

Université Mohamed Khider - Biskra

Faculté des Sciences ET de la Technologie

Département: Genie Electrique

Filière: Electronique

Réf: .....

جامعة محمد خيضر بسكرة

كلية العلوم و التكنولوجيا

قسم: الهندسة الكهربائية

فرع: إلكترونيك

المرجع: .....



## Mémoire

Présenté en vue de l'obtention du diplôme de magister en Electronique

Option : **Signaux et Communications**

# Etude de la fusion de modalités pour l'authentification en biométrie (visage, voix)

Par: **OUAMANE Abdelmalik**

Soutenu le: 07/12/ 2011

Devant le jury :

<b>SBAA Salim</b>	Maître de conférences A	Université de Biskra	Président
<b>BENYOUCEF Moussa</b>	Professeur	Université de Batna	Examineur
<b>BAARIR Zine Eddine</b>	Maître de conférences A	Université de Biskra	Examineur
<b>BENAKCHA Abdelhamid</b>	Maître de conférences A	Université de Biskra	Rapporteur



## Résumé

L'identification et/ou la vérification des visages possède plusieurs avantages sur les autres technologies biométriques : elle est naturelle, non intrusive et facile à utiliser. Les systèmes biométriques (unimodaux) permettent de reconnaître une personne en utilisant une seule modalité biométrique, mais ne peuvent pas garantir avec certitude une bonne identification. Alors la solution est la mise en place de systèmes biométriques multimodaux obtenus en fusionnant plusieurs systèmes de reconnaissance de visages. Dans ce présent travail, nous abordons plusieurs points importants concernant la biométrie multimodale. Tout d'abord, après avoir dressé un état de l'art de la reconnaissance de visages et étudié plusieurs méthodes pour sélectionner les meilleurs systèmes d'authentification de visages. Ensuite, nous