière progressive le seuil toléré du taux d'occultation des iris et examiner son influence à la fois sur la performance de la méthode (aux deux points de fonctionnement EER et FRR(FAR = 0,001)) et sur le nombre de comparaisons de tests. Les résultats sont illustrés par la Figure 2. 24 où la courbe représente le taux d'EER et l'histogramme représente le nombre de comparaisons gardées en fonction du seuil d'occultation. Généralement, l'EER diminue avec le seuil d'ocultation toléré. Néanmoins, il arrive qu'il subisse une légère augmentation comme nous pouvons le constater pour le seuil de 10%. Ceci peut s'expliquer par le fait que l'élimination des comprarison selon le critère d'occultation peut toucher à plus de compraraisons correctement classées qu'à des comparaisons mal classées. La Table 2. 10 fournit plus de détails sur les points où la pente de la courbe de la Figure 2. 24 varie le plus. Ces résultats montrent que l'on arrive à garder un taux d'erreur nul avec le seuil de 5% d'occultation. Cependant, pour atteindre ce seuil, le nombre de comparaisons gardées atteint à peine 4% du nombre initial de comparaisons (686 comparaisons contre 17156) avec, de plus, un nombre déséquilibré de comparaisons inter et intraclasses (566 comparaisons intraclasses contre 120 comparaisons interclasses). En augmentant progressivement le seuil d'occultation toléré, avec comme objectif de faire un compromis entre le taux d'EER que l'on veut minimiser et le nombre de comparaisons gardées que l'on veut maximiser, nous constatons d'après la **Table 2.10** que le seuil de 35% d'occultation tolérée répond le mieux à ce compromis. En effet, à ce seuil, l'EER diminue de près de deux tiers (1,3% contre 3,6% au seuil 100%), tout un gardant 74% du nombre initial de comparaisons. D'un autre côté, le fait d'imposer un seuil d'occultation plus strict de 20% divise en deux le nombre de comparaisons sans pour autant apporter un gain considérable sur le taux d'erreur comparé au seuil 35% (EER = 1% au seuil 20% contre EER = 1,3% au seuil 35%). Par ailleurs, si l'on augmente la tolérance de l'occultation en passant à un seuil de 55%, il est vrai que l'on garde la quasi-totalité des comparaisons initiales (96%) mais le taux d'erreurs diminue seulement de 0,9 (EER = 2,7% contre 3,6% au seuil 100%) et dépasse le double de l'EER au seuil 35% qui est égal à 1,3%.

Suite à ces éléments, nous optons pour garder le seuil de 35% qui revient à écarter les iris dont le taux d'occultation va au-delà de 35% (et par la suite, à éliminer, à partir de la liste de test, les comparaisons dans lesquelles ces iris interviennent).

La différence entre le seuil de 100%, qui correspond à garder tous les iris quelque soit leurs taux d'occultation, et le seuil de 35%, est illustrée par la **Figure 2.25** à travers les chevauchements des distributions inter et intraclasses correspondantes et par la **Figure 2.26** à travers les courbes DET en représentations linéaire et Log.

Seuil_occult	%EER	%FRR (FAR = 0,1%)	N. comp_tot	N. comp_intra
100%	3,6 (±0,3)	12,0	17156	8681
80%	3,6 (±0,3)	12,0	17130	8667
55%	2,7 (±0,3)	6,5	16479	8379
35%	1,3 (±0,2)	2,0	12730	6676
20%	1,0 (±0,3)	1,2	6754	3953
15%	0,1 (±0,1)	0,1	4216	2640
5%	0,0 (±0,0)	0,0	686	566
	0,04 0,035 0,03 0,025 E 0,02 0,015 0,015 0,01 0,005 0 100 90 80 70 60 5 Taux d'occut	10 14 14 12 10 10 10 10 10 10 10 10 10 10	Ncon b Nombre ce de con Noco d Noco Noco Noco Noco Noco Noco No	n

Table 2. 10 Résultats de l'algorithme de référence Log-Gabor sur les partitions de la base IV^2 selon le taux d'occultation des iris.

Figure 2. 24 EER et nombre de comparaisons de l'algorithme de référence Log-Gabor en fonction du taux d'occultation des iris de la base IV².



Figure 2. 25 Distributions inter et intra-classes des scores de la méthode Log-Gabor correspondant aux seuils 100% et 35% d'occultation des iris.



Figure 2. 26 Courbes DET associées à la méthode Log-Gabor correspondant aux seuils 100% et 35% d'occultation des iris : (à gauche) échelle linéaire et (à droite) échelle logarithmique.

Taux d'occultation de la comparaison entre deux iris

Ce critère de qualité est défini pour chaque paire d'iris à comparer dans la liste de test. Il s'agit de déterminer l'union des pixels marqués des masques des deux iris en question (ou l'intersection des pixels valides) puis de calculer la proportion de ces pixels par rapport au nombre total de pixels du masque. Le taux d'occultation entre deux iris peut être formulé selon l'équation (2. 13). Il est à noter que le critère du taux d'occultation des comparaisons est plus restrictif que celui des iris. En effet, d'après l'exemple de la **Figure 2. 27** et pour un même seuil de 35%, la comparaison faisant intervenir les deux iris en question serait gardée dans la liste de tests si le critère considéré était le « taux d'occultation des iris », puisque leurs taux d'occultations individuels (28,13% et 29,38%) sont inférieurs à 35%. Par contre, si on considère le « taux d'occultation de la comparaison » entre ces mêmes iris, cette comparaison serait éliminée de la liste de tests (avec 35,83% d'occultation > 35%) bien que les iris correspondants ne soient pas « trop » occultés (toujours par rapport au seuil de 35% que l'on a pris comme exemple).

$$Taux_occult_comp = Taux_occult(Mask(iris1) \cup Mask(iris2))$$
(2.13)



Figure 2. 27 Exemple de taux d'occultation d'une comparaison entre deux iris (haut) et illustration de leur masque union (bas).

Une étude comparable à celle du critère « taux d'occultation des iris » a été menée pour étudier la performance de la méthode en fonction du seuil d'occultation des comparaisons entre les iris. Les résultats sont reportés dans la **Table 2. 11** et schématisés sur la **Figure 2. 28** à travers la courbe d'EER et l'histogramme des nombres de comparaisons retenues. Ces résultats confirment ceux des deux paragraphes précédents dans le sens où l'élimination des iris trop occultés permet de réduire le taux d'EER d'une manière très considérable.

Le seuil que nous avons choisi pour le facteur de qualité « taux d'occultation des comparaisons entre les iris » s'élève aussi à 35% car il permet de réduire au tiers l'EER (passage de 3,6% à 1,2%) tout en gardant un nombre de comparaisons assez élevé et à la fois équilibré entre les comparaisons inter et intraclasses. A ce seuil, les comparaisons éliminées sont celles qui font intervenir des iris occultés à plus de 35% (taux individuel), plus celles dont les iris sont occultés à moins de 35% mais dont le taux d'occultation de la comparaison dépasse les 35%.

Seuil_occult	%EER	%FRR (FAR = 0,1%)	N. comp_tot	N. comp_intra
100%	3,6 (±0,3)	12,0	17156	8681
80%	3,5 (±0,3)	8,5	17107	8657
55%	2,4 (±0,3)	6,0	16307	8300
35%	1,2 (±0,2)	1,8	12202	6458
20%	0,7 (±0,3)	0,9	6157	3711
15%	0,1 (±0,1)	0,1	3726	2448
5%	0.0(+0.0)	0.0	582	491

Table 2. 11 Résultats de l'algorithme de référence Log-Gabor sur les partitions de la base IV^2 selon le taux d'occultation des comparaisons entre les iris.



Figure 2. 28 EER et nombre de comparaisons de l'algorithme de référence Log-Gabor en fonction du taux d'occultation des comparaisons entre les iris de la base IV².

d) Comparaison avec des méthodes existantes

Bien qu'il soit difficile de comparer la méthode de référence proposée avec les autres méthodes qui ont été développées auparavant - à cause de la diversité des bases d'images utilisées, des protocoles de test ainsi que des processus de la reconnaissance – il reste utile de situer sa performance par rapport à celles de méthodes connues développées jusque là. La **Table 2. 12** reporte les résultats de la méthode de référence proposée face à ceux d'autres travaux de l'état de l'art en mode vérification en spécifiant les bases d'iris utilisées ainsi que le nombre d'images associées.

Nous constatons que la méthode de référence sans facteur de qualité possède un EER égal à 3,6% qui se trouve en plein centre de la plage des EERs de la littérature se situant entre 0% et 7% environ. Par contre, le FRR = 12% au point où FAR = 0,1% est élevé comparé aux FRRs des autres méthodes qui ont spécifié le taux d'erreur à ce point de fonctionnement (compris entre 0% et 2%).

En introduisant la notion de facteur de qualité, nous sommes parvenu à réduire le taux d'EER et à atteindre 1,2% ce qui nous situe au début de l'intervalle des EERs des méthodes de l'état de l'art de la **Table 2. 12**. En même temps, le taux de FRR au point de fonctionnement FAR = 0,1% a chuté de 12% à 1,8% et se situe désormais dans l'intervalle [0%, 2%] des FRRs des autres méthodes. Tous ces éléments sont en faveur de la méthode de référence proposée, surtout après l'introduction des critères de qualité.

Algorithme	Base (Nbr. Images)	EER	(FAR, FRR)
[Masek 2003]	CASIA v1 (444)	5,8%	(0,02%;0,94%)
[Daugman 2006]	UAE national border-crossing system (632 500)	0%	(0%;0%)
[Khiari 2008]	CASIA v1 (444)	7,3%	(0,29%;0,94%)
[Krichen 2009]	ICE (2093) Casia v2 BioSecure CBS (240) IV ² (4026)	0,68% 1% 0,78%	(0,1%;1%) (0,1%;2%) (0,1%;1%)
[Rathgeb 2010]	CASIA V3 (-)	entre 1 et 2%	(-;-)
[Li 2012]	UBIRIS v2 (1000)	0,23%	(-;-)
[Chen 2014]	CASIA v1 (260) CASIA v3 (240) MMU v1 (179) JLUBRIRIS v1 (729)	3,23% 4,55% 7,16% 2,98%	(-;-) (-;-) (-;-) (-;-)
Méthode de référence sans facteur de qualité Khiari (2015)	IV ² (4026)	3,6%	(0,1%; 12,0%)
Méthode de référence avec facteur de qualité occult_iris Khiari (2015)	IV ² (4026)	1,3%	(0,1%;2,0%)
Méthode de référence avec facteur de qualité occult_comparaison Khiari (2015)	IV ² (4026)	1,2%	(0,1%;1,8%)

 Table 2. 12 Comparaison de la méthode de référence avec des méthodes existantes en mode vérification.

2.1.4 Approche de reconnaissance par l'iris retenue

Après avoir testé la méthode de référence Log-Gabor et l'avoir comparée avec d'autres méthodes de l'état de l'art, il était question de choisir quelle variante des critères de qualité il était plus judicieux de garder en vue d'une fusion avec la modalité du visage. Il est vrai que les performances sont assez proches pour les deux critères proposés, mais nous retenons la méthode Log-Gabor avec le critère « taux d'occultation de la comparaison des iris » car il présente un léger avantage par rapport à la méthode utilisant le critère « taux d'occultation des iris » (1,2% d'EER contre 1,3% d'EER).

Toutefois, dans un mode opérationnel, il serait légitime de mettre en cause l'intérêt de l'approche du rejet des images/comparaisons suivant l'un ou l'autre des critères de qualité proposés. En effet, cela induirait des demandes successives de réacquisition d'image tant que le seuil minimum de qualité n'a pas été atteint, ce qui risque d'incommoder l'utilisateur. C'est pourquoi nous allons proposer dans le chapitre 4, une solution basée sur la multi-biométrie, qui exploite les critères de qualité ici proposés pour pondérer les scores au moment de la fusion iris/visage sans pour autant rejetter les images.

2.2 Biométrie par le visage

Cette partie commence par décrire les principales méthodes de l'état de l'art en biométrie faciale. A l'issue de cette étude, nous nous intéresserons particulièrement aux méthodes combinées. Trois méthodes seront construites en se basant sur l'algorithme de la Pyramide Orientable – PO qui a été appliqué précédemment en biométrie par l'iris. Cette méthode sera couplée à trois méthodes de réduction de l'espace des caractéristiques en utilisant plusieurs niveaux de fusion. Nous allons recenser les limites des méthodes individuelles, puis nous allons présenter les résultats des méthodes combinées proposées en vue de choisir une première méthode de reconnaissance faciale comme candidate dans l'approche multimodale finale basée sur l'iris et le visage.

2.2.1 Etat de l'art

La reconnaissance faciale se présente comme étant une alternative intéressante parmi les solutions biométriques à cause de son caractère non invasif et de sa facilité d'utilisation, ce qui explique l'intérêt grandissant dont elle jouit auprès des organismes de recherche publics et privés. Cependant, la reconnaissance faciale doit affronter plusieurs difficultés provenant de facteurs indépendants de l'identité du visage, mais pouvant dégrader considérablement les taux de reconnaissance. Parmi ces facteurs, citons la variation de la pose qui peut engendrer des occultations de certaines parties du visage et des déformations par étirement ou compactage ; la variation de l'éclairage qui provoque l'apparition d'ombres accentuant ou masquant certaines caractéristiques faciales ; les expressions faciales qui produisent des changements sur la forme géométrique et la position des caractéristiques faciales ; les occultations partielles, qu'elles soient intentionnelles ou non ; l'intervalle de temps entre les prises de vues et le vieillissement. Malgré cela, les systèmes automatiques de reconnaissance faciale doivent rester invariants à toutes ces sources de variabilité.

Avant d'entrer dans le processus de la reconnaissance proprement dite, les images de visages doivent, d'abord, passer par une chaîne de traitements préalables, composée de trois étapes : la première est celle de la détection de visages, qui consiste à dire si l'image contient ou non un ou plusieurs visages ; la deuxième étape est celle de la localisation d'un visage dans laquelle se détermine la position exacte du visage dans l'image (**Figure 2. 29- b**) ; la troisième étape étant celle de la normalisation par rapport à la rotation et à l'échelle en se basant principalement sur la connaissance, a priori, des coordonnées des yeux (cf. **Figure 2. 29- c** à **e**). La normalisation consiste, d'abord, à pivoter le visage de telle sorte à avoir un axe interoculaire horizontal (**Figure 2. 29- c**). Le visage est, ensuite, remis à l'échelle de telle façon que toutes les images de visages soient à la même échelle et avec la même distance interoculaire (**Figure 2. 29- d**). A la fin de cette étape, un rognage est appliqué à l'image par rapport à la position des yeux, de telle sorte que seule la région d'intérêt du visage utile à sa reconnaissance soit gardée (**Figure 2. 29- e**). Toutes les images obtenues sont alors à la même taille.

De nombreuses méthodes ont été utilisées en biométrie faciale que ce soit pour des tâches de détection/localisation ou de reconnaissance. Dans notre étude, nous allons nous focaliser sur les méthodes de reconnaissance qui sont principalement classées en deux catégories : les approches locales et les approches globales.



Figure 2. 29 Etapes du prétraitement dans une chaîne de vérification biométrique de visage : (a) image initiale, (b) détection de la posistion du visage et des yeux, (c) image pivotée, (d) image redimensionnée, (e) région d'intérêt (visage) normalisée.

a) Les approches locales

Ce sont les plus anciennes. Elles se basent sur une analyse bas niveau de l'image, par exemple à travers des **caractéristiques géométriques** du visage telles que les yeux, le nez, la bouche [Kanade 1973] et la couleur de la peau [Singh 2003]. Les positions ainsi que les statistiques locales de ces caractéristiques géométriques peuvent également intervenir dans l'analyse de l'image. Ensuite, une correspondance des modèles locaux est engagée en respectant certaines contraintes géométriques.

L'analyse bas niveau peut également se faire à travers une **analyse fréquentielle**, notamment par **les méthodes d'analyse multirésolution**. Le point fort de ce type de méthodes est leur faculté à apporter une vision/information multi-niveaux sur le contenu de l'image. C'est pourquoi, les méthodes d'analyse multirésolution telles que les ondelettes de Gabor [Mellakh 2009-a], les transformées en ondelettes discrètes (Discrete Wavelet Transform – DWT) [Gumus 2010] [Chelali 2014], les Pyramides Orientables [Zhang X. 2005], ... ont été largement utilisées en reconnaissance faciale.

Parmi les approches locales, **les opérateurs LBP** [Harwood 1993] [Ojala 1996] ont aussi été largement utilisés en biométrie faciale [Hadid 2008]. Ils constituent des descripteurs non paramétriques simples et efficaces qui labélisent les pixels d'une image en seuillant le voisinage de chaque pixel par rapport à la valeur du pixel central et en considérant les résultats comme un nombre binaire appelé Local Binary Pattern. L'histogramme de l'image labélisée est ensuite calculé, donnant lieu à un opérateur LBP pouvant être exploité comme un descripteur de texture. Depuis son introduction à la reconnaissance de visage, la méthodologie LBP a été développée avec plusieurs variantes pour améliorer les performances d'analyse des images faciales [Huang D. 2011]. Ces variantes se sont concentrées sur plusieurs aspects de la LBP originale dont les plus connues sont la LBP^{RI} invariante à la rotation, le modèle uniforme LBP^{U2} et les modèles uniformes invariants à la rotation LBP^{RIU2}. Des études récentes ont montré que plus de 90% des modèles discriminants représentant les visages sont uniformes [Huang D. 2011].

D'autres méthodes locales à apprentissage adaptatif telles qu'Adaboost [Viola 2004] ont été utilisées en reconnaissance de visage [Yoon 2010] [Lee 2013] [Jiang 2015]. Le principe d'Adaboost consiste à promouvoir une séquence de classifieurs faibles pour former un classifieur fort avec des poids qui sont mis à jour dynamiquement selon les erreurs des apprentissages précédents – les classifieurs faibles agissant au niveau de chaque pixel de l'image issue de la différence entre deux images de visage à comparer.

Les méthodes locales sont robustes aux occlusions partielles mais vulnérables au bruit, aux changements d'expressions et de poses de tête, ce qui entraîne des imprécisions au niveau des caractéristiques [Yang 2002].

Cependant, des descripteurs plus récents à base de SIFT (Scale Invariant Feature Transform) [Lowe 2004] se sont montrés très efficaces et robustes face aux distorsions, aux changements de luminosité, d'échelle et d'orientation. L'extraction des caractéristiques SIFT se fait en quatre étapes. D'abord, on calcule la position d'éventuels points d'intérêt de l'image en détectant les maximas et minimas d'un ensemble de filtres DoG (Difference of Gaussians) appliqués à différentes échelles. Ensuite, ces positions sont redéfinies en écartant les points de bas contraste. Puis, une orientation est assignée à chaque point clé en se basant sur des caractéristiques locales de l'image. Finalement, un descripteur local de caractéristiques est calculé en chaque point clé. Ce descripteur est basé sur le gradient local de l'image, transformé selon l'orientation du point clé afin d'assurer une invariance par rapport à l'orientation. A chaque caractéristique correspond alors un vecteur qui identifie de manière distincte le voisinage autour du point clé. La robustesse de cette méthode et son pouvoir discriminant expliquent pourquoi elle est de plus en plus utilisée en détection et en reconnaissance de visage [Bicego 2006] [Barbu 2013].

b) Les approches globales

Elles consistent à utiliser le visage dans sa globalité. Généralement, une phase préalable d'analyse par des *méthodes statistiques* peut être réalisée afin d'obtenir une représentation et/ou une classification des visages à travers, par exemple, l'Analyse en Composantes Principales (*Principle Component Analysis* – PCA) comme utilisé dans [Slavković 2012], l'Analyse Linéaire Discriminante (*Linear Discriminant Analysis* – LDA) comme opéré par [Chen 2000] [Wang Z. 2015], l'Analyse en Composantes Indépendantes (*Independent Component Analysis* – ICA) comme réalisé dans [Draper 2003], les moments de Zernike comme dans [Chaari 2009], les classifieurs bayésiens comme procédé par [Liu 2005] ou encore les modèles déformables comme dans le travail de [Sauer 2011].

L'Analyse en Composantes Principales PCA ou Eigenfaces [Turk 1991] vise à obtenir une représentation compacte globale qui décorrèle les vecteurs de données des redondances afin de réduire l'espace de représentation. L'Analyse Discriminante Linéaire LDA ou Fisherfaces [Belhumeur 1997], à la différence de la PCA, cherche l'espace de projection maximisant la séparation linéaire des données, c'est-à-dire celui qui minimise les variations à l'intérieur de la même classe et maximise les variations entre les différentes classes. L'Analyse en Composantes Indépendantes ICA [Bartlett 2002], introduite comme généralisation de la PCA, vise à caractériser les données en les statistiquement indépendantes et en utilisant rendant des caractéristiques discriminantes dépassant les premier et second ordres. Les moments de Zernike [Zernike 1934] sont une autre variante des méthodes globales statistiques basés sur l'utilisation d'une famille de fonctions orthogonales, permettant la génération de descripteurs non-redondants. Ils sont très utilisés pour leur propriété d'invariance par rapport à la translation, à l'échelle et à la rotation.

La biométrie faciale a été également abordée par les **classifieurs Bayésiens** [Moghaddam 1996] qui sont appliqués sur des différences de paire d'images. Deux classes intra et extra personnelles sont définies à partir de ces différences sous l'hypothèse qu'elles sont issues de deux distributions gaussiennes discrètes. La probabilité qu'une différence d'images appartienne à la classe intra (c'est-à-dire que les deux images à l'origine de cette différence proviennent de la même personne) est alors donnée par la loi de Bayes.

Cootes et Taylor ont proposé en 2004 d'autres alternatives d'apprentissage statistique reposant sur **les modèles déformables**, que ce soit pour la détection de visage à travers des modèles statistiques de formes (*Active Shape Model – ASM*) ou pour la reconnaissance par des modèles statistiques de formes et de texture (*Active Appearance Models – AAM*). Le principe repose sur l'appariement d'un modèle statistique de forme et/ou d'apparence d'un objet avec une nouvelle image. Ce modèle est construit dans la phase d'apprentissage durant laquelle est fourni un ensemble d'images labélisées avec des coordonnées de points repères ou *landmarks* qui apparaissent sur toutes les images. Le modèle se déforme sur la nouvelle image, au fil des itérations, afin de représenter les variations de forme et/ou de texture et d'en synthétiser de nouveaux modèles dont les variations sont proches de celles de l'ensemble d'apprentissage [Cootes 2004].

Les principaux avantages des méthodes globales statistiques sont leur représentation compacte, leur mise en œuvre rapide et la complexité moyenne des calculs de base nécessaires dans la phase de *matching*. Par contre, elles sont sensibles aux variations d'éclairage, de pose et d'expression faciale. En effet, la moindre variation des conditions de l'environnement ambiant entraîne des changements inéluctables dans les valeurs des pixels qui sont traités directement.

Suite à la phase éventuelle d'analyse statistique, vient une phase de classification en utilisant, notamment, **des méthodes basées sur l'intelligence artificielle** (**IA**) telles que les Modèles de Markov Cachés (*Hidden Markov Models – HMM*) comme réalisé dans

[Takahashi 2010], les réseaux de neurones (*Neural Networks – NN*) comme dans [Kanade 2012], les Machines à Vecteurs de Support (*Support Vector Machines – SVM*) comme opéré par [Huang H. 2012] ou encore les Forêts aléatoires ou *Random Forests* comme réalisé dans [Dantone 2012]. Une description succincte de chacune de ces méthodes est présentée dans ce qui suit.

Les HMM [Baum 1966] sont des modèles statistiques dans lesquels le système représenté est supposé être un processus Markovien (la probabilité des états futurs, étant donné les états passés et l'état présent, ne dépend que de l'état présent et non pas des états passés : absence de « mémoire ») avec des états non observables/cachés mais dont les sorties qui en dépendent sont visibles. Comme chaque état a une distribution de probabilité sur les sorties possibles, la séquence de sorties générées par un HMM donne des informations sur la séquence d'états. En reconnaissance de visage, des fenêtres chevauchantes balayant l'image du visage sont approximées par des HMM qui génèrent une séquence d'observation de l'image considérée. Un modèle est construit en phase d'apprentissage pour chaque classe (chaque classe correspond à un individu). Etant donné une image de visage candidate, sa séquence d'observation est introduite comme entrée aux modèles de classes construits lors de l'apprentissage. La classe qui donne la plus grande probabilité de la séquence d'observation introduite est désignée comme l'identité probable de ce visage [Takahashi 2010]. Les calculs des HMM convergent rapidement, ce qui est pratique pour les applications en temps réel.

Les réseaux de neurones (Neural Networks – NN) sont des modèles de calcul inspirés par le système nerveux central du cerveau qui est capable de faire un apprentissage automatique et de reconnaitre les formes [McCulloch 1943]. Ils sont formés par des unités interconnectées appelées « neurones » organisées en couche d'entrée, couches cachées et couche de sortie. Un neurone connecté à un autre implique que l'activité de l'un influence directement celle de l'autre. Cette influence d'activité est modélisée sous forme d'un « poids » de connexion entre les deux neurones. L'apprentissage du réseau se fait à travers la modification de ces poids. Plusieurs variantes de NN ont été utilisées en reconnaissance faciale [Rowley 1996] [Lawrence 1997] telles que les perceptrons multicouches (MultiLayer Perceptron – MLP), les réseaux à Fonction Radiale de Base – RBF [Chelali 2014], les réseaux multicouches à rétro-propagation, et les réseaux flous [Patel 2012].

Les SVM [Cortes 1995] [Osuna 1997] sont un ensemble de techniques d'apprentissage supervisé destinées à résoudre des problèmes de classification et de régression, qui sont capable d'analyser des données et de reconnaitre des formes, en particulier les visages [Huang H. 2012]. Les SVM commencent par représenter les exemples d'apprentissage dans un espace caractéristique de grande dimension H. Ensuite, on recherche dans H l'hyperplan optimal pour séparer au mieux les exemples appartenant à des classes différentes tout en maximisant la distance de chaque classe par rapport à l'hyperplan. Les nouveaux exemples sont alors représentés dans ce même espace. La prédiction d'appartenance à une classe donnée est réalisée en se basant sur le côté de l'hyperplan dans lequel ils sont représentés. Il est à noter que les SVM sont considérés parmi les techniques les plus efficaces dans les problèmes de classification de formes. Leur performance est du même ordre, voire dépasse celle des réseaux de neurones.

Les **Random Forests** [Breiman 2001] sont des modèles prédictifs d'apprentissage supervisé utilisés pour la classification et la régression, notamment en reconnaissance faciale [Ghosal 2009]. Le principe consiste à agréger une multitude d'arbres décisionnels binaires – où les feuilles représentent les valeurs de la variable cible et les embranchements correspondent à des combinaisons de variables d'entrée qui mènent à ces valeurs – dans le but de générer la classe de sortie dominante parmi les classes des arbres individuels. La méthode peut être vue comme une amélioration des arbres décisionnels, de par leur inconvénient du surapprentissage, en introduisant l'idée de la sélection aléatoire des variables qui interviennent dans la construction des modèles. La sélection aléatoire d'un nombre réduit de prédicteurs potentiels à chaque étape, accroît significativement la variabilité en mettant en avant nécessairement d'autres variables. Chaque modèle de base est évidemment moins performant mais l'agrégation conduit finalement à de meilleurs résultats que ceux des arbres décisionnels non aléatoires.

Les approches globales basées IA ont l'avantage d'avoir un temps de calcul rapide en phase de test. Elles sont robustes au bruit et aux déformations de la région faciale.

c) Les approches hybrides

Les méthodes hybrides permettent d'associer les avantages des méthodes globales et locales en combinant la détection de caractéristiques géométriques (ou structurales) avec l'extraction de caractéristiques d'apparence locales. Elles permettent d'augmenter la stabilité de la performance de reconnaissance lors de changements de pose, d'éclairage et d'expressions faciales.

Parmi les approches hybrides, *les techniques modulaires* telles que la PCA modulaire [Pentland 1994] et la LDA modulaire [Jumahong 2011] procèdent par application de méthodes globales de manière modulaire en différentes régions faciales puis par combinaison, pour générer un modèle global alliant plusieurs modèles locaux, ce qui conduit à une plus grande robustesse aux changements d'éclairage, d'expressions faciales et aux occultations partielles.

D'autres variantes d'approches hybrides sont *les techniques d'appariement de graphes* telles que le Gabor-EBGM (Gabor Elastic Graph Matching) [Wiskott 1997] et le HOG-EBGM (Histograms of Oriented Gradients Elastic Bunch Graph Matching) [Albiol 2008] qui structurent des nœuds appelés « *jets* » représentant les caractéristiques locales de visages. Elles sont généralement meilleures que les autres techniques de reconnaissance en termes d'invariance par rapport à la rotation mais très coûteuses en de temps de calcul que ce soit pour la construction du modèle ou pour l'appariement des graphes.

Les approches hybrides peuvent également provenir d'une *combinaison de méthodes locales avec des méthodes globales*. Tel est le cas de l'algorithme Gabor/LDA proposé dans [Mellakh 2009-a] qui effectue une convolution avec des

ondelettes de Gabor orientées afin de créer des vecteurs contenant la localisation et la valeur d'amplitudes énergétiques locales ; ces vecteurs sont ensuite envoyés dans un algorithme LDA afin de réduire la dimension des données. Gumus et al. ont proposé, dans ce sens, un algorithme **DWT (Dicrete Wavelet Transform)/SVM** qui, après avoir extrait les caractéristiques faciales par une transformée d'ondelette discrète, les compare à l'aide d'un classifieur SVM [Gumus 2010].



Figure 2. 30 Catégorisation des méthodes principales de reconnaissance faciale.

La Figure 2. 30 fournit une synthèse des algorithmes principaux de reconnaissance faciale et la Table 2. 13 reporte les performances de quelques algorithmes de l'état de l'art telles que citées dans les publications correspondantes (en EER pour les algorithmes en mode vérification et en CMR – *Correct Match Rate* pour ceux en mode identification), en spécifiant les bases d'images utilisées.

Table 2.	13	Résultats	reportés	de	quelques	algorithmes	de	la	littérature	en	biométrie	faciale
jusqu'en 2	2015	5.										

Référence	Algorithme	Base (Nbr. Images)	Performance
[Bicego 2006]	SIFT	BANCA (1040)	EER = 6,35% et 10,51% selon l'ensemble de test
[Hadid 2008]	LBP	FERET (-)	CMR = 50% à 93%
[Ghosal 2009]	LBP/Gabor & Random Forest	FERET (488) IITKanpur (1656)	EER = 1,93% EER = 0,40% à 0,16%
[Gumus 2010]	DWT SVM	ORL (400)	CMR = 95,3%
[Slavkovic 2012]	PCA	ORL (230)	CMR = 97,5%
[Chelali 2014]	Haar DWT/Gabor & MLP/RBF NN	privée (200) ORL (200)	CMR = 95% à 97% CMR = 85% à 96%
[Wang Z. 2015]	DSLFDA Double Sparse Local Fisher Discriminant Analysis	ORL (400) Yale (165) CMU-PIE (1632)	CMR = 97,88% CMR = 81,47% CMR = 96,63%
[Jiang 2015]	LBP & Multiclass Adaboost ELM Extreme Learning Machine	ORL (-) Yale (-)	CMR = 93,61%

d) Approche adoptée

Au vu des méthodes citées dans l'état de l'art, nous avons voulu construire une méthode hybride afin de profiter à la fois des avantages des méthodes locales et de ceux des méthodes globales. Nous allons proposer des méthodes combinées se basant sur une méthode locale multirésolution : la Pyramide Orientable – PO. Cette méthode a déjà été appliquée à la première méthode biométrique par l'iris proposée dans le chapitre précédent et nous avons voulu investiguer ses performances en biométrie faciale.

Dans les parties suivantes, nous allons commencer par la présentation des résultats des méthodes individuelles qui seront par la suite utilisées dans les méthodes combinées proposées, en commençant par la PO.

2.2.2 Application de la Pyramide Orientable en biométrie faciale et limites

La Pyramide orientable est une méthode d'analyse multi-échelle multi-orientation qui a été initialement introduite dans [Simoncelli 1995]. Le principe de la PO ainsi que les contraintes de conception de son filtre sont présentés en détails dans § 2.2.1c)2.1.1c)i). Ses propriétés d'invariance à l'échelle et à la rotation constituent un point fort motivant pour son application dans plusieurs travaux d'analyse de la texture [Briand 2014], en particulier en biométrie faciale [Uçar 2014] [El Aroussi 2011]. Un exemple de décomposition d'une image de visage par une PO à 3 échelles et 4 orientations est fourni dans la **Figure 2. 31**.

Plusieurs expériences analogues à celles qui ont été réalisées pour l'analyse de la texture de l'iris (§ 2.1.2b)), ont été conduites pour explorer la meilleure manière de construire le vecteur caractéristique issu de la PO appliquée au visage. Une première expérience pour le calcul des caractéristiques consiste à prendre directement les intensités du filtrage de l'image entière à différentes orientations et échelles. Le vecteur caractéristique a alors une taille de l'ordre de $N \times (n+1)$ où N est le nombre de pixels de l'image de visage (dans notre cas 64^2) et n le nombre d'orientations de la PO. Une autre expérience divise d'abord l'image initiale de taille (64×64) en 64 blocs de taille (8×8), puis applique une PO sur chacun de ces blocs. Une première manière de procéder était de composer le vecteur caractéristique directement des intensités issues des filtrages par bloc. Une deuxième manière était de calculer des mesures statistiques telles que l'énergie, la variance ou l'entropie sur les 64 blocs filtrés pour aboutir à des vecteurs caractéristiques de taille égale au nombre de blocs, c'est à dire 64. Contrairement à l'analyse de la texture de l'iris qui a produit les meilleurs résultats avec une PO par bloc et un calcul d'énergie sur les blocs filtrés, les meilleurs résultats de l'analyse de la texture du visage par PO ont été obtenus par application de la PO sur l'image entière et en travaillant directement sur les intensités issues du filtrage (sans calcul de mesure statistique). Ceci peut s'expliquer par la différence de nature entre les deux textures traitées.



Figure 2. 31 Décomposition d'une image de visage par une PO à 3 échelles (s = 1..3) et à (n+1) = 4 orientations ($3\pi/4$, $\pi/2$, $\pi/4$ et 0).

D'autres tests ont été réalisés pour étudier les **paramètres de la PO**, à savoir, le nombre d'échelles et le nombre d'orientations. Comme dans le cas de l'analyse de la texture de l'iris, les paramètres optimums de la PO appliquée au visage sont de 3 échelles et 6 orientations. La **Figure 2. 32** montre les courbes DET (Detection Error Tradoff) relatives à l'application de la PO pour les 4 expériences IV² sur des images entières de taille (64×64). Les vecteurs caractéristiques composés de $o(10^4)$ valeurs d'intensité ($64^2\times(3+1)$) sont comparés par la distance L1. L'expérience 1 est conduite dans un environnement contrôlé avec une variation limitée d'expression. Les expériences 2 à 4 sont relatives à des environnements non contrôlés comportant respectivement des variations d'éclairage, de qualité et de session (cf. **Table 1. 5** Description des expériences de visages 2D (V.= variation, N.= nombre, img = image et comp. = comparaisons).**Table 1. 5**).



Figure 2. 32 Courbes DET de l'algorithme de la PO appliqué aux quatre expériences de la compagne IV².

Les résultats en EER issus de l'application de la PO sur les expériences de la campagne IV² sont donnés dans la **Table 2. 14**. L'algorithme de la PO est comparé aux cinq algorithmes ayant participé à la campagne d'évaluation IV² [Mellakh 2009-b]. Les deux premiers algorithmes IV² sont basés sur la PCA et présentent cependant des écarts entre les résultats. Ces écarts peuvent s'expliquer par des différences dans les phases de prétraitement, de normalisation, de comparaison ou de décision. Le troisième algorithme est basé sur une PCA modulaire appliquée localement sur les régions d'intérêt. Les deux derniers algorithmes font intervenir la LDA, le premier en application directe sur des images en niveaux de gris et le second sur une transformée de Gabor.

Participants	Exp1	Exp2	Exp3	Exp4
PCA1	6,7 (±0,8)	20,7 (±1,3)	20,1 (±1,6)	22,2 (±1,6)
PCA2	7,3 (±0,8)	21,6 (±1,4)	13,6 (±1,4)	16,3 (±1,4)
mod PCA	5,3 (±0,7)	20,7 (±1,4)	19,5 (±1,6)	20,5 (±1,5)
LDA1	3,7 (±0,6)	22,5 (±1,4)	21,7 (±1,7)	19,7 (±1,5)
LDA2/Gabor	4,2 (±0,6)	12,0 (±1,1)	8,3 (±1,1)	11,3 (±1,2)
PO	11,0 (±1,0)	23,0 (±1,4)	20,1 (±1,6)	23,9 (±1,7)

Table 2. 14 Comparaison entre l'algorithme de la PO et ceux de la compagne d'évaluation IV².

D'après la **Table 2. 14**, il est clair que, comparés aux autres méthodes, les résultats obtenus par l'algorithme de la PO sont nettement moins bons, en particulier pour l'expérience 1. En fait, les mauvais résultats de l'application de la PO pour la caractérisation des visages étaient attendus car la dimension excessivement grande de l'espace des caractéristiques de la PO et la redondance présente dans cet espace dégradent forcément les résultats de la reconnaissance. Pour dépasser ces limites, nous avons prévu de coupler la PO à des méthodes de réduction de l'espace caractéristique qui seront présentées dans le paragraphe suivant.

2.2.3 Méthodes de réduction de l'espace caractéristique

Les méthodes hybrides proposées reposent sur la combinaison de l'algorithme de la PO à des méthodes de réduction d'espace caractéristique. D'abord, nous explorons deux des méthodes projectives des plus connues en biométrie faciale à savoir la PCA et la LDA. Nous utilisons ensuite, à une méthode adaptative : Adaboost. Dans ce qui suit, seront présentés les concepts théoriques ainsi que les résultats de l'application des méthodes de réduction intervenant dans les méthodes hybrides proposées.

a) Concepts théoriques des méthodes de réduction

Cette partie est consacrée à l'explication des principes théoriques des méthodes de réduction d'espace caractéristique expérimentées : la PCA, la LDA et Adaboost.

i) PCA

L'Analyse en Composantes Principales (Principal Component Analysis – PCA), introduite par [Turk 1991] a pour but d'obtenir des vecteurs caractéristiques compacts et représentatifs d'une image donnée. Elle est basée sur la décomposition de l'image selon des directions de variations principales autour d'une image moyenne, correspondant aux vecteurs propres de plus grande énergie.

Les images faciales qui sont d'abord sous forme de matrices de taille $(p \times q)$ pixels, sont alignées et représentées par des vecteurs I_i de taille n $(n=p^*q)$ dans un espace $X = [x_1, x_2, ..., x_n]$ de grande dimension. Les vecteurs I_i formant les colonnes de la matrice d'entrée x peuvent être exprimés comme une combinaison linéaire dans une base orthogonale Φ de dimension m réduite (m << n):

$$x = \sum_{i=1}^{n} a_i x_i \approx \sum_{i=1}^{m} \alpha_i \phi_i$$
(2. 14)

 $x = [I_1, I_2, ..., I_L]$ est une matrice (*n*×*N*) de visages, *N* étant le nombre de visages.

L'espace de projection défini par la base orthogonale Φ est construit en résolvant l'équation (2.15):

$$C\phi = \phi\Lambda \tag{2.15}$$

où C la matrice de covariance de l'entrée x et \bar{x} le visage moyen, sont donnés par les équations (2.16) et (2.17).

$$C = \sum (x - \bar{x})(x - \bar{x})^T \qquad (2.16) \qquad \bar{x}(j) = \frac{1}{N} \sum_{i=1}^n I_i(j) \qquad (2.17)$$

Cette opération donne les vecteurs propres de taille n et les valeurs propres de C qui sont ensuite réordonnés selon l'importance de la direction de variation. Les vecteurs propres obtenus sont appelés *Eigenfaces* et sont utilisés en tant qu'espace de projection pour les images à traiter. La **Figure 2. 33** montre l'exemple des cinq premiers *Eigenfaces* de l'ensemble d'apprentissage IV².



Figure 2. 33 Les cinq premiers Eigenfaces calculés à partir de l'ensemble d'apprentissage IV².

ii) LDA

Le but de l'analyse linéaire discriminante (Linear Discriminent Analysis - LDA) est de rechercher, dans l'espace des données, les axes qui permettent de discriminer au mieux les différentes classes (plutôt que les axes qui décrivent au mieux les données comme pour la PCA) [Fisher 1936]. En d'autres termes, la projection des données sur ces axes vise à minimiser les variations à l'intérieur des classes, tout en maximisant les variations entre les classes. L'apprentissage de la LDA est effectué par l'analyse des matrices de dispersion inter et intra classes S_w et S_b (désignées par *within-class* et *between-class scatter matrices* en anglais).

Pour un problème à k classes (nombre de personnes dans la base de données), S_w et S_b sont définies par les **équations (2.18)** et **(2.19)**.

$$S_w = \sum_{j=1}^c \sum_{i=1}^{N_j} (x_i^j - \mu_j) (x_i^j - \mu_j)^T$$
(2.18)

avec x_i^j le i^{ème} échantillon de la classe j, μ_j la moyenne de la classe j, c le nombre de classes et N_j le nombre d'échantillons de la classe j.

$$S_b = \sum_{j=1}^{c} (\mu_j - \mu)(\mu_j - \mu)^{\mathrm{T}}$$
(2. 19)

avec μ la moyenne de tous les échantillons.

Le critère de Fisher est ensuite utilisé pour mesurer le pouvoir discriminant du sous espace Φ , maximisant le rapport du déterminant de la matrice de dispersion interclasse des échantillons projetés par le déterminant de la matrice de dispersion intraclasse selon l'équation (2. 20).

$$\Phi_{\rm opt} = \arg \max_{\Phi} \frac{|\Phi^{\rm T} S_b \Phi|}{|\Phi^{\rm T} S_w \Phi|}$$
(2. 20)

La matrice de projection optimale Φ_{opt} , qui maximise ce ratio, peut être obtenue en résolvant le problème des valeurs propres modélisé par l'équation (2. 21) :

$$S_b \Phi = S_w \Phi \Lambda_\Phi \tag{2.21}$$

où Λ_{Φ} est la matrice diagonale des valeurs propres des axes de l'espace Φ . Sous l'hypothèse que la matrice S_w est inversible (singulière), les colonnes de la matrice Φ sont constituées des vecteurs propres de la matrice $S_w^{-1} S_b$. Cette condition de singularité pose un problème pour les applications de reconnaissance de visage du moment que les données sont le plus souvent sous-représentées, au sens où la taille n des vecteurs images de la base d'apprentissage est très supérieure à leur nombre N (n = taille des vecteurs >> N = nombre des images). La matrice S_w est alors non inversible et on ne peut pas déterminer directement Φ .

Pour remédier à ce problème, Belhumeur et Kriegman d'une part et Swets et Weng par ailleurs côté ont proposé à la même année et de manière indépendante d'effectuer une PCA en amont de la LDA [Belhumeur 1996] [Swets 1996]. On commence par rechercher les axes principaux des données par la PCA. Puis on projette les données sur ces axes, ce qui permet de réduire la dimensionnalité du problème, de manière à ce que la nouvelle matrice de dispersion intra-classe soit inversible, et ceci en conservant au maximum la forme de la distribution initiale des données. On peut donc appliquer la LDA sur les données réduites. Les colonnes de la matrices de projection Φ , dont des exemples sont fournis en **Figure 2. 34**, sont appelés *Fisherfaces*.



Figure 2. 34 Les cinq premiers Fisherfaces calculés à partir de l'ensemble d'apprentissage IV².

iii) Adaboost

Le boosting est un principe qui consiste à construire un classifieur « fort » à partir d'une combinaison pondérée de classifieurs « faibles », c'est-à-dire donnant en moyenne une réponse meilleure qu'un tirage aléatoire. [Viola 2004] adapte ce principe en assimilant une caractéristique à un classifieur faible. L'apprentissage du classifieur faible consiste alors à trouver la valeur seuil de la caractéristique qui permet de mieux séparer les exemples (échantillons d'apprentissage) positifs correspondant à des images provenant d'une même personne, des exemples négatifs correspondant à des images provenant de personnes différentes. Le classifieur se réduit alors à un couple (caractéristique, seuil). L'algorithme est de type itératif adaptatif, à nombre d'itérations déterminé, où à chaque itération, l'algorithme met à jour le poids a_t (t étant le numéro de l'itération, avec $t \le p$ = nombre de pixels dans l'image) d'une caractéristique de manière dynamique selon les erreurs de classification de l'apprentissage issues de l'itération précédente. La caractéristique ayant le poids le plus fort sera ajoutée à la liste des caractéristiques sélectionnées aux itérations précédentes, et le tout va contribuer à la construction du classifieur fort final. La Figure 2. 35 illustre le principe général d'Adaboost. $(x_i)_{i=1..n} = (I_{1,i}, I_{2,i})_{i=1..n}$ sont les échantillons d'apprentissage formés par des paires d'images à comparer et n le nombre de comparaisons. Les étiquettes des échantillons d'apprentissage sont connues à l'avance : positif si $I_{1,i}$ et $I_{2,i}$ proviennent de la même personne et négatif sinon.



Figure 2. 35 Principe d'Adaboost.

La **Figure 2. 36** détaille le fonctionnement de la variante d'adaboost utilisée dans nos tests. L'entrée de l'algorithme est constituée, d'une part, des étiquettes y_i des échantillons $(x_i)_{i=1...n} = (I_{1,i}, I_{2,i})_{i=1...n}$, à savoir : 1 si $I_{1,i}$ et $I_{2,i}$ sont des visages de la même personne et -1 sinon ; d'autre part, de la matrice des filtres $(\Phi_j)_{j=1..p}$ constituant les caractéristiques des classifieurs faibles. Dans notre cas, les lignes de la matrice des filtres sont constituées chacune de la différence absolue entre $I_{1,i}$ et $I_{2,i}$ $(\Phi_j = |I_{1,i} - I_{2,i}|)$.

La première étape de l'algorithme consiste à calculer le meilleur classifieur faible h_j pour chaque filtre Φ_j . Chaque classifieur est caractérisé par son seuil. Ensuite, nous estimons les classes/étiquettes $(\hat{y}_i)_{i=1..n}$ des échantillons $(x_i)_{i=1..n}$ pour chacun de ces classifieurs faibles h_j . Après, nous calculons l'erreur e_j entre les classes réelles y_i et les classes estimées \hat{y}_l en tenant compte des poids $D_{i,l}$ (à l'itération t), initialisés à 1/n, des échantillons x_i . Le choix du meilleur classifieur faible $h_{min_{-t}}$ à l'itération t coïncidera avec celui qui donne l'erreur minimale $min(e_j)_{j=1..p}$. Cette erreur minimale sera, par la suite, utilisée dans la détermination du poids a_l du classifieur $h_{min_{-t}}$ qui traduit sa contribution dans le classifieur fort final. La dernière étape consiste à mettre à jour les poids des échantillons $D_{i,t}$ pour avoir les poids de l'itération suivante $D_{i,t+1}$ de telle sorte à favoriser les échantillons malclassifiés en augmentant leurs poids. En sortie de cette itération, nous obtenons : le meilleur classifieur faible $h_{min_{-t}}$, son poids de mélange a_t et les poids des échantillons $(D_{i,t+1})_{i=1..n}$ de l'itération suivante, s'il y a lieu.

Ces opérations seront répétées et les classifieurs faibles seront accumulés jusqu'à atteindre un nombre d'itérations T. Ce nombre d'itérations peut être fixé dès le départ ou bien dépendre d'un certain critère d'arrêt tel que des poids d'échantillons trop faibles, une erreur nulle du classifieur fort ou la fin de l'apprentissage de tous les classifieurs faibles.

Plus de détails sur l'algorithme et les équations d'Adaboost sont disponibles dans [Freund 1995] et [Viola 2004]. Cette technique a été souvent utilisée dans les applications récentes de biométrie faciale, par exemple dans [Yoon 2010] [Lee 2013] [Jiang 2015].


Figure 2. 36 Schéma fonctionnel de l'algorithme d'Adaboost

b) Résultats des méthodes de réduction

Suite à la présentation des concepts théoriques des méthodes expérimentées dans ce travail, nous présentons, maintenant, les résultats des méthodes simples expérimentées et appliquées séparément au problème de reconnaissance de visage. Puis, nous en déduisons les limites qui vont appuyer notre recherche pour la combinaison de méthodes.

i) PCA

Un premier paramètre à choisir après la détermination de l'espace de projection par la PCA est le choix de sa dimension optimale. Certains proposent de fixer de manière empirique la variance de cet espace entre 92% et 98% [Penev 2000] [Draper 2003]. D'autres réduisent la dimension de l'espace de projection en rapport avec la quantité d'information utile dans la nouvelle représentation. Par exemple, Mellakh et al. proposent une méthode qui se base sur l'étude du rapport entre l'énergie de l'information et l'énergie de l'information résiduelle [Mellakh 2006]. Nous avons choisi d'utiliser cette dernière méthode car elle offre un certain compromis entre la dimension retenue de l'espace de réduction à partir de ce rapport d'énergies et les performances de la méthode.

Notons E_{tot} , définie par l'équation (2. 22), l'énergie de l'espace de projection construit avec la totalité des n composantes principales et λ_i la $i^{\text{ème}}$ valeur propre, $i \in [1, n]$, n étant le nombre de pixels dans l'image.

$$E_{tot} = \sum_{i=1}^{n} \lambda_i^2$$
 (2. 22)

L'apport énergétique A_j de la $j^{\text{ème}}$ composante principale est donné par l'équation (2. 23).

$$A_j = \frac{\lambda_j^2}{E_{tot}}$$
(2.23)

Si on considère le rapport R_j entre l'énergie résiduelle à partir de la $j^{\text{ème}}$ composante principale et l'énergie totale E_{tot} donné par l'équation (2. 24) :

$$R_j = \frac{\sum_{i=j}^n \lambda_i^2}{E_{tot}}$$
(2. 24)

Le taux d'énergie résiduelle de la $j^{\text{ème}}$ composante principale est, alors, défini par le rapport T_j entre l'apport énergétique A_j et R_j et modélisé par l'équation (2. 25).

$$T_j = A_j / R_j = \frac{\lambda_j^2}{\sum_{i=j}^n \lambda_i^2}$$
 (2. 25)

Nous avons étudié l'impact de ce taux d'énergie résiduelle T_j en fonction de la variance de l'espace de projection sur l'ensemble d'apprentissage IV² (cf. **Figure 2. 37**). La variance Var_j de l'espace de projection au niveau de la $j^{\text{ème}}$ composante principale, aussi appelée inertie cumulée, est donnée par l'équation (2. 26) :

$$Var_j = \frac{\sum_{i=1}^j \lambda_i^2}{E_{tot}}$$
(2. 26)

L'évolution du taux d'énergie résiduelle T_j représentée en courbe logarithmique sur la **Figure 2. 37** conduit à une fréquence de coupure égale à 87% (trait rouge sur les figures ci-dessous) qui correspond au minimum de la courbe.



Figure 2. 37 Evolution du taux d'énergie résiduelle en fonction de la variance de l'espace de réduction de l'ensemble d'apprentissage IV^2 .



Figure 2. 38 Evolution du EER en fonction de la variance de l'espace de projection pour les quatre expériences de la base IV^2 .

Les résultats des évaluations sur les quatre expériences de la base IV² (cf. **Figure 2. 38** avec Exp1 : variation d'expressions, Exp2 : variation d'éclairage, Exp3 : variation de qualité et Exp4 : variation de session) montrent que l'utilisation des premiers axes principaux seuls (dont la variance est inférieure à 40%) n'est pas suffisante. Ceci est dû au fait que l'information la plus représentative ne se situe pas uniquement au niveau des premiers axes. D'un autre côté, les axes dont l'apport de variance est faible (donnant une variance supérieure à 90%) n'ont pas d'effet notoire sur les performances de la méthode. Ces résultats confirment la dimension optimale de l'espace de projection fixée à 87% selon la fréquence de coupure du taux d'énergie résiduelle.

Un autre paramètre qui a été étudié est le choix de la métrique de comparaison entre les projections des images sur l'espace réduit. Nous avons testé les métriques les plus utilisées dans la littérature [Moon 1998] [Draper 2003], à savoir : la distance L1, la distance Mahalanobis et la distance cosinus données par les **équations (2. 27)** à **(2. 29)**, avec $(x_i)_{i=1..r}$ et $(y_i)_{i=1..r}$ les coordonnées des projections des images dans l'espace réduit (rdimension de l'espace réduit).

$$D_{L1} = \sum_{i=1}^{r} |\mathbf{x}_i - \mathbf{y}_i|$$
 (2. 27)

$$D_{Cosinus} = -\frac{\sum_{i=1}^{r} x_i y_i}{\sqrt{\sum_{i=1}^{r} x_i^2 \sum_{i=1}^{r} y_i^2}}$$
(2.28)

$$D_{Mahalanobis} = -\sum_{i=1}^{r} \frac{1}{\sqrt{\lambda_i}} x_i y_i$$
(2. 29)

La **Table 2. 15** donne les résultats de la comparaison de ces métriques pour les quatre expériences IV². La distance cosinus donne les meilleurs résultats pour les expériences 1 et 2 (variation d'expression et variation d'éclairage). La distance Mahalanobis la suit de près puis la dépasse au niveau de l'expérience 4 (multisession). Tandis que la distance L1 ne fournit de bons résultats qu'au niveau de l'expérience 3 (éclairage). Au vu de ces résultats, nous garderons par la suite la distance cosinus qui paraît la plus stable pour l'algorithme de la PCA au fil des expériences.

 Table 2. 15 Performance en EER de la PCA en fonction de la métrique pour les expériences IV².

Métrique	Exp1	Exp2	Exp3	Exp4
L1	11,6 (±1,0)	24,8 (±1,4)	19,8 (±1,6)	24,5 (±1,7)
Cosinus	6,1 (±0,8)	21,9 (±1,4)	24,8 (±1,7)	21,9 (±1,6)
Mahalanobis	8,4 (±0,9)	22,5 (±1,4)	23,5 (±1,7)	19,4 (±1,5)

ii) LDA

Nous avons d'abord testé la LDA en utilisant la base d'apprentissage IV^2 pour se conformer au protocole IV^2 et pouvoir comparer les résultats des tests réalisés strictement dans les mêmes conditions que la PCA. Toutes les expériences de ce paragraphe utilisent la distance cosinus.

Les résultats présentés dans la **Table 2. 16** confirment le pouvoir discriminant de la LDA par rapport à la PCA comme cela a été montré dans [Belhumeur 1997] et ceci pour toutes les expériences sauf pour l'expérience 2 contenant des variabilités d'éclairage. Afin d'avoir plus d'explications quant à ce dernier résultat, nous nous sommes référés aux travaux de Martinez et Kak comparant la PCA et la LDA [Martinez 2001]. Ces travaux soulignent la dépendance étroite de la LDA au nombre et au caractère

représentatif des images d'apprentissage ; ce qui peut, dans certains cas, donner des résultats inattendus en faveur de la PCA.

Table 2. 16 Comparaison en EER de la performance des algorithmes de la PCA et de la LDA pourles expériences IV^2 avec l'ensemble d'apprentissage IV^2 .

Algorithme	Exp1	Exp2	Exp3	Exp4
PCA	6,1 (±0,8)	21,9 (±1,4)	24,8 (±1,7)	21,9 (±1,6)
LDA	5,2 (±0,7)	25,1 (±1,4)	18,5 (±1,5)	20,9 (±1,6)

iii) Adaboost

En testant la méthode d'Adaboost avec le protocole et l'ensemble d'apprentissage IV², nous avons noté que plus le nombre d'itérations du modèle d'Adaboost était grand, plus les performances augmentaient (cf. **Figure 2. 39**).



Figure 2. 39 Performance d'Adaboost en termes d'EER en fonction du nombre d'itérations du modèle d'apprentissage, pour l'expérience 1 d'IV².

A chaque itération, le modèle apprend un nombre de 64^2 caractéristiques. L'apprentissage devient aussitôt très coûteux en termes de temps de calcul dès qu'on dépasse les 400 itérations. Par exemple, le programme écrit sous Matlab7 et exécutant l'apprentissage par Adaboost des 156 images initiales de taille (64×64) nécessite $67h12min^1$.

D'après la **Table 2. 17**, Adaboost atteint un très bon EER = 3,7% dans l'expérience 1. Par contre, les performances chutent d'une manière drastique dans le reste des expériences incluant des variabilités d'éclairage, de qualité et de sessions. Ceci revient au fait qu'Adaboost est très dépendant de l'ensemble d'apprentissage qui, dans notre cas, a été capturé dans les mêmes conditions que celles de l'expérience 1.

 Table 2. 17 Résultats d'Adaboost sur les expériences IV².

Algorithme	Exp1	Exp2	Exp3	Exp4
Adaboost	3,7 (±0,6)	22,7 (±1,4)	44,3 (±2,0)	28,6 (±1,8)

¹ Caractéristiques techniques du PC : processeur Intel (R) Core (TM) I5, horloge de 2,67 GHz, RAM de 4Go, et un système d'exploitation Windows7 64 bits.

c) Limite des méthodes de réduction

Les résultats des méthodes de réduction d'espace sont montrés dans la Table 2. 18. Si on considère l'algorithme de la PCA3 que nous avons proposé au § 2.2.3-b)-i), en comparaison avec les deux autres PCA proposés dans la campagne IV² (PCA1 et PCA2), il est clair qu'il les dépasse dans la première expérience. Par contre, il est moins bon sur le reste des expériences. Cependant, il faut souligner que, contrairement à la PCA1, l'ensemble d'images d'apprentissage avec lequel la base de projection de la PCA2 a été construite, n'est pas celui indiqué par le protocole IV². En effet, il s'agissait d'un ensemble de 300 images issues de la base BANCA [Bailly-Bailliére 2003] (30 sujets, 10 images par sujet) et comprenant 3 qualités d'images différentes. Tandis que la méthode proposée a strictement suivi le protocole en se limitant à 52 images de 52 sujets capturés dans des conditions assez bonnes, ce qui n'est pas le cas des images de tests pour lesquelles plusieurs variations sont présentes. Le nombre relativement réduit des images d'apprentissage en plus de la différence entre les conditions d'acquisition des images d'apprentissage d'une part et celles des images de tests d'autre part, constituent des défis supplémentaires, ce qui explique les taux forts inférieurs observés pour les expériences 2, 3 et 4.

D'un autre côté, l'algorithme LDA2 que nous avons proposé au § 2.2.3-b)-ii) présente globalement des résultats en deçà de celui de la LDA1 de la campagne IV² bien qu'ayant été testés tous les deux avec la même base d'apprentissage d'IV².

Par ailleurs, Adaboost réalise la meilleure performance dans la première expérience. Toutefois, il se trouve en queue de liste dès qu'une variabilité non présente dans son ensemble d'apprentissage est introduite (expériences 2 à 4).

En outre, il est important de noter dans cette analyse, que les différences de résultats entre des méthodes identiques d'analyse de texture (LDA2 – LDA1 et PCA1 – PCA3) testées avec le même ensemble d'apprentissage peuvent très bien s'expliquer par des différences dans le prétraitement appliqué aux images en entrée. En effet, le protocole IV² fourni aux participants les images initiales avec la position des yeux comme métadonnées. Si l'on se réfère à la **Figure 2. 29** qui illustre les étapes de prétraitement d'un visage, il est clair que plusieurs différences peuvent avoir lieu entre les participants dans l'étape de la remise à l'échelle du visage (le choix de la distance intraoculaire, cf. **Figure 2. 29-d**) et dans la délimitation de la région d'intérêt – ROI (le choix du nombre de pixels de part et d'autre des yeux, cf **Figure 2. 29-e**). Ceci pourrait induire une différence dans la taille des images normalisées sur lesquels sont appliquées les méthodes d'analyse de la texture. Plus encore, même si les images obtenues par des candidats différents sont de même taille, il est fort probable que l'information contenue dans ces images ne soit pas la même à cause de la différence éventuelle entre les échelles et les ROI.

Description	Algorithmes	Exp1	Exp2	Exp3	Exp4
Campagne IV ²	PCA1	6,7 (±0,8)	20,7 (±1,3)	20,1 (±1,6)	22,2 (±1,6)
	PCA2	7,3 (±0,8)	21,6 (±1,4)	13,6 (±1,4)	16,3 (±1,4)
	mod PCA	5,3 (±0,7)	20,7 (±1,4)	19,5 (±1,6)	20,5 (±1,5)
	LDA1	3,7 (±0,6)	22,5 (±1,4)	21,7 (±1,7)	19,7 (±1,5)
	LDA/Gabor	4,2 (±0,6)	12,0 (±1,1)	8,3 (±1,1)	11,3 (±1,2)
Méthodes de	PCA3	6,1 (±0,8)	21,9 (±1,4)	24,8 (±1,7)	21,9 (±1,6)
réduction d'espace	LDA2	5,2 (±0,7)	25,1 (±1,4)	18,5 (±1,5)	20,9 (±1,6)
caractéristique	Adaboost	3,7 (±0,6)	22,7 (±1,4)	44,3 (±2,0)	28,6 (±1,8)

 Table 2. 18 Comparaison des méthodes de réduction d'espace avec les algorithmes de la campagne IV².

Nous pouvons conclure depuis les résultats de la **Table 2. 18** que l'utilisation des méthodes de réduction d'espace caractéristique s'avère assez robuste dans des environnements contrôlés, mais présente plusieurs limites lorsque l'ensemble d'apprentissage n'est pas de même nature que celui de test ou bien lorsque des défis tels que la variation d'éclairage, de qualité ou d'expression sont introduits. D'après les expériences présentées dans cette section, les méthodes projectives (PCA et LDA) souffrent d'une chute de performances dans les environnements non contrôlés. Adaboost, qui est une méthode à apprentissage adaptatif, est très dépendant de l'ensemble d'apprentissage et sensible aux valeurs aberrantes.

Les limites des méthodes de réduction d'espace caractéristique utilisées de manière indépendante, d'un côté, et les résultats non satisfaisants de la Pyramide Orientable qui souffre d'une dimension excessivement grande de son espace des caractéristiques ainsi que de la redondance présente dans cet espace, de l'autre côté, nous incite à envisager la solution de la combinaison d'algorithmes en vue de tirer profit de l'aspect complémentaire qui existe entre eux.

2.2.4 Proposition de méthodes combinées

Du moment que chaque algorithme permet de décrire des aspects différents du visage, et que des combinaisons pertinentes peuvent aboutir à des descriptions très précises, nous avons développé une série de méthodes de reconnaissance du visage associant la Pyramide Orientable (PO) aux trois méthodes de réduction d'espace caractéristique étudiées dans le paragraphe précédent : d'abord les deux méthodes projectives qui sont la PCA et la LDA [Khiari-Hili 2011] [Khiari-Hili 2012-a] ; ensuite, la méthode non linéaire de sélection et de classification : Adaboost [Khiari-Hili 2012-b]. Les différentes combinaisons proposées ont été réalisées suite à l'étude des différents niveaux de fusion d'algorithmes présentés dans le chapitre précédent (cf. § 1.3.3), au bout de laquelle nous nous sommes intéressés particulièrment à la fusion des caractéristiques au niveau de la PCA/PO et à la fusion des scores au niveau de la LDA/PO et Adaboost/PO. Les trois méthodes combinées proposées seront présentées dans ce qui suit à travers l'étude de l'influence de leurs paramètres. Certains de ces

paramètres sont relatifs aux méthodes individuelles qui les composent (par exemple le nombre d'échelles et d'orientations de la PO, la configuration des bandes de la PO, ...), d'autres sont communs aux méthodes d'analyse de la texture (comme le découpage de l'image) ou au processus de la reconnaissance (comme la distance) et d'autres encore à la démarche de la fusion (comme la pondération des scores par bande).

Tous les tests décrits dans cette section ont été strictement conduits dans les conditions décrites par le protocole IV^2 en utilisant exclusivement l'ensemble d'apprentissage IV^2 constitué de 52 personnes. Les images de visages normalisées sont de taille (64×64).

a) Couplage de la PO à des méthodes projectives

Les mauvais résultats obtenus par la PO sont dus à la dimension excessivement grande de l'espace des caractéristiques des images filtrées, ainsi qu'à la redondance présente dans cet espace. La PCA et la LDA sont des techniques à apprentissage, simples mais efficaces, permettant de réduire la dimension de l'espace des caractéristiques de manière à ne garder que les plus pertinentes sans pour autant perdre l'information utile.

i) Pyramide Orientable et PCA : PCA/PO [Khiari-Hili 2012-a]

La **Figure 2. 40** montre les étapes de construction du vecteur caractéristique par l'algorithme de la PCA/PO à travers l'exemple d'une PO à 3 échelles et 4 orientations. Les caractéristiques issues des bandes en sortie de la PO sont fusionnées pour former un vecteur de grande taille (de l'ordre de 10^4 dans notre cas). Ce vecteur sera par la suite projeté sur l'espace des *Eigenfaces* de la PCA, qui a été déterminé lors de l'apprentissage, pour avoir un vecteur caractéristique plus représentatif et très compact (de taille 52 dans notre cas = taille de l'espace des *Eigenfaces* = nombre des images de la base d'apprentissage).

La méthode PCA/PO comprend un certain nombre de paramètres dont l'influence sera présentée dans ce qui suit.

Influence du découpage de l'image entrée : Un premier ensemble de tests a été réalisé pour voir quelle était la meilleure entrée à introduire à la PO : une image entière ou des sous-blocs de l'image dont les intensités du filtrage seraient concaténées afin de former le vecteur caractéristique. Les résultats de la **Table 2. 19** liés à la PCA/PO montrent qu'il n'y a pas d'amélioration significative apportée par le partitionnement de l'image initiale en blocs ; c'est pourquoi nous avons gardé entières les images d'entrée.

Zone d'application de la PO	Exp1	Exp2	Exp3	Exp4
Image entière	5,3 (±0,7)	20,4 (±1,3)	18,3 (±1,5)	20,0 (±1,6)
Sous-blocs	5,8 (±0,8)	19,8 (±1,3)	18,8 (±1,5)	20,2 (±1,6)

Table 2. 19 Influence du découpage de l'image d'entrée sur l'algorithme PCA/PO.



Figure 2. 40 Génération du vecteur caractéristique final d'une image initiale par l'algorithme de la PCA/PO (à droite) et les 4 configurations testées des bandes passe-bas d'une PO à 3 échelles et 6 orientations (à gauche). Les bandes pleines sont gardées dans toutes les configurations.

Impact de l'inertie cumulée : Il s'agit d'un paramètre qui décide du nombre d'axes de projections (*Eigenfaces*) à garder. Aussi appelé variance de l'espace de projection au niveau de la $j^{\text{ème}}$ composante principale, l'inertie cumulée des j premiers *Eigenfaces* a été défini précédemment dans § 2.2.3b) par l'équation (2. 26).

Les résultats issus de la variation de ce paramètre dans l'algorithme combiné PCA/PO sont présentés dans la **Table 2. 20** et sont relatifs à l'expérience 1 du protocole IV^2 . Ces résultats montrent qu'il est mieux de garder la totalité des axes de projection de l'espace de réduction construit dans la phase d'apprentissage. Ceci revient au fait que l'entrainement de la PCA avec des images issues de 52 personnes (1 image/personne) implique un espace de réduction comprenant un nombre total de 52 axes de projections ; ce qui veut dire que la taille des vecteurs caractéristiques initialement de l'ordre de 10^4 à la sortie de la PCA en aval de la PO. Réduire davantage la taille de l'espace de projection conduirait indubitablement à la perte d'information utile.

Inertie cumulée	Nombre d'Eigenfaces	%EER	%FRR (FAR = 0,001)
100%	52	5,28 (±0,7)	23,6
95%	32	5,39 (±0,7)	26,0
90%	23	5,56 (±0,8)	32,8
85%	17	5,79 (±0,8)	31,0
80%	12	5,72 (±0,8)	41,0

Table 2. 20 Influence de l'inertie cumulée sur l'algorithme de la PCA/PO (Exp1).

Variation du nombre d'échelles et du nombre d'orientation de la PO: Les expériences menées confirment que les performances de la reconnaissance sont d'autant meilleures que les nombres d'orientations et d'échelles sont grands. La Figure 2. 41-a montre les courbes DET (en représentation logarithmique) issues de la variation du nombre d'échelles de la PO (1, 2 et 3 échelles), le nombre d'orientation étant fixé à 1, pour l'algorithme PCA/PO. D'après cette figure, l'échelle 3 (maximale pour des images 64×64) conduit à la meilleure performance. De même, la Figure 2. 41-b illustrant la variation du nombre d'orientations de la PO (2, 4 et 6), montre que le nombre de 6 orientations conduit aux meilleurs résultats pour un nombre d'échelles égal à 3.



Figure 2. 41 Impact de la variation (a) du nombre d'orientations – nombre d'échelles fixé à 3 – et (b) du nombre d'échelles – nombre d'orientations fixé à 6 – sur l'algorithme PCA/PO.

Fusion des caractéristiques et configuration des bandes de la PO : Un autre ensemble d'expériences a été conduit pour choisir, parmi les sous-bandes filtrées par la PO, celles à introduire comme entrée à l'algorithme de la PCA. Quatre configurations ont été étudiées. La configuration initiale (Config. 0) garde les images passe-haut (*HP*), passe-bandes (*BP_i*)_{*i*=0.*n*} et la dernière passe-bas résiduelle (*LP_n*). La première modification (Config. 1) apportée à la configuration initiale, ajoute la première image filtrée passe-bas (*LP₀*). La deuxième modification (Config. 2) était conduite par l'adjonction des images passe-bas intermédiaires (*LP_i*)_{*i*=1..*n*} à la configuration initiale. La troisième modification (Config. 3) a été réalisée en faisant participer toutes les images passe-bas de la PO (*LP_i*)_{*i*=0..*n*}. Les résultats de la **Table 2. 21** sont en faveur de la seconde configuration modifiée. Il est important de souligner que, dans la combinaison d'algorithmes PCA/PO, la fusion se fait au niveau des caractéristiques en concaténant les sorties des différents filtres (faisant partie de la configuration retenue) de la PO.

_						
	N° Config	Sous-bandes gardées	Exp1	Exp2	Exp3	Exp4
	0	HP, BP _i , LP _n , $(i = 0n)$	5,3 (±0,7)	20,4 (±1,3)	18,28 (±1,5)	18,3 (±1,6)
	1	HP, LP ₀ , BP _i , LP _n , $(i = 0n)$	5,5 (±0,7)	20,0 (±1,3)	18,49 (±1,5)	18,5 (±1,6)
	2	HP , BP _i , LP _i , (i = 1n)	5,0 (±0,7)	19,9 (±1,3)	17,37 (±1,5)	17,4 (±1,6)
	3	HP, BP _i , LP _i , $(i = 0n)$	5,3 (±0,7)	20,0 (±1,3)	17,60 (±1,5)	17,6 (±1,6)
-						

 Table 2. 21 Influence de configuration des sous-bandes de la PO sur l'algorithme PCA/PO.

Paramètres optimaux de la PCA/PO et comparaison aux méthodes individuelles

La méthode combinée PCA/PO ayant les paramètres optimaux consiste à filtrer des images entières par une PO à 3 échelles et 6 orientations. En plus de la bande passehaut HP et des bandes orientées ($BP_{i=0...n}$), les bandes passe-bas intermédiaires ($LP_{i=1...n}$) ont également été retenues à chaque niveau pour garder le maximum d'information utile (configuration n°2). L'algorithme procède par fusion au niveau des caractéristiques en concaténant les sous-bandes issues du filtrage de la PO et faisant partie de la configuration retenue. Un vecteur caractéristique de grande dimension (de l'ordre de 10^4) est obtenu et projeté sur la totalité des *Eigenfaces* construits dans la phase d'apprentissage de la PCA. Finalement, la comparaison se fait par la distance L1.

Avec ces paramètres, l'algorithme combiné de la PCA/PO réalise une amélioration considérable par rapport à la meilleure méthode individuelle qui le compose. Cette amélioration se traduit par une diminution de l'EER qui atteint les -20,5% dans le cas de l'expérience 4 (EER = 17,4% contre 21,9% pour PCA3, cf. **Table 2. 22**).

Algorithmes	Exp1	Exp2	Exp3	Exp4
PO	11,0 (±1,0)	23,0 (±1,4)	20,1 (±1,6)	23,9 (±1,7)
PCA3	6,1 (±0,8)	21,9 (±1,4)	24,8 (±1,7)	21,9 (±1,6)
PCA3/PO	5,0 (±0,7)	19,9 (±1,3)	17,4 (±1,5)	17,4 (±1,6)

Table 2. 22 Comparaison de la PCA/PO à ses méthodes individuelles.

ii) Pyramide Orientable et LDA : LDA/PO [Khiari-Hili 2011]

La deuxième méthode combinée proposée consiste à appliquer une LDA pour réduire l'espace caractéristique en sortie des filtres de la Pyramide Orientable. Etant donné que la LDA appartient à la même classe des méthodes projectives que la PCA, son couplage à la PO pour construire la méthode LDA/PO implique l'existence d'un certain nombre de paramètres en commun avec la méthode PCA/PO. Ces paramètres ont été définis dans le paragraphe précédent et concernent le découpage de l'image en entrée, l'impact de l'inertie cumulée, le nombre d'échelles et d'orientations de la PO ainsi que la métrique de comparaison. Dans le souci d'éviter les redondances dans la présentation de ces paramètres en commun et puisque les conclusions tirées à l'issue de la variation de ces paramètres pour la PCA/PO sont pour la plupart aussi valables pour la LDA/PO, nous allons nous limiter, dans cette partie, à donner les valeurs des paramètres optimaux correspondants. L'algorithme LDA/PO prend comme entrée des images entières non découpées en blocs. Ces images sont filtrées par une PO à 3 échelles et 6 orientations. Ensuite, une LDA est appliquée sur la sortie du filtrage par PO en gardant la totalité des Fisherfaces (c'est-à-dire 100% de l'inertie cumulée). Les vecteurs caractéristiques obtenus sont enfin comparés à l'aide de la distance cosinus. La suite de cette partie sera consacrée aux paramètres spécifiques à la méthode de la LDA/PO.

Influence de l'orientation de la bande de la PO : En optant pour l'utilisation de la LDA, une contrainte supplémentaire sur la taille des vecteurs propres issus des images filtrées s'est ajoutée car il n'était pas possible de calculer les matrices de dispersions avec des vecteurs d'aussi grande dimension (vecteur en sortie de la PO de l'ordre de $10^4 \times$ nombre d'images d'apprentissage = 3 images \times 52 sujets). C'est pour cette raison que l'étude de chaque bande orientée à différentes échelles a été envisagée au lieu d'analyser la totalité des sous-bandes à la fois.

La **Figure 2.** 42 montre les étapes de génération des vecteurs caractéristiques orientés (LDA/PO_B*i*) $_{i = 1 \ a \ 6}$ (des bandes n°*i* à la direction (*i*-1)π/6, $i = 1 \ a \ 6$) par l'algorithme de la LDA/PO à travers l'exemple d'une PO à 3 échelles et 6 orientations. Ceci se fait d'abord par une fusion au niveau des caractéristiques en sortie des différentes échelles d'une orientation donnée de la PO. Chaque vecteur caractéristique orienté possède alors une taille de l'ordre de 10³ (64²+32²+16²) et est ensuite introduit comme entrée de la LDA qui génère un vecteur caractéristique orienté de taille réduite (ici 52 = taille de l'espace des *Fisherfaces* = nombre de sujets de la base d'apprentissage). La comparaison se fait, ensuite, de manière indépendante pour chaque direction pour donner des scores propres à chacune de ces directions.

La **Table 2. 23** donne les taux d'EER correspondants aux bandes dans les directions $0, \pi/6, \pi/3, \pi/2, 2\pi/3$ et $5\pi/6$, et relatives aux quatre expériences de l'évaluation IV². Il est clair qu'il n'y a pas d'orientation qui soit meilleure que les autres pour tous les tests. Chacune d'entre elles a sa propre contribution. C'est pour cela qu'une stratégie de fusion des scores a été adoptée entre toutes les orientations afin d'exploiter cette complémentarité d'apports.



Figure 2. 42 Génération des vecteurs caractéristiques d'une image initiale par l'algorithme de la LDA/PO, avec une PO à 3 échelles et 6 orientations selon la configuration n°2 (bandes en vert).

Orientations	Exp1	Exp2	Exp3	Exp4
0	9,3 (± 1,0)	$25,6 (\pm 1,5)$	$20,1 (\pm 1,6)$	$25,6 (\pm 1,7)$
п/6	$11,4 (\pm 1,0)$	$26,4 (\pm 1,5)$	$19,4 (\pm 1,6)$	$25,9 (\pm 1,7)$
п/3	12,3 (± 1,1)	$23,5 (\pm 1,4)$	$17,6 (\pm 1,5)$	$24,0 (\pm 1,7)$
п/2	$10,2 (\pm 1,0)$	$21,5 (\pm 1,4)$	$16,3 (\pm 1,5)$	$23,2 (\pm 1,6)$
2п/3	$10,4 (\pm 1,0)$	$23,5 (\pm 1,4)$	$18,3 (\pm 1,5)$	22,9 (± 1,6)
5п/6	$10,8 (\pm 1,0)$	$24,7 (\pm 1,4)$	21,3 (± 1,6)	$25,6 (\pm 1,7)$

Table 2. 23 Influence de l'orientation de la bande dans l'application de la LDA sur une PO à 6orientations et 3 échelles dans les 4 expériences IV^2 .

Fusion des scores et configuration des bandes de la PO : Plusieurs stratégies de fusion ont été explorées dont la fusion des scores moyennant la somme, le produit et le max des scores obtenus pour chaque bande. Une amélioration qui atteint les 20,2% par rapport aux résultats de la meilleure bande ($\pi/2$), utilisée à elle seule, est observée après la fusion en diminuant l'EER initialement de 16,3% à 13% après la fusion des 6 bandes orientées dans l'expérience 3. Cependant, aucune généralisation n'est possible pour trancher de la meilleure stratégie de fusion car plusieurs paramètres entrent en jeu.

Par ailleurs, la **Figure 2. 42** montre qu'en plus des sous-bandes orientées, il existe d'autres sous-bandes dont une passe-haut et plusieurs passe-bas. La question s'est posée sur l'apport que pourrait avoir ces sous-bandes en tant qu'entrée à l'algorithme de la LDA. Les mêmes quatre configurations explorées dans l'algorithme PCA/PO ont été testées dans la LDA/PO et les résultats sont montrés dans la **Table 2. 24**. Par contre, à la différence de la PCA/PO qui opère par fusion au niveau des caractéristiques des sousbandes de la configuration retenue de la PO, la LDA/PO accomplit une fusion au niveau des scores en sortie des modules de comparaison relatifs à chaque bande (orientée, passebas et passe-haut) de la PO.

\mathbf{N}° onf	Bandes considérées	Exp1	Exp2	Exp3	Exp4
0	HP, BP _i , LP _n $(i = 0n)$	4,3 (±0,7)	17,3 (±1,3)	9,3 (±1,1)	14,8 (±1,4)
1	HP, LP ₀ , BP _i , LP _n $(i = 0n)$	4,2 (±0,7)	19,0 (±1,3)	10,7 (±1,2)	16,4 (±1,4)
2	HP , BP _i , LP _i (i =1n)	3,6 (±0,6)	17,5 (±1,3)	10,0 (±1,2)	14,5 (±1,4)
3	HP, BP _i , LP _i (i = 0n)	4,7 (±0,7)	19,7 (±1,3)	10,5 (±1,2)	16,6 (±1,4)

 Table 2. 24 Influence de configuration des sous-bandes de la PO sur l'algorithme LDA/PO.

Paramètres optimaux et comparaison aux méthodes individuelles

L'algorithme de la LDA/PO commence par filtrer des images entières par une PO à 3 échelles et 6 orientations. Les bandes traitées de la PO correspondent à la configuration n°2 comprenant : la bande passe-haut HP, les bandes orientées $(BP_{i=0...n})$ et les bandes passe-bas intermédiaires $(LP_{i=1...n})$. Ensuite, une fusion des caractéristiques est opérée en concaténant les sorties des différentes échelles pour chaque bande orientée et passebas de la PO. Un ensemble de 8 vecteurs caractéristiques par image, spécifiques à chaque bande (1 passe-haut, 6 orientées et 1 passe-bas), sont obtenus. La LDA est ensuite appliquée séparément sur chacun de ces vecteurs caractéristiques en gardant la totalité des *Fisherfaces* (100% d'inertie cumulée) donnant lieu à un vecteur de taille réduite (52). Dans la phase de test, le *matching* se fait en comparant de manière indépendante les vecteurs caractéristiques issus des 8 bandes via la distance Cosinus. 8 scores sont générés et fusionnés par la suite à travers la règle max pour construire le score final.

La **Table 2. 25** montre les résultats de l'algorithme LDA/PO appliqué avec ses paramètres optimums face aux méthodes individuelles qui la composent. L'intérêt de la combinaison d'algorithmes est notable car elle parvient à réduire l'EER de -45,9% par rapport à sa meilleure méthode individuelle dans le cas de l'expérience 3 en réalisant 10,0% contre 18,5% pour la LDA2.

Algorithmes	Exp1	Exp2	Exp3	Exp4
РО	11,0 (±1,0)	23,0 (±1,4)	20,1 (±1,6)	23,9 (±1,7)
LDA2	5,2 (±0,7)	25,1 (±1,4)	18,5 (±1,5)	20,9 (±1,6)
LDA2/PO	3,6 (±0,6)	17,5 (±1,3)	10,0 (±1,2)	14,5 (±1,4)

Table 2. 25 Comparaison de la LDA/PO à ses méthodes individuelles.

b) Couplage de la PO à une méthode adaptative : Adaboost/PO

Dans le but d'apporter une solution aux déficiences des méthodes de la PO et d'Adaboost quand elles sont appliquées séparément, nous avons opté pour l'utilisation d'Adaboost en tant que sélecteur adaptatif de caractéristiques et classifieur réduisant la taille de l'espace des caractéristiques de la PO et ce en ne gardant que les caractéristiques pertinentes [Khiari-Hili 2012-b].

Dans la suite, l'influence d'une liste non exhaustive de paramètres, liés à la stratégie d'application d'Adaboost sur les images filtrées par la PO, est d'abord présentée. Ensuite, une comparaison avec les algorithmes ayant participé à la campagne IV² 2007 est fournie.

i) Application d'Adaboost sur la Pyramide Orientable entière

Un premier lot d'expériences a été réalisé en appliquant Adaboost sur la totalité des caractéristiques issues du filtrage par PO et ce en sélectionnant celles qui sont les plus discriminantes. La **Figure 2. 43** illustre le processus de caractérisation et de classification d'un couple d'images candidates à comparer, à travers l'algorithme d'Adaboost appliqué sur la PO entière. La différence absolue des vecteurs caractéristiques VPO1 et VPO2 issus du filtrage par pyramide orientable, est introduite comme entrée au modèle d'Adaboost construit dans la phase d'apprentissage et constitué de t classifieurs faibles ($t \le p = \text{taille}(VPO1) = o(10^4)$) qui vont sélectionner les t caractéristiques les plus pertinentes. Le modèle Adaboost va alors générer la classe estimée \hat{Y} du couple d'images candidates à comparer. Si \hat{Y} est supérieure au seuil choisi, alors les deux images de visage correspondent à la même personne ; sinon les deux images proviennent de deux personnes différentes.

L'augmentation du nombre d'itérations du modèle d'Adaboost (c'est-à-dire du nombre de caractéristiques sélectionnées) améliore les taux de reconnaissance (diminue le taux d'EER) comme montré sur la **Figure 2. 44** avec une stabilité des performances observées au-delà des 300 itérations. Malheureusement, l'exécution des tests est coûteuse en termes de temps de calcul et les 400 itérations/caractéristique sont difficilement réalisables. A titre d'exemple, le programme écrit sous Matlab7 et exécutant l'apprentissage par Adaboost des caractéristiques filtrées par PO issues de 10 images initiales de taille (64×64) nécessite un temps de 46h30min².

D'après les résultats de la **Table 2. 26**, il est clair que la combinaison d'Adaboost et de la PO améliore considérablement les performances des méthodes individuelles (PO seule et Adaboost seule) avec une diminution du taux d'EER par rapport à la meilleure méthode individuelle pouvant atteindre les -29,8% dans le cas de l'expérience 3 en réalisant un EER de 14,1% contre 20,1% pour la PO. Les résultats sont acceptables comparés à ceux des autres algorithmes participant à la campagne IV² mais non encore suffisants. Dans le but d'améliorer cette méthode et en même temps, de contourner la restriction liée au temps de calcul, nous avons eu l'idée d'appliquer Adaboost séparément sur chaque bande de la PO.

² Caractéristiques techniques du PC : processeur Intel (R) Core (TM) I5, horloge de 2,67 GHz, RAM de 4Go, et un système d'exploitation Windows7 64 bits.



Figure 2. 43 Application de l'algorithme Adaboost/PO sur la PO entière (3 échelles, 6 orientations, Config.2) pour la caractérisation et la classification d'un couple d'images candidates à comparer.



Figure 2. 44 Performance de l'algorithme Adaboost/PO_entière en EER en fonction du nombre d'itérations dans le modèle Adaboost pour l'expérience 1 d'IV².

Algorithmes	Exp1	Exp2	Exp3	Exp4
РО	11,0 (±1,0)	23,0 (±1,4)	20,1 (±1,6)	23,9 (±1,7)
Adaboost	3,7 (±0,6)	22,7 (±1,4)	44,3 (±2,0)	28,6 (±1,8)
Ada/PO	5,0 (±0,7)	19,5 (±1,3)	14,1 (±1,4)	22,9 (±1,6)

Table 2. 26 Comparaison d'Adaboost/PO à ses méthodes individuelles.

ii) Application d'Adaboost sur chaque bande de la PO

Un ensemble de tests visant à déterminer l'existence d'une orientation privilégiée de la PO à injecter comme entrée à l'algorithme d'Adaboost a confirmé les résultats de l'algorithme combiné LDA/PO [Khiari-Hili 2012-a]. Ces résultats stipulent que les orientations de la PO étaient plutôt complémentaires et qu'aucune d'entre elles, utilisée seule, ne fournissait une information complète. Afin de profiter de cette complémentarité d'apport, nous avons opté pour une fusion des scores issus de l'application d'Adaboost sur chaque bande de la PO (dans notre cas, 8 bandes : 1 passe-haut VPO_H , 6 passe-bandes orientées (VPO_ B_i) i = 1.6 et 1 passe bas VPO_L sur la Figure 2.45) tout en mettant en œuvre plusieurs stratégies de fusion. La Figure 2. 45 montre les étapes de l'algorithme Adaboost/PO appliqué sur chaque sous-bande de la PO. Le vecteur caractéristique par bande d'une image candidate I_i (VPj_H, (VPOj_B_i) i = 1..6 et VPOj_L), en sortie de la PO, est construit de la même manière que dans l'algorithme LDA/PO montré en détails dans la Figure 2. 42. Les différences absolues entre les vecteurs caractéristiques par bande, issus des images candidates à comparer, sont calculées en parallèle. Chaque différence absolue est ensuite introduite comme entrée au modèle Adaboost correspondant à sa bande et ayant été calculé préalablement à la phase d'apprentissage. Selon la bande, le modèle Adaboost sélectionnera t_{H} , $(t_{Bi})_{i=1..6}$ ou t_{L} caractéristiques pertinentes et renverra en sortie la classe estimée \widehat{Y}_{H} , $(\widehat{Y}_{Bl})_{i=1,6}$ ou \widehat{Y}_{L} qui constituera un score par bande de la comparaison des images candidates I_1 et I_2 . Le score final \hat{Y} sera obtenu par la fusion des scores issus des différentes bandes de la PO.



Figure 2. 45 Application de l'algorithme Adaboost/PO sur chaque sous bande d'une PO à 3 échelles et 6 orientations (Config.2) pour la caractérisation et la classification d'un couple d'images candidates à comparer.

Tout d'abord, **le choix de la règle de fusion des scores** s'est posé. Plusieurs opérateurs usuels tels que la Somme, le Maximum, le Minimum et le Produit ont été testés. Les meilleurs résultats ont été atteints par la règle Somme. Ensuite, la fusion des scores par Somme Pondérée a été étudiée. Les détails de cette technique de fusion seront donnés plus loin dans ce paragraphe car elle dépend de plusieurs paramétres.

Une autre question s'est posée à propos du nombre de caractéristiques $(t_{-H}, (t_{-Bi})_{i=1..6}$ et t_{-L}) à garder par Adaboost à partir de chaque bande de la PO. Deux options ont été explorées : garder le même nombre de caractéristiques pour toutes les bandes ou alors pondérer le nombre des caractéristiques par bande.

Lorsqu'un **même nombre de caractéristiques** a été gardé sur chaque bande $(t_{-H} = (t_{-Bi})_{i=1..6} = t_{-L} = t/8)$, puis lorsqu'une fusion des scores par somme a été opérée (cf. **Table 2. 27** Sans pondération et **Table 2. 28** Tests 3.b), les performances se sont améliorées par rapport à l'application d'Adaboost sur la PO entière. De plus, cette méthode consomme désormais beaucoup moins de temps, ce qui a permis d'augmenter le nombre d'itérations/caractéristiques gardées au point de les doubler.

L'idée de procéder par pondération du nombre de caractéristiques par bande est venue suite à une analyse plus approfondie du comportement de l'algorithme d'Adaboost quand il était appliqué sur la PO entière (cf. Table 2. 28 Test 3.a). En effet, nous avons noté qu'Adaboost ne sélectionnait pas le même nombre de caractéristiques à partir des différentes bandes. Par exemple, il y avait beaucoup plus de caractéristiques sélectionnées à partir de la bande passe-bas qu'à partir de la bande passe-haut. Cette constatation n'annule en aucun cas la complémentarité d'apport des bandes mais lui rajoute une notion de différence de l'importance de chaque bande. D'où l'idée de la pondération du nombre de caractéristiques sélectionnées par bande. Une fois que le nombre total de caractéristiques souhaité est fixé, les poids sont attribués en se basant sur la distribution des caractéristiques trouvée dans le test 3.a comme suit : 2%; 12,75%; 13,5%; 10,25%; 18%; 15,25%; 18,5%; 9,75% respectivement pour les bandes passe-haut, passe-bandes orientées de 1 à 6 et passe-bas. Par exemple, si on se fixe un nombre total t de 800 caractéristiques à garder, Adaboost sélectionnera respectivement : $t_{-H} = 16$, $t_{-B1} =$ 102, $t_{B2} = 108$, $t_{B3} = 82$, $t_{B4} = 144$, $t_{B5} = 122$, $t_{B6} = 148$ et $t_{L} = 78$ caractéristiques à partir de chacune des bandes précitées comme montré dans la Table 2. 27.

La **Table 2. 28** Tests 3.c montre que la pondération apporte une amélioration sur presque toutes les expériences IV² en comparant les résultats avec ceux obtenus en gardant le même nombre de caractéristiques pour toutes les bandes (Table 2. 28 Tests 3.b). Cette amélioration atteint les -5,8% sur l'expérience 3 en réduisant l'EER de 13,7 à 12,9%.

	Sans ponderation du nombre de caracteristiques par bande							
Bande de la PO	HP	BP1	BP2	BP3	BP4	BP5	BP6	LP
Ratio nbr caract./bande	12,5 %	12,5%	12,5%	12,5%	12,5%	12,5%	12,5%	12,5%
Nbr caract./bande	t_H	t1	t82	t3	t4	t5	t6	t_L
Nbr total caract. t = 400	50	50	50	50	50	50	50	50
Nbr total caract. t = 800	100	100	100	100	100	100	100	100
		Ave	c pondératio	on du nombre	e de caracté	ristiques par	bande	
Bande de la PO	Η	B1	B2	B3	B4	B5	B6	L
Ratio nbr caract./bande	2 %	12,75%	13,5%	10,25%	18%	15,25%	18,5%	9,75%
Nbr caract./bande	t_H	t_ _{B1}	t82	t3	t84	t5	t6	t_L
Nbr total caract. t = 400	8	51	54	41	72	61	54	39
Nbr total caract. t = 800	16	102	108	82	144	122	148	78

Table 2. 27 Exemples de la répartition des caractéristiques sélectionnées par Adaboost sur chaque bande de la PO avec et sans pondération du nombre de caractéristiques.

A part la pondération au niveau des caractéristiques, une autre possibilité de pondération à explorer était celle des scores issus de chaque classifieur avant la fusion par somme. La règle de fusion des scores par somme pondérée que nous avons suivie est définie par l'équation (2. 30) où E_{sb} et W_{sb} sont respectivement l'EER et le poids associés aux scores de la bande sb [Su 2009], nb_{sb} étant le nombre total des bandes.

$$W_{sb} = \frac{1}{\frac{U}{E_{sb}}} \qquad \text{avec } U = \sum_{sb}^{nb_sb} \frac{1}{E_{sb}}, \sum_{sb}^{nb_sb} W_{sb} = 1 \text{ et } 0 < W_{sb} < 1$$
(2.30)

La fusion des scores a été testée pour le même nombre de caractéristiques par bande (**Table 2. 28** Tests 3.d) puis pour des nombres de caractéristiques pondérés par bande (**Table 2. 28** Tests 3.e). Les résultats montrent une amélioration allant jusqu'à 15,2% dans l'expérience 1 (3,9% contre 4,6%) en comparant la configuration 3.d à la fusion de scores non pondérés (**Table 2. 28** Tests 3.b). Les performances sont encore meilleures pour presque toutes les expériences avec la double pondération du nombre de caractéristiques et des scores par bande (**Table 2. 28** Tests 3.e).

En résumé, la méthode Adaboost/PO ayant la configuration optimale (3.e) s'obtient, d'abord, par le filtrage de l'image entière de taille (64×64) par une PO à 6 orientations et 3 échelles. Adaboost est ensuite appliqué sur chaque bande filtrée de la PO (1 passehaut, 6 passe-bandes orientées et 1 passe-bas) pour sélectionner un nombre optimal de caractéristiques sur chaque bande. Enfin, le score final est obtenu en fusionnant les scores issus des différentes classifieurs par la règle somme pondérée.

Table 2. 28 Comparaison entre les différentes configurations de l'algorithme Adaboost/PO (bleu) avec les algorithmes simples correspondants (vert). 3.a : Adaboost sur la PO entière. 3.b à 3.e : Adaboost sur chaque bande de la PO. Description : nombre de caractéristiques ; ratio = pondération du nombre de caractéristiques par bande ; s = fusion des scores par somme ; ws = fusion des scores par somme pondérée.

Numéro test	Description	Participants	Exp1	Exp2	Exp3	Exp4
1		Adaboost	3,7 (±0,6)	22,7 (±1,4)	44,3 (±2,0)	28,6 (±1,8)
2		РО	11,0 (±1,0)	23,0 (±1,4)	20,1 (±1,6)	23,9 (±1,7)
3. a	400(PO entière)	Ada/PO	5,0 (±0,7)	19,5 (±1,3)	14,1 (±1,4)	22,9 (±1,6)
3.b	400(8*50)s	Ada/PO	4,4 (±0,7)	18,8 (±1,3)	13,7 (±1,4)	22,4 (±1,6)
	800(8*100)s	Ada/PO	4,6 (±0,7)	18,6 (±1,3)	13,7 (±1,4)	21,3 (±1,6)
3.c	400(ratio)s	Ada/PO	4,5 (±0,7)	18,6 (±1,3)	13,0 (±1,3)	21,8 (±1,6)
	800(ratio)s	Ada/PO	4,7 (±0,7)	18,2 (±1,3)	12,9 (±1,3)	20,9 (±1,6)
3.d	400(8*50) ws	Ada/PO	3,7 (±0,6)	18,7 (±1,3)	13,5 (±1,3)	22,4 (±1,6)
	800(8*100)ws	Ada/PO	3,9 (±0,6)	18,5 (±1,3)	12,9 (±1,3)	21,0 (±1,6)
3.e	400(ratio) ws	Ada/PO	4,1 (±0,6)	18,6 (±1,3)	13,0 (±1,3)	21,7 (±1,6)
	800(ratio) ws	Ada/PO	4,1 (±0,7)	18,1 (±1,3)	12,9 (±1,3)	20,4 (±1,6)

2.2.5 Comparaison à la campagne IV² et méthode de reconnaissance faciale classique retenue

Suite à la présentation des trois méthodes combinant la PO à la PCA, la LDA et Adaboost, nous allons analyser leurs performances en comparaison à la campagne IV^2 . La meilleure méthode combinée classique de reconnaissance de visage sera retenue comme méthode candidate pour faire partie de l'approche multimodale finale de reconnaissance basée sur l'iris et le visage.

La **Table 2. 29** fournit un récapitulatif des méthodes de reconnaissance de visage présentées dans ce chapitre avec une comparaison aux algorithmes de la campagne IV² [Mellakh 2009-b].

Il est clair que les méthodes combinées apportent une amélioration considérable à la performance des méthodes individuelles utilisées séparément et ce pour la totalité des expériences. De plus, la supériorité de la LDA par rapport à la PCA s'est transposée aux méthodes combinées où la LDA/PO dépasse de loin la PCA/PO. Si on compare entre eux les trois algorithmes proposés, la LDA/PO réalise les meilleures performances, suivie d'Adaboost/PO et enfin de la PCA3/PO.

Maintenant, si on compare séparément chaque algorithme combiné proposé face aux résultats de la campagne IV^2 , nous commençons d'abord par la **PCA3/PO** qui dépasse pratiquement toutes les méthodes projectives de la campagne IV^2 pour les deux premières expériences. Par contre, elle se trouve dépassée par la PCA2 dans les deux dernières expériences contenant des variabilités de qualité et de sessions. Ceci est expliqué par l'ensemble d'apprentissage différent utilisé par la PCA2 de taille assez grande (300 images contre 52 pour la PCA3/PO) et contenant des variabilités analogues à celles des ensembles de tests dans les expériences 3 et 4.

Ces défis supplémentaires de taille réduite de l'ensemble d'apprentissage et des conditions différentes d'acquisition expliquent pourquoi PCA2 dépasse également la méthode combinée proposée **Adaboost/PO** au niveau de l'expérience 4. Mais ces défis n'ont pas empêché notre méthode d'être meilleure sur toutes les autres expériences. Il en est de même pour la PCA1 [Chaari 2009] qui se trouve dépassée par Adaboost/PO au niveau de toutes les expériences. D'un autre côté, comparée à la LDA1, à l'exception du premier scénario (expérience 1), Adaboost/PO est meilleure sur toutes les expériences. Elle reste, cependant moins performante que la LDA2/Gabor qui est le fruit de l'appariement d'une méthode projective (LDA) avec une méthode multi-échelles multi-orientation (Gabor).

Considérons maintenant l'algorithme combiné LDA3/PO. Ses résultats dépassent de loin ceux des algorithmes basés sur la PCA (PCA1, PCA2, PCA3, modPCA). Par ailleurs, comparée à la LDA2/Gabor qui s'appuie également sur la combinaison d'une LDA avec une méthode multi-résolution multi-orientation, la méthode LDA3/PO ici présentée est légèrement moins bonne pour les 3 dernières expériences. Néanmoins, elle est plus performante face à la première expérience. Il est important de souligner, par ailleurs, que, contrairement à la LDA3/PO ici proposée qui s'est strictement conformée au protocole IV² et dont l'ensemble d'apprentissage de taille réduite (156 images de 52 sujets) a été capturé dans des conditions assez bonnes, la LDA2/Gabor a utilisé un ensemble d'apprentissage différent comprenant plusieurs variabilités [Mellakh 2009-a]. La taille relativement grande de l'ensemble d'apprentissage en plus des conditions d'acquisition comparables des images d'apprentissage et ceux de test de la LDA2/Gabor expliquent ses meilleurs résultats dans les expériences 2, 3 et 4.

	Participants	Exp1	Exp2	Exp3	Exp4
	PCA1	6,7 (±0,8)	20,7 (±1,3)	20,1 (±1,6)	22,2 (±1,6)
Campagne IV ²	PCA2	7,3 (±0,8)	21,6 (±1,4)	13,6 (±1,4)	16,3 (±1,4)
	mod PCA	5,3 (±0,7)	20,7 (±1,4)	19,5 (±1,6)	20,5 (±1,5)
	LDA1	3,7 (±0,6)	22,5 (±1,4)	21,7 (±1,7)	19,7 (±1,5)
	LDA2/Gabor	4,2 (±0,6)	12,0 (±1,1)	8,3 (±1,1)	11,3 (±1,2)
	РО	11,0 (±1,0)	23,0 (±1,4)	20,1 (±1,6)	23,9 (±1,7)
Methodes individuelles	PCA3	6,1 (±0,8)	21,9 (±1,4)	24,8 (±1,7)	21,9 (±1,6)
muviduenes	LDA3	5,2 (±0,7)	25,1 (±1,4)	18,5 (±1,5)	20,9 (±1,6)
Máthadag aomhináng	PCA3/PO	5,0 (±0,7)	19,9 (±1,3)	17,4 (±1,5)	17,4 (±1,6)
proposóos	LDA3/PO	3,6 (±0,6)	17,5 (±1,3)	10,0 (±1,2)	14,5 (±1,4)
proposees	Adaboost/PO	4,1 (±0,7)	18,1 (±1,3)	12,9 (±1,3)	20,4 (±1,6)

Table 2. 29 Résultats comparatifs entre les algorithmes combinés proposés (bleu) avec les algorithmes individuels correspondants (vert) et ceux de la campagne d'évaluation IV^2 (noir).

En plus du travail présenté dans ce manuscrit, d'autres travaux combinant la LDA à la PO existent dans la littérature, mais traitent l'aspect identification plutôt que celui de la vérification. Tel est le cas du travail proposé dans [El Aroussi 2009]. Par contre, les résultats de reconnaissance qu'on y communique (98,75% sur la base ORL, 100% sur la base Yale et 91,72% sur un sous ensemble de la base FERET) ont été obtenus sur des

bases de test de tailles beaucoup plus réduites (40, 15 et 145 sujets) par rapport à celles de l'apprentissage, ce qui rend les résultats extrêmement dépendants des sujets présents dans ces ensembles. D'un autre côté, Zhang et Jia ont élaboré une étude statistique pertinente [Zhang X. 2005] permettant de démontrer la robustesse de leur méthode sur des bases d'images importantes aussi bien en taille qu'en variabilités (295 sujets de la base XM2VTS et 429 sujets de la base FERET).

En conclusion, les deux meilleures méthodes combinées proposées sont celles de la LDA3/PO qui est toujours classée en première ou en deuxième position comparée aux méthodes de l'évaluation IV², puis celle d'Adaboost/PO qui vient en troisième position. Il est clair que la méthode LDA2/Gabor reste à la tête des algorithmes pour les expériences 2 à 4. Néanmoins LDA3/PO la dépasse dans le premier scénario (avec 3,6% d'EER contre 4,2%) et la suit de près pour le reste des scénarios, malgré la taille réduite et le manque de variabilités de son ensemble d'apprentissage, contrairement à LDA2/Gabor [Mellakh 2009-a]. C'est pourquoi nous allons la retenir en tant que meilleure méthode classique de reconnaissance de visage dans la suite de nos travaux.

Conclusion

Ce chapitre constitué de deux grands volets a été consacré à l'étude des méthodes biométriques monomodales basées sur l'iris et le visage dans le but d'une fusion ultérieure dans une approche multimodale.

Dans le premier volet, nous avons choisi une méthode de référence basée sur les filtres Log-Gabor. Nous avons proposé deux critères de qualité basés sur l'occultation des iris que nous avons associés à la méthode de référence et qui nous ont permis de réduire au tiers son taux d'erreur.

A travers le deuxième volet de ce chapitre, nous avons proposé plusieurs approches de combinaison de méthodes, pour la reconnaissance de visage 2D, alliant la Pyramide Orientable (PO) avec, d'un côté, des méthodes projectives comme la PCA et la LDA et, d'un autre côté, une méthode de classification non linéaire qu'est Adaboost. La PO étant une méthode d'analyse spatio-fréquentielle, l'information est extraite des pixels de l'image en appliquant une famille de filtres passe-haut, passe-bas et passe-bande à différentes résolutions et à diverses orientations. Nous nous sommes intéressés aux paramètres de la PO comme les orientations et les résolutions mais aussi au choix des bandes. Nous avons observé que, parmi les trois algorithmes combinés, la LDA est capable d'extraire l'information la plus discriminante de la nouvelle représentation de la PO et ce suite à l'étude de plusieurs stratégies de fusion des scores issus des différentes bandes.

Les résultats des évaluations ont confirmé la robustesse de cette fusion face aux méthodes individuelles ascendantes. De plus, comparée aux algorithmes de la première campagne d'évaluation IV², cette méthode a montré sa supériorité par rapport à plusieurs algorithmes classiques comme la LDA, la PCA et la PCA modulaire. Lorsque les conditions de test sont dégradées, notre méthode est classée en deuxième position après l'algorithme LDA2/Gabor qui a eu l'avantage d'utiliser une base d'apprentissage non contrôlée, contrairement à la méthode proposée. La performance de la LDA3/PO est encore plus notable dans des conditions contrôlées où elle arrive en tête de liste.

Suite à ces résultats, nous avons décidé de retenir la LDA3/PO en tant que meilleure méthode classique de reconnaissance du visage. Toutefois, dans le souci de concevoir une méthode biométrique plus performante même dans des environnements non contrôlés, nous allons nous orienter, dans le chapitre suivant, vers une approche radicalement différente de reconnaissance de visage : l'approche bio-inspirée qui essaye d'émuler le fonctionnement du cerveau humain dans la tâche d'interprétation des images de visage et ce pour construire des systèmes automatiques de reconnaissance de visage.

Chapitre 3

Approches bio-inspirées appliquées en biométrie faciale

Intro	oducti	ion	
3.1	Cont	texte neuro-scientifique	
	3.1.1	La vision, de la rétine au cortex	
		a) Analyse de l'information visuelle dans la rétine	
		b) La perception des images par le cortex visuel	
		c) La perception des visages	
	3.1.2	2 Approches modélisant le mécanisme de la vision utilisées en r faciale	econnaissance
		a) Modélisation par la PCA	
		b) Modélisation par le traitement multirésolution	
		c) Modélisation par le traitement adaptatif à rétroaction	
		d) Modélisation par le filtre DoG	
		e) Travaux de Gosselin et Schyns	
		 a) Autres modélisations récentes appliquées en reconnaissar 	
	3.1.3	 Approches bio-inspirées retenues 	
3.2	Reco	onnaissance bio-inspirée par LBP/DoG	
	3.2.1	Concepts théoriques de la méthode LBP/DoG	
		a) Motifs binaires locaux – LBP	
		b) Différence de Gaussiennes – DoG	
	3.2.2	2 Méthode LBP/DoG	
	3.2.3	B Méthode LBP/DoG pondérée selon les travaux de Gosselin et S	Schyns136
		a) Sélection des régions d'intérêt	
		b) Pondération des ROIs et des échelles	
		c) Résultats expérimentaux et discussion	
3.3	Reco	onnaissance faciale bio-inspiré par Reconstruction Async	chrone – RA141
	3.3.1	Concepts théoriques de la RA	141
		a) Décomposition	
		b) Reconstruction	
	3.3.2	Prétraitement bio-inspiré par RA appliqué en reconnaissance :	faciale148
		a) RA avec critère fixe	
		b) RA avec criteres adaptatifs	
3.4	Appo	orts de la RA sur les meilleures méthodes proposées et bode unimodale finale du visage	t choix de la
	meth	noue uninouale infate un visage	
Cone	clusio	on	

Introduction

De nombreux chercheurs ont fait valoir que l'étude des voies et des comportements visuels est une méthode efficace pour la découverte de nouveaux algorithmes en vision artificielle. Marr a fait valoir en particulier que la biologie est une bonne source d'idées pour la conception d'ingénierie [Marr 1982]. D'autres chercheurs tels que Wandell pensent, à l'inverse, que les méthodes de rétroingénierie sont plus efficaces pour la compréhension de la neuroscience [Wandell 1995]. Quel que soit le sens du flux des intérêts, depuis la neuroscience à l'ingénierie ou dans le sens inverse, la simple présence de flux est une bonne raison au spécialiste en sciences de la vision et à l'ingénieur en imagerie pour se familiariser avec les questions biologiques, comportementales et computationnelles.

Dans le premier volet de ce chapitre nous commencerons par décrire le mécanisme de la vision dans le cerveau humain. Ensuite, nous en présenterons quelques modélisations par certains algorithmes fondamentaux utilisés en reconnaissance faciale. Trois de ces modélisations bio-inspirées seront exploitées dans les méthodes proposées qui font l'objet du second volet.

3.1 Contexte neuro-scientifique

Dans cette section, nous présentons les concepts neuroscientifiques de base pour la perception des images. En particulier, nous verrons comment certains algorithmes en biométrie faciale émulent le fonctionnement du cerveau humain dans la perception de visages.

3.1.1 La vision, de la rétine au cortex

De prime abord, rien n'est plus naturel que la vision. Pourtant, d'énormes difficultés techniques sont impliquées lors de ce processus : l'œil doit d'abord obtenir une image du monde extérieur qui doit se projeter sur la rétine en image inversée. Le signal électrique émis par les cellules photosensibles de la rétine doit ensuite être transmis via les deux nerfs optiques qui se subdivisent en deux branches : une branche qui rejoint le colliculus supérieur dirigeant les mouvements des yeux et de la tête, l'autre rejoint le corps genouillé latéral (ou CGL), premier relais des voies visuelles.



Figure 3. 1 Chemin visuel dans le cerveau humain.

Les cellules du corps genouillé latéral vont ensuite rejoindre leur cible principale : le cortex visuel primaire. C'est là que l'image va commencer à être reconstituée et envoyée vers différentes aires visuelles du cerveau pour prendre forme et être interprétée (cf. **Figure 3.1**).

a) Analyse de l'information visuelle dans la rétine

À la fin des années 50, Hubel et Wiesel ont émis une hypothèse qui fut dominante en neurosciences pendant quelques décennies : ils ont proposé que la fonction des aires visuelles primaires était de détecter et d'extraire les barres et contours des objets environnants [Hubel 1959]. À la fin des années soixante, des travaux importants effectués par Fergus Campbell et ses collaborateurs [Blakemore 1969] ont mené à une proposition différente, et aujourd'hui dominante, suggérant que le système visuel primaire réalisait sur les images naturelles une opération qui s'apparente à une analyse de Fourier locale au niveau des champs récepteurs.



Figure 3. 2 Réponses des cellules ganglionnaires à centre "ON" (à gauche) et à centre "OFF" (à droite) à un stimulus lumineux. Le jaune représente la zone du champ récepteur éclairée.

En 1980, Marr et Hildreth sont arrivés à la conclusion que ce sont les cellules ganglionnaires de la rétine et les cellules du corps genouillé latéral, désignées comme les cellules X, qui étaient les cellules sensibles au contraste dans le cerveau humain [Marr 1980]. Dans leurs recherches, Marr et Hildreth se sont appuyés sur des études menées par Kuffler et Barlow révélant que les cellules ganglionnaires de la rétine étaient des cellules possédant un antagonisme centre-périphérie de type centre ON – périphérie OFF, ou bien l'inverse [Kuffler 1953] [Barlow 1961]. Ces cellules transmettent le signal nerveux sous forme de potentiels d'action générés de façon spontanée. C'est la fréquence de décharge de ces potentiels d'actions qui est amplifiée ou diminuée par l'apparition de lumière dans les champs récepteurs des cellules ganglionnaires. La **Figure 3. 2** illustre les champs récepteurs de cellules ganglionnaires ainsi que leurs réponses à trois types de stimulus lumineux.



Figure 3. 3 (a) Cellules nerveuses de la rétine et (b) modalités de réponses en modes centre ON et centre OFF à des stimuli d'illumination centrale et (c) périphérique (extrait de [Dubuc 2013]).

Les processus à l'origine des réponses aux stimuli de type illumination centrale et périphérique sont montrés plus en détails dans la **Figure 3. 3**. Au départ, la lumière est captée par les photorécepteurs qui émettent en permanence un neurotransmetteur appelé glutamate, en l'absence de lumière. Si la lumière frappe les photorécepteurs au centre de leur champ récepteur, les cellules bipolaires à centre ON, connectées à ce centre illuminé, seront stimulées et vont être dépolarisées en augmentant la relâche de glutamate. Ceci va augmenter le taux de décharge du potentiel d'action des cellules ganglionnaires directement connectées et qui sont de même type ON. En revanche, la réaction inverse va se passer pour les cellules bipolaires à centre OFF connectées au centre illuminé, qui vont diminuer la relâche de glutamate, entraînant une diminution du taux de décharge du potentiel d'action des cellules ganglionnaires à centre OFF directement liées (**Figure 3. 3-b**). Le processus ainsi décrit, se trouve complètement inversé si la lumière frappe la périphérie du champ récepteur comme le montre la **Figure 3. 3-c**.

b) La perception des images par le cortex visuel

Le cortex visuel est constitué de deux principales parties qui se distinguent chacune par ses propres fonctions dans la tâche de la perception des images : le cortex visuel strié et le cortex visuel extrasié.

i) Le cortex visuel primaire (strié)

Aussi appelé cortex strié ou simplement V1, le cortex visuel primaire se situe dans la partie postérieure extrême du lobe occipital du cerveau (cf. **Figure 3. 4**). Les cellules du cerveau du cortex visuel primaire sont affectées à des tâches spécifiques de reconnaissance de certaines structures géométriques. C'est l'endroit du cerveau où la reconnaissance d'objet est la plus efficace [Logothetis 1999]. Elle se fait de manière très rapide et reste très robuste aux variations d'éclairage, de pose, de taille et d'angle de vue. Trois types de cellules sont présents dans le V1 : les cellules simples sensibles à l'axe d'orientation des stimuli, les cellules complexes sensibles au mouvement de l'axe et les cellules hypercomplexes sensibles aux bords et aux angles. Le cortex visuel sert à décrire ce que l'on voit avant de déléguer au cortex frontal le soin de choisir l'action la plus judicieuse.



Figure 3. 4 Aires visuelles et lobes cervicaux.

ii) Le cortex visuel extrasié

Le cortex visuel primaire envoie une forte proportion de ses connexions au cortex visuel secondaire (ou V2) dont les neurones ont des propriétés semblables à celles des neurones de l'aire visuelle primaire, mais qui s'en distinguent en répondant à des formes beaucoup plus complexes. L'analyse des stimuli visuels amorcée dans V1 et V2 se poursuit ensuite à travers deux grands systèmes corticaux de traitement de l'information visuelle comme illustré dans la **Figure 3. 5**. Le premier correspond à une voie ventrale qui s'étend vers le lobe temporal et serait impliquée dans la reconnaissance des objets : le « quoi » à travers une représentation consciente. Le second est une voie dorsale qui se projette vers le lobe pariétal et serait essentielle à la localisation de l'objet : le « où ». Cette voie serait essentiellement inconsciente. Par exemple, les personnes dont la voie du « quoi » est lésée, mais pas celle du « où », peuvent repérer et ramasser n'importe quel objet, mais elles ne savent pas ce qu'elles ramassent.

On pense aujourd'hui que la fonction principale de la voie dorsale, considérée essentiellement inconsciente, est de guider en temps réel les actions que nous dirigeons vers des objets. Dans le domaine de la vision, notre cerveau serait donc doté de deux circuits travaillant en parallèle : un circuit de l'action efficace en dehors de toute « conscience communicable » (par la voie dorsale), et un circuit analytique destiné à comparer ce que nous voyons à une « banque de données » préalablement acquise (par la voie ventrale en partant du cortex primaire).

Ce fonctionnement de nos centres visuels primitifs capables d'agir avec cohérence indépendamment de notre conscience pourrait expliquer les déplacements nocturnes des somnambules qui gardent généralement les yeux ouverts à ce moment là.



Figure 3. 5 Schéma de la répartition des différentes fonctions visuelles, intégratives et cognitives du cortex cérébral humain, à partir de [Birbaumer 2006].

c) La perception des visages

Des preuves neurologiques montrent un mécanisme spécifique pour le traitement des visages dans le cerveau humain. En électroencéphalographie (EEG), la présentation de l'image d'un visage engendre un potentiel évoqué en seulement 170ms environ [Bötzel 1995]. Comparé à la réponse à d'autres stimuli comme des véhicules, des animaux, des chaises, ... où l'activité neuronale visuelle peut atteindre les 320ms [Van Rullen 2004], cela donne un indice de la rapidité du traitement des informations visuelles qui permet la perception et la reconnaissance d'un visage. Plus encore, des études plus récentes, ont rapporté des effets de haut niveau bien plus précoces, pouvant apparaître avant 100ms post-stimulus. Les plus extrêmes suggèrent que la reconnaissance des visages [Seeck 1997] ou la discrimination du genre des visages [Mouchetant-Rostaing 2000] pourraient être réalisées avant même 50ms post-stimulus.

Le Dr. Doris Tsao, neuroscientifique à l'Université de Brême en Allemagne, a découvert, à travers l'imagerie par résonance magnétique fonctionnelle (fMRI), qu'un certain nombre de régions est fortement impliqué dans la perception des visages [Tsao 2008]. L'activation la plus prononcée des neurones pour une tâche liée au visage (identification, analyse des émotions, ...) se situe toujours dans l'aire fusiforme des visages ("Fusiform Face Area", FFA) qui constitue une partie du gyrus fusiforme au niveau de la jonction des lobes temporaux et occipitaux (cf. Figure 3. 4). Ces études vont dans le même sens que celles de Haxby et al. réalisées dans l'un des laboratoires de recherches du gouvernement des Etats-Unis [Haxby 2000] [Haxby 2010] dans lesquelles ils proposent un modèle décrivant les aires impliquées dans la perception des visages (cf. Figure 3. 6).



Figure 3. 6 Modèle du système neuronal distribué pour la perception de visage extrait de [Haxby 2010].

Ce modèle divise les aires du cerveau qui sont impliquées dans la perception de visage en un Système Noyau – les aires visuelles extrasiées occipito-temporales qui jouent un rôle central dans l'analyse visuelle des visages – et un Système d'Extension – systèmes neuronaux dont les fonctions ne sont pas essentiellement visuelles mais qui jouent des rôles critiques dans l'extraction de l'information à partir des visages. Dans le Système Noyau, une distinction est à souligner entre la représentation des caractères invariants qui sont critiques pour la détermination de l'identité faciale d'un côté et la représentation des caractéristiques variables qui sont critiques pour la compréhension des mimiques faciales telles que les expressions et le regard. Dans le Système d'Extension, trois aires du cerveau sont impliquées respectivement, dans la représentation des connaissances de la personne (connaissances biographiques, souvenirs, ...), dans la compréhension de l'action (y compris le regard et l'attention) et celle des émotions. La reconnaissance de visages familiers implique des codes visuels pour les individus familiers dans les aires fusiformes du Système Noyau et possiblement dans le cortex temporal antérieur, en parallèle avec l'activation automatique des connaissances de la personne et celle des réponses émotionnelles. La compréhension des expressions du visage impliquent des codes visuels dans le Sillon Temporal Supérieur – STS, en parallèle avec l'activation de représentations de programmes d'émotions et de programmes moteurs pour la production d'expressions. La perception du regard inclut également des codes visuels dans le STS, en parallèle avec l'activation d'aires cervicales pour le déplacement de l'attention et le contrôle oculomoteur.

Une lésion cérébrale de ces régions et en particulier de la FFA, peut entraîner une prosopagnosie (cf. **Figure 3. 7**), c'est-à-dire une incapacité à reconnaître les visages (avec une vision normale par ailleurs) [Sacks 2010].



Figure 3. 7 Illustration de la prosopagnosie (d'après [Wawrzyniak 2009]).



Figure 3. 8 Illusion d'optique Arbre aux dix visages.

Les recherches du Dr Tsao ont, par ailleurs, abouti à une conclusion connexe surprenante : *"des objets non-visages peuvent avoir certaines caractéristiques qui constituent des déclencheurs faibles des cellules spécifiques aux visages"*. Le mécanisme de la perception essaie de déterminer les régularités dans tous ces visages pour créer un prototype. Cependant, ce prototype n'est pas parfait : parfois, certains visages authentiques ne correspondent pas à ces régularités, et inversement, certains nonvisages peuvent parfois les satisfaire. En fait, les gens sont tellement habitués à voir des visages partout que la sensibilité à leur égard est devenue suffisamment élevée pour produire constamment des faux positifs, créant ainsi des illusions d'optique (**Figure 3**. 8). Selon le Dr Sinha du M.I.T. [Svoboda 2007], il est extrêmement bénéfique pour le cerveau de devenir performant dans la tâche de reconnaissance de visage et en même temps de ne pas être très strict dans ses critères d'inclusion. "Le coût de manquer un visage est plus élevé que le coût de déclarer un non-visage en être un" a-t-il dit.

3.1.2 Approches modélisant le mécanisme de la vision utilisées en reconnaissance faciale

Une approche à la réflexion sur les algorithmes de traitement d'images est d'adopter la vision artificielle. Il s'agit d'une initiative dans l'automatisation et l'intégration de modèles sur la perception visuelle aux algorithmes de traitement d'images [Sinha 2006]. Suivre cette approche n'impliquerait pas la limitation de la façon de penser, dans l'élaboration des algorithmes de traitement d'images, aux problèmes liés à la vision humaine. Au contraire, il est essentiel de rester ouvert à des algorithmes de traitement d'images n'ayant aucune contrepartie biologique. Dans cette section, cependant, nous allons limiter notre analyse à un sous-ensemble d'algorithmes qui sont liées à la vision humaine. Bien que n'étant pas forcément le plus large, ce sous ensemble est doté d'une importance d'autant plus forte qu'il existe de nombreuses applications potentielles aux algorithmes de traitement d'image émulant la performance humaine, notamment dans le domaine de la biométrie.

a) Modélisation par la PCA

En 1961, des recherches menées par le neuroscientifique Horace Barlow aboutissent à la conclusion que l'un des principaux objectifs de traitement du système visuel consiste à réduire la redondance de luminosité entre les points voisins d'une image naturelle [Barlow 1961]. Pour ce faire, il a élaboré un algorithme permettant de coder les images comportant des redondances statistiques par un code dont les composantes sont statistiquement indépendantes.

Plus tard en 1996, Olshausen & Field font l'hypothèse que le cortex visuel produit une représentation efficace des images naturelles en faisant une extraction statistiquement indépendante des structures qui s'y trouvent [Olshausen 1996] [Olshausen 1997]. L'algorithme qu'ils proposent produit des champs récepteurs qui ressemblent fortement aux champs récepteurs du cortex visuel primaire. Par ailleurs, ils parviennent à montrer que l'ICA (Independent Component Analysis) initialement introduite par Hérault et ses collaborateurs en 1985 et dont l'un des algorithmes a été proposé dans [Bell 1995] peut être vue comme une forme particulière de leur algorithme, et par la suite relèvent le défi d'éliminer la redondance statistique [Hérault 1985]. En reconnaissance faciale, l'ICA permet de minimiser les dépendances statistiques d'ordre élevé des données d'entrée. Ainsi, l'information pertinente est extraite à partir d'un ensemble riche de données en tentant de trouver les bases de projection sur lesquelles les données projetées sont statistiquement indépendantes entre-elles. Dans le cas particulier de distribution gaussienne et de statistiques du second ordre, l'ICA coïncide avec la PCA, qui constitue la méthode de référence en reconnaissance faciale. L'algorithme PCA présenté précédemment dans l'état de l'art (cf. § 2.2.3a)) peut être vu comme une modélisation des réponses des cellules nerveuses du cortex visuel.

b) Modélisation par le traitement multirésolution

Selon Campbell et Robson, la sensibilité au contraste des neurones du V1 ne découle pas de la sensibilité d'un seul mécanisme, mais de l'activité combinée de plusieurs mécanismes indépendants, nommés « canaux fréquentiels » [Campbell 1968]. Chaque mécanisme serait sélectivement sensible à une bande de fréquences spatiales, se comportant comme un filtre qui détecterait les variations de luminance sur une échelle spatiale particulière. Aussi, dans cette perspective théorique, la reconnaissance visuelle « ultra-rapide » suivrait une logique de traitement « *Coarse-to-Fine* » : l'information « grossière » issue des basses fréquences spatiales véhiculées rapidement permettrait une première reconnaissance que validerait ou non l'information « fine » issue des hautes fréquences spatiales plus tardives. C'est l'approche d'un traitement multirésolution. Ainsi, une transformée multirésolution peut être considérée comme une modélisation du fonctionnement non-linéaire complexe de neurones V1. Cette approche justifie que nous ayons choisi de nous intéresser à certains algorithmes multirésolution, notamment à celui des différences de Gaussiennes.

c) Modélisation par le traitement adaptatif à rétroaction

D'après les travaux réalisés dans [Olshausen 1996] [Thorpe 1996] [Thorpe 2001] [Van Rullen 2003], vu la manière avec laquelle le système visuel humain atteint un nombre aussi phénoménal de calculs en des délais aussi réduits (les tâches de reconnaissance sont pratiquement terminées dans les premières 100-200ms suivant le stimulus), une grande part de ce traitement se base essentiellement sur des mécanismes rétroactifs au niveau du cortex visuel.

De nouvelles découvertes sur la diversité des interactions corticales de haut en bas montrent que les aires du cortex fonctionnent comme des processeurs adaptatifs, faisant l'objet d'attention, d'attentes, et de tâches perceptives [Gilbert 2007]. Selon ces découvertes, les états du cerveau sont déterminés par les interactions entre les multiples aires corticales et la modulation des circuits intrinsèques par des connexions de rétroaction. Dans l'apprentissage perceptif, l'encodage et le rappel des informations apprises impliquent une sélection appropriée des entrées qui transmettent les informations sur le stimulus à distinguer. Sur la base de ces travaux, et comme le visage est un cas assez particulier de stimulus à distinguer, nous avons complété nos implémentations avec un certain nombre de versions adaptatives d'algorithmes de reconnaissance faciale.

d) Modélisation par le filtre DoG

Au milieu des années soixante, et dans le but de reproduire le fonctionnement des cellules ganglionnaires de la rétine, des physiologistes tels que Rodieck et Stone et Enroth-Cugell et Robson ont construit un modèle à Différence de Gaussiennes où la variance de la gaussienne centrale est inférieure à celle de la gaussienne périphérique et dont la résultante bidimensionnelle ressemble à « un chapeau Mexicain » [Rodieck 1965] [Enroth-Cugell 1966].

Plus tard, à la fin des années soixante dix, Marr et Ullman ont proposé un premier algorithme permettant la modélisation du fonctionnement des cellules ganglionnaires en utilisant un opérateur Laplacien donné par l'**équation (3. 1)** [Marr 1979].

$$\nabla^2 = \frac{\delta^2}{\delta x^2} + \frac{\delta^2}{\delta y^2}$$
(3. 1)

Selon cet algorithme, les changements d'intensité à une échelle donnée sont mieux détectés en trouvant, pour une image I, les passages par zéro de la fonction :

$$\nabla^2 G(x, y) * I(x, y)$$
 (3. 2)

Avec G une distribution Gaussienne à 2 dimensions donnée par l'équation (3. 3) :

G(x, y) = exp
$$\left(-\frac{x^2 + y^2}{2\pi\sigma^2}\right)$$
 (3. 3)

Mathématiquement parlant, ces opérations pourraient certes effectuer la détection nécessaire des changements d'intensité mais, en réalité, ce ne sont que des formules idéales sur le plan théorique. En effet, sur le plan pratique, il est à remarquer que la formule de ∇^2 est fort compliquée pour la phase d'implémentation. De plus, il est peu probable qu'il existe, dans les premières phases du système visuel, des mécanismes neuronaux qui puissent réaliser directement des opérations mathématiques aussi avancées.

Toutefois, à partir des connaissances sur les cellules ganglionnaires de la rétine et de leurs champs réceptifs, il serait raisonnable de penser que ces champs réceptifs puissent produire des sorties dont le type des unes ressemblerait à des fonctions Gaussiennes à pondération positive et le type des autres à pondération négative. Ainsi, un autre filtre multirésolution qui approxime de très près l'opérateur Laplacien ∇^2 , nommé Différence de Gaussiennes (DoG), aurait l'avantage sur le ∇^2 de permettre de visualiser concrètement le travail accompli par les neurones (cf. **Figure 3. 9**).

Dans une démonstration éminente de cette partie de leur théorie, Marr et Ullman présentent, simultanément, les sorties de filtres DoG ainsi que les sorties effectives des cellules X en réponse aux mêmes stimuli consistant en des lignes et des bords. La similitude entre ces sorties est très remarquable comme le montre la **Figure 3. 10** [Marr 1979].


Figure 3. 9 Approximation de l'opérateur Laplacien $\sigma^2 G$ (ligne continue) par une DOG (ligne en pointillés). Les deux profils sont très similaires (extrait de [Marr 1980]).



Figure 3. 10 Comparaison des estimations des réponses de cellules X à des enregistrements électrophysiologiques. La première ligne illustre l'estimation S de la réponse par le filtre, pour un bord isolé, une barre mince de largeur w (où w est la largeur de la région d'excitation centrale du champ récepteur), et une barre large de largeur 2,5*w. La deuxième ligne correspond à la partie positive de S et la quatrième ligne à l'opposé de sa partie négative. Les troisième et cinquième lignes correspondent aux réponses physiologiques (extrait de [Marr 1979]).

Sur la base de cette analogie associant le comportement du filtre DoG à celui des cellules X du cortex visuel, nous avons choisi de l'utiliser dans nos approches bioinspirées pour la reconnaissance de visage.

e) Travaux de Gosselin et Schyns

Dans le cadre du décryptage de la compréhension du mécanisme de traitement de l'information du visage par le cerveau humain, le centre de neuroimagerie cognitive du département de psychologie à l'université de Glasgow a mené plusieurs travaux visant à décrypter la manière avec laquelle le cerveau humain reconnaissait l'identité, le sexe, la race, l'âge et l'état émotionnel d'une personne [Gosselin 2001] [Schyns 2002].

i) Travaux de Gosselin et al.

Gosselin et al. ont proposé une nouvelle méthode nommée « méthode des Bubbles » [Gosselin 2001] qui révèle les régions du visage utiles à l'identification des personnes. Pour cela ils ont commencé par présenter aux participants d'une expérience (observateur humains) des visages avec des parties masquées et des parties aléatoirement révélées ; l'ensemble des essais couvrent cependant tout le visage.

Ils ont par la suite analysé la façon avec laquelle ces observateurs humains identifiaient une personne et son état émotionnel. Ils ont conclu que le cerveau humain traite l'information visuelle selon plusieurs fréquences (par exemple, les contours du visage sont des hautes fréquences tandis que la texture des cheveux ou de la peau sont des basses fréquences). Ainsi pour étudier séparément les informations contenues sur différentes échelles fréquentielles, Gosselin et Schyns ont décomposé les images en six bandes de fréquences, en utilisant une pyramide Laplacienne donnant lieu à six échelles dans le domaine spatial. Chacune des échelles est ensuite aléatoirement échantillonnée par des bulles ou *Bubbles* qui ne laissent transparaitre qu'une certaine surface (voir **Figure 3. 11**) ; la somme des surfaces révélées reste constante d'une échelle à l'autre.

L'image d'entrée **Figure 3. 11-a** est initialement décomposée en 6 bandes de fréquences spatiales contenant chacune une octave³: 128-64 (échelle 1), 64-32 (échelle 2), 32-16 (échelle 3), 16-8 (échelle 4), 8-4 (échelle 5) et 4-2 (échelle 6) cycles par visages⁴. Cette dernière (échelle 6) n'est cependant ni présentée ni échantillonnée car elle est considérée comme fond pour le reste des stimuli. Une fois que tous les observateurs ont répondu aux stimuli de l'expérience, il était possible d'analyser où était située l'information permettant aux observateurs de donner une réponse correcte. La **Figure 3.** 11-c montre ces aires pour chaque échelle et la **Figure 3. 11-d**, les parties du visage correspondantes. La **Figure 3. 11-e** est une reconstruction des 5 échelles de la **Figure 3.**

³ 1 octave = 1 doublement de la fréquence spatiale ; une bande de fréquences spatiales de *n* octaves s'étend donc d'une borne x à une borne x * 2^n .

⁴ Nombre de cycles par (la largeur de) visage : unité de mesure utilisée pour les fréquences spatiales dans le domaine de la reconnaissance de visages. Il s'agit du nombre de cycles par degré d'angle visuel dans une grille sinusoïdale [Fiset 2009].

11-d. Grace à ces expériences Gosselin et Schyns ont montré que les informations utiles à la tâche de la reconnaissance se concentrent dans les fréquences intermédiaires, ce qui se traduit par la petite taille et le grand nombre des bulles dans les hautes fréquences et la grande taille et le petit nombre des bulles dans les fréquences intermédiaires.



Figure 3. 11 Méthode des Bubbles : (a) image initiale, (b) images filtrées selon les différentes bandes de fréquences (échelles), (c) représentation des *bubbles* pour chaqur échelle constituant l'échantillonnage aléatoire, (d) parties du visage révélées pour chaque échelle et (e) visage résultant de la sommation des informations issues de chaque échelle (un stimulus) (figure extraite de [Gosselin 2001]).

ii) Travaux de Schyns et al.

Pour révéler les régions du visage utiles et les bandes de fréquences spatiales les plus efficaces pour la reconnaissance des personnes (menant au moins à 75% de reconnaissance correcte par les observateurs), Schyns et al. ont calculé la proportion de l'aire du visage efficace pour chacune des cinq échelles échantillonnées par Gosselin et Schyns ; leurs résultats ont montré que l'échelle 3 était la plus importante pour la tâche d'identification suivie de près par l'échelle 4. Leur étude a permis aussi de conclure à l'inefficacité de la 5^{ème} échelle (voir **Figure 3. 12**) en plus de la 6^{ème} déjà ignorée par Gosselin et Schyns [Schyns 2002].



Figure 3. 12 (a) Régions significatives pour la tâche de reconnaissance du visage et (b) histogramme indiquant la proportion des pixels significatifs dans les quatre premières échelles. Figure extraite de [Schyns 2002].

Les résultats de leurs expériences ont prouvé qu'il existe une relation entre les réponses des observateurs et les régions du visage car Schyns et al. ont remarqué que les participants utilisaient toujours les mêmes régions en fonction de la tâche qui leur était demandée. Par exemple, pour identifier l'état émotionnel d'une personne, les participants se focalisaient sur la bouche tandis que l'identification nécessitait de regarder les yeux en premier et ensuite la bouche et le nez. De manière plus générale, la majorité du visage serait utile pour l'identification mais les yeux auraient plus d'importance en étant présents à toutes les échelles, suivis de la bouche puis du nez.

f) Codage par ordre de classement - Rank Order Coding

A la fin du vingtième siècle, Thorpe et Gautrais ont démontré l'invalidité d'une hypothèse longtemps ancrée expliquant la rapidité avec laquelle le système visuel humain interprète les scènes [Thorpe 1998]. Selon cette hypothèse réfutée, l'information transmise par les neurones se fait à travers un codage du taux de décharge des cellules recevant le signal. Le degré de synchronisme entre les décharges des neurones afférents aurait une influence directe sur la vitesse de traitement. Dans leur analyse, Thorpe et Gautrais ont fourni des arguments mathématiques selon lesquels le codage du taux de décharge ne pouvait pas être à l'origine de la grande rapidité avec laquelle le système visuel humain pouvait interpréter les scènes complexes. En effet, même la transmission des signaux les plus simples de manière fiable requiert soit des périodes d'observation excessivement longues incompatibles avec la vitesse de traitement sensoriel, soit un nombre trop grand de neurones redondants, incompatibles avec les contraintes anatomiques imposées par les voies sensorielles. Thorpe et Gautrais ont affirmé que d'autres stratégies de codage sont nécessaires quand le traitement se produit de manière très rapide et ont par ailleurs émis une hypothèse soutenant que cette performance ne pouvait être atteinte qu'à travers l'existence d'un codage de l'ordre (rang) de décharge des cellules plutôt que du taux de leur décharge. L'utilisation d'un tel codage temporel repose sur l'asynchronisme de décharge à travers une population de neurones afférents.

Ce modèle, appelé Codage par ordre de classement permettrait de classer l'information selon son importance dans le but d'utiliser, en priorité, la plus significative. Il a été adopté dans plusieurs travaux, notamment dans celui de [Delorme 1999] qui l'a utilisé comme entrée pour le réseau de neurones SpikeNet dans des applications de reconnaissance de formes. De nombreux travaux autour du codage par ordre de classement ont suivi en utilisant des ondelettes à base de DoG [Van Rullen 2001] [Sen 2007] [Masmoudi 2010] pour réduire la quantité d'informations lors de l'encodage des images, avec un minimum de perte de qualité dans la phase de reconstruction. Les trois travaux précités diffèrent essentiellement par la méthode finale de reconstruction, ainsi que par la manière de coder les réponses positives et négatives. Le modèle de codage par ordre de classement a aussi été utilisé dans des applications de développement telles que le système de reconnaissance d'objets ROCIT [Reeves 2004]. Il a également été utilisé dans l'analyse de la texture pour la détection de défauts [Plantier 2002], l'extraction de caractéristiques à partir d'images de l'environnement routier [Mayr 2005] et la détection de formes [Chew 2008].

g) Autres modélisations récentes appliquées en reconnaissance de visage

Dans la quête de simuler l'habilité et la rapidité des humains à reconnaitre les visages, les travaux s'inspirant du mécanisme de la vision humaine ne cessent de se multiplier en exploitant parfois plusieurs facettes du traitement biologique de l'information sélective aux visages.

En 2011, Huang Y. et al. ont développé une version améliorée du modèle bio-inspiré proposé dans [Serre 2007] émulant le processus de reconnaissance d'objets (entre autres, les visages) dans le cortex visuel chez les primates [Huang Y. 2011]. Ce modèle construit un ensemble de caractéristiques invariantes à l'échelle et à la position dont les propriétés sont similaires à celles des cellules se situant au niveau du flux ventral du cortex visuel primaire. Le système proposé dans [Huang Y. 2011] vise à palier deux limites principales du système de Serre et al. : les faux appariements dus aux entrées trop abondantes et la sélection aveugle des caractéristiques due à son architecture à action directe. Pour résoudre ces limitations, Huang Y. et al. proposent les deux aspects suivants : 1) l'élimination des entrées non informatives en imposant des contraintes de parcimonie, 2) l'application d'une boucle de rétroaction au niveau intermédiaire de la sélection des caractéristiques. Le premier aspect lié à la parcimonie est motivé par les résultats de recherches neuroscientifiques publiées dans [Olshausen 1996] indiquant que, pour l'image naturelle, la stratégie de codage parcimonieux est suffisante pour expliquer les propriétés des champs récepteurs des cellules simples dans le cortex visuel primaire des mammifères. Le deuxième aspect lié à la rétroaction se rapporte aux recherches psychophysiques présentées dans § 3.1.2-c). Le système proposé dans [Huang Y. 2011] produit des résultats comparables à ceux d'algorithmes de la littérature et a l'avantage

de constituer un outil prometteur pour les applications temps-réel de catégorisation d'objets vu sa complexité de calcul réduite.

D'autres travaux s'inspirant du modèle de vision humain ont été réalisés en 2011 par [Vu 2011]. Dans leurs études, ils ont démontré que les cellules bipolaires de la rétine non seulement éliminent les variations d'éclairage et le bruit mais, en plus, améliorent les contours de l'image. Ils ont alors proposé une méthode de normalisation de l'illumination pour les images de visage en simulant, d'un côté, la performance des photorécepteurs qui agissent comme un filtre non linéaire d'adaptation à la lumière et, d'un autre côté, les cellules bipolaires de la couche plexiforme externe agissant tels des filtres passe-bande. Le système proposé opère par une série de compressions logarithmiques constituant le filtre d'adaptation à la lumière, suivie d'un filtrage par Différences de Gaussiennes (DoG). Un seuillage est appliqué à la fin en vue d'améliorer le contraste global de l'image qui a été réduit suite au filtrage DoG. Vu & Caplier ont testé leur méthode de prétraitement sur trois algorithmes de reconnaissance de visage basés sur la PCA, le LBP et les filtres de Gabor et ont produit des taux de reconnaissance atteignant les 100% tout en offrant un temps d'exécution minimal comparé à d'autres algorithmes de la littérature.

3.1.3 Approches bio-inspirées retenues

Au vu des travaux neurophysiologiques essayant de modéliser le mécanisme de la perception des visages par le système visuel humain et dont certains ont été pris en considération dans le chapitre précédent, nous allons, dans ce chapitre, nous intéresser particulièrement aux résultats issus de deux de ces travaux : les premiers sont ceux de Gosselin et Schyns s'articulant autour du choix et de l'importance des régions d'intérêt ainsi que des échelles de décomposition de l'image de visage en vue de la tâche d'identification [Gosselin 2001] [Schyns 2002]. Les seconds sont les fruits de travaux réalisés par Plantier et al. sur une méthode de renforcement du contenu informatif de l'image qui est une dérivée du codage par ordre de classement et que nous avons appliquée en tant que prétraitement sur les images de visage en vue d'une tâche d'authentification. Dans le souci d'avoir une modélisation imitant d'avantage le comportement du système visuel humain, les filtres Différence de Gaussiennes – DoG sont utilisés dans les deux modélisations choisies, constituant ainsi le troisième aspect bio-inspiré adopté dans la suite de ce chapitre.

3.2 Reconnaissance bio-inspirée par LBP/DoG

Bien qu'il n'existe pas de relation entre les motifs binaires locaux – LBP et le travail du système visuel humain, cette méthode est souvent utilisée en reconnaissance faciale [Huang D. 2011]. La première contribution de ce chapitre fait usage des LBP suite à un filtrage multirésolution par DoG. Les poids des régions d'intérêt du visage et des échelles de décomposition sont déterminés à partir des résultats des travaux neuroscientifiques de Gosselin et Schyns [Gosselin 2001] [Schyns 2002].

3.2.1 Concepts théoriques de la méthode LBP/DoG

Cette section est dédiée aux concepts théoriques des méthodes individuelles composant la méthode bio-inspirée LBP/DoG proposée. Elle décrit d'abord les motifs binaires locaux – LBP et ensuite la différence de Gaussiennes – DoG.

a) Motifs binaires locaux - LBP

Les opérateurs LBP introduits à l'origine par Harwood et al. [Harwood 1993] n'ont été réellement popularisés que trois ans plus tard par Ojala et al. pour l'analyse de la texture [Ojala 1996]. Ils sont ensuite devenus une technique très populaire en reconnaissance faciale [Huang D. 2011]. Il s'agit d'un descripteur non paramétrique simple mais efficace dont le principe général est de comparer le niveau de luminance d'un pixel avec les niveaux de ses voisins. Chaque pixel d'une image est étiqueté en seuillant un voisinage de taille (3×3) par rapport à la valeur du pixel central. Les résultats conduisent à des codes binaires appelés Motifs Binaires Locaux ou *Local Binary Patterns* – LBP (cf. **Figure 3. 13**).

Plus tard, Ojala et al. ont créé une extension du LBP de base permettant d'établir n'importe quelle valeur de rayon et de nombre de points du voisinage [Ojala 2002]. La notation $LBP_{P,R}$ dénote un LBP avec un voisinage de P points également espacés sur un cercle de rayon R (cf. Figure 3. 13). Cela peut être exprimé pour un pixel central $C(x_c, y_c)$ par l'équation (3. 4).



Figure 3. 13 Trois exemples de voisinages pour le calcul de motifs LBP avec (P; R) valant (8; 1), (12; 2,5) et (16; 4) en partant de la gauche vers la droite.

$$LBP_{P,R}(x_c, y_c) = \sum_{i=0}^{p-1} 2^i \delta(g_i - g_c)$$
(3.4)

i parcourt les P voisins du pixel central *C*, g_c et g_i sont respectivement les valeurs de niveaux de gris du pixel central et de ses voisins situés à un cercle de rayon *R*. Les pixels voisins sont de coordonnées $Rsin(2\pi p/P)$, $Rcos(2\pi p/P)$) si le pixel central *C* est de coordonnées (0, 0) (la valeur de niveau de gris des points du voisinage qui ne coïncident pas exactement avec un pixel réel est calculée par interpolation bilinéaire). $\delta(x)$ est égale à 1 si *x* est positif et 0 sinon. La **Figure 3. 14** montre un exemple de calcul d'un LBP de base.



Figure 3. 14 Exemple de calcul d'un LBP de base sur un voisinage (P=8; R=1).

Après avoir étiqueté une image par un opérateur LBP, son histogramme est ensuite calculé pour donner lieu à un descripteur de texture. Un opérateur $LBP_{P,R}$ produit 2^{P} motifs binaires différents à partir des P pixels de voisinage. Ceci conduit à un ensemble de 256 motifs dans le cas du LBP de base (P = 8; R = 1).

Depuis son utilisation en biométrie par le visage, la méthode du LBP a été développée avec une multitude de variantes afin d'améliorer sa performance dans l'analyse de l'image faciale [Huang D. 2011]. Ces variantes traitent différents aspects du LBP de base. Les plus connues d'entre elles sont les LBP à motifs uniformes (LBP^{U2}), les LBP invariants à la rotation (LBP^{RI}) et les LBP à motifs uniformes invariants à la rotation (LBP^{RIU2}) que nous allons détailler ci-dessous.

i) LBP invariants à la rotation (LBP^{RI})

Il s'agit d'une variante introduite par [Pietikäinen 2000] pour compenser l'effet de la rotation et selon laquelle, plusieurs motifs sont considérés identiques si l'un peut être obtenu par la rotation de l'autre suivant un certain angle. Par exemple, *11001011*, *11100101...10010111* sont considérés identiques et sont représentés par le motif *001011111* correspondant à la plus petite valeur d'entre eux (c'est-à-dire 47). Dans le cas d'un voisinage de 8 points (P = 8), le motif LBP^{RI}_(8,R) invariant à la rotation donne lieu uniquement à 36 classes différentes (contre 256 dans le cas du LBP_(8,R)).

ii) LBP à motifs uniformes(LBP^{U2})

Un LBP est désigné comme étant uniforme s'il contient, au plus, deux transitions de bits (0-1 ou 1-0) dans le code binaire circulaire [Mäenpää 2003]. Par exemple, 11111111 (0 transitions) et 01100000 (2 transitions) sont uniformes et représentent chacun une classe séparée. D'un autre côté, 11001001 (4 transitions) et 01010011 (6 transitions) sont non uniformes et sont groupés dans une même classe dite « de divers » (miscellaneous). Dans le cas de 8 voisins, cette répartition LBP^{U2} conduit à 59 classes (58 uniformes + 1 non uniforme) contre 256 dans le cas du LBP de base.

Des études récentes [Huang D. 2011] ont montré que plus que 90% des motifs discriminants dans la représentation de visages sont uniformes.

iii) LBP à motifs uniformes invariants à la rotation (LBP^{RIU2})

Cette variante de LBP obéit aux deux définitions précédentes à la fois [Ojala 2002]. Le code est calculé simplement par le comptage des « uns » dans le nombre binaire. Tous les autres motifs sont groupés dans la classe de divers et réduits à une seule valeur. Par exemple, 00111000 (2 transitions) et 10000011 (2 transitions) sont des motifs LBP^{RIU2} et sont représentés par la même classe « 3 », 00111110 (2 transitions) est un motif LBP^{RIU2} et est représenté par une autre classe « 5 »; tandis que 01000010 (4 transitions) et 00100001 (4 transitions) sont invariants à la rotation mais non uniformes et appartiennent alors à la classe de divers.

b) Différence de Gaussiennes - DoG

Les DoGs ont été introduites initialement comme une représentation mathématique de la forme des champs récepteurs des cellules ganglionnaires de la rétine [Rodieck 1965] (cf. § 3.2.1-b). Le profil de luminance de la DoG issu de la différence de deux Gaussiennes est montré sur la Figure 3. 15.



Figure 3. 15 Réponse fréquentielle d'une Différence de Gaussienne – DoG.

Dans le plan spatial, aussi appelé plan de l'image, la fonction de la DoG est donnée par l'équation (3.5).

$$DoG(x,y) = C_1 \exp\left(-\frac{x^2 + y^2}{2\pi a^2}\right) - C_2 \exp\left(-\frac{x^2 + y^2}{2\pi \sigma^2 a^2}\right)$$
(3.5)

Où x et y sont les coordonnées du pixel d'une image de taille $M \times M$; a est l'échelle de la DoG, la fréquence maximale des filtres étant proportionnelle à 1/a; σ représente la différence d'étendue spatiale entre les deux gaussiennes et la valeur choisie (telle que σ^2 = 2,25) correspond aux données neurophysiologiques sur les réponses des champs récepteurs des cellules ganglionnaires selon des travaux antérieurs sur la modélisation du système visuel humain [Schor 1983]. Les valeurs $C_1 = 1,8$ et $C_2 = 0,8$ sont déterminées de telle façon que la transformée de Fourier des DoG soit nulle pour la fréquence spatiale nulle (u = v = 0).

Pour passer d'une échelle à l'échelle suivante, une multiplication par 2 de l'échelle initiale permet l'augmentation octave par octave. En pratique, l'intervalle des échelles de décomposition dépend de la taille de l'image. Etant donné une échelle initiale a_1 , le nombre total d'échelles m_t est donné par l'équation (3. 6). Les échelles de la DoG sont des puissances de 2 appartenant à l'intervalle $[a_l, M/2]$, où a_1 peut être inférieur à 1. La valeur de l'échelle a_m relative à la $m^{ième}$ ondelette est donnée par **l'équation (3. 7)**.

$$m_t = \log_2(M/a_1)$$
 (3.6) $a_m = 2^{\log_2(a_1)+m-1}$ (3.7)

Ici, nous faisons un calcul direct dans le plan de Fourier en utilisant la formule analytique de la DOG (cf. équation (3. 8)).

$$DoG(u,v) = K \left[exp\left(-\frac{2(\pi a)^2}{\left(\frac{u}{l}\right)^2 - \left(\frac{v}{h}\right)^2} \right) - exp\left(-\frac{2(\pi a)^2}{\sigma^2 \left(\left(\frac{u}{l}\right)^2 - \left(\frac{v}{h}\right)^2 \right)} \right) \right]$$
(3.8)

Avec K = -1 une constante utilisée pour la normalisation des ondelettes dans le plan de Fourier, et ($h \times l$) la taille de l'image.

3.2.2 Méthode LBP/DoG

Depuis 2005, plusieurs études en biométrie faciale ont été menées avec succès en utilisant les opérateurs LBP [Huang D. 2011] réputés pour leur simplicité de calcul, leur tolérance aux changements d'éclairage et l'absence d'apprentissage. En voulant tirer profit des avantages des LBP d'un côté, et ayant pour but d'intégrer l'aspect bio-inspiré apporté par les DoG d'un autre côté, nous avons suivi la même stratégie de combinaison de méthodes entreprise dans le chapitre précédent en appliquant les LBP sur des images filtrées par des DoG.

Les différentes étapes de la méthode sont comme suit. Chaque image normalisée de taille (64×64) est décomposée par les filtres DoG selon plusieurs bandes de fréquence comme montré sur la **Figure 3. 16**. Le calcul se fait directement dans le plan de Fourier selon l'expression analytique de l'équation (3. 8) pour les différentes échelles.



Figure 3. 16 Exemple de représentation de quatre bandes de fréquence contenant chacune une octave d'une décomposition par DoG.

La Figure 3. 17-b montre l'exemple d'une telle décomposition pour une image normalisée de la base IV². Une fois les images filtrées, elles sont étiquetées par un opérateur LBP puis subdivisées en des sous-blocs non chevauchants (cf. Figure 3. 17-c). Ensuite, un calcul d'histogramme se fait pour chaque sous-bloc étiqueté et tous les histogrammes résultants sont concaténés en un seul histogramme constituant le vecteur caractéristique final (cf. Figure 3. 17-d). Cette représentation permet de capturer à la fois la texture locale – à travers l'utilisation des LBP et de la subdivision en blocs – et la forme globale du visage – à travers la concaténation des histogrammes.



Figure 3. 17 Exemple de décomposition par filtre DoG : (a) image originale, (b) images décomposées sur quatre échelles, (c) images LBP^{U2} de (b) subdivisées en blocs, (d) vecteur caractéristique final issu de la concaténation des histogrammes des blocs des images de (c).

A la fin, la comparaison entre deux histogrammes H_1 et H_2 de deux images de visage candidates se fait par la distance du χ^2 (khi deux). Cette mesure de similarité qui a été initialement proposée par Karl Pearson en 1900 a été fréquemment utilisée depuis, dans le domaine de la reconnaissance de visage [Huang D. 2011]. Sa formule est donnée par **l'équation (3. 9)**, *i* étant l'index d'éléments des histogrammes H_1 et H_2 .

$$\chi^{2}(H_{1}, H_{2}) = \sum_{i=0}^{255} \frac{(H_{1i} - H_{2i})^{2}}{H_{1i} + H_{2i}}$$
(3.9)

La méthode a été testée en faisant varier plusieurs paramètres dont certains sont liés au filtrage par DoG et d'autres, aux opérateurs LBP. Ces paramètres sont l'échelle initiale de la DoG, le nombre d'échelles NBO de la DoG (cf. **équation (3. 6)**), le motif LBP et la taille des blocs des histogrammes LBP.

L'influence de ces paramètres est illustrée par les **Table 3. 1** à **Table 3. 3** et la **Figure 3. 18**. Les meilleurs résultats sont surlignés en gras, et les paramètres retenus en vert.

Table 3. 1 Influence de la valeur de l'échelle initiale de la DoG sur la méthode LBP/DoG (LBP, NBO = 4, bloc (16×16)).

	Exp1	Exp2	Exp3	Exp4
a ₁ = 0,25	23,2 (±1,4)	33,3 (±1,6)	32,3 (±1,8)	33,1 (±1,8)
a ₁ = 0,5	3,9 (±0,6)	18,5 (±1,3)	8,6 (±1,1)	18,9 (±1,5)
$a_1 = 1$	3,3 (±0,6)	19,2 (±1,3)	5,7 (±0,9)	19,0 (±1,5)
$a_1 = 2$	3,5 (±0,6)	21,7 (±1,4)	5,8 (±0,9)	20,7 (±1,6)
$a_1 = 4$	3,9 (±0,6)	23,6 (±1,4)	6,1 (±0,9)	21,9 (±1,6)



Figure 3. 18 Influence du nombre d'échelles de la DoG sur la méthode LBP/DoG (LBP, bloc (16×16) , a1 = 0,5).

Table 3. 2 Influence du motif LBP sur la méthode LBP/DoG (NBO = 5, a₁ = 0,5, bloc (16×16)).

	Exp1	Exp2	Exp3	Exp4
LBP	3,5 (±0,6)	19,2 (±1,3)	6,3 (±1,0)	19,4 (±1,5)
LBP ^{RI}	6,0 (±0,8)	21,5 (±1,4)	17,0 (±1,5)	21,1 (±1,6)
LBP ^{U2}	3,2 (±0,6)	18,9 (±1,3)	6,1 (±0,9)	18,3 (±1,5)
LBP ^{RIU2}	6,3 (±0,8)	21,8 (±1,4)	18,1 (±1,5)	21,4 (±1,6)

Table 3. 3 Influence de la taille des blocs sur la méthode LBP/DoG (LBPU², NBO = 5, $a_1 = 0,5$).

	Exp1	Exp2	Exp3	Exp4
(8 × 8)	4,2 (±0,6)	18,4 (±1,4)	7,4 (±1,1)	18,1 (±1,6)
(16 × 16)	3,2 (±0,6)	18,9 (±1,3)	6,1 (±0,9)	18,3 (±1,5)
(32×32)	3,2 (±0,7)	21,0 (±1,3)	9,0 (±1,0)	20,4 (±1,5)

Les résultats de la **Figure 3. 18** viennent renforcer ceux de [Schyns 2002] selon lesquels l'information utile à la reconnaissance du visage s'arrête au niveau de la cinquième échelle. Les résultats de la **Table 3. 2** en faveur du motif LBP^{U2} confirment ceux de [Huang D. 2011] où il a été établi que la quasi majorité des motifs discriminants dans la représentation de visages (plus de 90%) était uniforme.

Au final, la meilleure configuration de paramètres est obtenue, d'abord, par l'application d'une DoG à 5 niveaux en partant de l'échelle 0,5. La fonction LBP^{U2} uniforme est ensuite appliquée sur les bandes filtrées qui sont découpées après en des blocs de taille (16×16). Les histogrammes sont calculés sur chaque bloc puis concaténés pour former le vecteur caractéristique final.

Une comparaison de la combinaison LBP/DoG face à l'application seule des LBP est fournie dans la **Table 3. 4** pour les quatre expériences IV².

Table 3. 4 Comparation	de la Methode LBP/Dou	J avec la methode LBP.

Algorithme	Exp1	Exp2	Exp3	Exp4
LBP ^{U2}	3,0 (±0,6)	22,3 (±1,4)	4,9 (±0,9)	18,5 (±1,5)
LBP ^{U2} /DoG	3,2 (±0,6)	18,9 (±1,3)	6,1 (±0,9)	18,3 (±1,5)

Ces résultats montrent que le filtrage DOG simplement ajouté à l'algorithme LBP ne présente pas beaucoup d'intérêt (bien qu'il améliore l'authentification dans certains cas jusqu'à 15,2% dans l'expérience 2 en diminuant l'EER de 22,3% à 18,9%). Mais, ce n'est pas le but recherché, car en réalité, plutôt que de procéder par variation des paramètres de la méthode de manière empirique, nous cherchons à nous appuyer sur des expérimentations réalisées dans le domaine neuro-physiologique [Gosselin 2001] [Schyns 2002] visant à comprendre le fonctionnement du cerveau humain dans sa tâche de reconnaissance de personnes ainsi que de leur sexe, âge et leurs émotions.

3.2.3 Méthode LBP/DoG pondérée selon les travaux de Gosselin et Schyns

La méthode proposée [El Ghanaoui 2013] ne reprend pas exactement les travaux de Gosselin et Schyns. En effet, nous avons décidé de remplacer le filtre Laplacien, utilisé par Gosselin et Schyns, par un filtre DoG car il modélise mieux la perception visuelle humaine au niveau de la rétine (cf. § 3.1.23.1.2d)). De plus, les échelles fréquentielles ne seront pas échantillonnées car notre objectif n'est pas de vérifier quelle région est importante pour quelle échelle comme fait par Gosselin et Schyns, mais plutôt d'adopter les résultats de leur travaux indiquant dès le départ les zones du visage et les fréquences contenant les informations les plus importantes à différentes fréquences (et de supposer qu'elles sont les plus discriminantes).

Le premier aspect bio-inspiré traité concerne la prise en compte du choix des régions d'intérêt – ROIs – tandis que le deuxième aspect concerne la prise en compte de l'importance des ROIs et des échelles de la DoG [El Ghanaoui 2013].

a) Sélection des régions d'intérêt

D'après les expériences réalisées par Gosselin et Schyns dans une tâche de reconnaissance faciale, les participants se sont concentrés sur des régions spécifiques à chaque échelle de décomposition du visage. Ces régions concernent principalement les yeux, la bouche, le nez et les zones restantes. A la première échelle, les participants se sont focalisés uniquement sur les yeux et la bouche ; à la $2^{\rm ème}$ échelle, ils ont rajouté le nez ; aux $3^{\rm ème}$ et $4^{\rm ème}$ échelles, les participants ont utilisé toutes les régions du visage dans leur tâche de reconnaissance. Ces expériences ont aussi permis d'écarter la $6^{\rm ème}$ échelle considérée comme fond et de conclure sur l'inefficacité de la $5^{\rm ème}$ échelle dans la tâche de reconnaissance.

La sélection des régions d'intérêts en fonction des échelles de décomposition se fait après le filtrage DOG et avant l'extraction des caractéristiques par l'operateur LBP uniforme. Elle est illustrée par la **Figure 3. 19** où les yeux sont représentés en jaune, le nez en rouge, la bouche en vert et le reste en gris. Les zones blanches ne sont pas prises en considération dans les échelles correspondantes.



Figure 3. 19 Sélection des régions d'intérêt sur l'image du visage selon les échelles de décomposition.

L'étiquetage des régions d'intérêt relatives aux yeux, nez, bouche et reste des zones a été testé selon plusieurs configurations dont certaines sont montrées dans la **Figure 3**. **20**. La configuration la plus à droite est celle qui donne les meilleurs résultats car elle assure une meilleure couverture des organes ciblés.



Figure 3. 20 Exemples de configurations pour l'étiquetage des ROI dans l'image de visage.

b) Pondération des ROIs et des échelles

La seconde application de l'approche bio-inspirée dans la méthode proposée consiste à pondérer l'importance de chaque ROI et de chaque échelle. Comme vu précédemment (cf. § 3.1.2e)), Gosselin et Schyns ont montré que les régions d'intérêts essentielles pour l'identification étaient dans l'ordre les yeux, la bouche et le nez. Par conséquent, un poids peut être défini pour chaque région en donnant plus d'importance aux yeux et à la bouche, en limitant le poids du nez qui n'a qu'un rôle secondaire, et en donnant un très faible poids au reste du visage. Mais ils ont aussi introduit la pondération des échelles en fonction de la quantité d'information qu'elles portent, comme illustré dans la **Figure 3**. 12 où la troisième échelle était la plus discriminante suivie de près par la quatrième, les deux premières échelles ayant moins d'influence sur la tâche de reconnaissance.

Pour tenir compte des poids dans le processus, nous avons proposé d'introduire des coefficients de pondération dans la formule de la distance χ^2 comme proposé aussi par López [López 2010]. La nouvelle formule de la distance χ^2 pondérée est donnée par l'équation (3. 10).

$$\chi^{2}_{\text{pondérée}}(H_{1}, H_{2}) = \sum_{s} w_{s} \sum_{r} w_{r} \chi^{2}(H_{1}, H_{2})$$
(3. 10)

Où w_r et w_s sont respectivement les poids de la région r et de l'échelle s ($r \in \{région_yeux, région_nez, région_bouche, autres_régions\}$ et $s \in \{1, 2, 3, 4\}$). L'affectation des poids a été établie d'une manière expérimentale comme indiqué sur la **Table 3.5**.

	N° échelle s	1	2	3	4
	Valeur échelle a	0,5	1	2	4
	Poids échelle w _s	0,05	0,1	0,4	0,45
Wr	Région_yeux	0,7	0,65	0,6	0,6
gion	Région_bouche	0,3	0,23	0,2	0,2
ds ré	Région_nez	-	0,12	0,12	0,12
Poi	Autres régions	-	-	0,08	0,08

 Table 3. 5 Poids des régions et des échelles dans la méthode LBP/DoG bio-inspirée.

En réalité, le calcul de la distance χ^2 pondérée entre deux histogrammes est équivalent au calcul de la distance χ^2 originale entre deux histogrammes pondérés suivant l'échelle et la région. Ceci se fait en affectant au préalable, les poids w_s et w_r à chaque histogramme élémentaire calculé sur un bloc donné, selon la région et l'échelle auxquels il fait référence.

Dans la pratique, cette solution est adoptée car elle est plus facile à implémenter. Elle consiste à stocker les vecteurs caractéristiques préalablement pondérés (composés de mini-histogrammes pondérés et concaténés) qui seront ensuite comparés par la distance χ^2 originale. La **Figure 3. 21** illustre les différentes étapes de la méthode finale LBP/DoG bio-inspirée en commençant par le filtrage DoG, la sélection des ROIs, l'application de l'opérateur LBP^{U2}, le calcul des histogrammes sur les blocs, l'affectation des poids aux histogrammes élémentaires et enfin la concaténation des histogrammes résultants en un seul histogramme final pondéré constituant le vecteur caractéristique. Dans la phase de test, les vecteurs caractéristiques sont comparés par la distance χ^2 originale [El Ghanaoui 2013].

c) Résultats expérimentaux et discussion

Dans ce paragraphe, nous étudions d'abord, l'influence de l'intégration des aspects bio-inspirés traités sur la performance de la méthode proposée. Ensuite, nous la comparons avec les méthodes de la campagne d'évaluation IV² que nous avons pris comme référence depuis le chapitre précédent.

Au cours de nos expériences, nous avons lancé, dans un premier temps, une série de tests dans le but de prouver l'importance de la sélection des ROIs sur les différentes échelles de la DoG plutôt que de considérer la totalité des blocs filtrés. Les résultats des deux premières lignes de la **Table 3. 6** montrent une amélioration significative des taux d'erreurs au niveau de toutes les expériences IV^2 avec un gain en EER de 22,5%, 15,7%, 33,7% et 14,3% respectivement aux expériences 1 à 4 (en réalisant 3,1%, 15,6%, 5,7%, 16,2%) par rapport aux taux d'erreurs initiaux sans sélection des ROIs (qui étaient à 4,0%, 18,5%, 8,6% et 18,9%).



Figure 3. 21 Etapes de la méthode LBP/DoG bio-inspirée : (a) image originale, (b) décomposition DoG sur quatre échelles, (c) sélection des ROIs, (d) application LBP^{U2} et subdivision en blocs, (e) calcul des histogrammes par blocs, (f) affectation des coefficients de pondération w_s et w_r , (g) concaténation des histogrammes pondérés, (h) distance χ^2 entre histogrammes H1 et H2.

Dans un second temps, nous avons investigué un autre aspect bio-inspiré qu'est l'effet de la pondération dans la distance χ^2 . La **Table 3. 6** lignes 2 à 5 illustre l'évolution des résultats en termes d'EER pour les quatre stratégies de pondération testées, à savoir : l'absence de pondération, la pondération des échelles, celle des ROIs, et enfin la pondération simultanées des ROIs et des échelles.

 Table 3. 6 Influence de l'approche bio-inspirée (sélection ROIs + pondération) sur la méthode

 LBP/DoG

Algorithme	Exp1	Exp2	Exp3	Exp4
LBP/DoG sans sélection des ROIs	3,2 (±0,6)	18,9 (±1,3)	6,1 (±0,9)	18,3 (±1,5)
LBP/DoG + sélection ROIs	$3,1 (\pm 0,7)$	15,6 (±1,2)	5,7 (±0,9)	16,2 (±1,4)
LBP/DoG + sélection ROIs + pondération échelles	3,3 (±0,6)	16,1 (±1,2)	5,7 (±0,9)	16,8 (±1,5)
LBP/DoG + sélection ROIs + podération ROIs	3,3 (±0,6)	14,7 (±1,2)	7,5 (±1,0)	14,7 (±1,4)
LBP/DoG + sélection ROIs + pondération (échelles et ROIs)	3,0 (±0,6)	14,9 (±1,2)	6,1 (±0,9)	14,8 (±1,4)

La **Table 3. 6** montre que les meilleurs résultats sont surtout observés avec les deux dernières configurations. Bien que la configuration alliant la sélection et la pondération des ROIs réalise les meileurs EER dans les expériences 2 et 4, elle dégrade toutefois les résultats des autres expériences par rapport à la configuration initiale (sans sélection des ROIs ni pondération), surtout l'expérience 3 dans laquelle la dégradation atteint les 23% d'EER. La dernière configuration alliant la sélection des ROIs plus la pondération à la fois des échelles et des régions donne alors le meilleur compromis pour l'ensemble des expériences. Elle produit les meilleurs résultats dans l'expérience 1 et parvient à réduire les taux d'EER de 21,2% et de 19,1% respectivement dans les expériences 2 et 4 (toujours par rapport à la configuration initiale), avec des résultats identiques dans l'expérience 3.

Pour évaluer la méthode LBP/DoG bio-inspirée finale, une comparaison face aux algorithmes de la campagne IV² [Mellakh 2009-b] sur la même base d'images et selon le même protocole de test est présentée dans la **Table 3.7**.

_					
	Participants	Exp1	Exp2	Exp3	Exp4
	PCA1	6,7 (±0,8)	20,7 (±1,3)	20,1 (±1,6)	22,2 (±1,6)
	PCA2	7,3 (±0,8)	21,6 (±1,4)	13,6 (±1,4)	16,3 (±1,4)
Campagne IV ²	mod PCA	5,3 (±0,7)	20,7 (±1,4)	19,5 (±1,6)	20,5 (±1,5)
	LDA1	3,7 (±0,6)	22,5 (±1,4)	21,7 (±1,7)	19,7 (±1,5)
	LDA2/Gabor	4,2 (±0,6)	12,0 (±1,1)	8,3 (±1,1)	11,3 (±1,2)
Méthode bio-inspirée proposée	LBP/DoG bio	3,0 (±0,6)	14,9 (±1,2)	6,1 (±0,9)	14,8 (±1,4)

Table 3. 7 Résultats comparatifs entre l'algorithme LBP/DoG bio-inspiré proposé (vert) et ceux de la compagne d'évaluation IV² (noir).

D'après les résultats de la table ci-dessus, l'algorithme proposé [El Ghanaoui 2013] donne les meilleures performances au niveau de l'expérience 1 comprenant de petites variations d'expressions, mais aussi au niveau de l'expérience 3 incluant des variations de qualité. Il est positionné au second rang dans les expériences 2 et 4 face à des changements d'éclairage et de sessions. Il est clair que la méthode proposée dépasse les méthodes classiques telles que la PCA et la LDA au niveau de toutes les expériences. En plus, comparée à la PCA modulaire qui fait également référence à la notion de régions d'intérêt, notre méthode est aussi plus performante sur toutes les expériences.

D'un autre côté, il est vrai que les méthodes LBP/DoG bio-inspiré et LDA/Gabor procèdent toutes les deux par analyse multi-résolution et se basent sur le principe de combinaison de méthodes. Toutefois, il convient de souligner que, contrairement à l'algorithme LBP/DoG bio-inspiré n'incluant aucun prétraitement des images, l'algorithme LDA/Gabor applique sur les images, avant d'extraire les caractéristiques, un lissage anisotropique réputé par son efficacité face à des variabilités de type éclairage et multisession. De même, l'algorithme LDA/Gabor comprend une phase d'apprentissage permettant d'obtenir l'espace de projection (*Fisherfaces*) qui sera ensuite utilisé dans la phase de test ; ce qui peut expliquer ses meilleurs résultats pour les expériences 2 et 4. La LBP/DoG, rappelons-le, travaille sans aucune étape d'apprentissage.

3.3 Reconnaissance faciale bio-inspiré par Reconstruction Asynchrone – RA

Dans cette partie, nous présentons la deuxième contribution de ce chapitre : renforcer le contenu informatif des images de visage en les prétraitant par une méthode appelée « Reconstruction asynchrone » – RA qui s'apparente à la famille du codage par ordre de classement, initialement proposé par Thorpe et Gautrais et dont le principe est décrit dans § 3.1.2f) [Thorpe 1998].

Le codage par ordre de classement par reconstruction asynchrone, que l'on propose d'utiliser comme prétraitement des images de visage, considère que l'information la plus importante est supposée être liée aux coefficients les plus élevés obtenus suite à la décomposition en ondelettes. Voilà pourquoi, la reconstruction de l'image est effectuée après le classement des coefficients d'ondelettes par ordre décroissant. Ainsi, le coefficient d'ondelette maximal, qui est classé au premier rang en reconstruction asynchrone, est supposé correspondre au premier neurone qui s'active en traitement visuel.

3.3.1 Concepts théoriques de la RA

Plusieurs expériences en psycho-vision ont été réalisées par [Plantier 1992] dans le laboratoire de recherche CERMA du Service de Santé des Armées françaises. Ces expériences ont soulevé des différences dans le comportement du système visuel humain face à une tâche donnée. Plus encore, en partant de l'hypothèse que, depuis la rétine, les cellules qui s'activent en premier sont celles dont les entrées ont les contrastes de luminance prépondérants, Plantier et al. ont conduit récemment de nouvelles expériences visant à évaluer les performances visuelles lors de tâches de détection et de reconnaissance de véhicules militaires et agricoles en utilisant des images reconstruites partiellement de manière « asynchrone » en se basant sur la notion de maximum de contraste de luminance [Plantier 2012]. Les résultats de l'évaluation n'ont montré aucune différence de performance de détection entre l'image originale et l'image partiellement reconstruite à partir de la carte de fréquences locales pour la tâche de détection. En effet, la suppression d'information à partir d'une image ne réduit pas nécessairement les performances de détection pourvu que les fréquences spatiales contenant le maximum de niveau d'énergie soient gardées dans cette image. Ces résultats confirment l'hypothèse que la reconstruction partielle basée sur les canaux représentant les plus hautes fréquences spatiales permet de conserver les fréquences les plus utiles au système visuel pour réaliser une tâche reconnaissance [Giraudet 2000]. Dans ce qui suit, cette nouvelle méthode de « reconstruction asynchrone » - RA sera présentée. Elle consiste en deux étapes principales : la décomposition et la reconstruction.

a) Décomposition

La décomposition se fait en utilisant la fonction DoG définie précédemment par l'équation (3. 8) à plusieurs échelles, en partant de la plus fine jusqu'à la plus grossière (*fine-to-coarse*). La **Figure 3. 22** présente une image de taille (64×64) décomposée sur 7 échelles ($m \in [1..7]$) en partant de l'échelle la plus fine $a_1 = 0.5$ jusqu'à l'échelle la plus grossière $a_7 = 32$. La valeur du coefficient de normalisation K = 1,7 est fixée en référence aux travaux antérieurs réalisés dans [Plantier 1992] permettant à l'image reconstruite d'être pratiquement égale à l'originale.



Figure 3. 22 Image originale et ses 7 décompositions *fine-to-coarse* après filtrage par une fonction DoG (échelle a_m allant de 0,5 à 32), m étant le niveau de décomposition.

La reconstruction peut par la suite se faire d'une manière totale ou partielle. Une reconstruction totale utilise toutes les échelles obtenues après l'étape de décomposition, tandis que la reconstruction partielle a pour but de garder uniquement l'information utile à la réalisation d'une tâche particulière. Le concept de l'information utile ou discriminante peut être défini en référence à la perception humaine. En effet, dans la phase de reconstruction, il est possible de garder uniquement un nombre de canaux spécifiques en concordance avec les canaux visuels humains. La sélection des canaux gardés ainsi que leur nombre seront étudiés dans la partie suivante.

b) Reconstruction

Une manière classique de reconstruire une image $I_0(x,y)$ est d'opérer une reconstruction coarse-to-fine [Marr 1976], ce qui est une approche globale puisque la transformation est produite de la même manière pour tous les pixels de l'image. Au début, le canal de plus faible fréquence (l'échelle a_l la plus élevée) est utilisé; puis additionné au canal ayant la fréquence immédiatement supérieure, et le processus d'addition des canaux de fréquence $C_m(x,y)_m = 1..mt$ continue selon l'équation (3. 11) permettant ainsi d'obtenir l'image reconstruite I_r au niveau r.

$$I_r = I_{r-1}(x, y) + C_r(x, y)$$
 (3.11)

Avec :

•
$$I_1(x, y) = C(x, y, 1),$$

- $C(x, y, m) = FT^{-1}(FT(I_o(x, y)) * DOGFT_{a_m}(u, v))$, u et v étant les fréquences spatiales le long des axes des abscisses et des ordonnées,
- l'échelle $a_m \in [a_1, M/2]$ et est donnée par $a_m = 2^{m_t \log_2(a_1) m}$,
- le niveau de reconstruction $r \in [1, m_t]$, avec r = 1 lorsque l'on utilise uniquement l'échelle la plus grossière ($m = m_t$) et $r = m_t$ lorsque l'on utilise le nombre total de décompositions.

Le problème d'une telle méthode est que l'information contenue dans les premiers niveaux de reconstruction est très pauvre comme montré sur la **Figure 3. 24**. Pour chaque échelle est également montrée l'image différence entre l'image originale et l'image reconstruite correspondante à cette échelle. Il est remarquable que l'image différence allant jusqu'au 3^{ème} niveau de reconstruction contienne toujours de l'information importante.

D'un autre côté, une image contient une grande quantité d'information répartie sur plusieurs bandes de fréquence. Etant donné que l'information essentielle relative à un pixel donné correspond aux plus grands coefficients de la transformée en ondelettes DoG à travers toutes les échelles de décomposition, la reconstruction asynchrone s'inspire du modèle de codage par ordre de classement en classant les coefficients dans un ordre décroissant. La **Figure 3. 23** illustre en détails le processus de la reconstruction asynchrone avec un exemple montrant la différence avec la reconstruction *coarse-to-fine*. $C(x, y, m)_{m = 1..mt}$ est une matrice tridimensionnelle contenant toutes les décompositions $C_m(x, y)$ d'une image initiale I_0 . Pour chaque pixel, les réponses provenant de tous les canaux décomposés sont rangés dans un ordre descendant, avec rang r = 1 pour la réponse la plus élevée et $r = m_t$ pour la réponse la plus faible. Les réponses maximales sont ainsi représentées par la matrice 3D $C_{sort}(x, y, r)_{-r = 1..mt}$ obtenue après le tri décroissant de $C(x, y, m)_{-m = 1..mt}$ selon sa $3^{\text{ème}}$ dimension. Au début, la réponse la plus élevée aussi appelée « 1^{er} maximum » est gardée pour reconstruire l'image finale, puis elle est additionnée au « second maximum » (c'est-à-dire les réponses venant en seconde place) et le processus continue jusqu'à atteindre un critère d'arrêt. L'image reconstruite $I_r(x, y)$ de manière asynchrone à un rang r est donnée par l'**équation (3. 12)**.

$$I_r(x, y) = I_{r-1}(x, y) + C_{sort}(x, y, r)$$
(3. 12)

Avec :

• $C_{\text{sort}}(x, y, m)_{m=1..m_t} = \text{SortD}(C(x, y, m)_{m=1..m_t})$

- $I_1(x, y) = C_{sort}(x, y, 1)$
- SortD est une fonction qui range les éléments d'une matrice 3D dans un ordre décroissant selon sa 3^{ème} dimension.

La matrice 3D $IndC_{sort}(x, y, m) = 1...mt$ de la **Figure 3. 23**, que l'on appellera « carte des indices locaux », associe à chaque réponse maximale de $C_{sort}(x_i, y_i, m_i)$, l'échelle de décomposition à partir de laquelle elle provient; c'est-à-dire que chaque réponse maximale de $C_{sort}(x_i, y_i, m_i)$ provient de l'échelle de décomposition $IndC_{sort}(x_i, y_i, m_i)$. Ainsi, contrairement à la reconstruction classique qui est une méthode globale, la reconstruction asynchrone est une méthode locale puisque les valeurs finales affectées aux pixels à un niveau m donné de la reconstruction de l'image ne proviennent pas des mêmes échelles de décomposition pour tous les pixels. D'un autre côté, il est vrai que les deux types de reconstruction procèdent pareillement par l'addition au fur et à mesure des canaux de fréquence (équations (3. 11) et (3. 12)). Cependant, la différence réside dans le contenu de ces canaux et donc dans la connotation de la variable r relative au rang de reconstruction : alors qu'en reconstruction coarse-to-fine, r = 1 utilise l'échelle de décomposition la plus grossière ($m = m_t$ dans C) et $r = m_t$ additionne tous les canaux de décomposition en partant de l'échelle la plus grossière jusqu'à arriver à l'échelle la plus fine $(m = m_t \text{ jusqu'à 1 dans } C)$; en reconstruction asynchrone, r = 1 utilise la réponse maximale (m = 1 dans C_{sort}) et $r = m_t$ additionne les canaux asynchrones en partant de la réponse maximale jusqu'à arriver à la réponse minimale (gm = 1 jusqu'à m_t dans C_{sort}).



Figure 3. 23 Schéma fonctionnel de la reconstruction asynchrone basée sur le codage par ordre de classement et comparaison avec la reconstruction *coarse-to-fine*.

Le pseudo-algorithme suivant décrit plus en détails le travail de la reconstruction asynchrone. Après le calcul de l'image centrée I_{center} d'une image initiale I, elle est décomposée par un filtre DoG. Les échelles de décomposition résultantes sont sauvegardées dans C qui est ensuite ordonnée selon sa troisième dimension dans C_{sort} . Puis, la reconstruction asynchrone est opérée en additionnant les canaux de C_{sort} jusqu'à atteindre une certaine condition d'arrêt. A la fin, la moyenne de l'image initiale est rajoutée pour obtenir l'image reconstruite finale I_r .

```
/* Calcul de l'image centrée */
I_{centrée}(x, y) = I(x, y) - moyenne(I(x, y))
/* Décomposition par un filtre DoG */
C(x, y, m) = 1..mt = DoG_decomposition (I_{centrée} (x, y))
/* Reconstruction asynchrone */
C_{sort} (x,y,m) m = 1..mt = SortD(C(x,y,m) m = 1..mt) /* Tri décroissant de C */
r = 1; /* Initialisation */
Iinterm(x,y) = C<sub>sort</sub> (x,y,1) /* Variable intermédiaire */
I1(x,y) = Iinterm(x,y) + movenne(I(x,y)) /* Initialisation */
Tant que (r \le m_t) et non (Critere (I_r))
                 r = r + 1;
                 I_{interm}(x, y) = I_{interm}(x, y) + C_{sort}(x, y, r)
                 I_r(x, y) = I_{interm}(x, y) + moyenne (I(x, y))
/* Le critère d'arrêt peut être par exemple un nombre fixe du rang r<sub>c</sub>
                                                                                         de
reconstruction ou un critère adaptatif comme il sera expliqué dans § 3.3.2 */
end
```

La **Figure 3. 24** montre les différentes étapes de la reconstruction asynchrone rang par rang mise côte à côte avec la reconstruction *coarse-to-fine*. Avec ce principe de reconstruction, dès les premières images de la reconstruction, les principales formes du visage sont visibles et ce, beaucoup plus rapidement que pour la reconstruction *coarse-tofine*. Cette observation est également appuyée par les images différence. Il est par ailleurs à souligner qu'en utilisant la totalité des canaux, les deux types de reconstructions reviennent au même. Ceci veut dire que l'avantage d'utiliser la reconstruction asynchrone est perçu uniquement lors d'une reconstruction partielle, notamment avec un nombre réduit de niveaux de reconstruction.



Figure 3. 24 Exemple comparatif entre reconstruction asynchrone (à droite) et reconstruction *coarse-to-fine* (à gauche) : (a) images reconstruites, (b) images des différences entre les images reconstruites et l'image originale.

3.3.2 Prétraitement bio-inspiré par RA appliqué en reconnaissance faciale

La biométrie est un domaine adéquat pour évaluer l'intérêt de la reconstruction asynchrone. En effet, plusieurs travaux basés sur le prétraitement par les filtres DoG dans ce domaine ont été publiés ces dernières années aussi bien pour des tâches de détection que pour des tâches de reconnaissance. Par exemple, dans [Wang H. 2005], il a été démontré que les caractéristiques sélectionnées en utilisant des filtres DoG décrivent beaucoup mieux le modèle du visage que les caractéristiques simples. D'autres travaux [FanC 2011] [Yuan 2013] ont apporté une amélioration considérable sur les taux de reconnaissance de visage par l'introduction d'une technique de normalisation basée-DoG sur un filtre homographique qui traite plusieurs sortes d'effets d'éclairage.

C'est dans ce sens que nous avons testé la reconstruction asynchrone en tant que technique de prétraitement en vue d'une tâche d'authentification de visage et ce en utilisant une méthode de référence pour la caractérisation des visages : l'Analyse en Composantes Principales – PCA.

Il est à noter que l'application de la reconstruction asynchrone dans la tâche de vérification de visage n'a pas pour but de réaliser les meilleurs résultats comparé à d'autres méthodes de reconnaissance de visages existants ; mais de savoir si cette nouvelle technique de prétraitement améliore les résultats de la vérification pour une méthode donnée. C'est pour cette raison que nous avons choisi la PCA, très utilisée en reconnaissance de visage, comme méthode de référence. Le processus d'évaluation par le test de la PCA sans prétraitement comme référence puis, par le couplage de la PCA à la méthode de prétraitement proposée, a été entrepris dans plusieurs travaux récents. Par exemple, Heseltine a testé plusieurs techniques de prétraitement classiques et a trouvé des EERs variant entre 24 et 60% [Heseltine 2005]. Zhang et Xie ont utilisé des types de prétraitement plus spécifiques et ont trouvé des EERs variant entre 6,5 et 24% avec un meilleur résultat (EER = 6,5%) correspondant à un filtrage DoG [Zhang J. 2012].

L'étude des différents aspects du prétraitement par reconstruction asynchrone a été réalisée sur l'expérience 1 avec 2725 images de la base IV² comme tests de référence, avec des images capturées dans un environnement contrôlé et incluant de légères variations d'expression et de pose.

a) RA avec critère fixe

Dans cette partie, nous allons étudier l'impact du prétraitement par reconstruction asynchrone sur la tâche de vérification de visage comparé à la reconstruction *coarse-tofine* et à deux autres techniques de prétraitement simples basées sur l'amélioration de contraste. La première consiste en un recadrage dynamique qui ajuste les valeurs d'intensité d'une image à niveaux de gris en les transformant dans un intervalle dont les bornes dépendent des valeurs des pixels de l'image elle-même. La borne inférieure est fixée de telle manière que 1% des intensités les plus basses soient saturées à zéro et la borne supérieure est fixée de telle sorte que 1% des intensités les plus élevées soient saturées à 255 (fonction *imadjust* du logiciel MATLAB). La seconde technique est basée sur le fait que le système visuel est sensible à la variation relative de la luminance. Basé sur cette hypothèse, le contraste de luminance d'une image I est relevé dans une image Jen utilisant une fonction logarithmique non linéaire définie par l'équation (3. 13) où les constantes k1 et k2 sont fixées respectivement à 4 et 1 selon les études réalisées dans [Plantier 1992].

$$J(x, y) = Log(k_1 * I(x, y) + k_2)$$
(3.13)

Les résultats des différentes techniques de prétraitement sont reportés dans la Table 3. 8 et montrent que la reconstruction asynchrone améliore les taux de vérification par rapport à l'utilisation de la PCA sans prétraitement d'environ 31% à 41% selon la taille de l'image. De plus, la RA s'avère plus performante que le recadrage dynamique et la transformée logarithmique. D'un autre côté, si l'on considère la reconstruction coarse-to-fine, les meilleurs résultats sont observés aux 2^{ème} et 3^{ème} niveaux de reconstruction où l'information la plus discriminante pour la tâche de vérification est gardée. Les taux d'erreurs subissent une augmentation considérable dans les niveaux de reconstruction suivants.

ille (64 eilleur:	1×64) et s résulta	(128×128). Les échelle ats sont surlignés en jau	es variant respectivemen ine et les seconds meilleu	it de 0,5 à 32 et de 0,5 à urs en bleu.	64. L
_		Image originale	Recadrage dynamique	Transformée Log	
	64×64	8,97 (±0,9)	10,55 (±1,0)	10,04 (±1,0)	
_	128×128	10,54 (±1,0)	9,89 (±1,0)	9,87 (±1,0)	

Table 3. 8 Impact en EER du prétraitement sur la performance de la PCA pour les images de
taille (64×64) et (128×128). Les échelles variant respectivement de 0,5 à 32 et de 0,5 à 64. Les
meilleurs résultats sont surlignés en jaune et les seconds meilleurs en bleu.

128×128	10,54 (±1,0)			9,89 (±1,0)			9,87 (±1,0)			
		Reconstruction coarse-to-fine								
	1 ^{er} niveau	2 ^{ème} niveau	3 ^{ème} niveau	4 ^{ème} niveau	5 ^{ème} niveau	6 ^{ème} niveau	7 ^{ème} niveau	8 ^{ème} niveau		
64×64	10,29 (±1,0)	6,31 (±0,8)	7,07 (±0,8)	7,53 (±0,9)	9,26 (±0,9)	9,08 (±0,9)	9,59 (±1,0)	-		
128×128	12,15 (±1,1)	8,60 (±0,9)	8,40 (±0,9)	9,00 (±0,9)	9,81 (±1,0)	10,25 (±1,0)	10,48 (±1,0)	10,45 (±1,0)		
	Méthode propose : Reconstruction asynchrone									
	1 ^{er} max	2 ^{ème} max	3 ^{ème} max	4 ^{ème} max	5 ^{ème} max	6 ^{ème} max	7 ^{ème} max	8 ^{ème} max		
64×64	6,27 (±0,8)	6,13 (±0,8)	7,39 (±0,9)	8,95 (±0,9)	9,93 (±1,0)	9,26 (±0,9)	9,59 (±1,0)	-		
128×128	5,97 (±0,8)	9,78 (±1,0)	10,24 (±1,0)	9,71 (±1,0)	10,72 (±1,0)	10,58 (±1,0)	10,37 (±1,0)	10,45 (±1,0)		

Par ailleurs, ce qui attire l'attention dans la Table 3. 8 et dans la Figure 3. 25, pour la reconstruction asynchrone, est que le fait de garder uniquement les réponses maximums du prétraitement asynchrone (c'est-à-dire uniquement le premier niveau de la RA) est suffisant pour avoir de bons résultats et même les meilleurs résultats (dans le cas des images (128×128). Ceci appuie l'hypothèse que la réponse maximale du filtre DoG contient déjà l'information la plus essentielle. Cela veut dire que les canaux suivants de la RA incluant les réponses de plus en plus faibles comportent de l'information inutile qui risque de dégrader les résultats de la vérification. L'augmentation des EERs observée au fur et à mesure que l'on rajoute les canaux asynchrones confirme cette déduction.



Figure 3. 25 Evolution de l'EER en fonction des niveaux de reconstruction coarse-to-fine (vert) et asynchrone (rouge) pour (a) des images de taille (64×64) et (b) des images de taille (128×128).

D'après ces résultats, un critère d'arrêt envisageable de la RA serait de fixer dès le départ, le rang de la reconstruction à 1 ; ce qui garantit à peu près d'avoir les meilleurs résultats quelque soit la taille de l'image. La question était ensuite de savoir s'il était possible d'améliorer d'avantage les résultats à travers un processus adaptatif qui permettrait de localiser de manière plus efficace l'information utile à la réalisation d'une tâche spécifique, dans notre cas, la reconnaissance de visage.

b) RA avec critères adaptatifs

La méthode présentée dans le paragraphe précédent a besoin de connaître à l'avance à quel rang la RA devrait s'arrêter, ce qui n'est pas évident à estimer. Pour cette raison, au lieu de fixer à l'avance le nombre de canaux asynchrones à garder pour la reconstruction de toutes les images, l'idée était de procéder par une approche adaptative en cherchant le contenu le plus informatif d'une image. De cette manière, chaque image aurait son propre rang de reconstruction asynchrone adaptable à son contenu informatif et en concordance avec un certain critère adaptatif. Dans ce sens, la variation de trois critères adaptatifs a été explorée : l'écart-type, l'entropie et le pourcentage des contours perdus. Les équations de ces critères sont données par les **équations (3. 14)** à **(3. 17)** suivantes.

$$Critère_STD_m = (STD_o - STD_m)/STD_o$$
(3. 14)

Avec STD_{θ} l'écart-type de l'image originale et STD_m celui de l'image reconstruite au rang m.

Critère_ENT_m =
$$|ENT_o - ENT_m|/ENT_o$$
 (3. 15) $ENT(I) = -\sum_{i=1}^{n} P_i \log_2(P_i)$ (3. 16)

Avec ENT_o l'entropie de l'image originale et ENT_m celle de l'image reconstruite au rang m, n étant le nombre de niveaux de gris dans l'image I (généralement n = 256) et P_i la probabilité du niveau de gris i.

$$Critère_Contours_m = (edge_o - (edge_o \cap edge_m))/edge_o$$
(3. 17)

Avec $edge_0$ les contours dans l'image originale et $edge_m$ ceux de l'image reconstruite au rang *m*, edge étant une fonction du logiciel MATLAB utilisant la méthode de *canny*.

Une fois qu'un critère est choisi, il est calculé à chaque niveau de reconstruction pour chaque image. Ensuite, pour chaque niveau, la valeur moyenne du critère choisi est calculée sur les 2725 images reconstruites et soustraite de la moyenne du critère des images originales. Le comportement du critère introduit à travers les niveaux de reconstruction est montré sur la **Figure 3. 26** par l'évolution de la différence entre les moyennes des critères des images reconstruites et des images originales de taille (128×128) pour les deux types de reconstruction *coarse-to-fine* (bleu) et asynchrone (rouge).



Figure 3. 26 Evolution des critères adaptatifs - (a) écart-type, (b) entropie et (c) contours perdus - en fonction des niveaux de reconstructions *coarse-to-fine* (bleu) and asynchrone (rouge) sur une base de 2725 images.

Toutes les courbes de ce dernier graphique sont décroissantes puisqu'elles reflètent la différence de caractéristiques entre les images originales et les images reconstruites qui converge vers zéro. L'interprétation de la **Figure 3. 26** ainsi que l'étude des critères adaptatifs qui sera présentée dans ce qui suit seront données pour des images de taille (128×128) plutôt que celles de taille (64×64) pour des raisons d'illustration. En effet, doubler la taille des images permet d'augmenter de 1 le nombre d'échelles de la DoG (8 au lieu de 7) qui varient ainsi de 0,5 à 64 selon l'**équation (3. 6)**. Ceci étant dit, l'analyse qui suivra est aussi valable pour les images de taille (64×64).

Les indicateurs de toutes les courbes sont en faveur de la RA qui permet de retrouver beaucoup plus rapidement les caractéristiques de l'image originale. Cette propriété reste vraie aussi bien pour les descripteurs des différences de niveaux de gris tels que l'entropie et l'écart-type que pour les descripteurs d'information locale comme le critère des contours perdus.

Le but était ensuite de chercher le seuil auquel les reconstructions coarse-to-fine et asynchrone devaient s'arrêter, et le niveau auquel l'image reconstruite contenait l'information optimale pour la réalisation d'une tâche donnée. A cette fin, nous avons pris un ensemble de sept valeurs de seuils pour chaque critère comme référence. Plus la valeur du seuil est faible (5%), plus l'image reconstruite est proche de l'image originale (c'est-à-dire sans prétraitement). Inversement, plus la valeur du seuil est élevée (95%), plus l'image reconstruite est différente de l'image originale. Les résultats des expériences sont donnés dans la **Table 3.9** pour les trois critères écart-type (STD), Entropie et Contours perdus. Alors qu'en reconstruction coarse-to-fine les taux d'erreurs décroissent au fur et à mesure que les images reconstruites se rapprochent des images originales, dans la RA, les taux d'erreurs minimaux sont obtenus dès les seuils les plus élevés, indépendamment du critère choisi. En outre, un écart significatif est observé entre les EERs correspondants aux seuils successifs; en particulier pour la reconstruction coarse-to-fine et pour le critère Contours pour la RA. Ce qui est également à remarquer à partir de la **Table 3.9** est qu'il n'existe aucun critère qui décrit au mieux le contenu informatif de l'image lors du prétraitement par RA. En effet, les trois critères produisent les taux minimaux d'erreurs dès le seuil le plus élevé (95%).

			2				
				10.54 (±1.0)			
	Reconstruction coarse-to-fine						
	95%	80%	65%	50%	35%	20%	5%
STD	28,02 (±1,5)	15,24 (±1,2)	8,56 (±0,9)	12,90 (±1,1)	9,21 (±0,9)	9,77 (±1,0)	10,56 (±1,0)
Contours	13,34(±1,1)	9,63 (±1,0)	10,36 (±1,0)	10,25 (±1,0)	10,50 (±1,0)	10,57 (±1,0)	10,45 (±1,0)
Entropie	12,15 (±1,1)	12,15 (±1,1)	15,09 (±1,2)	8,90 (±0,9)	11,51 (±1,0)	8,50 (±0,9)	9,89 (±1,0)

Table 3. 9 Impact du prétraitement adaptatif sur la performance de la PCA (en EER).

	Méthode proposée : Reconstruction asynchrone										
	95%	80%	65%	50%	35%	20%	5%				
STD	5,97 (±0,8)	5,97 (±0,8)	5,97 (±0,8)	9,78 (±1,0)	10,27 (±1,0)	9,67 (±1,0)	10,60 (±1,0)				
Contours	5,97 (±0,8)	12,22 (±1,1)	$6,82 (\pm 0,8)$	10,51 (±1,0)	10,53 (±1,0)	10,55 (±1,0)	9,64 (±1,0)				
Entropie	5,97 (±0,8)	5,97 (±0,8)	5,97 (±0,8)	5,97 (±0,8)	5,97 (±0,8)	9,84 (±1,0)	10,62 (±1,0)				

Pour mieux comprendre les résultats de la **Table 3. 9**, nous avons réalisé une étude plus poussée aux seuils extrêmes des trois critères proposés. Nous avons calculé la répartition des 2725 images prétraitées à travers les différents niveaux de reconstruction. Le calcul des statistiques s'est fait comme suit : le nombre de canaux (ou rangs) nécessaires pour atteindre le critère d'arrêt de la reconstruction (*coarse-to-fine* ou asynchrone) est d'abord calculé pour chaque image. Puis, un histogramme donnant le nombre d'images pour chaque rang est présenté dans la **Table 3. 10**. Il est à rappeler que le premier niveau de reconstruction correspond aux réponses fréquentielles les plus basses dans le cas de la reconstruction *coarse-to-fine*, alors que dans le cas de la RA, cela désigne la réponse maximale du filtre DoG.

	Reconstruction coarse-to-fine						Reconstruction asynchrone						
	Critère	STD Entropie Contours				STD Entropie			Contours				
	Seuils	95%	5%	95%	5%	95%	5%	95%	5%	95%	5%	95%	5%
	EER	28,02 (±1,5)	10,56 (±1,0)	12,15 (±1,1)	9,89 (±1,0)	13,34 (±1,1)	10,45 (±1,0)	5,97 (±0,8)	10,60 (±1,0)	5,97 (±0,8)	10,62 (±1,0)	5,97 (±0,8)	9,64 (±1,0)
Rangs reconst.	1	1100	0	2725	0	2511	0	2725	0	2725	0	2725	0
	2	1625	0	0	0	186	0	0	0	0	1	0	0
	3	0	0	0	0	5	0	0	0	0	2165	0	0
	4	0	0	0	631	22	0	0	0	0	559	0	0
	5	0	0	0	2093	1	0	0	411	0	0	0	0
	6	0	688	0	0	0	0	0	2162	0	0	0	0
	7	0	1958	0	0	0	0	0	133	0	0	0	44
	8	0	79	0	1	0	2725	0	19	0	0	0	2681
Moy(Rangs reconst.)		1,60	6,78	1,00	4,77	1,10	8,00	1,00	5,91	1,00	3,21	1,00	7,98
STD(Rangs reconst.) 0.		0,49	0,48	0,00	0,43	0,38	0,00	0,00	0,47	0,00	0,41	0,00	0,13

Table 3. 10 Reconstruction Asynchrone versus Reconstruction *coarse-to-fine* : histogrammes de la répartition des images sur les rangs de reconstruction nécessaires pour atteindre les seuils extrêmes des critères d'arrêt adaptatifs : *STD*, *Contours* et *Entropie*.

Si l'on considère la reconstruction coarse-to-fine, nous pouvons noter à partir de la **Table 3. 10** que la reconstruction est produite avec un nombre dispersé d'images le long des rangs de reconstruction nécessaires aux critères d'arrêt puisque les écart-types des rangs dépassent les 0,38; excepté pour *Entropie_95%* où toutes les images sont reconstruites en utilisant uniquement l'échelle la plus grossière (ce qui explique les mauvais résultats avec EER = 12,15) et *Contours_5%* où toutes les images sont reconstruites en utilisant la totalité des huit échelles, ce qui explique l'EER très proche de celui des images originales sans aucun prétraitement.

D'un autre côté, les meilleurs résultats de la RA sont obtenus en utilisant uniquement la première réponse maximale de la DoG indépendamment du critère d'arrêt. Par conséquent, engager un processus adaptatif afin de sélectionner l'information la plus pertinente à la vérification de visage s'avère inutile. Ceci reste également valable pour les images de taille (64×64) où la RA peut s'arrêter dès le premier rang quoique les meilleurs résultats soient obtenus au deuxième rang, vu le faible écart entre les deux EERs (6,27% contre 6,13%, cf. **Table 3. 8**).

Aussi, ce qui attire l'attention est le fait de produire les meilleurs résultats aux seuils les plus grands (95%) ce qui paraît contradictoire à première vue car à ces seuils les images reconstruites sont supposées être les plus différentes des images originales. Pour mieux comprendre ce qui est derrière les meilleurs résultats d'authentification en utilisant seulement le premier canal asynchrone en dépit des seuils élevés des critères (95%), nous avons exploré l'origine des coefficients se trouvant sur chaque canal asynchrone, modélisée par la carte des indices locaux de l'image (cf. $IndC_{sort}$ de la **Figure 3. 23**). Nous commençons d'abord par montrer sur la **Figure 3. 27** l'exemple d'une image reconstruite et les différents niveaux de sa carte des indices locaux donnant la répartition des échelles de décomposition sources à travers les rangs de la RA. Les échelles de décompositions sont représentées en niveaux de gris de telle sorte que les pixels colorés du noir au blanc proviennent respectivement de l'échelle la plus grossière $(a_8 = 64)$ à l'échelle la plus fine $(a_1 = 0,5)$.



Figure 3. 27 Représentation en niveaux de gris de la carte des indices locaux $IndC_{sort}$ d'une image reconstruite de manière asynchrone.

Il est clair d'après la **Figure 3. 27** qu'au niveau des premiers canaux asynchrones, les réponses maximales proviennent principalement d'échelles grossières décrivant les formes. Ceci étant dit, il existe un nombre important de pixels provenant des échelles de décomposition les plus fines pour capturer les détails. La proportion des échelles grossières et fines est inversée en allant du premier (réponses maximales) au dernier (réponses minimales) canal asynchrone.

Cette constatation nous a incités à examiner la répartition des échelles de décomposition sur les canaux asynchrones de toutes les images reconstruites de la base (cf. **Figure 3. 28**). Les statistiques de cette figure sont calculées à partir de l'analyse des cartes des indices locaux des 2725 images de la base donnant lieu à 2725×128×128 pixels. Ces statistiques montrent que le premier canal asynchrone, correspondant aux réponses maximales et comportant les données les plus informatives, est composé de trois échelles de décomposition intermédiaires (échelle 16, échelle 8 et échelle 4) contenant environ 85% de l'information; alors que les canaux asynchrones correspondant aux réponses les plus faibles se composent essentiellement d'échelles extrêmes (échelle 0,5 et échelle 64). Le fait d'avoir un nombre réduit d'échelles au niveau du premier canal asynchrone vient appuyer une hypothèse énoncée dans [Watson 2005] après la prise de mesures sur le cortex strié du macaque. Cette hypothèse indique qu'il est possible d'associer uniquement un petit nombre de canaux de fréquences (quatre à six) à chaque emplacement du champ visuel humain.





Maintenant que nous avons analysé les meilleurs résultats de la RA reportés dans la **Table 3. 10**, essayons de mettre l'accent sur ses pires résultats, spécialement avec les critères *STD* et *Entropie* où les EERs vont au-delà des 10%. En effet, il est clair à partir de la **Table 3. 10** que les images sont très dispersées entre les nombres de canaux asynchrones (rangs de reconstruction) nécessaires pour atteindre ces critères puisque les écart-types associés dépassent les 0,4. Ceci revient à dire que les images d'enrôlement et de test à comparer sont susceptibles d'être reconstruites avec un nombre très différent de canaux asynchrones et donc avoir une apparence très différente. Cet inconvénient disparaitrait si le nombre de canaux asynchrones gardés pour la reconstruction était fixé pour toutes les images.



Figure 3. 29 Exemple d'image de la base IV^2 reconstruite de manière asynchrone en utilisant le 1^{er} et le 4^{em} rang de reconstruction avec les images différence correspondantes.

La **Figure 3. 29** illustre des exemples issus d'une image où les images reconstruites sont montrées au premier et au quatrième rang de la RA. Les images différence avec l'image originale sont aussi fournies pour mettre l'accent visuellement sur l'information perdue. L'application du prétraitement asynchrone à l'authentification du visage montre que peu d'information peut être suffisante pour obtenir une bonne reconstruction de l'image à condition que l'information essentielle y soit gardée et l'information inutile éliminée.

En conclusion de l'étude de la RA à critères adaptatifs, nous retenons que les meilleurs taux de vérification de visage sont toujours obtenus lorsque l'on garde uniquement le premier canal asynchrone contenant les réponses maximales indépendamment du critère adaptatif choisi. Ces résultats montrent l'inutilité du processus adaptatif en vue de déterminer le meilleur rang de RA lors d'une tâche de vérification de visage, puisqu'il suffit de travailler avec un rang fixe de RA : r = 1.

3.4 Apports de la RA sur les meilleures méthodes proposées et choix de la méthode unimodale finale du visage

L'étude de la RA réalisée jusque là a utilisé la PCA comme méthode de référence. Le but maintenant est de voir son intérêt sur les deux meilleures méthodes (classique : LDA/PO et bio-inspirée : LBP/DoG bio) proposées dans ce travail de thèse.

Considérons d'abord la méthode LBP/DoG bio-inspirée. Prétraiter les images de manière asynchrone pour les présenter comme entrée à cet algorithme revient à faire une décomposition DoG, puis un tri asynchrone pour ne garder que les réponses maximales situées sur le premier canal asynchrone dans la phase de prétraitement et éliminer toutes les autres réponses ; ensuite à refaire une décomposition DoG pour ne trouver sur chaque échelle qu'un petit nombre de coefficients significatifs correspondant aux réponses maximales que nous avions gardés dans la phase prétraitement. Cette démarche enlèverait tout l'intérêt de l'aspect bio-inspiré de la méthode LBP/DoG puisque cette dernière s'appuie sur les résultats des travaux de [Gosselin 2001] [Schyns 2002] par la méthode des bubbles (cf. § 3.1.2e)) dont le but est de déterminer l'importance des régions d'intérêt du visage et celle des échelles de la DoG et ce en conservant tous les coefficients sur les échelles de décomposition considérées. C'est pour cette raison que l'étude de l'intérêt de la RA va se limiter à la meilleure méthode classique proposée : la LDA/PO.

Une manière envisageable de tirer profit de la RA est de prétraiter les images avant de les passer par l'algorithme de la LDA/PO. Cependant, comme cette méthode inclut déjà un filtrage par la PO, nous avons raisonné de manière analogue à celle de la méthode LBP/DoG bio-inspirée, dans le sens où les coefficients éliminés par le prétraitement asynchrone – qui ne conserve que les coefficients maximums à travers les échelles – enlèveraient certainement de l'information utile à l'analyse multirésolution de la PO. Ce raisonnement est confirmé par des tests de la LDA/PO sur l'expérience 1 comprenant des images capturées dans un environnement contrôlé où les taux d'erreurs ont grimpé de 3,6% à 4,6% une fois que les images en entrée ont été prétraitées par RA.

Une deuxième manière d'exploiter la RA est de fusionner les scores issus de la méthode LDA/PO appliquée sur des images sans prétraitement, avec les scores de la PCA appliquée sur des images prétraitées de manière asynchrone. Cette proposition rejoint l'approche de combinaison entreprise dans le chapitre précédent. De plus, nous avons pensé que la méthode d'analyse locale par pyramide orientable (PO) à travers les filtrages successifs (passe-haut, passe-bas et passe-bandes à plusieurs niveaux) était susceptible d'engendrer une perte d'information au niveau global qui serait utile à la reconnaissance de visage. L'idée était alors d'injecter de l'information globale par l'utilisation d'une méthode d'analyse globale qui est la PCA. La fusion se fait par la suite en sommant les scores de la méthode LDA/PO d'un côté et ceux de la PCA avec images prétraitées de manière asynchrone de l'autre. Les résultats de la **Table 3. 11** montrent l'intérêt de cette fusion pour toutes les expériences.

Table 3.	11	Comparaison	des	meilleures	méthodes	de	reconnaissance	de	visage	proposée	es.
----------	----	-------------	-----	------------	----------	----	----------------	----	--------	----------	-----

Algorithmes	Exp1	Exp2	Exp3	Exp4
LBP/DoG bio	3,0 (±0,6)	14,9 (±1,2)	6,1 (±0,9)	14,8 (±1,4)
LDA/PO	3,6 (±0,6)	17,5 (±1,3)	10,0 (±1,2)	14,5 (±1,4)
LDA/PO + PCA_pretBio	3,2 (±0,6)	16,4 (±1,2)	10,0 (±1,2)	13,9 (±1,3)

La **Table 3. 11** fournit également une comparaison avec la méthode LBP/DoG bioinspirée basée sur la pondération des régions d'intérêts et des échelles selon des travaux dans le domaine neuroscientifique. Cette dernière méthode réalise les meilleures performances sur les trois premières expériences (Exp1 : environnement contrôlé avec petites expressions, Exp2 : variation d'éclairage et Exp3 variation de la qualité). D'un autre côté, l'intérêt du prétraitement asynchrone est surtout perçu au niveau de l'expérience 4 comprenant des images multisession.

Mis à part les avantages de la méthode LBP/DoG bio-inspirée au niveau de la performance, elle présente un autre avantage de taille : l'absence d'apprentissage. Cette propriété est fondamentale lors du choix d'une méthode parmi plusieurs autres, surtout lorsqu'il est prévu de la coupler par la suite à une autre (celle de l'iris). En effet, cela simplifie considérablement le processus des tests comparé à l'utilisation de méthodes nécessitant un apprentissage. Au vu de ces résultats, nous retenons la LBP/DoG bio-inspirée comme méthode unimodale de reconnaissance du visage.

Conclusion

Ce chapitre commence par présenter des résultats neuroscientifiques visant à mieux comprendre le processus de la perception d'images par le cerveau humain. Nous avons vu que les cellules sensibles au contraste dans le cerveau humain sont les cellules ganglionnaires de la rétine et les cellules du corps genouillé latéral qui conduisent l'information au cortex visuel primaire -V1, siège principal de la reconnaissance d'objet.

En particulier, la perception des visages se situe dans l'aire fusiforme des visages de la zone V1.

Par ailleurs, certaines approches utilisées en reconnaissance faciale pourraient modéliser le fonctionnement du cortex visuel. Notamment, le travail des neurones du V1 peut être modélisé par une analyse multirésolution ; l'antagonisme centre-périphérie des cellules ganglionnaires par le filtre DoG; la représentation statistiquement indépendante réalisée au niveau de la rétine par l'algorithme de la PCA; et les interactions de rétroaction entre les différentes aires corticales par un mécanisme adaptatif.

A l'issue de cette étude, nous avons proposé deux approches bio-inspirées de reconnaissance faciale s'appuyant sur des algorithmes émulant le fonctionnement du cerveau humain dans la perception des visages :

La première repose sur les travaux neuroscientifiques de Gosselin et Schyns qui déterminent les régions d'intérêt utiles à l'identification de visages et leur importance ainsi que celle des échelles de décomposition de l'image. L'approche proposée commence par une analyse multirésolution à l'aide de filtres DoG ; ensuite, par le calcul des opérateurs LBP^{U2} uniformes sur les bandes filtrées. Nous avons démontré que la considération seule des régions pertinentes et la pondération à la fois des régions et des échelles selon les résultats de Gosselin et Schyns apportaient une amélioration notable sur les performances de la reconnaissance de visage. Cette méthode a fait preuve de robustesse non seulement dans un environnement contrôlé, mais aussi face à des défis d'éclairage et de qualité. Elle est classée parmi les deux meilleures en comparaison avec la campagne d'évaluation IV².

La deuxième approche bio-inspirée proposée pour la première fois en biométrie faciale, à notre connaissance, fait usage d'une nouvelle méthode de prétraitement appelée reconstruction asynchrone – RA qui s'apparente aux modèles de codage par ordre de classement et qui sert à renforcer le contenu informatif des images [Plantier 2012]. Après la décomposition des images par un filtre DoG, cette méthode reconstruit chaque pixel en utilisant les réponses fréquentielles maximales et en ignorant les échelles des réponses les plus faibles, émulant ainsi le traitement bas niveau du système visuel humain. Ce processus est réalisé au niveau local et rend la méthode très intéressante dans le sens où les différents pixels d'une image reconstruite de manière asynchrone ne proviennent pas des mêmes échelles. La comparaison de la RA avec la reconstruction *coarse-to-fine* et d'autres techniques de prétraitement sur la PCA, comme méthode de référence de reconnaissance faciale, a démontré l'intérêt de la RA.

L'exploration d'un processus adaptatif pour la RA, à travers des critères d'arrêt adaptables au contenu des images (l'écart-type, l'entropie et le pourcentage des contours perdus), a confirmé l'intérêt d'utiliser la reconstruction asynchrone adaptative par rapport à la reconstruction *coarse-to-fine* adaptative dans la récupération des caractéristiques de l'image originale. Les statistiques ont toutefois démontré que les meilleurs taux de vérification étaient toujours obtenus lorsque la RA s'arrêtait au premier rang pour toutes les images, où seules les réponses maximales sont gardées et ce, indépendamment du critère adaptatif. Ces résultats ont alors prouvé l'inutilité d'un processus adaptatif en vue de déterminer le meilleur rang de RA dans le cadre d'une tâche de vérification de visage.

L'intégration de l'aspect asynchrone dans la méthode LDA/PO + PCA proposée dans le chapitre précédent a amélioré les taux de vérification. Cependant, la méthode LBP/DoG bio-inspirée détient les meilleures performances et a l'avantage de ne pas nécessiter de phase d'apprentissage. C'est pourquoi, nous allons la conserver en tant que méthode unimodale finale de reconnaissance de visage en vue de la fusionner, dans le chapitre suivant, avec la méthode Log-Gabor de reconnaissance de l'iris.
Chapitre 4

Biométrie multimodale basée sur l'iris et le visage : Approches et expérimentation

Intro	oduction	.161
4.1	Etat de l'art des méthodes multimodales par fusion de scores basées sur l'iris et le visage	.161
4.2	Structures des bases multimodales IV ² expérimentées	.164
4.3	Méthode bimodale par l'iris droit et le visage	.168
	4.3.1. Normalisation des scores monomodaux	. 169
	4.3.2. Fusion de scores par des méthodes usuelles de l'état de l'art	.170
	4.3.3. Approche de fusion par somme pondérée basée sur le taux d'occultation de la comparaison des iris – ICORws	.174
	 a) Description de la méthode b) Résultats expérimentaux et discussion 	$.174 \\ 176$
	 4.3.4. Impact de l'agrandissement de la base d'apprentissage des paramètres de fusion – BApf 	. 186
4.4	Méthode multimodale par l'iris droit, le visage et l'iris gauche	.189
	4.4.1. Fusion multimodale par les règles usuelles de l'état de l'art	. 189
	4.4.2. Comparaison à la méthode biomodale par l'iris droit et le visage	. 193
	4.4.3. Comparaison à la campagne IV ²	. 194
4.5	Impact du rejet des comparaisons selon le taux d'occultation des iris	.197
Con	clusion	.199

Introduction

La multimodalité offre, à travers la complémentarité des traits biométriques utilisés, de nombreux avantages et, par conséquent, une reconnaissance biométrique plus précise. A partir des méthodes biométriques monomodales basées sur l'iris et le visage que nous avons introduites dans les chapitres 2 et 3, nous allons construire, dans ce chapitre, des méthodes biométriques multimodales à double et à triple échantillons biométriques. Le niveau de fusion adopté étant celui des scores, nous commençons par mener une étude bibliographique sur les méthodes multimodales ayant opéré une fusion des scores de l'iris et du visage. Nous fournirons également les détails caractéristiques de notre corpus de donnée multimodal IV² et de notre protocole de test. Ensuite, nous présenterons notre première approche bimodale basée sur la fusion de l'iris droit et du visage. Nous proposerons dans le cadre de cette premiere approche, une nouvelle règle de fusion de scores basée sur un critère de qualité de l'iris. Après, nous introduirons notre deuxième approche multimodale fusionnant les scores de l'iris droit et gauche avec celui du visage. Enfin, nous étudierons l'influence du rejet des comparaisons en fonction de la qualité des iris sur les deux approches multimodales proposées.

4.1 Etat de l'art des méthodes multimodales par fusion de scores basées sur l'iris et le visage

Cette section est dédiée à la présentation de travaux de l'état de l'art portant sur la fusion de scores entre les modalités de l'iris et du visage. Plusieurs d'entre eux [Wang Y. 2003] [Zhang Z. 2007] [Connaughton 2012] ont utilisé des méthodes de référence dans leurs systèmes unimodaux, plus précisément : la méthode de la PCA en reconnaissance du visage et l'algorithme de Daugman (à des modifications près) en reconnaissance de l'iris.

En 2003, Wang et al. présentent un système bimodal combinant des visages issus de deux capteurs différents et des iris. La fusion se fait ensuite au niveau des scores par trois méthodes. La première est une simple somme pondérée dont les poids sont calculés à partir des performances des systèmes unimodaux. Dans les deux autres méthodes, les scores concaténés de l'iris et du visage sont considérés comme étant un vecteur caractéristique 2D. La deuxième méthode consiste à y appliquer une LDA afin de déterminer un séparatrice linéaire entre les échantillons authentiques et imposteurs. Dans la troisième méthode, un réseau de neurones à noyaux RBF, ayant été entraîné pour chaque sujet, classifie chaque vecteur de scores 2D en tant qu'authentique ou imposteur. Il a été conclu par [Wang Y. 2003] que toutes les méthodes de fusion de scores proposées améliorent les résultats des systèmes monomodaux et que la méthode de fusion par les *NN-RBF* réalise les meilleurs taux de vérification.

En 2006, Shu et Ding proposent une technique de normalisation de scores pour fusionner les scores d'iris, de visage et de signature provenant d'une base chimérique. Cette technique transforme les scores unimodaux en des probabilités *a posteriori* qui

sont fusionnées par des techniques communes de fusion de scores. Les auteurs montrent que leur technique de normalisation dépasse celles du min-max et de Tanh, les meilleures performances étant obtenues avec la fusion par SVM [Shu 2006].

En 2007, Zhang et al. considèrent le problème de la fusion des modalités de l'iris et du visage en utilisant un même capteur proche-infrarouge. Une fois les scores unimodaux obtenus, ils sont normalisés par la règle min-max et ensuite fusionnés par les règles somme et produit. Les auteurs réalisent les meilleures performances par la somme des scores du visage, avec ceux de l'iris droit et de l'iris gauche [Zhang Z. 2007].

En 2008, Morizet et Gilles proposent un système bimodal Iris / Visage qu'ils testent sur une base de données chimérique. Les systèmes unimodaux du visage et de l'iris opèrent respectivement par Log-Gabor/PCA et Log-Gabor. Avant la fusion, les auteurs procèdent d'abord par normalisation zScores (cf. § 1.3.5) puis considèrent les scores imposteurs comme étant du bruit blanc gaussien. Ils y appliquent alors un débruitage par ondelettes pour supprimer les scores imposteurs et renforcer les scores authentiques. La fusion est ensuite opérée par une somme pondérée où les poids sont spécifiques à chaque utilisateur. Ces poids sont déterminés de telle sorte à optimiser une fonction de coût dépendant de la distance de séparation et des coefficients de dissymétrie ("skewness") et d'aplatissement ("kurtosis") des distributions des scores imposteurs [Morizet 2008] [Morizet 2009].

En 2009, Poh et al. organisent une campagne d'évaluation entre plusieurs techniques de fusion de scores basées sur la qualité [Poh 2009]. Dans le souci d'éliminer le biais entre les différents algorithmes unimodaux des participants, les auteurs ont fourni un ensemble de scores biométriques d'iris, de visage et d'empreinte digitale avec un score de qualité d'acquisition associé à chaque échantillon. Les deux meilleurs systèmes parmi 22 participants sont ceux de Maurer et Baker [Maurer 2008] et de Poh et al. [Poh 2007]. Le système soumis par Maurer et Baker possède une architecture de fusion basée sur les réseaux bayésiens. Ces réseaux permettent de définir et de modéliser des dépendances statistiques entre plusieurs variables telles que les scores, les estimations de qualité correspondantes et les variables globales de décision. Cette approche a été initialement testée pour la fusion de données biométriques de voix et d'empreinte digitale, mais a été étendue pour intégrer le visage, l'iris et l'empreinte digitale et pouvoir participer à la campagne d'évaluation. Le système soumis par Poh et al. a opéré une fusion par somme suite à une normalisation basée sur les taux de log-vraisemblance [Maurer 2008].

En 2010, Zhang et al. construisent une nouvelle technique de fusion entre images d'iris et de visages provenant des mêmes sujets. Les images des deux modalités sont acquises par des capteurs différents de qualité relativement faible. Dans la phase d'apprentissage, les auteurs utilisent une analyse de corrélation canonique pour construire une cartographie statistique associant un iris à son visage. Dans la phase de test, les auteurs utilisent la cartographie statistique par régression qui, étant donné les échantillons de visage et d'iris d'un candidat, retourne une liste réduite des correspondances les plus probables d'enrôlement. A la fin, une normalisation min-max et une fusion de scores par somme est réalisée uniquement entre les échantillons du candidat et les échantillons d'enrôlement de la liste réduite [Zhang X. 2010].

En 2012, Connaughton et al. proposent un système d'identification par le visage et l'iris à partir d'un unique capteur Iris On the Move (IOM) en lumière proche infrarouge. Lors de la phase de test, plusieurs échantillons d'un candidat donné sont comparés à tous les échantillons de tous les sujets enrôlés. Suite à une normalisation par la méthode du min-max, [Connaughton 2012] proposent deux méthodes de fusion. La première se fait au niveau rang par la méthode du Borda count, un système de vote pondéré où les scores de toutes les comparaisons d'un échantillon d'un candidat donné sont d'abord triés dans un ordre décroissant (le premier rang correspond alors au meilleur score de cet échantillon). Chaque échantillon du candidat traité attribue des votes pondérés aux v premiers rangs de sa liste, où le poids est inversement proportionnel au rang. Après, le vote de tous les échantillons d'un candidat, le sujet d'enrôlement collectant le plus de votes est élu comme étant la meilleure correspondance avec le candidat traité. La deuxième méthode de fusion se fait au niveau des scores par une somme pondérée. Leurs expériences ont montré que bien que la fusion au niveau rang par la méthode Borda count produisait de bons résultats, c'est la fusion au niveau des scores par la règle somme pondérée qui donne les meilleurs résultats.

En 2013, Wang et al. combinent les traits provenant d'iris gauche et droit, et des visages capturés en lumière visible avec des visages capturés en lumière infrarouge thermique. La caractérisation des iris se fait par les filtres Log-Gabor et celle des visages par une PCA fonctionnelle multi-niveaux. Une fusion des caractéristiques est d'abord opérée entre les iris gauche et droit et entre les visages visible et thermique. Deux vecteurs caractéristiques sont alors obtenus : un pour chacune des modalités iris et visage. Chaque comparaison donne lieu à deux scores unimodaux qui sont ensuite normalisés par la règle min-max et fusionnés par la règle norme triangulaire [Wang N. 2013].

Suite à cette étude bibliographique, nous remarquons que la majorité des systèmes proposés testent leurs méthodes avec les règles usuelles de normalisation et de fusion, soit : min-max, zScores et tanh en normalisation ; et max, min, somme, produit et somme pondérée en fusion. Certains travaux se limitent à ces méthodes usuelles et d'autres les comparent à de nouvelles méthodes qu'ils proposent. Nous allons adopter la même démarche dans la présentation de nos résultats expérimentaux en commençant d'abord par l'exposition des résultats des méthodes usuelles. Les plus pertinentes seront ensuite comparées à la méthode de fusion proposée.

D'autre part, il est important de rappeler que vu le manque d'information sur le protocole d'évaluation multimodal IV^2 et vu la non disponibilité des listes de comparaison multimodales, nous avons été amené à créer notre propre protocole et nos propres listes de comparaison pour tester les approches multomodales proposées. Nous nous sommes alors référés au protocole mis au point par le NIST (National Institute of Standards and Technology) pour générer les scores de similarité et conduire nos évaluations [Snelick 2003]. Le principe général de ce protocole consiste à :

- générer des mesures unimodales de similarité à partir de listes de comparaisons créées depuis des listes de références alignées où les données ont été mises en correspondance afin de créer des individus réels ou virtuels – les mesures de similarité doivent être converties dans un format commun (distances ou scores) –,
- normaliser les mesures de similarité pour les transformer dans un intervalle commun,
- fusionner les listes des mesures unimodales de similarité en une seule liste de fusion multimodale.
- calculer les statistiques de performance à partir des distributions des authentiques et des imposteurs.

4.2 Structures des bases multimodales IV² expérimentées

Dans cette section, nous allons détailler la démarche avec laquelle nous avons créé notre corpus de données à partir des sous-bases d'iris et de visage IV². Lorsque les images d'iris et de visages étaient considérées de manière séparée, la génération des scores à partir des listes de comparaisons des expériences IV² était systématique car un score était attribué entre chaque image de test et chaque image d'enrôlement d'une comparaison donnée de la liste. En revanche, comme les listes de comparaisons des deux modalités de l'iris et du visage étaient produites de manière indépendante, il était impossible de mettre en correspondance les scores issus de chaque modalité pour une comparaison donnée. Il fallait donc créer de nouvelles listes de comparaisons unimodales qui soient alignées, en vue de générer des scores à fusionner et ce, pour chaque expérience IV² (Exp1 : environnement contrôlé avec petites expressions, Exp2 : variation de l'éclairage, Exp3 : variation de la qualité, Exp4 : multisession).

Plus encore, il existe des sujets pour lesquels l'une ou l'autre des modalités est manquante ou alors dont le nombre d'images est insuffisant. Il fallait alors remonter au listes de références unimodales (associant chaque image à son identité) afin d'en générer de nouvelles listes où chaque instance d'image d'iris est mise en correspondance avec une instance d'image de visage de la même personne. Certes, cette démarche réduit de manière sensible la taille des listes de comparaisons par rapport à celles des listes unimodales. Cependant, nous allons voir à la fin de cette partie que les caractéristiques du corpus de données obtenu sont comparables voir meilleures que celles des bases de données multimodales réélles de la littérature.

Avant la mise en correspondance des images de plusieurs modalités, il est important de connaître en détails la constitution de la base d'images dont on dispose. La **Figure 4.** 1 montre les histogrammes relatifs aux distributions du nombre d'images par sujet pour chaque liste de référence des modalités iris et visage de la base IV² (au total 12 listes : listes d'enrôlement et de test pour les expériences iris droit, iris gauche, visage Exp1, visage Exp2, visage Exp3 et visage Exp4).

Ce qui est à remarquer à partir de la **Figure 4.1** est que les sujets disposent en général d'un nombre d'images d'iris beaucoup plus élevé que celui de visage. Nous

sommes alors contraints, lors du choix du nombre d'images par personne dans les listes de références multimodales, de nous référer à la modalité du visage (présentant le moins d'images par personne) lors de la construction de notre base de test réelle.



Figure 4. 1 Distributions du nombre d'images par sujets dans les listes de références originales des modalités de l'iris (à gauche) et du visage (à droite) de la base IV².

Mais avant, il fallait commencer par la construction de la base d'apprentissage chimérique des paramètres de fusion (cf. § 1.4.1a)). Sachant que nous disposons d'un ensemble d'apprentissage disjoint constitué de 52 sujets ayant chacun 3 images de visage, nous avons alors extrait à partir des bases d'iris droit et gauche le même nombre de sujets (ayant chacun 3 images) que nous leur avons assigné de manière arbitraire afin de créer notre base d'apprentissage multimodale chimérique. A partir de cette base, nous avons généré les listes de comparaisons multimodales d'apprentissage dont les caractéristiques sont données dans la **Table 4. 1**. Cette base sera utilisée aussi bien dans l'entraînement de la méthode bimodale Iris droit / Visage que celle de la méthode multimodale Iris droit / Visage / Iris gauche. Il est à rappeler que le partitionnement des données entre base d'apprentissage et de test se fait par personne et non pas par image [Allano 2009] [Morizet 2009]. C'est pourquoi les identités utilisées dans la base d'apprentissage seront éliminées de la base de test.

Par ailleurs, dans le cadre de l'étude de l'influence de l'agrandissement de la base d'apprentissage et du changement de protocole de test sur la performance de la méthode bimodale Iris_Visage (Exp1), nous avons généré des listes de comparaisons où les nombres de comparaisons intra et interclasses sont égales (protocole IV²) et d'autres listes où le nombre de comparaisons interclasses est largement supérieur à celui des comparaisons intra-classes (protocole de Lausanne).

Table 4. 1 Caractéristiques de la base d'apprentissage chimérique des paramètres de fusion et des listes de comparaisons d'apprentissage associées.

	Nb. sujets / modalité	Nb. img / sujet / modalité	Nb. comparaisons intra-classes / modalité	Nb. comparaisons totales / modalité
Nb. comp inter = Nb. comp intra	52	3	156	312
Nb. comp inter >> Nb. comp intra	52	3	156	12090

Une fois la base d'apprentissage créée, nous construisons la base de test en nous référant à la démarche présentée dans [Snelik 2003]. Les étapes sont traduites par le pseudo-algorithme suivant où le principe de la génération des listes de comparaisons est également décrit (cf. **Figure 4. 2**).

I. Création des listes de références uni et multimodales : Entrées : Listes de références originales d'enrôlement et de test des expériences IV² à partir desquelles les identités de l'apprentissage ont été éliminées. 1. Pour chaque modalité : Retenir les identités ayant au moins n nombre d'images / personnes (voir les valeurs prises par n en dans la Table 4. 2) 2. Extraire les identités en communs entre les modalités à apparier, ayant au moins n nombre d'images / personnes. L'intersection doit se faire de manière indépendante entre listes d'enrôlement et listes de test suivant les équations (4.1) et (4.2) où m est nombre total de modalités à fusionner. $Identit\acute{es_enrol_{multi}} = \bigcap_{modalit\acute{e}=1} identit\acute{es_enrol_{uni}}(modalit\acute{e})$ (4.1) $Identit\acute{e}s_test_{multi} = \bigcap_{i=1}^{m} identit\acute{e}s_test_{uni}(modalit\acute{e})$ (4.2)3. Pour chaque identité retenue : Pour chaque modalité : Sauvegarder les noms de n images et leur identité dans les listes de références unimodales alignées Sauvegarder les noms de (n × m) images alignées de toutes les modalités et leur identité dans la liste de référence multimodale finale. Sorties : Listes de références unimodales alignées (enrôlement et test) Liste de référence multimodale (enrôlement et test) II. Création des listes de comparaisons uni et multimodales suivant le protocole IV² (nombre de comparaisons intra-classes : n comp intra = n comp inter : nombre de comparaisons interclasses): Entrées : Listes de références unimodales alignées (enrôlement et test) Liste de référence multimodale (enrôlement et test) 1. Générer les index de toutes les comparaisons possibles entre les images d'enrôlement et les images de test d'une modalité biométrique donnée (protocole de Lausanne). 2. Extraire les index des comparaisons intra-classes à garder dans les listes de comparaisons finales. 3. Extraire les index des comparaisons interclasses et en tirer de manière aléatoire n comp intra comparaisons interclasses à garder dans les listes de comparaisons finales. 4. Concaténer les index des comparaisons intra et interclasses à garder dans les listes de comparaisons finales (protocole IV²). 5. Pour chaque modalité : Générer les couples (image enrôlement, image test) correspondant aux index de l'étape II.4 et les sauvegarder dans les listes de comparaisons unimodales. Sauvegarder les comparaisons entre m images alignées d'enrôlement et m images alignées de test provenant de toutes les modalités et correspondant aux index de l'étape II.4 dans une liste de comparaison multimodale. Sorties : Listes de comparaisons unimodales alignées (protocole IV²) Liste de comparaison multimodale (protocole IV²) Liste de comparaisons unimodales alignées Iris D/Visage Expl (protocole Lausanne)

A l'issue de cet algorithme, nous avons pu générer les bases et les listes de comparaisons de test relatives aux deux approches multimodales expérimentées dans ce chapitre : Iris droit / Visage et Iris droit / Visage / Iris gauche. Il est à rappeler que le nombre de sujets relatif à la fusion Iris droit / Visage / Iris gauche est inférieur à celui de la fusion Iris droit / Visage car certains sujets ne disposent pas des trois traits biométriques à la fois. La comparaison des caractéristiques des corpus de données générés à partir de la base IV² (cf. **Table 4. 2**) à celles des bases de données réelles existantes (cf. **Chapitre 1**, **Table 1. 3**), montre qu'en termes de nombre de sujets, les corpus présentés sont comparables voire dépassent ceux des autres bases.

Expérience	Nb. Img /sujet enrôlement /modalité [n]	Nb. Img /sujet test /modalité [n]	Nb. total sujets retenus /modalité	Nb. total images retenues /modalité	Nb. comparaisons intra-classes /modalité	Nb. comparaisons totales /modalité
		Liste	es bimodales Iris_	_D / Visage		
Exp1	3	3	248	1290	1638	3276
Exp2	3	3	245	1251	1548	3096
Exp3	3	3	227	1083	1206	2412
Exp4	4	5	77	486	580	1155
		Listes mul	timodales Iris_D	/ Visage / Iris_G		
Exp1	3	3	233	1131	1297	2594
Exp2	3	3	229	1095	1225	2450
Exp3	3	3	206	942	973	1946
Exp4	4	5	63	432	441	882

Table 4. 2 Caractéristiques des bases de test réelles expérimentées et des listes de comparaisons de test correspondantes.

L'étape suivante consiste alors à faire passer les listes de comparaisons à travers les méthodes unimodales correspondantes pour générer les scores. Ces scores unimodaux seront ensuite normalisés puis fusionnés pour produire les scores multimodaux finaux. La **Figure 4. 2** montre le processus de génération de deux échantillons de scores multimodaux (un score intra-classes et un score interclasses) à partir des listes de comparaisons unimodales alignées qui ont été générées en suivant les étapes du pseudoalgorithme précédent. Sur cette figure sont entourées d'une même couleur chaque association d'images d'iris et de visage d'une même personne ainsi que leur référence. Les étapes de normalisation / fusion seront décrites en détails dans les sections suivantes.



Figure 4. 2 Illustration du processus de la mise en correspondance d'échantillons de listes de référence unimodales jusqu'à la génération d'échantillons de scores multimodaux Iris_D / Visage.

4.3 Méthode bimodale par l'iris droit et le visage

Cette partie est consacrée à l'étude de la performance de l'approche bimodale combinant les traits de l'iris droit et du visage. Le choix s'étant porté sur la fusion au niveau des scores, une étape préalable de normalisation des scores est nécessaire car les sorties des modules monomodaux risquent de se situer dans des domaines différents.

4.3.1. Normalisation des scores monomodaux

Les distributions des scores authentiques et imposteurs issus des méthodes monomodales sont montrées dans la Figure 4.3.



Figure 4. 3 Distributions originales des scores authentiques et imposteurs des méthodes monomodales de l'iris droit (à gauche) et du visage (à droite).

On remarque que ces distributions sont très différentes en termes de plage de variation (scores iris dans [1,7;7,3] et scores visage dans [0,0021;0,0159]) et de forme. L'étape de normalisation est alors nécessaire pour convertir les scores des deux modalités dans un intervalle commun et être ainsi compatibles dans les étapes subséquentes de la fusion. Nous avons vu dans le premier chapitre que parmi les techniques communes de normalisation, min-max, zScores et Tanh étaient les plus efficaces si elles sont combinées aux méthodes de fusion de scores usuelles – somme, produit, maximum [Jain 2005]. C'est pourquoi nous allons utiliser ces techniques dans la suite de nos expériences.

Les résultats de la normalisation sont montrés dans la **Figure 4. 4**. Comme nous pouvons le constater, les formes des distributions des authentiques et des imposteurs restent inchangées par rapport aux distributions originales de la **Figure 4. 3**. Seuls leurs intervalles ont changé. La normalisation min-max a transformé les scores des deux modalités de l'iris et du visage dans l'intervalle [0 1], la normalisation zScores les a centrés autour de 0 dans l'intervalle [-1,6 ; 5,3] et la normalisation Tanh autour de 0,5 dans l'intervalle [0,48 ; 0,52].



Figure 4. 4 Distributions normalisées des scores authentiques et imposteurs des méthodes monomodales de l'iris droit (à gauche) et du visage (à droite) par les techniques de normalisation Min-Max, zScores et Tanh. Les échelles des scores authentiques et imposteurs sont différentes pour faciliter la visualisation.

4.3.2. Fusion de scores par des méthodes usuelles de l'état de l'art

Dans cette partie, les méthodes de normalisations étudiées dans le paragraphe précédent sont associées aux méthodes communes de fusion de scores qui sont le maximum, le minimum, le produit, la somme et la somme pondérée. La **Table 4. 3** en reporte les résultats pour la combinaison de l'iris droit avec le visage de l'expérience 1 (environnement contrôlé avec variation limitée d'expression) des ensembles IV^2 et ce au point de fonctionnement EER avec l'intervalle de confiance à 90% et au point de fonctionnement FRR lorsque FAR = 1% (avec les meilleurs résultats surlignés en gras). Les résultats des autres expériences seront donnés dans la partie suivante.

Dans la **Table 4. 3** on constate qu'il y a de grandes différences de performances selon les associations de méthodes de normalisation/fusion. Trois classes peuvent être distinguées : la classe « Correcte » (cases en blanc et gris) désigne les associations diminuant l'EER des modules monomodaux de plus de la moitié ; la classe « Moyenne » (cases en bleu) désigne les associations donnant des résultats qui restent dans l'ordre de ceux des modules monomodaux ou les améliorent de peu et la classe « Catastrophique » (cases en orange) relative aux associations qui dégradent sensiblement les résultats monomodaux. En particulier, la méthode produit donne de mauvais résultats lorsqu'elle est associée à zScores (EER = 5,09%). En effet, le produit génère des problèmes lorsque les scores sont soit proches de zéro soit très grands. Dans notre cas, un grand nombre de scores sont proches de zéro suite à la normalisation zScores, ce qui donne de si mauvais résultats. Cependant, une fois associé à la normalisation Tanh, le produit donne de très bons résultats (les meilleurs avec EER = 0.45% et FRR (FAR = 1%) = 0.29%) qui sont identiques à ceux obtenus par l'association Tanh/Somme. Selon [Allano 2009], les performances équivalentes du produit et de la somme pour la normalisation Tanh sont dues au fait que Tanh transforme les scores dans un intervalle de taille très réduite (entre 0,48 et 0,52) et que sur un intervalle aussi petit le produit est quasi-linéaire donc son comportement est quasi-équivalent à celui la somme. La Figure 4.5 montre les distributions des scores authentiques et imposteurs ainsi que les courbes DET relatives à la règle produit associées aux deux normalisations zScores et Tanh donnant respectivement les pires et les meilleurs résultats. Les distributions des scores normalisés des modules monomodaux correspondants sont également montrées pour illustrer les transformations qui ont suivi la fusion. Idéalement, la dististribution des authentiques serait translatée vers la droite et celle des imposteurs vers la gauche, de telle manière à obtenir un recouvrement nul entre les deux distributions. Avec la normalisation Tanh, les deux distributions se sont translatées vers la gauche suite au produit, avec une translation plus prononcée de la distribution des imposteurs de telle sorte que leur recouvrement s'est réduit. Dans la normalisation zScores, le produit a inversé l'emplacement des deux distributions de telle sorte que la distribution des imposteurs se soit trouvée à droite de celle des authentiques, avec un recouvrement plus prononcé que ceux des modules monomodaux. Cette inversion provient du fait que le produit de deux scores imposteurs négatifs donne un score imposteur positif.

		Ľ	EK (70)	$\mathbf{K}(70) \qquad \mathbf{FKK}(\mathbf{FAK} = 170)$		
	Iris_D (Log-Gabor)		75 (±0,77)	4,74		
	Visage_Exp1 (LBP/DoG bio)		69 (±0,77)	7,39		
		Fusion	multimoda	le Iris_D / Visag	ge_Exp1	-
Fus. \ Norm.	Min-Max		zScores		Tanh	
	EER (%)	FRR (FAR = 1%)	EER (%)	FRR (FAR = 1%)	EER (%)	FRR (FAR = 1%
Max	2,68 (±0,66)	5,55	1,06 (±0,42)	1,20	0,83 (±0,37)	0,74
Min	3,90 (±0,79)	4,77	3,48 (±0,75)	5,37	3,70 (±0,77)	4,34
Produit	1,28 (±0,46)	1,33	5,09 (±0,89)	11,20	0,45 (±0,27)	0,29
Somme	1,23 (±0,45)	1,30	1,20 (±0,44)	1,26	0,45 (±0,27)	0,30
Somme	1.12 (+0.43)	1.17	1.01 (+0.41)	1.01	0.66 (+0.33)	0.57

Table 4. 3 Association des méthodes de normalisation aux méthodes de fusion de scores usuellesdans l'approche bimodale Iris_D / Visage_Exp1.

FFD (0/.)

Méthodes monomodales

FDD (**FAD** = 10()

Par ailleurs, la règle de fusion minimum n'améliore pas voire dégrade les performances des modules monomodaux. En effet, en prenant le minimum entre deux scores, si l'un des modules est mauvais, le score authentique va être très faible et risque plus d'être rejeté.

L'apport de la règle maximum est sensible surtout en association avec Tanh et zScores contrairement à l'association avec min-max présentant un FRR de 5,55% au point FAR = 1%, ce qui est du même ordre que celui des méthodes monomodales.

La règle de fusion Somme donne de bons résultats pour les trois normalisations et améliore toujours les résultats par rapport à la meilleure approche monomodale seule (résultats recalculés sur les seules images retenues) en diminuant l'EER d'un taux variant entre 67% et 87,8% et le FRR au point FAR = 1% d'un taux entre 72,6% et 93,7%. On remarque que les normalisations min-max et zScores sont à peu près équivalentes une fois combinées avec la somme (EERs de 1,23% et 1,20%) et que Tanh donne les meilleures performances (EER = 0,45%). Ceci peut s'expliquer par le caractère quasilinéaire de la tangente hyperbolique qui attribue moins d'importance aux scores extrêmes susceptibles de dégrader les performances de la méthode finale.

La somme pondérée consiste à attribuer des poids à chaque modalité. Ces poids sont inversement proportionnels aux taux d'erreurs des méthodes monomodales calculés hors ligne sur la base d'apprentissage et sont donnés par l'équation (4. 3) qui est analogue à l'équation (2. 30) donnant les poids des scores par bande de la méthode Adaboost/PO proposée dans le chapitre 2. ω_j et E_j sont respectivement le poids et l'EER de la modalité j, n étant le nombre total de modalités. Si nous comparons les performances de la somme pondérée à celles de la somme qui est un cas particulier de somme pondérée est meilleure sauf en association avec Tanh où la somme donnait les meilleurs résultats.

$$\omega_{j} = \frac{1}{E_{j}} \qquad \text{avec } U = \sum_{j=1}^{n} \frac{1}{E_{j}}, \sum_{j=1}^{n} \omega_{j} = 1 \text{ et } 0 < \omega_{j} < 1$$
(4.3)

En conclusion, les associations de méthodes de normalisation/fusion n'ont pas toutes le même intérêt. Les associations appartenant aux classes « Catastrophique » et « Moyenne » (cases en orange et en bleu de la **Table 4. 3**) seront écartées de la suite de notre étude. Généralement, la normalisation min-max est la plus sensible aux valeurs aberrantes et donne de moins bons résultats que les autres techniques de normalisation. En revanche, les meilleures associations proviennent généralement d'une normalisation Tanh – dont la robustesse et l'efficacité a été démontrée dans [Jain 2005] – ou d'une normalisation zScores. La somme est une des méthodes de fusion les plus performantes [Kittler 1998] [Ross 2003]. La somme pondérée arrive généralement à améliorer légèrement les performances de la somme si ses poids ont été appris sur une base d'entraînement caractéristique de la statistique de la base de test.

4. BIOMETRIE MULTIMODALE BASÉE SUR L'IRIS ET LE VISAGE



Figure 4. 5 Performances de la règle de fusion par produit associée aux normalisations par Tanh (à gauche) et zScores (à droite) sur l'approche bimodale Iris_D/Visage_Exp1. En haut sont montrées les distributions normalisées des scores des approches monomodales correspondantes, au milieu : les courbes DET (échelles linéaire et logarithmique) et en bas : les distributions des scores authentiques et imposteurs de l'approche bimodale.

4.3.3. Approche de fusion par somme pondérée basée sur le taux d'occultation de la comparaison des iris – ICORws

Nous avons vu suite à la comparaison de la performance des règles de fusion dans le paragraphe précédent mais aussi dans l'état de l'art que la fusion par somme donne toujours de bons résultats (et parfois les meilleurs). Nous avons également vu que l'amélioration des résultats de la somme par la somme pondérée était dépendante de la base d'entraînement sur laquelle ont été appris les poids. Nous avons alors pensé à introduire sur la somme un autre type de pondération qui soit indépendant de la base d'apprentissage et qui s'applique en ligne, de manière conditionnée et en fonction d'un certain critère de qualité.

a) Description de la méthode

Etant donné que la modalité de l'iris est plus fiable que celle du visage aux bas FAR, il convient de lui attribuer plus de poids. Seulement, cette fiabilité est remise en question lorsque les images des iris sont de mauvaise qualité. Le taux d'occultation de la comparaison entre les iris est un cas particulier de facteur de qualité que nous avons introduit au chapitre 2 (cf. § 2.1.3c)-ii)) et dont l'expression est reformulée par l'équation (2.12). Il s'agit d'abord de déterminer l'union des pixels occultés (marqués dans *Mask*) de deux iris à comparer dans la phase de test. La proportion de ces pixels par rapport au nombre total de pixels de l'iris traduit le taux d'occultation de la comparaison des iris que l'on désignera dans la suite de ce manuscrit par ICOR (*the Irises Comparison's Occultation Rate*).

$$ICOR = \frac{nb_pixels_marqués(Mask(iris1) \cup Mask(iris2))}{nb_pixels_tot}$$
(4. 4)

Une fois l'ICOR déterminé, le calcul du score multimodal final S se fait selon l'équation (4.5) suivante où $S_n(m)$ désigne le score normalisé de la modalité m.

$$\begin{cases} S = 0,5 \times S_n(iris) + 0,5 \times S_n(visage), & si \, ICOR \le Seuil \\ S = (1 - ICOR) \times S_n(iris) + ICOR \times S_n(visage), & si \, ICOR > Seuil \end{cases}$$
(4.5)

Le *Seuil* à partir duquel sera appliquée une pondération différente de celle d'une somme simple (poids de 0,5 pour les deux modalités) est déterminé de manière empirique et le choix de sa valeur sera discuté un peu plus loin dans cette partie. Ainsi, en se référant à l'**équation (4.5)**, si par exemple la comparaison en ligne entre un iris candidat et un iris enrôlé présente un taux d'occultation ICOR = 0,4 et que le seuil à partir duquel on applique la pondération en ligne est *Seuil* = 0,3, le score de la modalité iris aurait un poids ω_{iris} = 0,6 et celui du visage un poids ω_{visage} = 0,4. La **Figure 4.6** montre les étapes de génération de scores multimodaux par la règle de fusion proposée ICORws (*weighted sum based on the Irises Comparison's Occultation Rate*) à partir des traits biométriques de l'iris et du visage d'un candidat revendiquant une identité enrôlée.



Figure 4. 6 Schéma fonctionnel de l'authentification multimodale par l'iris et le visage avec la règle de fusion de scores ICORws par pondération des scores en ligne selon le taux d'occultation de la comparaison des iris.

En plus de la pondération en ligne, il est également intéressant d'injecter dans notre règle de fusion une pondération hors ligne correspondant à la somme pondérée usuelle. Les poids de cette somme sont calculés selon l'équation (4. 3) dès la phase d'apprentissage à partir de la performance des méthodes unimodales de l'iris et du visage sur la base d'apprentissage des paramètres de fusion (BApf). Ces poids sont ensuite ajustés dans la phase de test suivant la mesure de qualité de l'ICOR comme schématisé sur la **Figure 4. 7**. Les résultats issus de cette double pondération sont donnés dans la **Table 4. 4** et l'intérêt de cette cette démarche sera discuté dans le paragraphe suivant.



Figure 4. 7 Schéma fonctionnel du module de fusion en ligne et hors ligne des scores d'une m éthode d'authentification multimodale par l'iris et le visage (règle de fusion : ws + ICORws).

b) Résultats expérimentaux et discussion

La **Table 4.** 4 donne les résultats de la fusion des deux méthodes proposées par pondération en ligne ICORws et par pondération à la fois en ligne et hors ligne [ws + ICORws] comparées aux règles de fusion faisant partie des associations de méthodes de normalisation/fusion intéressantes vues dans le paragraphe précédent. Les résultats sont présentés aux deux points de fonctionnement que sont l'EER et le FRR pour un taux de FAR de 0,01 qui constituent des points de mesure courants. Il est vrai que les méthodes de normalisation les plus intéressantes sont celles de zScores et Tanh, mais nous allons présenter dans la **Table 4.** 4 les résultats de la normalisation Min-Max associée aux règles de fusion somme et somme pondérée (ws) pour des fins de comparaison avec les méthodes proposées. Nous montrons également, sur les **Figure 4.** 8 et **Figure 4. 9**, les courbes DET en échelles linéaires et logarithmiques relatives aux associations les plus intéressantes pour avoir plus d'information sur le comportement de la méthode en fonction du seuil de décision. **Table 4. 4** Influence de l'association des méthodes de normalisation aux méthodes de fusion de scores sur les performances de l'approche bimodaleIris_D / Visage pour les quatre expériences IV² (EER en %, intervalle de confiance de confiance à 90% et FRR au point FAR = 1%).

		Méthodes unimodales										
	Exp1 [env. co	ontrôlé +petite expr]	Exp2 [v. éclai	rage]	Exp3 [v. qua	lité]	Exp4 [multise	ssion]				
	EER (%)	FRR (FAR = 1%)	EER (%)	FRR (FAR = 1%)	EER (%)	FRR (FAR = 1%)	EER (%)	FRR (FAR = 1%)				
Iris_D	3,75 (±0,77)	4,74	3,46 (±0,76)	4,81	2,93 (±0,80)	4,35	1,52 (±0,84)	1,94				
Visage	3,69 (±0,77)	7,39	16,17 (±1,54)	38,24	6,45 (±1,16)	17,69	12,44 (±2,26)	29,42				
			F	Fusion multimod a	ale Iris_D /	Visage						
Normalisation/Fusion	EER (%)	FRR (FAR = 1%)	EER (%)	FRR (FAR = 1%)	EER (%)	FRR (FAR = 1%)	EER (%)	FRR (FAR = 1%)				
Min-Max, Somme	1,23 (±0,45)	1,30	4,10 (±0,83)	7,47	1,70 (±0,61)	2,64	3,28 (±1,22)	5,44				
Min-Max, ICORws [Seuil = 0]	0,88 (±0,38)	0,73	2,58 (±0,66)	4,28	1,32 (±0,54)	1,80	2,07 (±0,97)	2,22				
Min-Max, ws	1,12 (±0,34)	1,17	2,96 (±0,71)	4,63	1,23 (±0,52)	1,46	2,45 (±1,06)	3,35				
Min-Max, ws + ICORws [Seuil = 0]	0,83 (±0,37)	0,63	1,92 (±0,57)	2,75	1,12 (±0,50)	1,16	1,64 (±0,87)	1,91				
zScores, Max	1,06 (±0,42)	1,20	1,63 (±0,53)	2,33	1,45 (±0,57)	1,94	1,69 (±0,88)	2,04				
zScores, Somme	1,20 (±0,44)	1,26	3,31 (±0,75)	5,85	1,49 (±0,57)	2,01	2,84 (±1,14)	3,92				
zScores, ICORws [Seuil = 0]	0,51 (±0,29)	0,36	2,45 (±0,65)	3,02	0,78 (±0,42)	0,75	1,17 (±0,73)	1,31				
zScores, ws	1,01 (±0,41)	1,01	2,58 (±0,66)	3,60	1,10 (±0,49)	1,18	1,86 (±0,92)	2,38				
zScores, ws + ICORws [Seuil = 0]	0,41 (±0,26)	0,31	2,07 (±0,60)	2,41	0,63 (±0,37)	0,54	0,96 (±0,67)	0,95				
Tanh, Produit	0,45 (±0,27)	0,29	1,47 (±0,50)	1,73	0,94 (±0,46)	0,83	1,44 (±0,81)	1,66				
Tanh, Somme	0,45 (±0,27)	0,30	1,36 (±0,48)	1,62	1,05 (±0,48)	1,12	1,55 (±0,85)	1,72				
Tanh, ICORws [Seuil = 0,5]	0,30 (±0,22)	0,20	1,24 (±0,46)	1,34	0,99 (±0,47)	0,99	1,55 (±0,85)	1,72				
Tanh, ws	0,66 (±0,33)	0,57	1,61 (±0,53)	1,92	0,79 (±0,42)	0,68	1,21 (±0,75)	1,28				
Tanh, ws + ICORws [Seuil = 0,5]	0,49 (±0,29)	0,16	1,38 (±0,49)	1,54	0,70 (±0,40)	0,47	1,22 (±0,75)	1,29				
N. comp_tot		3276		3096		2412		1155				
N. comp_intra	p_intra 1638			1548		1206	580					



Figure 4. 8 Courbes DET en échelle linéaire représentant la performance de l'approche bimodale Iris_D/Visage pour les quatre expériences IV² en fonction des associations des méthodes de Normalisation/Fusion et comparaison avec les approches monomodales correspondantes.



Figure 4. 9 Courbes DET en échelle logarithmique représentant la performance de l'approche bimodale Iris_D/Visage pour les quatre expériences IV² en fonction des associations des méthodes de Normalisation/Fusion et comparaison avec les approches monomodales correspondantes.

i) Choix du Seuil d'ICORws

Suite à la variation du *Seuil* à partir duquel la pondération en ligne prend effet dans les règles de fusion basées sur l'ICOR (*Seuil* = $n \times 0,1$ avec n = 0 à 10), nous montrons dans la **Table 4. 4** les résultats relatifs aux seuils optimaux. Nous remarquons d'après la **Table 4. 4** que la valeur du *Seuil* optimal dépend de la technique de normalisation associée. En effet, avec Min-Max et zScores, c'est le *Seuil* zéro qui donne les meilleures performances pour les règles de fusion ICORws et ws + ICORws, alors qu'avec Tanh la valeur optimale du *Seuil* est à 0,5.

Le Seuil nul dans la formule d'ICORws correspond à l'élimination de la condition de la pondération en ligne. Ceci veut dire que la pondération des modalités de l'iris avec ω_{iris} = (1-ICOR) et du visage avec $\omega_{ivisage}$ = ICOR se fait systématiquement quelque soit la valeur d'ICOR. D'un autre côté, un Seuil =1 dans la formule d'ICORws correspond au cas particulier de la somme, où la pondération en ligne ne se fait pour aucune comparaison quelque soit son ICOR. Plus on s'approche du Seuil nul, plus on minimise la restriction de la pondération en ligne. Inversement, plus on s'approche du Seuil = 1, plus ICORws est équivalente à la règle somme simple.

Il est alors intéressant de voir de plus près la proportion des comparaisons des listes de test IV² qui seraient affectées par la pondération en ligne en fonction du *Seuil* d'ICOR. D'après la **Figure 4.10**, nous constatons qu'à partir du seuil 0,6, ICORws est à peu près équivalente à la somme car très peu de comparaisons sont affectées par la pondération. Ceci est également vrai pour l'expérience 4 au seuil 0,5.

Maintenant, essayons de comprendre pourquoi le *Seuil* zéro, impliquant une pondération en ligne quelque soit le taux d'ICOR, donne les meilleurs résultats pour Min-Max et zScores et pas pour Tanh. Il est utile de rappeler que Tanh donnait déjà les meilleurs résultats une fois associée à Somme avec un EER de 0,45% (cf. **Table 4. 3**). L'association Tanh/Somme serait alors celle qui a le moins besoin d'être modifiée en vue d'être améliorée : peu de rectifications sur l'affectation des poids sont suffisantes pour l'améliorer et un peu trop de modifications risquent de la dégrader.



Figure 4. 10 Influence de la variation du *Seuil* d'COR sur la proportion des comparaisons affectées par la pondération en ligne.



Figure 4. 11 Influence de la variation du Seuil de l'association Tanh/ICORws sur les distributions des scores de l'approche bimodale Iris_D/Visage pour les quatre expériences IV².

Pour investiguer d'avantage dans ce sens, nous présentons dans la **Figure 4. 11** une comparaison entre les distributions des scores suite à une normalisation Tanh et une fusion par ICORws avec *Seuil* = 0 et *Seuil* = 0,5. Les distributions des scores suite à l'association Tanh/Somme sont également montrées en tant que référence.

D'après la **Figure 4. 11**, le *Seuil* de 0,5 est meilleur que le *Seuil* nul pour toutes les expériences sauf l'expérience 4, dans laquelle le nombre de comparaisons affectées par la pondération en ligne est négligeable. Dans les expériences 1 à 3, nous notons surtout une transformation de la forme des distributions des authentiques (et parfois aussi de celle des imposteurs) en appliquant un *Seuil* nul dans ICORws. En effet, la distribution des authentiques ayant initialement une forme proche d'une gaussienne subit une

dégradation suite à l'application d'ICORws à seuil nul en présentant une queue droite plus allongée la déviant d'une forme gaussienne. Par contre, l'application d'ICORws avec un *Seuil* = 0,5 rapproche d'avantage les sores des authentiques (de la fusion par somme) vers une distribution gaussienne. Ceci peut expliquer pourquoi les résultats se dégradent suite à l'application d'ICORws à *Seuil* nul et s'améliorent avec un seuil de 0,5, étant donné que la normalisation Tanh est performante sous l'hypothèse d'une distribution gaussienne [Morizet 2009].

ii) Comparaison aux méthodes monomodales

Nous pouvons tout d'abord remarquer à partir de la **Table 4. 4** que les performances des règles de fusion proposées basées sur l'ICOR surpassent largement celles des approches monomodales de l'iris et du visage, et ce quelque soit la technique de normalisation. Cette constatation est confirmée par les **Figure 4. 12** et **Figure 4. 13** où sont montrés en détails les résultats des meilleures combinaisons de méthodes de normalisation/fusion face aux approches monomodales en termes de courbes DET et de recouvrement entre les distributions authentiques et imposteurs.

Nous notons également qu'en général, la fusion de l'iris et du visage améliore d'une manière significative la performance de la modalité du visage et que ce gain en performances est plus prononcé en considérant des images de visages acquises sous un éclairage non contrôlé. Ces résultats vont dans le même sens que ceux qui ont été publiés par le NIST suite à l'épreuve MBGC [NIST MBGC 2009] [Phillips 2009].

iii) Analyse des résultats par expérience

Nous remarquons d'après les **Figure 4. 8** et **Figure 4. 9** qu'en général, ce sont les couples de normalisation/fusion Tanh/ICORws, Tanh/ws+ICORws et Zscores/ws+ICORws qui sont les concurrents pour la première place au niveau des différentes expériences IV². Les meilleurs résultats inscrits en gras dans la **Table 4. 4** confirment cette observation aux deux points de fonctionnement EER et FRR(FAR=1%).

Dans l'expérience 1 (visages en environnement contrôlé avec variation limitée d'expression), l'association Tanh/ICORws présente la courbe DET ayant le plus faible FRR pour la plupart des valeurs de FAR.

Dans l'expérience 2 (visages avec variation d'éclairage), nous observons que toutes les associations faisant intervenir la normalisation Tanh ont des performances très proches avec un léger avantage à la règle de fusion ICORws pour les faibles FAR jusqu'à 4% environ qui se transforme en faveur de la règle ws+ICORws pour les FAR plus élevés.

Dans l'expérience 3 (visages avec variation de la qualité), l'association Tanh/ws+ICORws offre une courbe DET ayant la plus grande vitesse de convergence. Cette méthode est donc la plus performante aux FAR élevés. Cependant, ce sont les associations zScores/ICORws et zScores/ws+ICORws qui priment aux faibles FAR.

Enfin, dans l'expérience 4 (visages capturés en plusieurs sessions), nous notons que c'est l'association zScores/ws+ICORws qui réalise les meilleures performances sur la plupart de l'intervalle des FAR. Nous notons également que pour cette expérience, la pondération en ligne ne confère aucune valeur ajoutée en cas d'association avec la normalisation Tanh (courbe de Tanh/ICORws confondue avec celle de Tanh/Somme et courbe de Tanh/ws+ICORws confondue avec celle de Tanh/ws) étant donné le nombre négligeable des comparaisons affectées par la pondération dans ce cas (*Seuil* = 0,5).

iv) Comparaison aux règles somme et somme pondérée

La règle de fusion proposée par ICORws peut être vue comme une amélioration de la règle somme et celle de [ws+ICORws] comme une amélioration de la somme pondérée. C'est pourquoi nous allons, dans un premier temps, considérer la somme et la somme pondérée en tant que références respectivement pour les règles ICORws et [ws+ICORws]. Les cases en jaune de la **Table 4. 4** désignent les cas où la règle de fusion par ICORws surpasse la somme, et la règle [ws+ICORws] surpasse la somme pondérée, pour une même technique de normalisation.

Il est clair que les méthodes proposées dépassent la somme et la somme pondérée au niveau de la quasi-totalité des associations. Le gain en EER atteint les 37% avec la normalisation Min-Max, les 58,8% avec zScores et les 33,3% avec Tanh qui réalisait déjà de très bonnes performances. D'un autre côté, le gain en FRR au point où le FAR est à 1% par les règles de fusion proposées basées sur l'ICOR atteint 59,2%, 71,4% et 71,9% respectivement avec les normalisations Min-Max, zScores et Tanh.

Une autre comparaison intéressante serait entre la pondération en ligne par la règle ICORws et la pondération hors ligne par la règle somme pondérée. Nous constatons d'après la **Table 4.** 4 que Min-Max/ICORws réalise de meilleures performances que Min-Max/ws sur les expériences 1, 2 et 4. Quant à zScores/ICORws, elle surpasse zScores/ws au niveau de toutes les expériences. Si nous complétons notre analyse par l'observation du comportement des courbes DET de la **Figure 4. 9**, l'avantage d'ICORws par rapport à la somme pondérée est confirmé pour les expériences 1 et 4, mais a tendance à s'inverser aux FAR élevés pour les expériences 2 et 3. De manière analogue, Tanh/ICORws prime sur Tanh/ws au niveau de l'expérience 1 quelque soit le FAR et au niveau des expériences 2 et 3 aux FAR faibles (l'expérience 4 n'est pas concernée par la comparaison car ICORws est équivalente à Somme à cause du nombre négligeable des comparaisons de cette expérience qui sont affectées par la pondération d'ICORws au vu du choix du *Seuil* d'ICORws à 0,5 pour la normalisation Tanh).

Par ailleurs, nous notons qu'en général, les associations faisant intervenir une pondération à la fois en ligne et hors ligne par la règle [ws+ICORws] améliore les résultats de la pondération en ligne dans les cas où, à la base, la somme pondérée améliore ceux de la somme. Dans le cas contraire, c'est toujours la règle ICORws qui réalise les meilleures performances.

Cette étude comparative montre la performance des méthodes de fusion proposées basées sur le taux d'occultation de la comparaison des iris, notamment aux FAR faibles. Ceci permet d'avoir une plus grande sécurité qu'avec les autres règles de fusion, en faisant moins d'erreurs sur l'acceptation d'individus qui ne devraient pas l'être.



Figure 4. 12 Performances de la règle de fusion par pondération en ligne ICORws proposée, associée à la normalisation Tanh sur la méthode bimodale Iris_D/Visage pour les expériences 1 et 2. Pour chaque expérience sont montrées en haut : les distributions normalisées des scores des méthodes monomodales correspondantes, au milieu : les courbes DET (échelles linéaire et logarithmique), en bas : les distributions des scores authentiques et imposteurs de la méthode bimodale.



Figure 4. 13 Performances de la règle de fusion par pondération en ligne et hors ligne proposée [ws+ICORws] associée à la normalisation zScores sur la méthode bimodale Iris_D/Visage pour les expériences 3 et 4. Pour chaque expérience sont montrées en haut : les distributions normalisées des scores des méthodes monomodales correspondantes, au milieu : les courbes DET (échelles linéaire et logarithmique), en bas : les distributions des scores authentiques et imposteurs de la méthode bimodale.

4.3.4. Impact de l'agrandissement de la base d'apprentissage des paramètres de fusion – BApf

Etant donné le nombre réduit des images (156) et des comparaisons (312) de la base d'apprentissage des paramètres de fusion (BApf), nous avons voulu étudier l'effet de l'agrandissement de cette base sur les résultats de la reconnaissance. Cette étude sera conduite sur la fusion bimodale entre les iris droits et les visages de l'expérience 1 (environnement contrôlé avec de petites expressions).

Rappelons tout d'abord que les comparaisons ont été générées en respectant le protocole IV², c'est-à-dire en ayant des nombres de comparaisons intraclasses et interclasses équivalents (BApf0).

En vue d'agrandir la BApf, deux manières ont été envisagées. La première est de garder le même ensemble d'apprentissage et de conserver toutes les comparaisons interclasses possibles entre les images d'entraînement (BApf1), plutôt que d'en sélectionner aléatoirement un sous-ensemble dont le nombre est égal à celui des comparaisons intraclasses (protocole IV²).

La deuxième manière est de changer de BApf en choisissant un sous-ensemble d'iris gauches la base IV² et un sous-ensemble de visages relatifs à l'expérience 3 (variation de la qualité) pour constituer une nouvelle BApf (BApf2).

La **Table 4. 5** résume les caractéristiques des 3 BApfs que nous allons comparer dans ce paragraphe. Nous précisons que les deux premières BApf sont chimériques et que la troisième est réelle.

	Nb. sujets / modalité	Nb. img / sujet / modalité	Nb. total images retenues /modalité	Nb. comparaisons intra-classes / modalité	Nb. comparaisons totales / modalité
BApf0	52	3	156	156	312
BApf1	52	3	156	156	12090
BApf2	206	3 enrôl. + 3 test	942	973	1946

 Table 4. 5 Caractéristiques des bases d'apprentissages des paramètres de fusion et des listes de comparaisons d'apprentissage associées

Nous présentons dans la **Table 4. 6** les résultats de l'approche bimodale Iris_D/visage de l'expérience 1 suite à la variation de la BApf pour les mêmes associations étudiées au paragraphe précédent. Nous remarquons à partir de cette table que dans la majorité des cas, les résultats restent constants ou alors subissent une faible dégradation suite à l'agrandissement de la BApf, notamment avec les règles de fusion proposées.

Les cas où BApf1 et BApf2 améliorent les résultats trouvés initialement par BApf0 sont marqués en jaune dans la **Table 4.6** et leurs courbes DET sont montrées sur la **Figure 4.14**. Cette amélioration est surtout notable avec la règle de normalisation Min-Max qui est la plus sensible au nombre de comparaisons de la base d'apprentissage, même si ces derniers proviennent de conditions différentes de celles du test (courbes en rouge de la **Figure 4. 14**). Le gain en EER atteint les 41% et le gain en FRR au point où le FAR est à 1% atteint les 64% par rapport aux résultats avec BApf0.

Une amélioration est également notée au niveau de la normalisation zScores lorsqu'elle est combinée aux règles de fusion somme, ICORws et somme pondérée. En particulier un gain de l'ordre de 40% d'EER est perçu au niveau de l'association zScores/ws suite à l'agrandissement de la BApf. L'observation des courbes en bleu de la **Figure 4. 14** indique un comportement équivalent pour les deux bases BApf1 et BApf2 aux faibles FARs avec un avantage de la BApf1 aux FARs plus élevés. L'augmentation du nombre de comparaisons interclasses (distribution des imposteurs) de la BApf0 apporte donc le plus d'avantage sur la méthode zScores, ce qui a du sens car les paramètres de fusion (moyenne et écart-type) de zScores se calculent sur la distribution totale des scores (authentiques et imposteurs étendus) à partir du même ensemble d'images que BApf0.

Table 4. 6 Performances des associations de méthodes de normalisation/fusion suite àl'agrandissement de la BApf avec l'approche bimodale Iris_D/Visage Exp1.

	BApf0		BApf1		BApf2	
Normalisation/Fusion	EER (%)	FRR à (FAR = 1%)	EER (%)	FRR à (FAR = 1%)	EER (%)	FRR à (FAR = 1%)
Min-Max, Somme	1,23 (±0,45)	1,30	1,22 (±0,45)	1,29	1,16 (±0,43)	1,21
Min-Max, ICORws	0,88 (±0,38)	0,73	1,43 (±0,48)	1,95	1,23 (±0,45)	1,58
Min-Max, ws	1,12 (±0,34)	1,17	1,01 (±0,41)	1,03	0,66 (±0,33)	0,42
Min-Max, ws + ICORws	0,83 (±0,37)	0,63	1,33 (±0,47)	1,78	1,19 (±0,44)	1,38
zScores, Max	1,06 (±0,42)	1,20	1,09 (±0,42)	1,30	1,50 (±0,49)	2,40
zScores, Somme	1,20 (±0,44)	1,26	0,97 (±0,40)	0,97	1,01 (±0,41)	1,03
zScores, ICORws	0,51 (±0,29)	0,36	0,42 (±0,26)	0,24	0,45 (±0,27)	0,23
zScores, ws	1,01 (±0,41)	1,01	0,61 (±0,32)	0,36	0,58 (±0,31)	0,39
zScores, ws + ICORws	0,41 (±0,26)	0,31	0,47 (±0,28)	0,24	0,67 (±0,33)	0,40
Tanh, Produit	0,45 (±0,27)	0,29	0,45 (±0,27)	0,29	0,42 (±0,26)	0,33
Tanh, Somme	0,45 (±0,27)	0,30	0,45 (±0,27)	0,30	0,45 (±0,27)	0,33
Tanh, ICORws	0,30 (±0,22)	0,20	0,30 (±0,22)	0,20	0,32 (±0,23)	0,12
Tanh, ws	0,66 (±0,33)	0,57	0,65 (±0,33)	0,57	0,90 (±0,38)	0,84
Tanh, ws + ICORws	0,49 (±0,29)	0,16	0,53 (±0,29)	0,29	0,85 (±0,37)	0,75

Fusion multimodale Iris_D / Visage Exp1

La règle de normalisation Tanh reste pratiquement stable suite à l'augmentation du nombre de comparaisons interclasses (scores imposteurs) dans BApf1, ce qui était prévisible étant donné que ses paramètres sont calculés sur la distribution des authentiques (moyenne et écart-type) qui elle, n'a pas changé entre BApf0 et BApf1. Nous remarquons également qu'avec la normalisation Tanh, BApf2 dégrade la performance des associations faisant intervenir la somme pondérée car la base d'apprentissage BApf2 comporte des images d'iris gauche et des images de visage comprenant une variation d'éclairage et donc n'est pas caractéristique de la base de test composées d'iris droit et de visages capturés en environnement contrôlé.



Figure 4. 14 Courbes DET en échelles linéaire (haut) et logarithmique (bas) représentant l'amélioration de la performance d'associations de méthodes de normalisation/fusion par l'agrandissement de la BApf avec l'approche bimodale Iris_D/Visage Exp1.

En conclusion, nous retenons que l'agrandissement de la base d'apprentissage est avantageux surtout pour la normalisation Min-Max avec le moins de restrictions sur les différences entre la BApf et la base de test. La normalisation zScores tire spécialement profit de l'augmentation du nombre de comparaisons interclasses sur une BApf caractéristique de la base de test. Enfin, la normalisation Tanh est la plus indépendante de l'agrandissement de la BApf, sauf avec une fusion par somme pondérée qui dépend de la nature de BApf.

Une solution qui serait profitable à toutes les normalisations serait de travailler avec des techniques de ré-échantillonnage telles que la validation croisée (tirage aléatoire sans remise) ou le bootstrap (tirage aléatoire avec remise), tout en gardant toutes les comparaisons interclasses générées. Le principe général de ces techniques consiste à considérer la base de données complète (l'union de la Bapf0 et base de test) et à la partager aléatoirement et un certain nombre de fois en deux : une base d'apprentissage et une base de test. Pour chaque découpage, on réalise l'apprentissage des paramètres de fusion et le test de performances. Ensuite, on calcule la moyenne et l'écart-type sur les différents tirages. Ainsi la taille de la base d'apprentissage est augmentée sans réduire celle du test et les résultats obtenus sont indépendants du choix de la base d'apprentissage [Allano 2009] [Morizet 2009].

L'application de ces techniques constitue l'une des principales perspectives de ce travail en vue de généraliser les résultats obtenus et avoir des mesures de performances statistiquement plus significatives.

4.4 Méthode multimodale par l'iris droit, le visage et l'iris gauche

Après avoir proposé, dans le paragraphe précédent, une approche bimodale fusionnant les traits de l'iris et du visage en prenant comme échantillon d'iris, celui de droite, nous allons maintenant étendre cette fusion pour prendre en compte à la fois les deux échantillons d'iris en plus du visage.

4.4.1. Fusion multimodale par les règles usuelles de l'état de l'art

La fusion basée sur le taux d'occultation des comparaisons de l'iris ICORws n'étant pas encore généralisée à une méthode prenant en compte les deux échantillons d'iris droit et gauche, les résultats de la fusion de l'approche multimodale ici présentés se limitent aux associations usuelles de normalisation/fusion de l'état de l'art.

La **Table 4.** 7 montre, en pourcentages, les performances de ces associations pour les quatre expériences IV^2 au point de fonctionnement EER à 90% d'intervalle de confiance, au point FRR où FAR est à 1%, mais aussi au point FRR où FAR est à 0,1% vu les faibles taux d'erreurs. On remarque une très nette augmentation des performances grâce à l'utilisation de la multimodalité comparée à la meilleure méthode unimodale utilisée individuellement (à part pour l'expérience 4 comprenant un nombre réduit de comparaisons et où l'iris gauche réalise déjà des taux d'erreurs presque nuls - cf. **Annexe B-Exp4**). Le gain en EER atteint les 92,4% dans Exp1, 93,6% dans Exp2, 90,3% dans Exp3 et 100% dans Exp4 par rapport à la meilleure modalité. Par ailleurs, il est à noter que, compte tenu des taux d'erreurs très faibles et parfois nuls obtenus par la fusion multimodale, la taille des listes de comparaisons de test devient trop petite pour refléter de manière précise la performance de chaque association. C'est d'ailleurs ce que révèlent les intervalles de confiance à 90% dont la moitié dépasse, dans certains cas, les EERs faibles auxquels ils sont associés.

Nous allons quand même présenter à titre indicatif une comparaison entre les différentes associations de normalisation/fusion à partir de la Table 4.7 et de la Figure 4. 15. Généralement, la normalisation Tanh combinée aux règles somme et produit donne les meilleurs résultats au niveau de toutes les expériences, avec un léger avantage pour la règle somme dans l'expériences 2 (variation de l'éclairage) et pour la règle produit dans l'expérience 1 (environnement contrôlé avec de petites expressions) où toutes les associations ont des courbes DET très entremêlés reflétant les performances équivalentes des différentes associations. Dans l'expérience 3 (variation de la qualité), les courbes de Tanh/Somme et Tanh/Produit sont les plus en dessous et sont complètement superposées. Par ailleurs, on note l'absence de l'expérience 4 (multisession) de la Figure 4. 15 car cette dernière donne une représentation logarithmique de l'évolution du FRR en fonction du FAR, ce qui est impossible avec des taux de FRR nuls. C'est pourquoi nous montrons dans la Figure 4. 16, les distributions des scores des quatre associations donnant des taux d'erreurs nuls à tous les points de fonctionnement de la Table 4. 7. Ces associations sont zScores/Somme pondérée, Tanh/Somme, Tanh/Produit et Tanh/Somme pondérée. Cette figure montre l'absence de recouvrement entre les distributions des imposteurs et celles des authentiques pour les quatre associations (ce qui reflète les taux d'erreurs nuls de la Table 4.7), avec une distance de séparation légèrement supérieure pour les associations faisant intervenir la normalisation Tanh. Il reste cependant difficile de décider laquelle de ces associations de méthodes de normalisation/fusion de scores est la meilleure.

Table 4. 7 Comparaison de la performance des associations de méthodes de normalisation/fusion de scores sur l'approche multimodale Iris droit / Visage / Iris gauche pour les quatre expériences IV² (EER en %, intervalle de confiance à 90% et FRRs aux points FAR = 1% et 0,1%).

	Exp1		Exp2	Exp2			Exp3			Exp4		
	EER (%)	FRR à 1% FAR	FRR à 0,1% FAR	EER (%)	FRR à 1% FAR	FRR à 0,1% FAR	EER (%)	FRR à 1% FAR	FRR à 0,1% FAR	EER (%)	FRR à 1% FAR	FRR à 0,1% FAR
Iris_D	2,76 (±0,75)	3,29	6,23	2,20 (±0,69)	2,51	21,95	2,47 (±0,82)	3,26	6,14	1,59 (±0,98)	1,86	2,63
Visage	3,47 (±0,84)	6,11	23,70	14,31 (±1,65)	40,12	63,58	6,14 (±1,27)	18,88	46,04	13,65 (±2,69)	41,27	64,68
Iris_G	3,02 (±0,78)	4,34	11,40	3,26 (±0,83)	4,87	12,50	3,15 (±0,92)	4,96	9,32	0,23 (±0,37)	0,00	1,39

Méthodes unimodales

Fusion multimodale Iris_D / Visage / Iris_G

Normalisation/Fusion	EER (%)	FRR à 1% FAR	FRR à 0,1% FAR	EER (%)	FRR à 1% FAR	FRR à 0,1% FAR	EER (%)	FRR à 1% FAR	FRR à 0,1% FAR	EER (%)	FRR à 1% FAR	FRR à 0,1% FAR
Min-Max, ws	0,33 (±0,26)	0,25	0,70	0,50 (±0,313	0,35	0,80	0,41 (±0,34)	0,21	0,52	0,23 (±0,37)	0,00	0,35
zScores, Somme	0,31 (±0,25)	0,29	1,17	0,64 (±0,38)	0,51	1,82	0,51 (±0,38)	0,24	0,72	0,23 (±0,37)	0,00	1,97
zScores, ws	0,33 (±0,26)	0,15	0,59	0,33 (±0,27)	0,26	0,55	0,26 (±0,27)	0,19	0,52	0,00 (±0,00)	0,00	0,00
Tanh, Produit	0,21 (±0,21)	0,12	0,29	0,20 (±0,21)	0,00	0,24	0,24 (±0,26)	0,07	0,52	0,00 (±0,00)	0,00	0,00
Tanh, Somme	0,31 (±0,25)	0,15	0,38	0,14 (±0,18)	0,00	0,16	0,24 (±0,26)	0,10	0,52	0,00 (±0,00)	0,00	0,00
Tanh, ws	0,36 (±0,27)	0,22	0,45	0,31 (±0,26)	0,10	0,95	0,31 (±0,29)	0,06	0,83	0,00 (±0,00)	0,00	0,00
N. comp_tot		2594			2450			1946	-		882	
N. comp_intra		1297			1225			973			441	



Figure 4. 15 Courbes DET (échelle logarithmique) représentant la performance de l'approche multimodale Iris_D/Visage/Iris_G en fonction des associations des méthodes de Normalisation/Fusion et comparaison avec les approches monomodales correspondantes.



Figure 4. 16 Distribution des scores authentiques et imposteurs des associations de méthodes de normalisation/fusion réalisant des taux d'erreurs nuls avec l'approche multimodale Iris_D/Visage/Iris_G sur l'expérience 4.

4.4.2. Comparaison à la méthode biomodale par l'iris droit et le visage

Suite à l'analyse précédente démontrant l'intérêt de l'approche multimodale Iris_D/Visage/Iris_G par rapport aux approches unimodales et donnant les meilleures associations de méthodes de normalisation/fusion pour chaque expérience, nous allons comparer, dans ce paragraphe, l'approche multiomdale Iris_D/Visage/Iris_G avec l'approche bimodale Iris_D/Visage.

Avant de présenter les résultats entre les deux fusions (cf. **Figure 4. 18**), il est utile de rappeler que le nombre de comparaisons avec lequel a été testée la fusion multimodale est inférieur à celui de la fusion bimodale en raison de l'existence de sujets disposant d'images d'iris droit et de visage mais qui n'ont pas d'images d'iris gauche dans la base IV².

Par ailleurs, étant donné que ce ne sont pas les mêmes associations de méthodes de normalisation/fusion qui donnent les meilleurs résultats pour les deux approches proposées, nous allons considérer la meilleure association de la méthode Iris_D/Visage/Iris_G qui est Tanh/Somme. Puis nous allons comparer ses performances sur la méthode Iris_D/Visage/Iris_G (traits pleins de la Figure 4. 17) à ses performances sur la méthode Iris_D/Visage en un premier temps (traits interrompus), ensuite aux performances de la meilleure association Iris_D/Visage (traits en pointillés). Chaque expérience ayant une couleur différente sur la Figure 4. 17, la comparaison se fait entre les trois courbes de traits continu, interrompu et en pointillé de la même couleur pour une expérience donnée pami les 12 courbes relatives aux quatre expériences.



Figure 4. 17 Comparaison de la performance de l'approche multimodale Iris_D/Visage/Iris_G à celle de l'approche bimodale Iris_D/Visage pour les quatre expériences IV².

Ces deux séries de courbes montrent l'intérêt de la fusion multimodale incluant les deux échantillons de l'iris en plus du visage. Non seulement les résultats de la même association Tanh/Somme sont meilleurs sur la méthode Iris_D/Visage/Iris_G mais, en plus, ils surpassent celles des meilleures associations de normalisation/fusion sur la méthode bimodale Iris_D/Visage. A part l'expérience 1 où une équivalence est ponctuellement notée au niveau de l'EER, le gain en EER par rapport à la meilleure association est de 88,7%, 61,9% et 100% respectivement pour les expériences 2 à 4 et le gain en FRR où FAR est à 1% est de 25%, 100%, 81,5% et 100% respectivement des des expériences 1 à 4. La fusion multimodale par double iris et par le visage s'avère plus avantageuse et son intérêt est d'autant plus grand qu'elle ne nécessite pas d'introduire un capteur supplémentaire ni de développer de nouveaux algorithmes de caractérisation et de comparaison.

4.4.3. Comparaison à la campagne IV²

L'approche multimodale ici proposée étant basée sur des méthodes de normalisation et de fusion de scores de l'état de l'art, le but de ce paragraphe n'est pas d'évaluer les méthodes de normalisation ou de fusion en soi, mais plutôt de comparer la combinaison des approches monomodales proposées dans le cadre de cette thèse avec d'autres approches multimodales ayant été testées sur la base IV².

En 2007, l'évaluateur de la campagne IV^2 – Thalès Reseach and Technology – publie le rapport d'évaluation des algorithmes biométriques participants au projet IV^2 [Colineau 2007]. A partir des méthodes monomodales participant à IV^2 , l'évaluateur a construit plusieurs méthodes multimodales suite à des normalisations et des fusions de scores par des méthodes de l'état de l'art. C'est à ces méthodes multimodales que nous allons nous comparer. Mais avant, il est important de noter plusieurs différences entre les expériences multimodales conduites dans la campagne IV^2 et celles qui sont présentées dans cette étude.

Il est vrai que nous travaillons sur la même base de données, cependant les échantillons des modalités considérées ne sont pas les mêmes. Tandis que nous fusionnant les iris droit et gauche avec le visage 2D, les méthodes multimodales IV^2 fusionnent l'iris droit avec le visage 2D et 3D. Plusieurs différences s'en suivent notamment en ce qui concerne le nombre de sujets disposant à la fois des trois échantillons biométriques. De plus, nous ne disposons pas des listes de comparaisons d'apprentissage et d'évaluation des expériences multimodales IV^2 , c'est pourquoi nous avons créé nos propres listes. Il en résulte plusieurs dissimilitudes, en particulier, dans le partitionnement de la base et la signification statistique des taux d'erreurs qui dépend de la taille de cette base. Les caractéristiques indiquées dans le rapport technique IV^2 et relatives aux expériences multimodales de la campagne IV^2 sont présentées dans la **Table 4. 8** côte à côte avec celles des bases testées par nos algorithmes. Ces bases de données sont relatives à deux séries de tests dont l'une se déroule en monosession (Test04) et l'autre en multisession (Test03). Nous constatons d'après cette table que dans

les expériences monosessions, notre base d'apprentissage (BApf) est plus réduite en nombre de personnes et en nombre de comparaisons que celles de la campagne IV^2 (BApf04) contrairement à notre base de test qui est plus grande que celle d' IV^2 . Ce qui défavorise nos algorithmes pendant la comparaison. Par contre nous disposons de bases (apprentissage et test) plus réduites que celle de la campagne IV^2 dans les expériences multisessions.

Table 4. 8 Comparaison entre les caractéristiques des bases de données expérimentées en fusion multimodale des iris droit et gauche et du visage 2D avec celles de la campagne IV^2 en fusion multimodale des visages 2D et 3D et de l'iris droit.

	Nb. total sujets retenus /modalité	Nb. comparaisons totales /modalité		
	Expériences	monosessions		
BApf	52	312		
Exp1	233	2594		
Exp2	229	2450		
Exp3	206	1946		
BApf04 [Campagne IV ²]	88	712		
Test04 [Campagne IV ²]	176	1831		
	Expériences	s multisessions		
BApf	52	312		
Exp4	63	882		
BApf03 [Campagne IV ²]	-	3690		
Test03 [Campagne IV ²]	-, (≤77)	1470		

Les résultats de la comparaison des performances des méthodes multimodales sont illustrés par la **Figure 4. 18**. Chaque graphique relatif à la campagne IV² correspond à la variation des associations de règles de fusion/normalisation dans une méthode multimodale donnée. Les détails de ces associations sont donnés pour chaque méthode en **Annexe C**. Les quatre graphiques de la campagne IV² correspondent à des fusions multimodales faisant intervenir les mêmes méthodes biométriques par le visage 3D et par l'iris. Les tests (a) et (b) diffèrent en leurs méthodes biométriques par le visage 2D.

Cette figure montre bien que les performances réalisées par notre approche sont comparables voire surpassent celles de la campagne IV². Dans les expériences mono sessions notre approche réalise des EERs compris dans l'intervalle $[0,14 (\pm 0,18)\%; 0,24 (\pm 0,26)\%]$ avec, au point où FAR est à 1%, des FRRs dans [0% et 0,12%]; alors que les deux approches multimodales (a) et (b) d'IV² donnent respectivement des EERs dans les intervalles $[0,16 (\pm 0,20)\%; 0,70(\pm 0,50)\%]$ et $[0 (\pm 0,10)\%; 0,51(\pm 0,40)\%]$ et des FRRs dans les intervalles [0%; 0,20%] et [0%; 0,40%]. Ainsi, notre approche est classée en seconde position entre les deux approches d'IV² pour les expériences mono sessions et ce malgré la taille réduite de la base d'apprentissage et le nombre plus élevé de comparaisons de test.


Figure 4. 18 Comparaison de la performance de l'approche proposée par fusion de l'iris droit et gauche et du visage à ceux des participants à la campagne IV² par fusion du visage 2D, du visage 3D et de l'iris droit (extraits de [Colineau 2007]).

Dans les expériences multisessions, les EERs des méthodes (a) et (b) d'IV² sont compris respectivement dans les intervalles $[0,42 (\pm 0,40)\%; 0,80(\pm 0,50)\%]$ et $[0,14 (\pm 0,20)\%;$ $0,75(\pm 0,50)\%]$ et les FRR entre [0,05%; 0,40%] et [0%; 0,53%] au point où FAR est à 1%. Bien que notre méthode réalise des taux d'erreurs nuls à la fois en EER et en FRR(FAR=1%), et que la base sur laquelle ont été appris ses paramètres de fusion soit beaucoup plus réduite que celle de la campagne IV² (312 contre 3690), nous concluons que notre méthode est plus performante que celles d'IV² avec certaines réserves à cause du nombre plus réduit des comparaisons de test (882 contre 1470).

4.5 Impact du rejet des comparaisons selon le taux d'occultation des iris

Après avoir présenté, dans la partie précédente, une étude détaillée sur les performances de nos approches multimodales à deux et à trois échantillons biométriques, nous allons, dans ce paragraphe étudier l'influence du rejet des comparaisons en fonction des taux d'occultation de leurs iris. Ce critère de qualité a déjà été utilisé au chapitre 2 (cf. § 2.1.3c)-ii)) pour étudier l'impact du rejet sur l'approche monomodal ede l'iris. Nous avions conclu à ce moment que le seuil optimal à partir duquel il convenait de rejeter les iris « trop » occultés était de 35% d'occultation et c'est ce même seuil que nous allons adopter dans cette étude.

Nous montrons dans la **Figure 4. 19** les courbes DET comparant les performances des deux approches multimodales proposées avant (lignes en pointillés) et après l'élimination (lignes continues) des comparaisons incluant des iris occultés à plus de 35% pour toutes les expériences IV². La variation de l'EER peut également être observée dans la **Figure 4. 20**.



Figure 4. 19 Influence du rejet des comparaisons suivant le taux d'occultation des iris sur les performances des approches multimodales proposées à deux (à gauche) et à trois échantillons biométriques (à droite).

Dans la méthode bimodale Iris_D/Visage, le gain en EER varie entre 49,2% et 80,6% dans les expériences 1 à 3. En revanche, une augmentation de 6,25% d'EER est observée

au niveau de l'expérience 4 suite au rejet des iris trop occultés, avec un passage de l'EER de 0,96 (±0,67)% à 1,02 (±0,74)%. Il est vrai qu'au vu des intervalles de confiance et de l'allure des deux courbes à comparer sur l'expérience 4, nous pouvons dire que les performances sont équivalentes avant et après le rejet. Cependant, nous nous attendions à ce qu'elles s'améliorent et non pas à ce qu'elles restent équivalentes. Pour comprendre nous nous sommes tournés vers les ces résultats. autres associations de normalisation/fusion. Suite à l'analyse de leurs comportements après le rejet, nous avons trouvé qu'elles ont toutes subi une petite chute de performances, même les associations qui, par définition, sont indépendantes du taux d'occultation des iris. A titre d'exemple, l'EER de zScores/Somme est passé de 2,84 (±1.14)% à 3,09 (±1.29)% induisant une augmentation de 8,8% supérieure à celle de l'association zScores/ws+ICORws de la méthode bimodale proposée (6,25%).

Par ailleurs, nous remarquons que la méthode multimodale Iris_D/Visage/Iris_G parvient à annuler les FRRs à partir d'un FAR de 0,13% pour les expériences 1 et 2, et quel que soit le FAR pour les expériences 3 et 4 (aucun recouvrement entre les distributions authentiques et imposteurs).

Maintenant, analysons l'impact du rejet des comparaisons sur la constitution des listes de test. La **Figure 4. 20** montre la variation du nombre de comparaisons intra et interclasses ainsi que celle de l'EER suite au rejet des comparaisons incluant des iris occultés à plus de 35%. Nous constatons à partir de ce graphique, que le nombre des éliminations est trop élevé, surtout dans la méthode multimodale à triple échantillons biométriques où, par exemple, elles affectent plus que le tiers du total des comparaisons dans l'expérience 4.



Figure 4. 20 Influence du rejet des comparaisons suivant le taux d'occultation des iris sur l'EER et sur la composition des bases de test des approches multimodales proposées à deux (à gauche) et à trois échantillons biométriques (à droite).

Nous concluons à partir de cette étude, que plutôt que de rejeter une comparaison faisant intervenir un iris ayant un certain seuil d'occultation, il serait plus judicieux de tirer profit de la multimodalité en exploitant le reste des échantillons disponibles, c'està-dire le visage et éventuellement l'autre iris afin d'apporter une réponse à la requête candidate et d'éviter de demander à l'utilisateur une deuxième série d'acquisitions. Ainsi, l'information sur le taux d'occultation servira à décider lesquels des échantillons biométriques seraient considérés pendant le test, et le système aura une architecture séquentielle lui permettant d'avoir beaucoup plus de probabilité de fournir une réponse à l'utilisateur.

Une deuxième manière envisageable d'exploiter les critères de qualité basés sur l'occultation des iris serait de les intégrer dans les règles de fusion par exemple pour pondérer les scores comme nous l'avions réalisé avec la règle ICORws proposées dans le cadre d'une architecture en parallèle.

Conclusion

Dans ce chapitre, nous avons combiné les deux approches monomodales par l'iris (méthode Log-Gabor) et par le visage (méthode bio-inspirée LBP-DoG) proposés dans les chapitres 2 et 3 pour construire deux approches multimodales à deux et à trois échantillons biométriques. La première approche prend en considération l'échantillon de l'iris droit et celui du visage, alors que la deuxième approche intègre, en plus, l'échantillon de l'iris gauche. Pour effectuer nos tests, nous avons utilisé la base multimodale réelle IV². En plus de sa taille convenable (environ 300 sujets), cette base a l'avantage de présenter plusieurs variabilités.

Les deux approches proposées reposent sur une fusion de scores opérée suite à une phase de normalisation. Plusieurs méthodes usuelles de normalisation et de fusion de scores présentées dans le chapitre 1 ont été comparées. De plus, nous avons proposé, dans l'approche multimodale à deux échantillons, une nouvelle règle de fusion qui est une adaptation de la règle Somme pondérée. La règle de fusion proposée repose sur une pondération de scores qui dépend du critère de qualité du taux d'occultation des comparaisons des iris (ICOR). La pondération se fait alors en ligne pendant la phase de test, mais nous avons également prévu la possibilité de la combiner à la pondération hors ligne usuelle par somme pondérée. La comparaison avec les autres règles de fusion a montré que les meilleures performances sont toujours réalisées par des associations de normalisation/fusion faisant intervenir l'une de nos règles de fusion basées sur l'ICOR.

Par ailleurs, les expériences ont prouvé que la fusion de scores entre modalités améliore très sensiblement les résultats obtenus avec l'une ou l'autre des modalités. De manière analogue, la fusion multimodale par double échantillons d'iris est plus performante que par échantillon unique d'iris. La comparaison de notre méthode multimodale basée sur la fusion des iris droit et gauche et du visage avec les méthodes multimodales à triple échantillons de la campagne IV² indique des performances comparables ou en faveur de notre méthode.

Nous avons aussi considéré la signification statistique des résultats en envisageant des améliorations à apporter sur le protocole de test par l'introduction de techniques de ré-échantillonnage de bases de données.

Enfin, nous avons discuté le cas du rejet des comparaisons de test suivant le critère de qualité de l'occultation de l'iris. Nous avons conclu, à partir du nombre élevé des comparaisons éliminées et des contraintes de ré-acquisition de données biométriques qui s'en suivraient, qu'il était plus judicieux d'exploiter ce critère de qualité en vue d'améliorer les performances sans pour autant faire d'élimination. Dans ce sens, deux procédés peuvent être envisagés. Le premier fait partie des contributions de ce chapitre par l'introduction de la règle de fusion de scores ICORws basée sur le critère du taux d'occultation de la comparaison des iris. Le deuxième procédé fait partie des travaux que nous avons l'intension de réaliser dans le cadre d'un système biométrique multimodal à architecture séquentielle, où le critère de qualité servira pour la décision des modalités qui interviendraient dans le processus de la reconnaissance lors de la phase de test.

Conclusion générale et perspectives

Au cours de ce travail de thèse, nous nous sommes intéressés à la biométrie multimodale par l'iris et le visage. Nous avons commencé par situer le contexte du sujet de thèse en introduisant des notions de base sur les systèmes biométriques en général et les outils permettant leur évaluation. Nous avons, ensuite, présenté les différents aspects des systèmes biométriques multimodaux de par leurs architectures, leurs sources d'information et leurs niveaux de fusion. Suite à des études bibliographiques portant en particulier sur la biométrie monomodale par l'iris et par le visage et sur la multimodalité iris/visage, nous avons développé de nombreuses approches personnelles de reconnaissance d'individus pour chacune des deux modalités. La reconnaissance faciale a été particulièrement abordée selon des approches classiques et des approches bio-inspirées émulant le fonctionnement du cerveau humain. Toutes ces approches ont été construites par des combinaisons d'algorithmes et ont confirmé l'intérêt du scénario multi-algorithmes face aux méthodes individuelles ascendantes. Les résultats de la meilleure approche de chaque modalité ont été retenus en tant que références pour l'évaluation de différentes stratégies de fusion multimodale au niveau des scores. Ce niveau de fusion a été choisi car il offre le meilleur compromis entre richesse d'information et facilité d'implémentation. Les méthodes proposées ont été évaluées sur la base multimodale réelle IV² capturée dans des environnements variables voire dégradés et en suivant un protocole bien précis fourni dans le cadre de la campagne d'évaluation IV². Une étude comparative des méthodes proposées face aux algorithmes participants à la campagne IV² a été engagée et a prouvé la compétitivité de nos algorithmes qui arrivent dans plusieurs cas à se positionner en tête de liste.

Cette thèse apporte donc une série de contributions dans le domaine de la biométrie mono et multimodale qui peuvent être répertoriées en quatre parties.

La première partie dédiée à la biométrie par l'iris offre **une première contribution** par le développement d'une méthode d'authentification par l'iris basé sur les filtres Log-Gabor réalisé à travers la réunion de plusieurs modules de référence de l'état de l'art. Cette méthode a prouvé ses performances sur la base d'images d'iris IV² en réalisant de faibles taux d'EER de 3,6%. **La deuxième contribution** de cette partie est l'introduction de deux critères de qualité basés sur le taux d'occultation des iris pour améliorer la performance de la méthode initiale. Grâce à ces critères de qualité, le taux d'EER a pu être réduit au tiers avec 1,2%.

La deuxième partie a été dédiée à la biométrie faciale 2D par le biais d'approches classiques. Cette partie comprend **trois contributions** de cette thèse à travers la proposition de trois méthodes d'authentification basées chacune sur la combinaison de deux algorithmes classiques. Les trois méthodes proposées font toutes intervenir la méthode d'analyse spatio-fréquentielle de la Pyramide Orientable (PO) qui est combinée par la suite, à des méthodes projectives : la PCA (Analyse en Composantes Principales), la LDA (Analyse Linéaire Discriminante), ou une méthode adaptative de classification non linéaire « Adaboost ». Plusieurs stratégies de fusion de caractéristiques et de scores issus des différentes bandes passe-haut, passe-bas et passe-bande orientées de la PO ont été étudiées. Au bout de cette étude, nous avons vu que la LDA permettait une extraction plus efficace de l'information discriminante contenue dans la PO. Les trois approches classiques proposées on été évaluées sur la base IV² à travers des expériences en environnements non contrôlés comprenant des variations d'éclairage, de qualité et de session. Elles ont prouvé leur performance face à d'autres méthodes ayant participé à la campagne d'évaluation IV² en réalisant des taux d'égales d'erreurs entre 3,6% et 5% dans un environnement contrôlé et entre 10% et 20% dans des environnements non contrôlés, la meilleure approche classique étant celle de la LDA/PO.

Dans la troisième partie, où sont proposées deux contributions majeures de cette thèse, nous avons abordé le problème de la reconnaissance faciale 2D par des approches bio-inspirées. Suite à la description du mécanisme de la perception de visage dans le cerveau humain, nous en avons présenté quelques modélisations par des algorithmes fondamentaux utilisés en biométrie faciale. Puis, nous avons élaboré deux approches bio-inspirées s'appuyant sur certaines de ces modalisations. La première approche se base sur une décomposition multi-échelles par filtre DoG (Différence de Gaussiennes) des images de visage, suivie de l'application d'un opérateur LBP^{U2} (Local Binary Patterns) à motifs uniformes sur les bandes filtrées. Nous avons ensuite prouvé que l'exploitation des résultats issus des travaux neuroscientifiques de Gosselin et Schyns apportait une amélioration significative sur le taux de la reconnaissance. Ces résultats nous ont orientés dans le choix des régions pertinentes et dans la pondération à la fois des régions et des échelles de décomposition utiles à l'identification d'un visage. Nous avons démontré la robustesse de la méthode (LBP/DoG-bio) qui arrive à surpasser toutes les approches classiques proposées dans la partie précédente et qui parvient à se classer au premier ou au second rang (selon l'expérience) en comparaison avec les algorithmes de la campagne IV² (3% d'EER en environnement contrôlé et entre 6,1% et 14,9% d'EER en environnement non contrôlé présentant des défis d'éclairage, de qualité et de session).

La deuxième approche bio-inspirée consiste à appliquer une méthode de codage par ordre de classement dans la phase de prétraitement pour renforcer le contenu informatif des images de visage. La méthode en question, appelée Reconstruction Asynchrone (RA), a été développée par Justin Plantier de l'Institut de Recherche Biomédicale des Armées, en France, dans le cadre d'études sur la compréhension du système visuel humain et a été appliquée en biométrie faciale pour la première fois dans le cadre de ce travail de thèse. Cette méthode consiste d'abord à décomposer le signal en fréquences spatiales à l'aide d'un filtre DoG. Puis, les réponses fréquentielles provenant de tous les canaux décomposés sont rangées dans un ordre descendant dans des canaux asynchrones. Chaque pixel est ensuite reconstruit en gardant les canaux donnant les réponses maximales et en ignorant les canaux de réponses les plus faibles, émulant ainsi le traitement bas niveau du système visuel humain. Ce processus est réalisé au niveau local et rend la méthode très intéressante dans la mesure où on ne garde pas les mêmes échelles de décomposition lors de la reconstruction des différents pixels d'une image donnée. La comparaison de la RA avec la reconstruction coarse-to-fine et d'autres techniques de prétraitement a montré l'intérêt de la RA.

Dans la quatrième partie de cette thèse, les meilleures méthodes monomodales par l'iris (Log-Gabor) et par le visage (LBP/DoG-bio) ont été associées pour aboutir à l'objectif final de ce sujet de thèse : la conception de méthodes multimodales. Deux contributions émanent de cette partie à travers l'élaboration de deux approches multimodales par le biais de techniques classiques de normalisation et de fusion de scores par combinaison. La première approche multimodale associe les échantillons de l'iris droit et du visage. Cette approche comprend la principale contribution de cette partie à travers la proposition d'une nouvelle règle de fusion par une somme pondérée faisant intervenir un critère de qualité basé sur le taux d'occultation de la comparaison des iris (ICOR : the Irises Comparison's Occultation Rate). L'application de cette règle de fusion se fait en ligne pendant la phase de test, mais nous avons également proposé de la coupler avec la somme pondérée usuelle dont les poids sont déterminés hors ligne lors de la phase d'apprentissage. La deuxième contribution consiste à avoir développé une approche multimodale à triple échantillons : l'iris droit, l'iris gauche et le visage. Les deux approches multimodales proposées ont été testées sur la base multimodale réelle IV^2 avec des expériences en environnements contrôlés et non contrôlés d'acquisition de visage. Plusieurs aspects de la multimodalité ont été discutés et l'intérêt du scénario multimodal par multi-biométries a été démontré par rapport aux biométries individuelles. Par ailleurs, nous avons illustré l'efficacité de la règle de fusion basée sur l'ICOR comparée aux règles de fusion usuelles dans l'approche [Iris droit / Visage] qui arrive à produire des taux d'égales erreurs faibles entre 0,3% et 1,2% selon la variabilité présente dans l'expérience. Nous avons également prouvé l'intérêt du scénario multiéchantillons à travers l'intégration des deux échantillons de l'iris en plus du visage dans l'approche multimodale [Iris droit / Visage / Iris gauche]. Cette approche réalise des résultats comparables voire meilleurs que ceux des algorithmes multimodaux à triple échantillons de la campagne d'évaluation IV² avec des taux d'égales erreurs nuls ou quasi-nuls entre 0 et 0,3% selon l'expérience.

Les perspectives de ce travail sont nombreuses notamment en biométrie bio-inspirée et en biométrie multimodale :

Tout d'abord, grâce à l'initiation au domaine de la neuroscience nous avons eu l'occasion d'élaborer des approches bio-inspirées s'appuyant sur des méthodes qui essayent de modéliser le fonctionnement du cerveau humain dans la tâche de la perception de visage. Nous avons l'intention d'explorer d'avantage ce domaine passionnant en élaborant d'autres approches bio-inspirées au niveau de l'extraction et de la sélection de caractéristiques biométriques basées par exemple sur une représentation parcimonieuse de l'information et intégrant des processus adaptatifs à rétroaction [Olshausen 1996] [Van Rullen 2003]. D'un autre côté, les expériences multimodales réalisées ont ouvert plusieurs perspectives. La considération de la signification statistique des résultats nous a déjà révélé la nécessité d'apporter des améliorations sur les protocoles de test par l'introduction de techniques de ré-échantillonnage de bases de données telles que la validation croisée ou le bootstrap. L'expérimentation sur des bases de données plus riches en nombre et en modalités telles que MBGC [NIST MBGC 2010] et Biosecure [Ortega-Garcia 2010] sera également envisagée. De plus, la fusion au niveau des scores peut se faire sur des bases de scores telles que la base NIST BSSR1 [NIST BSSR1 2011].

Par ailleurs, nous avons engagé une discussion au **Chapitre 4** autour de l'alternative du rejet des comparaisons suivant un critère de qualité. A l'issue de cette discussion, et dans le souci de remédier au nombre trop élevé d'éliminations qui risquent de compromettre la praticité du système biométrique multimodal, nous nous sommes orientés vers deux solutions possibles. La première consiste à intégrer les critères de qualité dans le processus de la reconnaissance tel que nous l'avions fait à travers la règle de fusion ICORws basée sur le taux d'occultation de la comparaison des iris. Ce critère reste encore à généraliser pour couvrir des méthodes faisant intervenir les deux échantillons de l'iris. D'autres critères de qualité liés à l'iris (le mauvais focus) ou au visage (l'orientation) ou aux deux à la fois (l'éclairage, le flou et l'énergie spectrale) sont à explorer [Bharadwaj 2014] [Phillips 2013]. Une deuxième solution pour exploiter les critères de qualité en vue d'amélioration de la performance de la reconnaissance sans pour autant avoir recours au rejet des comparaisons serait de travailler dans le contexte d'une architecture séquentielle dans laquelle les mesures de qualité serviraient à traiter les cas où l'une des modalités présenterait une signature de qualité insuffisante.

Une autre perspective intéressante de ce travail serait de prendre en compte la spécificité de chaque utilisateur en proposant des méthodes de normalisation/fusion de scores adaptées à chaque utilisateur à la place de simples règles de combinaison de scores qui restent identiques quelque soit l'individu. Cet axe de recherche a connu plusieurs contributions notamment par les travaux de [Jain 2002] [Fierrez 2006] [Poh 2013] et reste encore à explorer.

D'autre part, les approches biométriques proposées dans cette thèse ont été conçues pour opérer en mode vérification. Il serait intéressant de travailler en mode identification en utilisant des classifieurs plus complexes notamment ceux qui émanent de l'intelligence artificielle tels que les réseaux de neurones et les arbres décisionnels.

En conclusion de ce travail de recherche, nous pensons que la biométrie multimodale est la solution ultime pour la construction de systèmes de reconnaissance de haute précision en vue d'être déployés à grande échelle. Bien que plus lourde à mettre en œuvre car nécessitant plusieurs chaînes d'évaluations biométriques, la voie à suivre est d'exploiter pleinement les apports des diverses sources d'informations biométriques multimodales

Références bibliographiques

- [Acuity 2009] Acuity Market Intelligence. The Future of Biometrics Market Analysis, Segmentation & Forecasts, Insight into the Trends, Drivers & Opportunities that will Shape the Industry through 2020. Rapport technique. Août 2009. 3
- [Adelson 1985] E.H. Adelson, J.R. Bergen. Spatiotemporal energy models for the perception of motion. Journal of the Optical Society of America, 1(2): 284-299, 1985. 47
- [Albiol 2008] A. Albiol, D. Monzo, A. Martin, J. Sastre, A. Albiol. Face recognition using HOG-EBGM. Pattern Recogn. Lett., 29(10): 1537-1543, 2008. 79
- [Allano 2009] L. Allano. La Biométrie multimodale : stratégies de fusion de scores et mesures de dépendance appliquées aux bases de personnes virtuelles. Thèse de doctorat, Université d'Evry Val d'Essonne, Institut National des Télécommunications, 2009. 14, 165, 171, 189
- [Annis Fathima 2014] A. Annis Fathima, S.Vasuhi, N.T. Naresh Babu, V. Vaidehi, M.T. Teena. Fusion Framework for Multimodal Biometric Person Authentication System. IAENG International Journal of Computer Science, 41(1), 2014. 20
- [Bailly-Baillière 2003] E. Bailly-Baillière, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mari éthoz, J. Matas, K. Messer, V. Popovici, F. Porée, B. Ruiz, J.P. Thiran. *The BANCA Database and Evaluation Protocol*. Proc. 3rd Inter Conference on Audio-and Video-Based Biometric Person Authentication – AVBPA'03, Springer, LNCS(2688) : 625-638, Guildford, UK, 2003. 29, 93
- [Barbu 2013] T. Barbu. Unsupervised SIFT-based Face Recognition Using an Automatic Hierarchical Agglomerative Clustering Solution. Proceedia Computer Science, Elsevier, 22: 385-394, 2013. 76
- [Barlow 1961] H. B. Barlow. Possible principles underlying the transformation of sensory messages. Sensory Communication, 217-234, 1961. 115, 121
- [Bartlett 2002] M. Bartlett, J. Movellan, T. Sejnowski. Face recognition by independent component analysis. IEEE Transactions on Neural Networks, 13(6):1450-1464, 2002. 77
- [Baum 1966] L.E. Baum, T. Petrie. Statistical Inference for Probabilistic Functions of Finite State Markov Chains. The Annals of MathematicalStatistics, 37(6): 1554–1563, 1966. 78
- [Belhumeur 1996] P.N. Belhumeur, D.J. Kriegman. What is the set of images of an object under all possible lighting conditions ?. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Proceedings – CVPR'96, 270 – 277, Juin 1996. 86
- [Belhumeur 1997] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman. Eigenfaces vs fisherfaces: Recognition using class specific linear projection. IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), 19(7):711-720, Juillet 1997. 77, 91
- [Bell 1995] A. J. Bell, T. J. Sejnowski. An information maximization approach to blind separation and blind convolution. Neural Computation, 7: 1129-1159, 1995. 121
- [Ben Soltana 2010] W. Ben Soltana, D. Huang, M. Ardabilian, L. Chen, B.C Ben Amar. Comparison of 2D/3D Features and Their Adaptive Score Level Fusion for 3D Face Recognition. 3D Data Processing, Visualization and Transmission - 3DPVT, Paris, France. 2010. 22
- [Bharadwaj 2014] S. Bharadwaj, M. Vatsa, R. Singh. Biometric quality: a review of fingerprint, iris, and face. EURASIP Journal on Image and Video Processing. Springer. 2014(34). 2014. 204

- [Bhattacharyya 1943] A. Bhattacharyya. On a measure of divergence between two statistical populations defined by their probability distributions. Bulletin of the Calcutta Mathematical Society, 35: 99–109, 1943. 46
- [Bicego 2006] M. Bicego, A. Lagorio, E. Grosso, M. Tistarelli. On the Use of SIFT Features for Face Authentication. IEEE Conference on Computer Vision and Pattern Recognition Workshops, 17-22, Juin 2006. 76, 80
- [Bigun 1997] E.S. Bigun, J. Bigun, B. Duc, S. Fischer. Expert conciliation for multimodal person authentication systems using Bayesian statistics. Proceedings of First International Conference on AVBPA, Crans-Montana: 291–300, Switzerland, 1997. 21
- [Biosecure 2009] Biosecure, A biometric reference system for iris OSIRIS version 2.01, Octobre, 2009 [En ligne]. Disponible sur http: /svnext.it-sudparis.eu/svnview2eph/refsyst/IrisOsiris/doc/howto.pdf. 45
- [Birbaumer 2006] N. Birbaumer, R. F., Schmidt. Biologische Psychologie. 6th ed. Springer Berlin Heidelberg. 2006. 118
- [Blakemore 1969] C. Blakemore, F. W. Campbell. On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. Journal of Physiology, 203: 237-260, Londres, 1969. 115
- [Boles 1998] W. Boles, B. Boashash. A human identification technique using images of the iris and wavelet transform. IEEE Trans. on Signal Processing, 46(4):1185–1188, Avril 1998. 43
- [Bolle 2000] R. M. Bolle, S. Pankanti, N. K. Ratha. Evaluation techniques for biometrics-based authentication systems (FRR). Proc. of Int. Conf. on Pattern Recognition. 2:831-837. 2000. 205
- [Bolle 2004] R. M. Bolle, N. K. Ratha, S. Pankanti. Error analysis of pattern recognition systems the subsets bootstrap. Computer Vision and Image Understanding, 93(1):133, Janvier 2004. 205
- [Bötzel 1995] K. Bötzel, S.Schulze, S. R. Stodieck, Scalp topography and analysis of intracranial sources of face-evoked potentials. Exp. Brain. Res., 104(1): 135-143, 1995. 118
- [Bowyer 2008] K.W. Bowyer, K.P. Hollingsworth, P.J. Flynn. ImageUnderstanding for Iris Biometrics: ASurvey. Computer Vision and Image Understanding, Elsevier, 110(2): 281-307, New York, May, 2008. 45
- [Bowyer 2013] K.W. Bowyer, K.P. Hollingsworth, P.J. Flynn. A survey of iris biometrics research: 2008-2010. Handbook of Iris Recognition, Springer, Burge & Bowyer, 2013. 45
- [Breiman 2001] L. Breiman. Random Forests. Machine Learning, 45(1): 5-32. Octobre 2001. 79
- [Briand 2014] T. Briand, J. Vacher, B. Galerne, J. Rabin. The Heeger & Bergen Pyramid Based Texture Synthesis Algorithm. IPOL: Image Processing On Line, 2014. 4 (2014) : 276–299, 2014. 81
- [Campbell 1968] F. W. Campbell, J. G. Robson. Application of Fourier analysis to the visibility of gratings. Journal of Physiology, 197: 551-566, 1968. 122
- [CASIA-V1 2007] CASIA-V1 [En ligne]. Specification of CASIA Iris Image Database(ver 1.0), Chinese Academy of Sciences, http://www.nlpr.ia.ac.cn/english/irds/irisdatabase.htm, Mars 2007. 56
- [CASIA-IrisV1 2010] CASIA-IrisV1 [En ligne]. Disponible sur http://biometrics.idealtest.org/dbDetailForUser.do?id=1. Chinese Academy of Sciences' Institute of Automation (CASIA). 2010. 56
- [Chaari 2009] A. Chaari. Nouvelle approche d'identification dans les bases de données biométriques basée sur une classification non supervisée. Thèse de doctorat : Sciences pour l'ingénieur et Informatique. Evry : Université d'Evry-Val d'Essonne, Tunis : Université de la Manouba. 2009. 76, 110

- [Chelali 2014] F.Z. Chelali, A. Djeradi. Face Recognition Using MLP and RBF Neural Network with Gabor and Discrete Wavelet Transform Characterization: A Comparative Study. Mathematical Problems in Engineering, Hindawi Publishing Corporation, 2014. 75, 78, 80
- [Chen 2000] L.F. Chen, H.Y. Mark Liao, M.T. Ko, J.C. Lin, G.J. Yu. A new LDA-based face recognition system which can solve the small sample size problem. Pattern Recognition, 33(10): 1713–1726, Octobre 2000. 76
- [Chen 2014] Y. Chen, Y. Liu, X. Zhu, H. Chen, F. He, Y. Pang. Novel Approaches to Improve Iris Recognition System Performance Based on Local Quality Evaluation and Feature Fusion. The Scientific World Journal, 2014. 46, 73
- [Chew 2008] L. W. Chew, L. M. Ang, K. P. Seng. Survey of Image Compression Algorithms in Wireless Sensor Networks. Int. Symp. On Information Technology, 4, Kuala Lumpur, Août 2008. 128
- [Chibelushi 1999] C. Chibelushi, S. Gandon, J. Mason, F. Deravi, R. Johnston. Design issues for a Digital Integrated Audio-Visual Database. Proc. IEE Colloquium on IAVPRSC, 1999. 29
- [Colineau 2007] J. Colineau, J. d'Hose. Evaluation d'algorithmes biométriques : Détection de la position des yeux, Iris, Visage 2D, Visage 3D, Multimodalité. Rapport technique final du Projet IV², Novembre 2007. 32, 35, 194, 196, 210
- [Connaughton 2012] R. Connaughton. Fusion of face and iris biometrics using a stand-off video sensor. Mémoire partiel de thèse, Université de Notre Dame, Indiana, USA. 2012. 161, 163
- [Cootes 2004] T.F. Cootes, C.J. Taylor. Statistical Models of Appearance for Computer Vision. Technical report, Univ. of Manchester, 2004. 77
- [Cortes 1995] C. Cortes, V. Vapnik. Support-Vector Networks. Machine Learning, 20(3): 273-297. September 1995. 78
- [Crihalmeanu 2007] S. G. Crihalmeanu, A. Ross, S. Schuckers, L. Hornak. A Protocol for Multibiometric Data Acquisition, Storage and Dissemination. Rapport technique. West Virginia University, Lane Department of Computer Science and Electrical Engineering, 2007. 29
- [Dantone 2012] M. Dantone, J. Gall, G. Fanelli, L. Van Gool. Real-time Facial Feature Detection using Conditional Regression Forests. IEEE Conference on Computer Vision and Pattern Recognition - CVPR'12: 2578-2585, Rhode Island, USA. 2012. 78
- [Daugman 1993] J. Daugman. High confidence visual recognition of persons by a test of statistical independence. IEEE Trans. on Pattern Analysis and Machine Intelligence, 15(11): 1148-1161. Novembre 1993. 41, 54, 62, 63
- [Daugman 1994] J. Daugman. Biometric personal identification system based on iris analysis. U.S. Patent No. 5,291,560, Mars 1994. 41
- [Daugman 2001] J. Daugman. Statistical Richness of Visual Phase Information: Update on Recognizing Persons by Iris Patterns. International Journal of Computer Vision, 45(1): 25-38. Octobre 2001. 39, 41
- [Daugman 2004] J. Daugman. How Iris RecognitionWorks. IEEE Trans. on Circuits and Systems for Video Technology, 14(1): 21-30, Janvier 2004. 42
- [Daugman 2006] J. Daugman. Probing the Uniqueness and Randomness of Iris Codes: Results From 200 Billion Iris Pair Comparisons. Proc. of the IEEE, 1927-1935, Novembre 2006. 41, 46, 73
- [Daugman 2007] J. Daugman. New Methods in Iris Recognition. IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics, 37(5): 1167-1175, 2007. 32, 45
- [Dinerstein 2007] S. Dinerstein, J. Dinerstein, D. Ventura. Robust multi-modal biometric fusion via multiple SVMs. IEEE Int. Conf. on Systems, Man and Cybernetics. 1530-1535. Montréal, 2007. 20
- [Delorme 1999] A. Delorme, J. Gautrais, R. Van Rullen, S. J. Thorpe. SpikeNET: A simulator for modeling large networks of integrate and fire neurons. Neurocomputing, 26-27: 989-996,1999. 128

- [Dorizzi 2006] B. Dorizzi, S. Garcia-Salicetti, L. Allano. Multimodality In Biosecure : Evaluation On Real Vs. Virtual Subjects. In : Proceedings of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing – ICASSP, 2006. 27
- [Draper 2003] B.A. Draper, K. Baek, M. S. Bartlett, J.R. Beveridge. Recognizing faces with PCA and ICA. Computer Vision and Image Understanding, Special Issue on Face Recognition, 91(1-2): 115-137, Juillet-Août 2003. 76, 89, 91
- [Dubuc 2013] B. Dubuc, P. Robert, D. Paquet, A. Daigen. Le cerveau à tous les niveaux Les détecteurs sensoriels la vision, 2013. [En ligne]. Disponible sur http://lecerveau.mcgill.ca/flash/a/a_02/a_02_cl_vis/a_02_cl_vis/a_02_cl_vis.html. 116
- [Dumas 2005] B. Dumas, C. Pugin, J. Hennebert, D. Petrovska-Delacretaz, A. Humm, F. Evequoz, R. Ingold, D. Von Rotz. MyIdea - multimodal biometrics database, description of acquisition protocols. Proc. COST-275 Workshop on Biometrics on the Internet, Rapport technique, 2005. 29
- [El Aroussi 2009] M. El Aroussi, M. El Hassouni, S. Ghouzali, M. Rziza, D. Aboutajdine. Novel face recognition approach based on steerable pyramid feature extraction. IEEE Int. Conf. On Image Processing – ICIP, 4165 – 4168, Caire, Novembre 2009. 110
- [El Aroussi 2011] M. El Aroussi, M. El Hassouni, S. Ghouzali, M. Rziza, D. Aboutajdine. Local appearance based face recognition method using block based steerable pyramid transform. Signal Processing, Elsevier, 91(1): 38–50, 2011. 81
- [El Ghanaoui 2013] A. El Ghanaoui, N. Khiari-Hili, C. Montagne, S. Lelandais : Bio-Inspired Face Authentication using Multiscale LBP. Int. Conf. Bio-inspired Syst. and Signal Process Barcelone, 2013. 136, 138, 140
- [Enroth-Cugell 1966] C. Enroth-Cugell, J. G. Robson. The contrast sensitivity of retinal ganglion cells of the cat. Journal Physiol. 187: 517-552, 1966. 123
- [Fakhar 2015] K. Fakhar, M.El Aroussi, M.N. Saidi, D. Aboutajdine. Applying the Upper Integral to the Biometric Score Fusion Problem in the Identification Model. Information. 6(3): 494-504, 2015. 20
- [FanC 2011] N. FanC., Y. ZhangF. Homomorphic filtering based illumination normalization method for face recognition. Journal Pattern Recognition Letters. 32: 1468-1479, 2011. 148
- [Fisher 1936] R. A. Fisher. The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics, 7 (2): 179–188, 1936. 85
- [Fierrez 2006] J. Fierrez-Aguilar. Adapted Fusion Schemes for Multimodal Biometric Authentication. Mémoire de these. Université Polytecnique de Madrid, 2006. 16, 17, 204
- [Fierrez 2007] J. Fierrez, J. Ortega-Garcia, D. Torre Toledano, J. Gonzalez-Rodriguez. Biosec baseline corpus : A multimodal biometric database. Pattern Recogn., 40(4) :1389-1392, 2007. 29
- [Fierrez 2010] J. Fierrez, J. Galbally, J. Ortega-Garcia, M. R. Freire, F. Alonso-Fernandez, D. Ramos, D. T. Toledano, J. Gonzalez-Rodriguez, J. A. Siguenza, J. Garrido-Salas, E. Anguiano, G. Gonzalez-de-Rivera, R. Ribalda, M. Faundez-Zanuy, J. A. Ortega, V. Cardeñoso-Payo, A. Viloria, C. E. Vivaracho, Q. I. Moro, J. J. Igarza, J. Sanchez, I. Hernaez, C. Orrite-Uruñuela, F. Martinez-Contreras, J. J. Gracia-Roche. *BiosecurID: a multimodal biometric database*. In Pattern Analysis and Applications, 13(2): 235-246, Mai 2010. 29
- [Field 1987] D.J. Field. Relations between the statistics of natural images and theresponse properties of cortical cells. Journal of Optical Society of America, 4(12), Décembre 1987. 44, 52, 53, 54
- [Fiset 2009] D. Fiset., F. Gosselin, . L'information visuelle efficace pour la reconnaissance des visages. In E. Barbeau, S. Joubert, & O. Felician (Eds.), Traitement et reconnaissance des visages: du percept à la personne, 145:164, 2009. 125
- [Flom 1987] L. Flom, A. Safir. Iris recognition system. U.S. Patent 4,641,349, 1987. 41

- [Freund 1995] Y. Freund, R.E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. Computational Learning Theory: Eurocolt '95, Springer, 23-37, 1995. 87
- [Garcia-Salicetti 2003] S. Garcia-Salicetti, Charles Beumier, Gérard Chollet, Bernadette Dorizzi, Jean Leroux les Jardins, Jan Lunter, Yang Ni, Dijana Petrovska-Delacrétaz. *BIOMET: A multimodal person authentication database including face, voice, fingerprint, hand and signature modalities.* In AVBPA, 845-853, 2003. 29
- [Ghosal 2009] V. Ghosal, P. Tikmani, P. Gupta. Face Classification Using Gabor Wavelets and Random Forest. Canadian Conference on Computer and Robot Vision, 68-73, 2009. 79, 80
- [Gilbert 2007] C.D. Gilbert, M. Sigman. Brain states: top-down influences in sensory processing. Neuron 54: 677-696, 2007. 122
- [Giraudet 2000] G. Giraudet, C. Roumes, J. Plantier. *Diagnostic spatial frequency for object recognition in natural images*. Perception, 29: 82-83, 2000. 142
- [Gosselin 2001] F. Gosselin, P.G. Schyns. Bubbles: a technique to reveal the use of information in recognition tasks. Vision Research, Elsevier. 41:2261–2271, 2001. 125, 126, 129, 136, 156
- [Gu 2005] H. Gu, Y. Zhuang, Y. Pan. An iris recognition method based on multi-orientation features and nonsymmetrical SVM. Journal of Zhejiang University Science, 6A(4):428–432, 2005. 44
- [Gumus 2010] E. Gumus, N. Kilic, A. Sertbas, O.N. Ucan. Evaluation of face recognition techniques using PCA, wavelets and SVM. Expert Systems with Applications, 37(9): 6404– 6408, Septembre 2010. 75, 80
- [Hadid 2008] A. Hadid, G. Zhao, T. Ahonen, M. Pietikäinen. Face Analysis using Local Binary Patterns. In: M. Mirmehdi, X. Xie, J. Suri (ed.) Handbook of Texture Analysis, 347-373. Imperial College Press, Londres, 2008. 75, 80
- [Hampel 1986] F. Hampel, P. Rousseeuw, E. Ronchetti, W. Stahel. *Robust Statistics : The* Approach Based on Influence Functions. John Wiley & Sons, 1986. 24
- [Harwood 1993] D. Harwood, T. Ojala, M. Pietikäinen, S. Kelman, S. Davis. Texture classification a by center-symmetric auto-correlation, using Kullback discrimination of distributions. Rapport technique. Computer Vision Laboratory, Center for Automation Research, University of Maryland, College Park, Maryland. CAR-TR-678, 1993. 75, 130
- [Haxby 2000] J.V. Haxby, E.A. Hoffman, M.I. Gobbini. The distributed human neural system for face perception. Trends in Cognitive Sciences, 4(6): 223-233, 2000. 119
- [Haxby 2010] J.V Haxby, M.I. Gobbini. Distributed neural systems for face perception. In A. Calder, G. Rhodes, M. Johnson, J.V. Haxby (Eds), Oxford handbook of face perception, Oxford University Press, 93-110. New York, 2010. 119
- [Hérault 1985] J. Hérault, C. Jutten, B. Ans. Détection de Grandeurs Primitives dans un Message Composite par une Architecture de Calcul Neuromimétique en Apprentissage non Supervisé. In: Actes du 10^{ème} colloque GRETSI, 1017-1022, 1985. 121
- [Heseltine 2005] T. D. Heseltine. Face Recognition: Two-Dimensional and Three-Dimensional Techniques. Thèse de doctorat, The University of York, Department of Computer Science, 2005. 148
- [Hollingsworth 2011] K.P. Hollingsworth, K.W. Bowyer, P.J. Flynn. Improved iris recognition through fusion of hamming distance and fragile bit distance. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(12): 2465–2476, 2011. 45
- [Huang D. 2011] D. Huang, C. Shan, M. Ardabilian, Y. Wang, L. Chen. Local Binary Patterns and Its Application to Facial Image Analysis: A Survey. IEEE Trans. On Systems, Man, and Cybernetics, 41(6):765-781, 2011. 75, 76, 129, 130, 131, 133, 134, 135
- [Huang H. 2012] H. Huang, H. Liu, G. Liu. Face Recognition Using Pyramid Histogram of Oriented Gradients and SVM. Advances in information Sciences and Service Sciences – AISS, 4(18), Octobre 2012. 78

- [Huang Y. 2011] Y. Huang, K. Huang, D. Tao, T. Tan, X. Li. Enhanced biologically inspired model for object recognition. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics 41:1668-1680, 2011. 128
- [Hubel 1959] D. H. Hubel, T. N. Wiesel. Receptive fields of single neurons in the cat's striate cortex. Journal of Physiology, 148: 574-591, 1959. 115
- [Huber 1981] P. Huber. Robust Statistics. JohnWiley & Sons, 1981. 22
- [IREX II IQCE 2011] IREX II IQCE. Iris Quality Calibration and Evaluation: Performance of Iris Image Quality Assessment Algorithms. NIST Interagency Report 7820. 2011 [En ligne]. Disponible sur http://www.nist.gov/itl/iad/ig/irexii.cfm. 66
- [Jain 1999] A. K. Jain, R. Bolle, S. Pankanti, editors. Biometrics: Personal Identification in Networked Society. Kluwer Academic Publishers, 1999. 3
- [Jain 2002] A.K. Jain, A. Ross. Learning user-specific parameters in a multibiometric system. Proceedings of Int. Conf. on Image Processing, 57–60, New York, 2002. 22, 204
- [Jain 2004] A.K. Jain, A. Ross, S. Parabhakar. An Introduction to Biometric Recognition. IEEE Trans. on Circuits And Systems For Video Technology, 14(1), Janvier 2004. 4, 13, 16
- [Jain 2005] A.K. Jain, K. Nandakumar, A. Ross. 2005. Score Normalization in Multimodal Biometric Systems. Pattern Recognition, 38: 2270-2285. 18, 19, 21, 22, 23, 25, 169, 172
- [Jain 2007] A. K. Jain, P. Flynn, A. Ross. Handbook of Biometrics. Springer, 2007. 2, 9
- [Jiang 2015] Y. Jiang, Y. Shen, Y. Liu, W. Liu. Multiclass AdaBoost ELM and Its Application in LBP Based Face Recognition. Mathematical Problems in Engineering, Hindawi Publishing Corporation, 2015. 76, 80, 87
- [Jumahong 2011] H. Jumahong, W. Liu, C. Lu. A new rearrange modular two-dimensional LDA for face recognition. IEEE, Int. Conf. on Machine Learning and Cybernetics – ICMLC, 1: 361-366, 2011. 79
- [Kanade 1973] T. Kanade. Picture Processing System by Computer Complex and Recognition of Human Faces. Thèse de doctorat, Kyoto University, Novemvre 1973. 75
- [Kanade 2012] A. Kanade, A. Khare. Face Recognition: Based on Neural Network Approach. Int. Journal of Engineering and Innovative Technology (IJEIT), 1(5), 2012. 78
- [Karasaridis 1996] A. Karasaridis, E.P. Simoncelli. A Filter Design Technique for Steerable Pyramid Image Transforms. IEEE, Inter. Conf. on Acoustics, Speech and Signal Processing (ICASSP), 4: 2387-2390, Atlanta, Mai 1996. 47, 49
- [Khiari 2008] N. Khiari, H. Mahersia, K. Hamrouni. Iris Recognition Using Steerable Pyramids. 1st IEEE Int. Conf. on Image Processing, Theory, Tools and Applications – IPTA'08, Sousse, 2008. 6, 39, 54, 61, 73
- [Khiari 2011] N. Khiari-Hili, S. Lelandais, C. Montagne, K. Hamrouni. Reconnaissance de visages 2D sur la base IV² par les Pyramides orientables et la LDA. 7^{ème} atelier de travail sur le Traitement et Analyse de l'Information : Méthodes et Applications (TAIMA), Hammamet, 2011. 94, 99
- [Khiari 2012-a] N. Khiari-Hili, S. Lelandais, C. Montagne, K. Hamrouni. A Combining Approach for 2D Face Recognition Application on IV² Database. 3rd Int. Conf. on Image Processing Theory, Tools and Applications – IPTA. Istanboul, 2012. 94, 95, 105
- [Khiari 2012-b] N. Khiari-Hili, S. Lelandais, C. Montagne, K. Hamrouni. Boosting Steerable Features For 2d Face Recognition On Iv² Database. Biosignals 2012, 2nd Int. Special Session on Multivariable Processing for Biometric Systems (MPBS). Vilamoura, 2012. 94, 102
- [Kittler 1998] J. Kittler, M. Hatef, R.P. Duin, J.G. Matas. On combining classifiers. IEEE Trans. PAMI, 20(3): 226–239, 1998. 20, 172
- [Kovesi 2015] P. Kovesi. What Are Log-Gabor Filters and Why Are They Good? [En ligne]. Disponible sur http://www.csse.uwa.edu.au/~pk/research/matlabfns/. Page citée en 2003. 53

- [Krichen 2007] E. Krichen, S. Garcia-Salicetti, B. Dorizzi. A New Probabilistic Iris Quality Measure for Comprehensive Noise Detection. IEEE Conf. on Biometrics: Theory, Applications and Systems – BTAS, Washington DC, Septembre 2007. 32
- [Krichen 2009] E. Krichen, S. Garcia-Salicetti, B. Dorizzi. A New Phase-Correlation-Based Iris Matching for Degraded Images. IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics, 39 (4): 924-934, Août 2009. 45, 46, 73
- [Kuffler 1953] S. W. Kuffler. Discharge patterns and functional organization of mammalian retina. Journal of Neurophysiology, 16, 37-68, 1953. 115
- [Kumar 2014] S. G. Kumar, J. Devi. A Multimodal SVM Approach for Fused Biometric Recognition. Int. Journal of Computer Science and Information Technologies – IJCSIT, 5(3): 3327-3330, 2014. 20
- [Lawrence 1997] S. Lawrence, C.L. Giles, Ah Chung Tsoi, A.D. Back. Face recognition: a convolutional neural-network approach. IEEE Trans. on Neural Networks, 8(1): 98-113, 1997. 78
- [Lease 2009] D. R. Lease, R. Guess, S. Lovaas, E. Solveggio. Biometric Authentication. In S. Bosworth and M. E. Kabay (Ed.), Computer Security Handbook (Chapter 29), 4th Edition, John Wiley, New Jersy, 2009. 2
- [Lee 2013] Y. Lee, D.K. Han, H. Ko. Reinforced AdaBoost Learning for Object Detection with Local Pattern Representations. The Scientific World Journal, 2013. 76, 87
- [Lelandais 2007] S. Lelandais. Rapport final du projet IV² : Identification par l'Iris et le Visage via la Vidéo. Rapport technique. Novembre 2007. 31, 33
- [Li 2012] P. Li, X. Liu, N. Zhao. Weighted Co-occurrence Phase Histogram for Iris Recognition. Pattern Recognition Letters, 33(8):1012-1018. June 2012. 46, 73
- [Lim 2001] S. Lim, K. Lee, O. Byeon, T. Kim. Efficient Iris Recognition through Improvement of Feature Vector and Classifier. ETRI Journal, 23(2): 61–70, 2001. 44
- [Liu 2005] C. Liu. Robust face detection using discriminating feature analysis and Bayes classifier. Proc. SPIE 5779, Biometric Technology for Human Identification II, 164, April 05, 2005. 76
- [Logothetis 1999] N. Logothetis. Vision : A window on consciousness. Scientific American, 281: 44-51, 1999. 117
- [López 2010] L. S. López. Local Binary Patterns applied to Face Detection and Recognition. Rapport technique. Novembre 2010. 137
- [Lowe 2004] D.G. Lowe. Distinctive image features from scale-invariant keypoints. Int. Journal Comput. Vision 60: 91–110, 2004. 76, 45
- [Mäenpää 2003] T. Mäenpää, M. Pietikäinen. Multi-scale Binary Patterns for Texture Analysis. In J. Bigun, T. Gustavsson (eds) SCIA 2003. LNCS, 2749:885-892. Springer, Heidelberg, 2003. 131
- [Manvelyan 2015] S. Manvelyan. Gallery Eyes Your beautiful eyes [En ligne]. Disponible sur http://www.surenmanvelyan.com/eyes/your-beautiful-eyes/?wppa-album=5&wppacover=0&wppa-occur=1. Consultée en 2015. 38
- [Marr 1976] D. Marr, T. Poggio. Cooperative computation of stereodisparity. Science, 194(4262): 283-287, 1976. 143
- [Marr 1979] D. Marr, S. Ullman. Directional selectivity and its use in early visual processing. MIT AI, Memo 524, 1979. 47, 123, 124
- [Marr 1980] D. Marr, E. Hildreth. Theory of Edge Detection. Proc. of the Royal Society of London, Series B, Biological Sciences, 207(1167): 187-217, 1980. 115, 124
- [Marr 1982] D. Marr. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. Freeman, New York, 1982. 114

- [Martin 1997] A. Martin, G. Doddington, T. Kamm, M. Ordowsk, M. Przybocki. The DET Curve in Assessment of Detection Task Performance. In Proc. Eurospeech'97, 1895–1898, Rhodes, 1997. 12
- [Martinez 2001] A.M Martinez, A.C Kak. *PCA versus LDA*. Pattern Analysis and Machine Intelligence, 23 :228-233, 2001. 91
- [Masek 2003] L. Masek. Recognition of human iris patterns for biometric identification. Rapport de Master, University of Western Australia, 2003, [En ligne]. Disponible sur http://www.csse.uwa.edu.au/~pk/studentprojects/libor/. Page mise à jour le 30 Novembre 2003. 44, 45, 46, 47, 54, 56, 60, 61, 62, 63, 65, 73
- [Masmoudi 2010] K. Masmoudi, M. Antonini, P. Kornprobst, L. Perrinet. A novel bio-inspired static image compression scheme for noisy data transmission over low-bandwidth channels. 35th ICASSP, Dallas, 2010. 128
- [Maurer 2008] D. E. Maurer, J. P. Baker. Fusing multimodal biometrics with quality estimates via a bayesian belief network. In Pattern Recognition, 41: 821-832. Elsevier Science Inc., New York, Mars 2008. 162
- [Mayr 2005] C. Mayr, R. Schüffny. Image Pulse Coding Scheme applied to Feature Extraction. Proc. Image and Vision Computing, Dunedin, 2005, 49-54. 128
- [McCool 2012] C. McCool, S. Marcel, A. Hadid, M. Pietikäinen, P. Matějka, J. Černocký, N. Poh, J. Kittler, A. Larcher, C. Lévy, D. Matrouf, J.F. Bonastre, P. Tresadern, T. Cootes. *Bi-Modal Person Recognition on a Mobile Phone: using mobile phone data.* in IEEE ICME Workshop on Hot Topics in Mobile Mutlimedia, 2012. 29
- [McCulloch 1943] W.S. McCulloch, P. Walter. A Logical Calculus of Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics 5 (4): 115–133. Décembre 1943. 78
- [Mellakh 2006] M.A. Mellakh, D. Petrovska-Delacretaz, B. Dorizzi. Using signal/residual information of eigenfaces for pca face space dimensionality characteristics. 18th Int. Conf. on Pattern Recognition (ICPR), 574-577, 2006. 89
- [Mellakh 2009-a] M.A. Mellakh. Reconnaissance des visages en conditions dégradées. Thèse de doctorat : Sciences et Ingénierie. Institut National des Télécommunications, Université d'Evry-Val d'Essonne, 2009. 75, 79, 110, 111
- [Mellakh 2009-b] A. Mellakh, A. Chaari, S. Guerfi, J. Dhose, J. Colineau, S. Lelandais, D. Petrovska-Delacrétaz, B. Dorizzi. 2D Face Recognition in the IV² Evaluation Campaign. Advanced Concepts for Intelligent Vision Systems, Springer, LNCS 5807: 24-32, 2009. 34, 83, 109, 140
- [Messer 1999] K. Messer, J. Matas, J. Kittler, J. Luettin, G. Maitre. XM2VTSDB: The extended M2VTS database. Proc. 2nd Int. Conf. on Audio and Video-Based Biometric Person Authentication, Washington, 1999. 29
- [Moghaddam 1996] B. Moghaddam, C. Nastar, A. Pentland. A Bayesian similarity measure for direct image matching. 13th Int. Conf. on Pattern Recognition, 2: 350–358, 1996. 77
- [Moon 1998] H. Moon, P.J. Phillips. Analysis of PCA-based face recognition algorithms. Empirical Evaluation Techniques in Computer Vision, 1998. 91
- [Morizet 2008] N. Morizet, J. Gilles. A new adaptive combination approach to score level fusion for face and iris biometrics combining wavelets and statistical moments. In Proc. of the 4th Int. Symposium on Advances in Visual Computing - ISVC'08, Part II, 661-671, Springer-Verlag, Heidelberg, Berlin, 2008. 162
- [Morizet 2009] N. Morizet. Reconnaissance Biométrique par Fusion Multimodale du Visage et de l'Iris. Mémoire de thèse. Ecole Nationale Supérieure des Télécommunications (Télécom ParisTech). 2009. 30, 162, 165, 182, 189
- [Mouchetant 2000] Y. Mouchetant-Rostaing, M. H. Giard, S. Bentin, P. E. Aguera, J. Pernier, Neurophysiological correlates of face gender processing in humans. Eur J Neurosci, 12(1), 303-310, 2000. 119

- [Murakami 2015] T. Murakami, K. Takahashi. Information-theoretic performance evaluation of likelihood-ratio based biometric score fusion under modality selection attacks. IEEE 7th Int. Conf. on Biometrics Theory, Applications and Systems – BTAS'15, 1–8, 2015. 21
- [NIST BSSR1 2011] *NIST BSSR1* [En ligne]. Disponible sur http://www.nist.gov/itl/iad/ig/biometricscores.cfm. Page créée le 7 Novembre 2011. 204
- [NIST ICE 2011] ICE Iris Challenge Evaluation [En ligne]. Disponible sur http://www.nist.gov/itl/iad/ig/ice.cfm. Page créée le 24 Janvier 2011. 32
- [NIST FRGC 2006] *NIST FRGC* [En ligne]. Disponible sur http://www.nist.gov/itl/iad/ig/frgc.cfm. 2006. Page créée le 7 Décembre 2010. 29
- [NIST MBGC 2009] NIST Portal Challenge Problem: Multiple Biometric Grand Challenge Preliminary Results of Version 2, 2009. Dispinible sur http://biometrics.nist.gov/cs_links/face/mbgc/2009/PORTAL_V2_FINAL.pdf. 182
- [NIST MBGC 2010] *NIST MBGC* [En ligne]. Disponible sur http://www.nist.gov/itl/iad/ig/mbgc.cfm. Page créée le 1 Décembre 2010. 29, 204
- [Ojala 1996] T. Ojala, M. Pietikinen, D. Harwood. A comparative study of texture measures a with classification based on feature distributions. Pattern Recognition 29: 51-59. 1996. 75, 130
- [Ojala 2002] T. Ojala, M. Pietikäinen, T. Mäenpää, 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. In IEEE Trans. Pattern Analysis and Machine Intelligence, 24(7), 971-987. 130, 131
- [Olshausen 1996] B. Olshausen, D. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature, 381(6583): 607-609, 1996. 121, 122, 128, 203
- [Olshausen 1997] B. Olshausen, D. Field. Sparse coding with an overcomplete basis set : a strategy employed by V1. Vision Research, 37: 3311-3325, 1997. 121
- [Oppenheim 1981] A. Oppenheim, J. Lim. The importance of phase in signals. In Proc. of the IEEE 69(5), 529-541, 1981. 63
- [Ortega-Garcia 2003] J. Ortega-Garcia, J. Fierrez-Aguilar, D. Simon, J. Gonzales, M. Faundez-Zanuy, V. Espinosa, A. Satue, I. Hernaez, J. J. Igarza, C. Vivaracho, D. Escudero, Q. Isaac Moro. MCYT baseline corpus : a bimodal biometric database. IEEE Proc.-Vis. Image Signal Process., 150(6) :395401, Décembre 2003. 29
- [Ortega-Garcia 2010] J. Ortega-Garcia, J. Fierrez, F. Alonso-Fernandez, J. Galbally, M.R. Freire, J. Gonzalez-Rodriguez, C. Garcia-Mateo, J-L Alba-Castro, E Gonzalez-Agulla, E. Otero-Muras, S. Garcia-Salicetti, L. Allano, B. Ly-Van, B. Dorizzi, J. Kittler, T. Bourlai, N. Poh, F. Deravi, M. Ng, M. Fairhurst, J. Hennebert, A. Humm, M. Tistarelli, L. Brodo, J. Richiardi, A. Drygajlo, H. Ganster, F.M. Sukno, S.K. Pavani, A. Frangi, L. Akarun, A. Savran. *The Multiscenario Multienvironment BioSecure Multimodal Database (BMDB)*. IEEE Trans. on Pattern Analysis and Machine Intelligence. 32(6):1097-111, Juin 2010. 29, 204
- [Osuna 1997] E. Osuna, R. Freund, F. Girosi. Training support vector machines: an application to face detection. Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'97): 130 - 136. San Juan, Puerto Rico, 1997. 78
- [Patel 2012] R. Patel, N. Rathod, A. Shah. Comparative Analysis of Face Recognition Approaches: A Survey. Int. Journal of Computer Applications, 55(17): 50-61. Novembre 2012. 78
- [Penev 2000] P Penev, L Sirovich. *The global dimensionality of face space*. Face and Gesture, 264-270, 2000. 89
- [Pentland 1994] A. Pentland, B. Moghaddam, T. Starner. View-based and modular eigenspaces for face recognition. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. Proceedings - CVPR'94: 84-91, 1994. 79
- [Petrovska-Delacretaz 2008] D. Petrovska-Delacretaz, S. Lelandais, J. Colineau, L. Chen, B. Dorizzi, M. Ardabilian, E. Krichen, M.A. Mellakh, A. Chaari, S. Guerfi, J. D'Hose, B. BenAmor. The IV2 Multimodal Biometric Database (Including Iris, 2D, 3D, Stereoscopic, and Talking Face Data), and the IV2- 2007 Evaluation Campaign. 2nd IEEE Inter. Conf. on

Biometrics: Theory, Applications and Systems (BTAS), Washington DC, USA, Septembre-Octobre 2008. 29, 30, 33, 64

- [Phillips 2005] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, W. Worek. Overview of the face recognition grand challenge. in Proc. of the 2005 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition – CVPR'05, 1: 947-954, Washington, DC,2005. 29
- [Phillips 2008] P. J. Phillips, K. W. Bowyer, P. J. Flynn, X. Liu, W. T. Scruggs. The Iris Challenge Evaluation 2005. IEEE 2nd Int. Conf. on Biometrics: Theory, Applications and Systems (BTAS 08), Arlington, 2008. 32
- [Phillips 2009] P. J. Phillips, P. J. Flynn, J. R. Beveridge, W. T. Scruggs, A. J. O'Toole, D. Bolme, K. W. Bowyer, B. A. Draper, G. H. Givens, Y. Man-Lui, H. Sahibzada, J. A. Scallan III, S. Weimer. Overview of the Multiple Biometrics Grand Challenge. Advances in Biometrics, Lecture Notes in Computer Science, 5558(2009): 705-714. 2009. 29, 182
- [Phillips 2013] P.J. Phillips, J.R. Beveridge, D.S. Bolme, B.A. Draper, G.H. Given, Y.M. Lui, S. Cheng, M.N. Teli, H. Zhang. On the existence of face quality measures. In Proc. of IEEE Int. Conf. on Biometrics: Theory, Applications and Systems, 1-8, Washington DC September 2013. 204
- [Pietikäinen 2000] M. Pietikäinen, T. Ojala, Z. Xu. Rotation-invariant texture classification using feature distributions. In Pattern Recognition 33: 43–52, 2000. 131
- [Plantier 1992] J. Plantier, J. P. Menu. Visual contrast and image quantification from wavelet theory. In 14th Annu. Int. Conf. of IEEE Eng. in Med. and Biol. Soc., 4: 1586-1587, Paris, 1992. 141, 142, 149
- [Plantier 2002] J. Plantier, L. Boutté, S. Lelandais. Defect detection on inclined planes using the shape from texture method and the Delaunay triangulation. EURASIP Journal on App. Sig. Proc., Special Issue on Applied Visual Inspection, (7): 659-666, 2002. 128
- [Plantier 2012] J. Plantier, S. Buffat, C. Roumes. Background influence on object detection in natural scenes. European Conf. on Visual Perception, Aghero, September 2012. 141, 158
- [Poh 2005] N. Poh, S. Bengio. Can chimeric persons be used in multimodal biometric authentication experiments ? In 2nd Int. Machine Learning and Multimodal Interaction Workshop (MLMI'05), LNCS 3869: 87:100, 2005. 27
- [Poh 2006] N. Poh, S. Bengio. Using chimeric users to construct fusion classifiers in biometric authentication tasks : An investigation. In IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), Mai 2006. 27, 204
- [Poh 2007] N. Poh, J. Kittler, T. Bourlai. Improving biometric device interoperability by likelihood ratio-based quality dependent score normalization. In 1st IEEE Int. Conf. on Biometrics: Theory, Applications, and Systems (BTAS 2007), 1-5, Septembre 2007. 162
- [Poh 2009] N. Poh, T. Bourlai, J. Kittler, L. Allano, F. Alonso-Fernandez, O. Ambekar, J. Baker, B. Dorizzi, O. Fatukasi, J. Fierrez, H. Ganster, J. Ortega-Garcia, D. Maurer, A. Salah, T. Scheidat, C. Vielhauer. *Benchmarking quality-dependent and cost-sensitive score-level multimodal biometric fusion algorithms*. In IEEE Trans. on Information Forensics and Security, 4: 849-866, Décembre 2009. 162
- [Poh 2013] N. Poh, A. Ross, W. Li, J. Kittler. A User-Specific and Selective Multimodal Biometric Fusion Strategy by Ranking Subjects, Pattern Recognition 46(12): 3341-3357, 2013. 204
- [Prabhakar 2002] S. Prabhakar, A.K. Jain. Decision-level fusion in fingerprint verification. Pattern Recogn. 35(4): 861–874, 2002. 21
- [Proença 2010] H. Proença, L. A. Alexandre. Iris recognition: Analysis of the error rates regarding the accuracy of the segmentation stage. In Image and Vision Computing, 28 (2010): 202–206, 2010. 32
- [Rathgeb 2010] C. Rathgeb, A. Uhl. Privacy Preserving Key Generation for Iris Biometrics, Communications and Multimedia Security: LNCS 6109: 191-200, 2010. 45, 46, 73

- [Reeves 2004] P. C. Reeves, B. D. Farkas, J.J. Jones, A. I. Gonzales. ROCIT: A Visual Object Recognition Algorithm Based on a Rank-Order Coding Scheme. Rapport technique, Sandia National Laboratories, SAND2004-2129, 2004. 128
- [Rodieck 1965] R. W. Rodieck, J. Stone. Analysis of receptive fields of cat retinal ganglion cells. Journal Neurophysiol. 28: 833-849, 1965. 123, 132
- [Ross 2003] A. Ross, A. K. Jain. *Information Fusion in Biometrics*. Pattern Recognition Letters, Special Issue on Multimodal Biometrics, 24(13): 2115-2125, Septembre 2003. 20, 22, 172
- [Ross 2006] A. Ross, K. Nandakumar, A. K. Jain. Handbook of Multibiometrics. Springer, 2006. 19
- [Roy 2008] K. Roy, P. Bhattacharya. Improving Features Subset Selection Using Genetic Algorithms for Iris Recognition. Artificial Neural Networks in Pattern Recognition, LNCS 5064: 292-304, 2008. 45, 46
- [Rowley 1996] H.A. Rowley, S. Baluja, T. Kanade. Neural Network-Based Face Detection. Proc. of IEEE Conf. on Computer Vision and Pattern Recognition – CVPR'96), 203-208. Californie, 1996. 78
- [Sacks 2010] O. Sacks, Face-Blind: Why are some of us terrible at recognizing faces?. The New Yorker, page 36, 30 August 2010. 120
- [Sauer 2011] P. Sauer, T.F. Cootes, C.J. Taylor. Accurate Regression Procedures for Active Appearance Models. British Machine Vision Conference (BMVC), 1-11, Dundee, UK, 2011. 76
- [Scheidat 2007] T. Scheidat, C. Vielhauer. Analyzing a multimodal biometric system using real and virtual users. In : Proc. of SPIE Security, Steganography, and Watermarking of Multimedia Contents IX, Février 2007. 27
- [Schor 1983] C. M. Schor, I. Wood. Disparity range for local stereopsis as a function of luminance spatial frequencies. Vision Research 23(12):1649-1654, 1983. 132
- [Schyns 2002] P.G. Schyns, L. Bonnar F. Gosselin. Show me the features!. Understanding Recognition from the Use of Visual Information. In Psychological Science. 13(5):402-409. 2002. 125, 126, 127, 129, 135, 136, 156
- [Seeck 1997] M. Seeck, C. M. Michel, N. Mainwaring, R. Cosgrove, H. Blume, J. Ives, T. Landis, D. L. Schomer. Evidence for rapid face recognition from human scalp and intracranial electrodes. Neuroreport, 8(12), 2749-2754, 1997. 118
- [Sen 2007] B. Sen, S. Furber. Maximising information recovery from rank-order codes. Proc. of SPIE Defense and Security Symposium, 6570(65700C): 1-12, Florida, 2007. 128
- [Serre 2007] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, T. Poggio. Robust object recognition with cortex-like mechanisms. IEEE Trans. Pattern Anal. Mach. Intell., 29(3):411-426, Mars 2007. 128
- [Shu 2006] C. Shu, X. Ding. *Multi-Biometrics Fusion for Identity Verification*. Int. Conf. on Pattern Recognition, 493-496, 2006. 162
- [Simoncelli 1995] E.P. Simoncelli, W.T. Freeman. The Steerable Pyramid : A Flexible Architecture For Multi-scale Derivative Computation. 2nd IEEE Inter. Conf. on Image Processing (ICIP), 3: 444-447, Washington, Octobre 1995. 47, 49, 81
- [Singh 2003] S.K. Singh, D.S. Chauhan, M. Vatsa, R. Singh. A Robust Skin Color Based Face Detection Algorithm. Tamkang Journal Of Science And Engineering, 6: 227-.234, 2003. 75
- [Sinha 2006] P. Sinha, B. Balas, Y. Ostrovsky, R. Russell. Face Recognition by humans: Nineteen results all computer vision researchers should know about. Proc. of the IEEE, 94(11):1948-1962. 2006. 121
- [Slavković 2012] M. Slavković, D. Jevtić. Face recognition using eigenface approach. Serbian Journal of Electrical Engineering, 9(1): 121-130, 2012. 76, 80

- [Snelick 2003] R. Snelick, M. Indovina, J. Yen, A. Mink. Multimodal biometrics: issues in design and testing. In : Proc. of the 5th int. conf. on Multimodal Interfaces – ICMI, 68-72, Vancouver, 2003. 163, 166
- [Son 2004] B. Son, H. Won, G. Kee, Y. Lee. Discriminant iris feature and support vector machines for iris recognition. Int. Conf. on Image Processing, 2: 865-868, 2004. 44
- [Su 2009] Y. Su, S.G. Shan, X.L. Chen, W. Gao. Hierarchical ensemble of global and local classifiers for face recognition. IEEE Trans. on Image Processing, 18(8): 1885-1896, 2009. 108
- [Svoboda 2007] E. Svoboda. Faces, Faces Everywhere. The New York Times, 13 February 2007. 121
- [Swets 1996] D. Swets, J. Weng. Using discriminant eigenfeatures for image retrieval. IEEE Trans. on PAMI 18(8): 831–836, 1996. 86
- [Takahashi 2010] Y. Takahashi, A. Tamamori, Y. Nankaku, K. Tokuda. Face recognition based on separable lattice 2-D HMM with state duration modeling. IEEE Int. Conf. on Acoustics Speech and Signal Processing (ICASSP' 2010): 2162 – 2165. Texas, Mars 2010. 78
- [Thoonsaengngam 2006] P. Thoonsaengngam, K. Horapong, S. Thainimit, V. Areekul. Efficient iris recognition using adaptive quotient thresholding. Int. Conf. on Biometrics, LNCS 3832: 472-478, Springer, Janvier 2006. 45
- [Thorpe 1996] S. Thorpe, D. Fize, C. Marlot. Speed of processing in the human visual system. Nature, 381(6582): 520–522, Juin. 1996. 122
- [Thorpe 1998] S. J. Thorpe, J. Gautrais. Rank Order Coding. In J. Bower (Ed.), Computational Neuroscience: Trends in Research. Plenum Press, 113-118, New York, 1998. 127, 141
- [Thorpe 2001] S. Thorpe, M. Fabre-Thorpe. Seeking categories in the brain. Science, 291(5502): 260-263, Janvier 2001. 122
- [Tieniu 2005] T. Tieniu, S. Zhenan. Note on CASIA-IrisV3 [Intenet]. Disponible sur http://www.cbsr.ia.ac.cn/english/Iris%20Databases.asp. Center for Biometric and Security Research, Institute of Automation, Chineese Academy of Science, 2005, Chine. 56
- [Tissé 2002] C.L. Tissé, L. Martin, L. Torres, M. Robert. Person identification technique using human iris recognition. In Proc. of Vision Interface, 294–299, 2002. 44
- [Tissé 2003] C.L. Tissé. Contribution à la vérification biométrique de personnes par reconnaissance de l'iris. Thèse de doctorat, Université de MontpellierII, 2003, France. 15
- [Tsao 2008] D. Y. Tsao, M. S. Livingstone, Mechanisms of face perception. Annu Rev Neurosci. 31: 411-437, 2008. 119
- [Turk 1991] M. Turk, A.P. Pentland. *Eigenfaces for recognition*. Journal of Cognitive Neuroscience, 3(1):71-86, 1991. 77, 84
- [Uçar 2015] A. Uçar. Color Face Recognition Based on Steerable Pyramid Transform and Extreme Learning Machines. The Scientific World Journal. Hindawi Publishing Corporation, 2014. 81
- [Van Rullen 2001] R. Van Rullen, S. J. Thorpe. *Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex*. Neural Comp. 13: 1255-1283, 2001. 128
- [Van Rullen 2003] R. Van Rullen, C. Koch. Visual selective behavior can be triggered by a feedforward process. Journal of Cognitive Neuroscience, 15(2):209-217, Février 2003. 122, 203
- [Van Rullen 2004] R. Van Rullen, S.J. Thorpe. Perception, decision, attention visuelles: ce que les potentiels evoques nous apprennent sur le fonctionnement du systeme visuel. Dans B. Renaut, Hermes (Ed), L'imagerie fonctionnelle electrique et magnetique: ses applications en sciences cognitives, 95-121, Paris, 2004. 118
- [Viola 2004] P. Viola, M. J. Jones. Robust Real-Time Face Detection. Int. Journal of Computer Vision 57(2):137-154, 2004. 76, 86, 87
- [Vu 2011] N. S., Vu, A. Caplier. Biologically Inspired Processing for Lighting Robust Face Recognition. Dans Dr. Jucheng Yang (Édit.), State of the art in Biometrics. ISBN: 978-953-

307-489-4, InTech, 2011. Disponible sur : http://www.intechopen.com/books/state-of-the-art-in-biometrics/biologically-inspired-processing-for-lighting-robust-face-recognition. 129

- [Wandell 1995] B. A. Wandell. Foundations of Vision. Sinauer Associates Inc.1st edition ,Mai 1995. 114
- [Wang H. 2005] H. Wang, P. Li, T. Zhang. Proposal of Novel Histogram Features for Face Detection, In Singh S. et al. (Eds.), Pattern Recog. and Image Anal,. Springer, Heidelberg. LNCS(3687), 334-343, 2005. 148
- [Wang N. 2013] N. Wang, Q. Li; A. A. Abd El-Latif; J. Peng; X. Niu. Multibiometrics Fusion for Identity Authentication: Dual Iris, Visible and Thermal Face Imagery. Int. Journal of Security and Its Applications. 7(3): 33-44. Mai 2013. 163
- [Wang Y. 2003] Y. Wang, T. Tan, A. Jain. Combining face and iris biometrics for identity verification. Proc. of 4th Int. Conf. on Audio and Video-Based Authentication – AVBPA, 805-813, Guildford, Juin 2003. 20, 21, 161
- [Wang Z. 2015] Z. Wang, Q. Ruan, G. An. Face Recognition Using Double Sparse Local Fisher Discriminant Analysis. Mathematical Problems in Engineering. 2015. 76, 80
- [Watson 2005] A. Watson, A. Ahumada. A standard model for foveal detection of spatial contrast. Journal of Vision, 5: 717–740, 2005. 154
- [Wawrzyniak 2009] H. Wawrzyniak. *Prosopagnosia*. 2009. [En ligne] Disponible sur https://www.flickr.com/photos/zakos/favorites. 120
- [Wildes 1997] R.P. Wildes. Iris recognition: An emerging biometric technology. Proc. of the IEEE, 85(9):1348-1363, Septembre 1997. 38, 42
- [Wiskott 1997] L. Wiskott, J.M. Fellous, N. Krüger, C. von der Malsburg. Face recognition by elastic bunch graph matching. IEEE Trans. Patt. Anal. Mach. Intell. 19(7): 775–779, 1997. 79
- [Yang 2002] M. Yang, D.J. Kriegman, N. Ahuja. Detecting Faces in Images: A Survey. IEEE Trans. on Pattern Analysis and Machine Intelligence, 24(1), Janvier 2002. 76
- [Yoon 2010] J. Yoon, D. Kim. Frontal face classifier using AdaBoost with MCT features. Proc. of the Int. Conf. on Control, Automation, Robotics and Vision – ICARCV, 2084–2087, Singapore, December 2010. 76, 87
- [Yu 2007] L. Yu, D. Zhang, K. Wang. The relative distance of key point based iris recognition. Pattern Recognition, 40(2): 423–430, 2007. 45, 46
- [Yuan 2013] X. Yuan, Y. Meng, X. Wei. Illumination Normalization Based on Homomorphic Wavelet Filtering for Face Recognition. Journal Inf. Science and Eng. 29: 579-594, 2013. 148
- [Zernike 1934] F. Zernike. Diffraction theory of the cut procedure and its improved form, the phase contrast method. Physica, 1: 689-704, 1934. 77
- [Zhang X. 2005] X. Zhang, Y. Jia. Face Recognition Based on Steerable Feature and Random Subspace LDA. Inter. workshop on Analysis and Modelling of Faces and Gestures – AMFG, 3723(2):170-183, Pekin, Octobre 2005. 75, 111
- [Zhang X. 2010] X. Zhang, Z. Sun, T. Tan. Hierarchical fusion of face and iris for personal identification. In 20th Int. Conf. on Pattern Recognition – ICPR'2010, 217-220, Août 2010. 163
- [Zhang Z. 2007] Z. Zhang, R. Wang, K. Pan, S. Li, P. Zhang. Fusion of near infrared face and iris biometrics. In S.W. Lee, S. Li, editors, Advances in Biometrics, Springer, Heidelberg, LNCS(4642): 172-180, 2007. 161, 162
- [Zhang J. 2012] J. Zhang, X. Xie, (2012). A Study on the Effective Approach to Illumination-Invariant Face Recognition Based on a Single Image. W.-S. Zheng et al., (Eds.), Springer, Heidelberg, CCBR, LNCS(7701): 33-41, 2012. 148
- [Zhu 2000] Y. Zhu, T. Tan, Y. Wang. Biometric personal identification based on iris patterns. Proc. of the 15th Int. Conf. on Pattern Recognition, 801–804, 2000. 43

Liste des publications

Article de journal

Nefissa KHIARI HILI, Sylvie LELANDAIS, Christophe MONTAGNE, Corinne ROUMES, Kamel HAMROUNI, Justin PLANTIER. *Bio inspired image enhancement derived from a rank order coding model.* IET Image Processing (Accepté le 25 Novembre 2015).

Conférences internationales

- Ayoub EL GHANAOUI, Nefissa KHIARI HILI, Christophe MONTAGNE, Sylvie LELANDAIS. Bio-Inspired Face Authentication using Multiscale LBP. Int. Conf. on Bio-inspired Systems and Signal Processing – BioSignals, Barcelone, 2013.
- Nefissa KHIARI HILI, Sylvie LELANDAIS, Christophe MONTAGNE, Kamel HAMROUNI. A Combining Approach for 2D Face Recognition Application on IV² Database. 3rd Int. Conf. on Image Processing Theory, Tools and Applications IPTA, Istanboul 2012.
- Nefissa KHIARI HILI, Sylvie LELANDAIS, Christophe MONTAGNE, Kamel HAMROUNI. Boosting Steerable Features For 2d Face Recognition On Iv² Database. 2nd Int. Special Session on Multivariable Processing for Biometric Systems – MPBS, Vilamoura, 2012.
- Nefissa KHIARI HILI, Sylvie LELANDAIS, Christophe MONTAGNE, Kamel HAMROUNI. Reconnaissance de visages 2D sur la base IV² par les Pyramides orientables et la LDA. 7^{ème} Atelier de travail sur le Traitement et Analyse de l'Information, Méthodes et Applications – TAIMA, Hammamet, 2011.
- Nefissa KHIARI, Hela MAHERSIA, Kamel HAMROUNI. Iris Recognition Using Steerable Pyramids. 1st IEEE Inter. Conf. on Image Processing, Theory, Tools and Applications – IPTA, Sousse, 2008.

ANNEXE A : Intervalles de confiance pour l'estimation des erreurs d'un système biométrique

Les performances des systèmes biométriques sont souvent évaluées sur des bases de données qui ne sont pas complètement représentatives de la réalité. Non seulement ces bases de données sont de taille réduite, mais en plus, les protocoles d'évaluation nécessitent leur découpage en deux sous-ensembles (apprentissage et test) ne permettant pas d'avoir une valeur significative de performance avec une seule valeur d'erreur [Bolle 2000] [Bolle 2004]. En effet, les tests de performance réalisés sur des découpages différents de la base de données n'aboutissent pas aux mêmes taux d'erreurs. Plutôt que de présenter la valeur ponctuelle du taux d'erreurs, il est alors plus approprié d'estimer un encadrement de la valeur du taux d'erreurs grâce à un intervalle de confiance permettant de quantifier son degré de certitude (ou d'incertitude). Par exemple, un intervalle de confiance à 90% a une probabilité égale à 0,9 de contenir la vraie valeur du taux d'erreurs, avec un risque de se tromper de seulement 10%. Plus cet intervalle est étroit, meilleure sera la précision de ce taux d'erreurs, plus il est large, moins bonne sera la précision du taux d'erreurs et par conséquent, moins bonne sera sa signification statistique.

Une solution est de répéter un certain nombre de fois le découpage en base de test et base d'apprentissage afin de pouvoir estimer l'erreur en généralisation [Bolle 2004]. Dans les cas où le découpage de la base ne se fait qu'une seule fois, il existe des méthodes paramétriques pour estimer cet intervalle de confiance. L'estimation paramétrique des taux d'erreurs implique un certain nombre d'hypothèses parmi lesquelles on suppose que les valeurs du taux d'erreurs à un seuil fixé ont une forme normale [Bolle 2004]. Cette hypothèse permet de déterminer l'intervalle de confiance à partir des percentiles associés à la loi normale. Par exemple, l'intervalle de confiance à 90% pour le taux de faux rejets $FRR(\tau)$ à un seuil τ donné, relativement à une valeur mesurée $\widehat{FRR}(\tau)$ à partir de N_i scores imposteurs, correspond au percentile à a = 0,90 centré de la loi normale, c'est-à-dire à la valeur de la variable binomiale $Z_{a/2} = Z_{0,45} = 1,645$ (cf. équations A.1 et A.2) [Bolle 2000].

$$FRR(\tau) \in \left[\widehat{FRR}(\tau) - 1,645\,\widehat{\sigma}_{l}(\tau), \quad \widehat{FRR}(\tau) + 1,645\,\widehat{\sigma}_{l}(\tau)\right] \tag{A.1}$$

avec :

$$\widehat{\sigma}_{i}(\tau) = \sqrt{\frac{\widehat{FRR}(\tau)\left(1 - \widehat{FRR}(\tau)\right)}{N_{i}}}$$
(A.2)

Pour des intervalles de confiance à 95%, il suffit de remplacer 1,645 = $Z_{0,45}$ par 1,96 = $Z_{0,475}$ (cf. ligne et colonne en vert de la **Table A.1**).

z 0.	00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0 0.	0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1 0.	0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2 0.	0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3 0.	1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4 0.	1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5 0.	1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6 0.	2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7 0.	2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8 0.	2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9 0.	3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0 0.	3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1 0.	3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2 0.	3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3 0.	4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4 0.	4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5 0.	4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6 0.	4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7 0.	4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
										0.4706
1.8 0.	4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4700
1.8 0. 1.9 0 .	4641 4713	0.4649 0.4719	0.4656 0.4726	0.4664 0.4732	0.4671 0.4738	0.4678 0.4744	0.4686 0.4750	0.4693 0.4756	0.4699 0.4761	0.4768
1.8 0. 1.9 0 . 2.0 0.	4641 4713 4772	0.4649 0.4719 0.4778	0.4656 0.4726 0.4783	0.4664 0.4732 0.4788	0.4671 0.4738 0.4793	0.4678 0.4744 0.4798	0.4686 0.4750 0.4803	0.4693 0.4756 0.4808	0.4699 0.4761 0.4812	0.4768 0.4767 0.4817
1.8 0. 1.9 0. 2.0 0. 2.1 0.	4641 4713 4772 4821	0.4649 0.4719 0.4778 0.4826	0.4656 0.4726 0.4783 0.4830	0.4664 0.4732 0.4788 0.4834	0.4671 0.4738 0.4793 0.4838	0.4678 0.4744 0.4798 0.4842	0.4686 0.4750 0.4803 0.4846	0.4693 0.4756 0.4808 0.4850	0.4699 0.4761 0.4812 0.4854	0.4767 0.4817 0.4857
1.8 0. 1.9 0. 2.0 0. 2.1 0. 2.2 0.	4641 4713 4772 4821 4861	0.4649 0.4719 0.4778 0.4826 0.4864	0.4656 0.4726 0.4783 0.4830 0.4868	0.4664 0.4732 0.4788 0.4834 0.4871	0.4671 0.4738 0.4793 0.4838 0.4875	0.4678 0.4744 0.4798 0.4842 0.4878	0.4686 0.4750 0.4803 0.4846 0.4881	0.4693 0.4756 0.4808 0.4850 0.4884	0.4699 0.4761 0.4812 0.4854 0.4887	0.4767 0.4817 0.4857 0.4890
1.8 0. 1.9 0. 2.0 0. 2.1 0. 2.2 0. 2.3 0.	4641 4713 4772 4821 4861 4893	0.4649 0.4719 0.4778 0.4826 0.4864 0.4896	0.4656 0.4726 0.4783 0.4830 0.4868 0.4898	0.4664 0.4732 0.4788 0.4834 0.4871 0.4901	0.4671 0.4738 0.4793 0.4838 0.4875 0.4904	0.4678 0.4744 0.4798 0.4842 0.4878 0.4906	0.4686 0.4750 0.4803 0.4846 0.4881 0.4909	0.4693 0.4756 0.4808 0.4850 0.4884 0.4911	0.4699 0.4761 0.4812 0.4854 0.4887 0.4913	0.4768 0.4767 0.4817 0.4857 0.4890 0.4916
1.8 0. 1.9 0. 2.0 0. 2.1 0. 2.2 0. 2.2 0. 2.3 0. 2.4 0.	4641 4713 4772 4821 4861 4893 4918	0.4649 0.4719 0.4778 0.4826 0.4826 0.4864 0.4896 0.4920	0.4656 0.4726 0.4783 0.4830 0.4868 0.4898 0.4898	0.4664 0.4732 0.4788 0.4834 0.4831 0.4901 0.4901	0.4671 0.4738 0.4793 0.4838 0.4835 0.4904 0.4927	0.4678 0.4798 0.4798 0.4842 0.4878 0.4906 0.4929	0.4686 0.4750 0.4803 0.4846 0.4881 0.4909 0.4931	0.4693 0.4756 0.4808 0.4850 0.4884 0.4911 0.4932	0.4699 0.4761 0.4812 0.4854 0.4887 0.4913 0.4934	0.4767 0.4817 0.4857 0.4890 0.4916 0.4936
1.8 0. 1.9 0. 2.0 0. 2.1 0. 2.2 0. 2.3 0. 2.3 0. 2.4 0. 2.5 0.	4641 4713 4772 4821 4861 4893 4918 4938	0.4649 0.4719 0.4778 0.4826 0.4864 0.4896 0.4920 0.4940	0.4656 0.4726 0.4783 0.4830 0.4868 0.4898 0.4992 0.4941	0.4664 0.4732 0.4788 0.4834 0.4871 0.4901 0.4925 0.4943	0.4671 0.4738 0.4793 0.4838 0.4875 0.4904 0.4927 0.4945	0.4678 0.4798 0.4798 0.4842 0.4878 0.4906 0.4929 0.4946	0.4686 0.4750 0.4803 0.4846 0.4881 0.4909 0.49031 0.4948	0.4693 0.4756 0.4808 0.4850 0.4884 0.4911 0.4932 0.4949	0.4699 0.4761 0.4812 0.4854 0.4887 0.4913 0.4934 0.4951	0.4767 0.4817 0.4857 0.4890 0.4916 0.4936 0.4952
1.1.8 0. 1.1.9 0. 2.10 0. 2.11 0. 2.12 0. 2.13 0. 2.14 0. 2.15 0. 2.15 0.	4641 4713 4772 4821 4861 4893 4918 4938 4953	0.4649 0.4719 0.4778 0.4826 0.4826 0.4896 0.4896 0.4920 0.4920 0.4955	0.4656 0.4726 0.4783 0.4830 0.4868 0.4898 0.4992 0.4941 0.4956	0.4664 0.4732 0.4788 0.4834 0.4871 0.4901 0.4925 0.4943 0.4957	0.4671 0.4738 0.4793 0.4838 0.4875 0.4904 0.4927 0.4945 0.4959	0.4678 0.4798 0.4798 0.4842 0.4878 0.4906 0.4929 0.4946 0.4960	0.4686 0.4750 0.4803 0.4846 0.4881 0.4909 0.4931 0.4931 0.4948 0.4961	0.4693 0.4756 0.4808 0.4850 0.4884 0.4911 0.4932 0.4949 0.4962	0.4699 0.4761 0.4812 0.4854 0.4887 0.4913 0.4934 0.4951 0.4963	0.4767 0.4817 0.4857 0.4890 0.4916 0.4936 0.4952 0.4964
1.8 0. 1.9 0. 2.0 0. 2.1 0. 2.2 0. 2.3 0. 2.4 0. 2.5 0. 2.5 0. 2.6 0.	4641 4713 4772 4821 4861 4893 4918 4938 4953 4965	0.4649 0.4719 0.4778 0.4826 0.4826 0.4896 0.4920 0.4940 0.4955 0.4966	0.4656 0.4726 0.4783 0.4830 0.4868 0.4898 0.4992 0.4922 0.4941 0.4956 0.4967	0.4664 0.4732 0.4788 0.4834 0.4871 0.4901 0.4925 0.4943 0.4957 0.4968	0.4671 0.4738 0.4793 0.4838 0.4875 0.4904 0.4927 0.4945 0.4959 0.4969	0.4678 0.4798 0.4798 0.4842 0.4878 0.4906 0.4929 0.4946 0.4960 0.4970	0.4686 0.4750 0.4803 0.4846 0.4881 0.4909 0.4931 0.4931 0.4948 0.4961	0.4693 0.4756 0.4808 0.4850 0.4884 0.4911 0.4932 0.4949 0.4949 0.4962 0.4972	0.4699 0.4761 0.4812 0.4854 0.4887 0.4913 0.4934 0.4951 0.4963 0.4973	0.4767 0.4817 0.4857 0.4890 0.4916 0.4936 0.4952 0.4964 0.4974
1.8 0. 1.9 0. 2.0 0. 2.1 0. 2.2 0. 2.2.3 0. 2.2.4 0. 2.2.5 0. 2.2.6 0. 2.2.7 0. 2.2.8 0.	4641 4713 4772 4821 4861 4893 4918 4938 4953 4965 4974	0.4649 0.4719 0.4778 0.4826 0.4826 0.4896 0.4896 0.4920 0.4940 0.4955 0.4966 0.4975	0.4656 0.4726 0.4783 0.4830 0.4868 0.4898 0.4922 0.4941 0.4956 0.4956 0.4967 0.4976	0.4664 0.4732 0.4788 0.4834 0.4871 0.4901 0.4925 0.4943 0.4957 0.4968 0.4977	0.4671 0.4738 0.4793 0.4838 0.4875 0.4904 0.4927 0.4945 0.4959 0.4969 0.4977	0.4678 0.4798 0.4798 0.4842 0.4878 0.4906 0.4929 0.4946 0.4960 0.4970 0.4978	0.4686 0.4750 0.4803 0.4846 0.4881 0.4909 0.4909 0.4931 0.4948 0.4961 0.4971 0.4979	0.4693 0.4756 0.4808 0.4850 0.4884 0.4911 0.4932 0.4949 0.4949 0.4962 0.4972 0.4979	0.4699 0.4761 0.4812 0.4854 0.4887 0.4913 0.4934 0.4951 0.4963 0.4973 0.4980	0.4767 0.4817 0.4857 0.4890 0.4916 0.4936 0.4936 0.4952 0.4964 0.4974 0.4981
1.18 0. 1.19 0. 2.0 0. 2.1 0. 2.2.1 0. 2.2.2 0. 2.2.3 0. 2.13 0. 2.2.4 0. 2.15 0. 2.2.5 0. 2.2.6 0. 2.2.8 0. 2.19 0.	4641 4773 4772 4821 4861 4893 4918 4938 4953 4965 4974 4981	0.4649 0.4719 0.4778 0.4826 0.4826 0.4896 0.4920 0.4920 0.4955 0.4955 0.4975 0.4982	0.4656 0.4726 0.4783 0.4830 0.4868 0.4898 0.4992 0.4922 0.4941 0.4956 0.4967 0.4967 0.4982	0.4664 0.4732 0.4788 0.4834 0.4871 0.4901 0.4925 0.4943 0.4957 0.4968 0.4977 0.4983	0.4671 0.4738 0.4793 0.4838 0.4875 0.4904 0.4927 0.4945 0.4959 0.4969 0.4967 0.4984	0.4678 0.4798 0.4798 0.4842 0.4878 0.4906 0.4929 0.4946 0.4960 0.4970 0.4978 0.4984	0.4686 0.4750 0.4803 0.4803 0.4846 0.4881 0.4909 0.4909 0.4931 0.4948 0.4961 0.4971 0.4979 0.4985	0.4693 0.4756 0.4808 0.4850 0.4884 0.4911 0.4932 0.4932 0.4949 0.4962 0.4972 0.4979 0.4985	0.4699 0.4761 0.4812 0.4854 0.4887 0.4913 0.4934 0.4951 0.4963 0.4963 0.4980 0.4986	0.4767 0.4817 0.4857 0.4890 0.4916 0.4936 0.4952 0.4964 0.4974 0.4981 0.4986
1.18 0. 1.19 0. 2.10 0. 2.11 0. 2.12 0. 2.13 0. 2.14 0. 2.15 0. 2.15 0. 2.16 0. 2.17 0. 2.18 0. 2.19 0. 3.10 0.	4641 4713 4772 4821 4861 4893 4918 4938 4953 4965 4974 4981 4987	0.4649 0.4719 0.4778 0.4826 0.4864 0.4896 0.4920 0.4940 0.4955 0.4966 0.4975 0.4982 0.4987	0.4656 0.4726 0.4783 0.4830 0.4868 0.4898 0.4922 0.4922 0.4941 0.4956 0.4967 0.4967 0.4976 0.4982 0.4987	0.4664 0.4732 0.4788 0.4834 0.4871 0.4901 0.4925 0.4943 0.4957 0.4968 0.4977 0.4983 0.4988	0.4671 0.4738 0.4793 0.4875 0.4875 0.4904 0.4927 0.4945 0.4959 0.4969 0.4969 0.4977 0.4984 0.4988	0.4678 0.4798 0.4798 0.4842 0.4878 0.4906 0.4929 0.4946 0.4960 0.4970 0.4978 0.4978 0.4989	0.4686 0.4750 0.4803 0.4803 0.4846 0.4881 0.4909 0.4931 0.4931 0.4948 0.4961 0.4971 0.4979 0.4985 0.4989	0.4693 0.4756 0.4808 0.4850 0.4884 0.4911 0.4932 0.4949 0.4949 0.4942 0.4972 0.4972 0.4975 0.4985	0.4699 0.4761 0.4812 0.4854 0.4887 0.4913 0.4934 0.4951 0.4963 0.4973 0.4980 0.4986 0.4990	0.4767 0.4817 0.4857 0.4890 0.4916 0.4936 0.4952 0.4964 0.4974 0.4981 0.4986 0.4990
1.18 0. 1.19 0. 2.10 0. 2.11 0. 2.12 0. 2.13 0. 2.14 0. 2.15 0. 2.15 0. 2.15 0. 2.15 0. 2.15 0. 2.16 0. 2.17 0. 2.18 0. 2.19 0. 3.10 0. 3.10 0.	4641 4713 4772 4821 4861 4893 4938 4938 4953 4965 4974 4981 4981 4987 4990	0.4649 0.4719 0.4778 0.4826 0.4826 0.4864 0.4896 0.4920 0.4940 0.4940 0.4955 0.4955 0.4955 0.4975 0.4982 0.4987 0.4991	0.4656 0.4726 0.4783 0.4830 0.4868 0.4898 0.4922 0.4922 0.4941 0.4956 0.4956 0.4967 0.4976 0.4982 0.4987 0.4991	0.4664 0.4732 0.4788 0.4834 0.4871 0.4901 0.4925 0.4943 0.4957 0.4968 0.4977 0.4983 0.4988 0.4991	0.4671 0.4738 0.4793 0.4838 0.4875 0.4904 0.4927 0.4945 0.4959 0.4959 0.4969 0.4977 0.4984 0.4988 0.4992	0.4678 0.4798 0.4798 0.4842 0.4878 0.4906 0.4929 0.4946 0.4960 0.4970 0.4978 0.4984 0.4984 0.4989 0.4992	0.4686 0.4750 0.4803 0.4803 0.4846 0.4881 0.4909 0.4931 0.4931 0.4948 0.4961 0.4971 0.4979 0.4985 0.4989 0.4982	0.4693 0.4756 0.4808 0.4850 0.4884 0.4911 0.4932 0.4949 0.4949 0.4962 0.4972 0.4979 0.4985 0.4989 0.4992	0.4699 0.4761 0.4812 0.4854 0.4887 0.4913 0.4934 0.4951 0.4963 0.4973 0.4980 0.4986 0.4990 0.4993	0.4767 0.4817 0.4857 0.4890 0.4990 0.4936 0.4936 0.4952 0.4964 0.4981 0.4981 0.4986 0.4990 0.4993
1.18 0. 1.19 0. 2.10 0. 2.11 0. 2.12 0. 2.13 0. 2.13 0. 2.14 0. 2.15 0. 2.16 0. 2.17 0. 2.18 0. 2.19 0. 3.10 0. 3.11 0.	4641 4773 4772 4821 4861 4893 4918 4918 4938 4953 4965 4974 4981 4987 4990 4993	0.4649 0.4719 0.4778 0.4826 0.4826 0.4896 0.4920 0.4920 0.4920 0.4955 0.4955 0.4955 0.4966 0.4975 0.4982 0.4987 0.4991 0.4993	0.4656 0.4726 0.4783 0.4830 0.4868 0.4898 0.4922 0.4922 0.4941 0.4956 0.4967 0.4967 0.4982 0.4987 0.4981 0.4994	0.4664 0.4732 0.4788 0.4834 0.4871 0.4901 0.4901 0.4925 0.4943 0.4957 0.4968 0.4977 0.4983 0.4988 0.4991 0.4994	0.4671 0.4738 0.4793 0.4838 0.4875 0.4904 0.4927 0.4945 0.4959 0.4959 0.4969 0.4969 0.4984 0.4988 0.4988 0.4992 0.4994	0.4678 0.4798 0.4798 0.4842 0.4878 0.4906 0.4929 0.4946 0.4960 0.4960 0.4970 0.4978 0.4984 0.4989 0.4992 0.4994	0.4686 0.4750 0.4803 0.4803 0.4846 0.4881 0.4909 0.4931 0.4931 0.4948 0.4971 0.4971 0.4975 0.4985 0.4985 0.4992	0.4693 0.4756 0.4808 0.4850 0.4884 0.4911 0.4932 0.4932 0.4949 0.4949 0.4972 0.4979 0.4985 0.4989 0.4992 0.4995	0.4699 0.4761 0.4812 0.4854 0.4887 0.4913 0.4934 0.4951 0.4963 0.4963 0.4980 0.4980 0.4980 0.4990 0.4993 0.4995	0.4767 0.4817 0.4857 0.4890 0.4916 0.4936 0.4952 0.4964 0.4974 0.4981 0.4981 0.4986 0.4990 0.4993 0.4995
1.18 0. 1.19 0. 2.0 0. 2.1 0. 2.2 0. 2.3 0. 2.3 0. 2.3 0. 2.5 0. 2.5 0. 2.6 0. 2.6 0. 2.7 0. 2.8 0. 2.9 0. 3.1 0. 3.1 0. 3.2 0. 3.3 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.	4641 4713 4772 4821 4861 4893 4938 4938 4953 4965 4974 4981 4987 4990 4993 4995	0.4649 0.4719 0.4778 0.4826 0.4826 0.4896 0.4920 0.4920 0.4940 0.4955 0.4955 0.4966 0.4975 0.4982 0.4982 0.4987 0.4993 0.4995	0.4656 0.4726 0.4783 0.4830 0.4868 0.4898 0.4922 0.4922 0.4941 0.4956 0.4967 0.4967 0.4976 0.4987 0.4987 0.4991 0.4994 0.4995	0.4664 0.4732 0.4788 0.4834 0.4871 0.4901 0.4925 0.4943 0.4957 0.4968 0.4977 0.4983 0.4988 0.4991 0.4994 0.4996	0.4671 0.4738 0.4793 0.4875 0.4875 0.4904 0.4927 0.4945 0.4959 0.4969 0.4969 0.4984 0.4988 0.4988 0.4992 0.4994 0.4996	0.4678 0.4798 0.4798 0.4842 0.4878 0.4906 0.4929 0.4946 0.4960 0.4970 0.4978 0.4978 0.4989 0.4989 0.4992 0.4994 0.4996	0.4686 0.4750 0.4803 0.4803 0.4846 0.4881 0.4909 0.4931 0.4948 0.4961 0.4971 0.4979 0.4985 0.4985 0.4989 0.4992 0.4994	0.4693 0.4756 0.4808 0.4850 0.4884 0.4911 0.4932 0.4949 0.4949 0.4949 0.4972 0.4972 0.4979 0.4985 0.4989 0.4995 0.4995	0.4699 0.4761 0.4812 0.4854 0.4887 0.4913 0.4934 0.4951 0.4963 0.4973 0.4980 0.4980 0.4986 0.4990 0.4993 0.4995 0.4995 0.4996	0.4767 0.4817 0.4857 0.4890 0.4916 0.4936 0.4952 0.4964 0.4974 0.4981 0.4986 0.4990 0.4993 0.4995 0.4997
1.18 0. 1.19 0. 2.10 0. 2.11 0. 2.12 0. 2.13 0. 2.14 0. 2.15 0. 2.16 0. 2.17 0. 2.18 0. 2.19 0. 3.10 0. 3.2 0. 3.3 0. 3.3 0. 3.3 0.	4641 47713 4772 4821 4861 4893 4918 4938 4953 4953 4965 4974 4981 4987 4990 4993 4995 4997	0.4649 0.4719 0.4778 0.4826 0.4826 0.4864 0.4896 0.4920 0.4920 0.4920 0.4920 0.4955 0.4955 0.4966 0.4975 0.4982 0.4987 0.4993 0.4995 0.4997	0.4656 0.4726 0.4783 0.4830 0.4868 0.4898 0.4922 0.4922 0.4941 0.4956 0.4967 0.4976 0.4976 0.4987 0.4987 0.4991 0.4994 0.4995 0.4997	0.4664 0.4732 0.4788 0.4834 0.4871 0.4901 0.4925 0.4943 0.4957 0.4968 0.4977 0.4983 0.4988 0.4991 0.4994 0.4996 0.4997	0.4671 0.4738 0.4793 0.4838 0.4875 0.4904 0.4927 0.4945 0.4945 0.4959 0.4959 0.4969 0.4977 0.4984 0.4988 0.4992 0.4994 0.4996 0.4997	0.4678 0.4798 0.4798 0.4842 0.4878 0.4906 0.4929 0.4946 0.4960 0.4970 0.4978 0.4978 0.4984 0.4984 0.4989 0.4992 0.4994 0.4996 0.4997	0.4686 0.4750 0.4803 0.4803 0.4846 0.4881 0.4909 0.4909 0.4931 0.4948 0.4961 0.4971 0.4979 0.4985 0.4989 0.4989 0.4992 0.4994 0.4996 0.4997	0.4693 0.4756 0.4808 0.4850 0.4884 0.4911 0.4932 0.4949 0.4949 0.4962 0.4972 0.4979 0.4985 0.4985 0.4992 0.4995 0.4995 0.4997	0.4699 0.4761 0.4812 0.4854 0.4887 0.4913 0.4934 0.4951 0.4963 0.4963 0.4980 0.4980 0.4986 0.4990 0.4993 0.4995 0.4995 0.4997	0.4767 0.4817 0.4857 0.4890 0.4996 0.4936 0.4936 0.4952 0.4964 0.4974 0.4981 0.4986 0.4990 0.4993 0.4995 0.4997 0.4998
1.18 0. 1.19 0. 1.9 0. 2.0 0. 2.1 0. 2.2 0. 2.2 0. 2.3 0. 2.3 0. 2.5 0. 2.5 0. 2.5 0. 2.6 0. 2.7 0. 2.8 0. 2.8 0. 3.1 0. 3.1 0. 3.1 0. 3.3 0. 3.3 0. 3.3 0. 3.4 0. 3.5 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.	4641 4713 4772 4821 4861 4893 4918 4938 4953 4965 4974 4987 4987 4990 4993 4995 4997 4998	0.4649 0.4719 0.4778 0.4826 0.4826 0.4896 0.4920 0.4920 0.4920 0.4920 0.4955 0.4955 0.4966 0.4975 0.4996 0.4997 0.4993 0.4995 0.4998	0.4656 0.4726 0.4783 0.4830 0.4868 0.4898 0.4922 0.4922 0.4941 0.4956 0.4967 0.4967 0.4987 0.4987 0.4991 0.4991 0.4995 0.4997 0.4998	0.4664 0.4732 0.4788 0.4834 0.4871 0.4901 0.4901 0.4925 0.4943 0.4957 0.4968 0.4977 0.4988 0.4991 0.4994 0.4994 0.4997 0.4998	0.4671 0.4738 0.4793 0.4838 0.4875 0.4904 0.4927 0.4945 0.4959 0.4959 0.4969 0.4969 0.4988 0.4998 0.4994 0.4994 0.4996 0.4998	0.4678 0.4798 0.4798 0.4878 0.4906 0.4906 0.4929 0.4946 0.4960 0.4970 0.4970 0.4978 0.4989 0.4989 0.4992 0.4994 0.4996 0.4997 0.4998	0.4686 0.4750 0.4803 0.4803 0.4846 0.4909 0.4931 0.4948 0.4951 0.4971 0.4979 0.4985 0.4985 0.4989 0.4999 0.4994 0.4996 0.4998	0.4693 0.4756 0.4808 0.4800 0.4884 0.4911 0.4932 0.4932 0.4949 0.4949 0.4972 0.4972 0.4979 0.4985 0.4995 0.4995 0.4996 0.4998	0.4699 0.4761 0.4812 0.4854 0.4854 0.4913 0.4934 0.4934 0.4951 0.4963 0.4973 0.4980 0.4980 0.4990 0.4995 0.4995 0.4995 0.4997 0.4998	0.4767 0.4817 0.4857 0.4857 0.4890 0.4916 0.4936 0.4952 0.4964 0.4974 0.4981 0.4986 0.4990 0.4993 0.4993 0.4995 0.4997 0.4998
1.18 0. 1.19 0. 2.0 0. 2.1 0. 2.2.1 0. 2.2.2 0. 2.3 0. 2.3 0. 2.4 0. 2.5 0. 2.5 0. 2.6 0. 2.7 0. 2.8 0. 2.9 0. 3.10 0. 3.2 0. 3.3 0. 3.4 0. 3.5 0. 3.6 0.	4641 4713 4772 4821 4861 4893 4938 4938 4953 4953 4965 4974 4987 4987 4990 4993 4995 4995 4995 4998	0.4649 0.4719 0.4778 0.4826 0.4826 0.4896 0.4920 0.4940 0.4955 0.4955 0.4955 0.4966 0.4975 0.4982 0.4982 0.4987 0.4993 0.4995 0.4995 0.4995 0.4998	0.4656 0.4726 0.4783 0.4830 0.4830 0.4868 0.4992 0.4922 0.4922 0.4941 0.4956 0.4967 0.4967 0.4976 0.4987 0.4981 0.4994 0.4995 0.4997 0.4998 0.4999	0.4664 0.4732 0.4732 0.4788 0.4834 0.4871 0.4901 0.4925 0.4943 0.4957 0.4968 0.4957 0.4968 0.4977 0.4983 0.4988 0.4991 0.4994 0.4994 0.4995 0.4997 0.4998 0.4999	0.4671 0.4738 0.4793 0.4875 0.4875 0.4904 0.4927 0.4945 0.4959 0.4969 0.4969 0.4977 0.4984 0.4988 0.4992 0.4994 0.4996 0.4997 0.4998 0.4999	0.4678 0.4798 0.4798 0.4842 0.4878 0.4906 0.4929 0.4946 0.4960 0.4970 0.4970 0.4978 0.4978 0.4989 0.4989 0.4992 0.4994 0.4994 0.4997 0.4997 0.4998 0.4999	0.4686 0.4750 0.4803 0.4803 0.4846 0.4881 0.4909 0.4931 0.4948 0.4961 0.4971 0.4979 0.4985 0.4985 0.4989 0.4992 0.4994 0.4996 0.4997 0.4998 0.4999	0.4693 0.4756 0.4808 0.4808 0.4850 0.4884 0.4911 0.4932 0.4949 0.4949 0.4949 0.4972 0.4979 0.4985 0.4985 0.4989 0.4995 0.4996 0.4997 0.4998 0.4999	0.4699 0.4761 0.4812 0.4854 0.4854 0.4913 0.4934 0.4934 0.4951 0.4963 0.4973 0.4980 0.4980 0.4990 0.4993 0.4995 0.4995 0.4995 0.4997 0.4998 0.4999	0.4767 0.4817 0.4857 0.4890 0.4990 0.4936 0.4936 0.4936 0.4952 0.4964 0.4981 0.4981 0.4986 0.4990 0.4993 0.4995 0.4995 0.4998 0.4998 0.4999
1.18 0. 1.19 0. 2.0 0. 2.1 0. 2.2 0. 2.2 0. 2.2 0. 2.2 0. 2.2 0. 2.2 0. 2.1 0. 2.2 0. 2.2 0. 2.5 0. 2.7 0. 2.8 0. 2.9 0. 3.10 0. 3.2 0. 3.3 0. 3.3 0. 3.4 0. 3.5 0. 3.6 0. 3.7 0.	4641 47713 47713 4821 4821 4861 4893 4918 4938 4953 4965 4974 4981 4987 4990 4993 4995 4995 4997 4998 4998 4999	0.4649 0.4719 0.4778 0.4826 0.4826 0.4896 0.4920 0.4920 0.4920 0.4955 0.4925 0.4955 0.4966 0.4975 0.4995 0.4997 0.4993 0.4997 0.4998 0.4998 0.4999	0.4656 0.4726 0.4783 0.4830 0.4868 0.4898 0.4992 0.4922 0.4941 0.4956 0.4967 0.4967 0.4987 0.4987 0.4981 0.4991 0.4995 0.4995 0.4997 0.4999 0.4999	0.4664 0.4732 0.4788 0.4834 0.4871 0.4901 0.49025 0.4943 0.4957 0.4968 0.4977 0.4983 0.4988 0.4991 0.4994 0.4994 0.4997 0.4998 0.4999 0.4999	0.4671 0.4738 0.4793 0.4838 0.4875 0.4904 0.4927 0.4945 0.4945 0.4959 0.4969 0.4969 0.4994 0.4994 0.4994 0.4997 0.4998 0.4999 0.4999 0.4999	0.4678 0.4798 0.4798 0.4878 0.4906 0.4906 0.4929 0.4946 0.4960 0.4970 0.4970 0.4978 0.4984 0.4989 0.4992 0.4994 0.4999 0.4997 0.4998 0.4999 0.4999	0.4686 0.4750 0.4803 0.4803 0.4846 0.4909 0.4931 0.4948 0.4951 0.4971 0.4971 0.4979 0.4985 0.4985 0.4999 0.49994 0.49994 0.49995 0.49999 0.49999	0.4693 0.4756 0.4808 0.4800 0.4884 0.4911 0.4932 0.4949 0.4949 0.4962 0.4979 0.4979 0.4985 0.4995 0.4995 0.4995 0.4995 0.4997 0.4998 0.4999 0.4999	0.4699 0.4761 0.4812 0.4854 0.4887 0.4913 0.4934 0.4951 0.4963 0.4963 0.4980 0.4980 0.4980 0.4990 0.4993 0.4995 0.4995 0.4997 0.4998 0.4999 0.4999	0.4767 0.4817 0.4817 0.4857 0.4890 0.4916 0.4936 0.4936 0.4952 0.4964 0.4974 0.4981 0.4981 0.4986 0.4990 0.4993 0.4995 0.4995 0.4997 0.4998 0.4998 0.4999 0.4999

Table A. 1 Table de la loi normale centrée réduite – Table Z.

ANNEXE B : Performances de la méthode multimodale par l'iris droit, le visage et l'iris gauche

Nous montrons, dans cette annexe, les performances de l'association de normalisation/fusion Tanh/Somme sur la méthode multimodale Iris_D/Visage/Iris_G pour les quatre expériences IV². Pour chaque expérience sont montrées en haut : les distributions normalisées des scores des méthodes monomodales correspondantes, au milieu : les courbes DET (échelles linéaire et logarithmique) et en bas : les distributions des scores authentiques et imposteurs de la méthode multimodale.



Figure B. 1 Performances de l'association Tanh/Somme sur la méthode multimodale Iris_D/Visage/Iris_G pour l'expérience 1.



Figure B. 2 Performances de l'association Tanh/Somme sur la méthode multimodale Iris_D/Visage/Iris_G pour l'expérience 2.



Figure B. 3 Performances de l'association Tanh/Somme sur la méthode multimodale Iris_D/Visage/Iris_G pour l'expérience 3.



Figure B. 4 Performances de l'association Tanh/Somme sur la méthode multimodale Iris_D/Visage/Iris_G pour l'expérience 4.

ANNEXE C : Description des algorithmes multimodaux de la campagne IV²

Dans cette annexe, nous nous référons au rapport technique [Colineau 2007] pour reporter les performances liées à l'évaluation multimodale de la campagne d'évaluation IV² lancée en 2007. Nous commençons d'abord par les algorithmes monomodaux qui ont servi à la conception des algorithmes multimodaux. Il s'agit de quatre algorithmes dont un est relatif à la modalité de l'iris [INT (EK) – Alg1], deux à celle du visage 2D [IBISC (AC) et INT (AM)] et un dernier à celle du visage 3D [TRT (JHD)]. Ensuite, nous présentons les caractéristiques et les résultats de la fusion multimodale opérée.

C.1 Algorithmes Monomodaux

Igorithmes multimodaux de la campagne d'evaluation 1v ⁻ .									
	Iris	Visa	ge 2D	Visage 3D					
	INT (EK) – Alg1	IBISC(AC)	INT (AM)	TRT (JHD)					
Algorithme	phase de Gabor	LDA	LDA + Gabor	ICP2 (Iterative Closest Point)					
Distance	corrélation locale et globale de la phase de Gabor + mesures de qualité par GMM	L1	cosinus	-					
Normalisation	min-max	log	min-max	log					

Table C. 1 Description des algorithmes monomodaux participant dans la conception desalgorithmes multimodaux de la campagne d'évaluation IV^2 .

Table C. 2 Résultats expérimentaux des algorithmes monomodaux des modalités de l'iris et du visage 3D ayant participé dans la conception des algorithmes multimodaux de la campagne d'évaluation IV².

	Iris	Visage 3D sans expression				
Algorithme	INT (EK) – Alg1	TRT (JHD) intra-session	TRT (JHD) inter-session			
moyenne intra	0,646	0,599	0,802			
écart-type intra	0,152	0,049	0,071			
moyenne inter	0,185	0,350	0,445			
écart-type inter	0,047	0,083	0,127			
EER	0,020 (±0,002)	0,020 (±0,005)	0,038 (±0,011)			
FRR (FAR=0,001)	0,040	-	-			
FRR (FAR=0,1)	-	0,003	0,014			

	Exp1 intra-session		Exp2 intra-session, éclairage		Exp3 intra-session, qualité		Exp4 inter-session	
Algorithme	IBISC (AC)	INT (AM)	IBISC (AC)	INT (AM)	IBISC (AC)	INT (AM)	IBISC (AC)	INT (AM)
moyenne intra	0,607	0,886	0,406	0,788	0,564	0,768	0,477	0,786
écart-type intra	0,124	0,083	0,136	0,119	0,157	0,116	0,122	0,117
moyenne inter	0,254	0,437	0,240	0,442	0,399	0,436	0,300	0,450
écart-type inter	0,075	0,159	0,080	0,154	0,132	0,159	0,092	0,149
EER	0,037 (±0,006)	0,042 (±0,007)	0,255 (±0,014)	0,120 (±0,011)	0,272 (±0,019)	0,125 (±0,014)	0,198 (±0,015)	0,113 (±0,012)
FRR (FAR=0,1)	0,021	0,021	0,310	0,138	0,498	0,170	0,320	0,126

Table C. 3 Résultats expérimentaux des algorithmes monomodaux de la modalité du visage 2D

 ayant participé dans la conception des algorithmes multimodaux de la campagne d'évaluation IV².

C.2 Algorithmes multimodaux

Expériences multi-session

Table C. 4 Description et performances des algorithmes multimodaux de la campagne IV^2 pour
l'expérience multi-session Test 03a.

Algorithmes unimodaux fusionnés	Iris [INT (EK) – Alg1] Visage 2D [INT (AM)] Visage 3D [TRT (JHD)]						
Algorithme multimodal	TRT1	TRT2	TRT3	TRT4	INT1	INT2	INT3
Règle de fusion	produit	produit	somme	somme	moyenne	-	-
Normalisation	min-max	min-max + gaussienne	Tanh	Bayes + fenêtres de Parzen	Bayes + gaussienne	Bayes 3D	Bayes + GMM
EER	0,0053 (±0,004)	0,0046 (±0,004)	0,0046 (±0,004)	0,0042 (±0,004)	0,0080 (±0,005)	0,0056 (±0,005)	0,0054 (±0,004)
FRR (FAR=0,01)	0,0021	0,0005	0,0027	0,0013	0,0040	0,0040	0,0018

Table C. 5 Description et performances des algorithmes multimodaux de la campagne IV^2 pour
l'expérience multi-session Test 03b.

Algorithmes unimodaux fusionnés	Iris [INT (EK) – Alg1] Visage 2D [IBISC (AC)] Visage 3D [TRT (JHD)]						
Algorithme multimodal	TRT1	TRT2	TRT3	TRT4	INT1	INT2	INT3
Règle de fusion	produit	produit	moyenne	moyenne	moyenne	-	-
Normalisation	min-max	min-max + gaussienne	Tanh	Bayes + fenêtres de Parzen	Bayes + gaussienne	Bayes 3D	Bayes + GMM
EER	0,0040 (+0.004)	0,0059 (+0.005)	0,0053 (+0.004)	0,0063 (+0.005)	0,0041 (+0.004)	0,0075	0,0014 (+0.002)
FRR (FAR=0,01)	0,0027	0,0000	0,0027	0,0053	0,0000	0,0053	0,0013

Expériences mono-session

Algorithmes unimodaux fusionnés	Iris [INT (EK) – Alg1] Visage 2D [INT (AM)] Visage 3D [TRT (JHD)]						ID)]
Algorithme multimodal	TRT1	TRT2	TRT3	TRT4	INT1	INT2	INT3
Règle de fusion	produit	produit	somme	somme	moyenne	-	-
Normalisation	min-max	min-max + gaussienne	Tanh	Bayes + fenêtres de Parzen	Bayes + gaussienne	Bayes 3D	Bayes + GMM
EER	0,0035 (±0,003)	0,0070 (±0,005)	0,0023 (±0,003)	0,0016 (±0,002)	0,0022 (±0,003)	0,0041 (±0,003)	0,0023 (±0,003)
FRR (FAR=0,01)	0,0008	0,0005	0,0015	0,0000	0,0000	0,0020	0,0010

Table C. 6 Description et performances des algorithmes multimodaux de la campagne IV^2 pour
l'expérience mono-session Test 04a.

Table C. 7 Description et performances des algorithmes multimodaux de la campagne IV^2 pour
l'expérience mono-session Test 04b.

Algorithmes unimodaux fusionnés	nnés Iris [INT (EK) – Alg1] Visage 2D [IBISC					e 3D [TRT (J]	HD)]
Algorithme multimodal	TRT1	TRT2	TRT3	TRT4	INT1	INT2	INT3
Règle de fusion	produit	produit	somme	somme	moyenne	-	-
Normalisation	min-max	min-max + gaussienne	Tanh	Bayes + fenêtres de Parzen	Bayes + gaussienne	Bayes 3D	Bayes + GMM
EER	0,0015 (±0,002)	0,0022 (±0,003)	0,0016 (±0,002)	0,0000 (±0,001)	0,0020 (±0,002)	0,0051 (±0,004)	0,0012 (±0,002)
FRR (FAR=0,01)	0,0000	0,0000	0,0000	0,0000	0,0000	0,0049	0,0010



Titre : Biométrie multimodale basée sur l'iris et le visage

Mots clés : Biométrie multimodale, Biométries par l'iris et par le visage, Extraction/sélection de caractéristiques, Méthodes bio-inspirées, Méthodes multi-algorithmes, Normalisation/Fusion de scores.

Résumé : Cette thèse vise à apporter une contribution dans le domaine de la biométrie par l'iris (l'une des modalités les plus précises et difficiles à pirater) et le visage (l'une des modalités les moins intrusives et les moins coûteuses).

A travers ce travail, nous abordons plusieurs aspects importants de la biométrie mono et multimodale. Nous commençons par dresser un état de l'art sur la biométrie monomodale par l'iris et par le visage et sur la multimodalité iris/visage, avant de proposer plusieurs approches personnelles de reconnaissance d'individus pour chacune des deux modalités. Nous abordons, en particulier, la reconnaissance faciale par des approches classiques reposant sur des combinaisons d'algorithmes et des approches bioinspirées émulant le mécanisme de la vision humaine. Nous démontrons l'intérêt des approches bio-inspirées par rapport aux approches classiques à travers deux méthodes. La première exploite les résultats issus de travaux neuroscientifiques indiquant l'importance des régions et des échelles de décomposition utiles à l'identification d'un visage.

La deuxième consiste à appliquer une méthode de codage par ordre de classement dans la phase de prétraitement pour renforcer le contenu informatif des images de visage. Nous retenons la meilleure approche de chacune des modalités de l'iris et du visage pour concevoir deux méthodes biométriques multimodales. A travers ces méthodes, nous évaluons différentes stratégies classiques de fusion multimodale au niveau des scores. Nous proposons ensuite une nouvelle règle de fusion de scores basée sur un facteur de qualité dépendant du taux d'occultation des iris. Puis, nous mettons en avant l'intérêt de l'aspect double échantillons de l'iris dans une approche multimodale.

L'ensemble des méthodes proposées sont évaluées sur la base multimodale réelle IV² capturée dans des environnements variables voire dégradés et en suivant un protocole bien précis fourni dans le cadre de la campagne d'évaluation IV². Grâce à une étude comparative avec les algorithmes participants à la campagne IV², nous prouvons la compétitivité de nos algorithmes qui arrivent dans plusieurs cas à se positionner en tête de liste.

Title : Multimodal biometrics based on iris and face

Keywords : Multimodal biometrics, Iris and face biometrics, Feature extraction/selection, Bio-inspired methods, Multi-algorithm methods, Score normalization/fusion

Abstract : This thesis aims to make a contribution in the field of biometrics based on iris (one of the most accurate and hard to hack biometrics) in conjunction with face (one of the cheapest and less intrusive biometrics).

Through this work, we discuss several important aspects of unimodal and multimodal biometrics. After an overview on unimodal and multimodal biometrics based on iris and face, we propose several personal approaches of biometric authentication using each single trait. Particularly, we address facial recognition first with conventional approaches based on combined algorithms, then with bio-inspired approaches emulating the human vision mechanism. We demonstrate the interest of bio-inspired approaches over conventional approaches through two proposed methods. The first one exploits the results of neuroscientific work indicating the relevant regions and scales in a face identification task. The second consists in applying a rank order coding method at the preprocessing step so as to enhance the information content of face images.

We keep the best unimodal approach of iris and face recognition to design two multimodal biometric methods. Through these methods, we evaluate different classic strategies of multimodal score-level fusion. Afterwards, we propose a new score-level fusion rule based on a quality metric according to irises occultation rates. Then, we point out the interest of the double-sample iris aspect in a multimodal approach.

All the proposed methods are evaluated on the real multimodal IV^2 database captured under variable to degraded environments, and following a specific protocol provided as part of the IV^2 evaluation campaign. After a comparative study with the participant algorithms in the IV^2 campaign, we prove the competitiveness of our algorithms witch outperform most of the participant ones in the IV^2 campaign in many experiments.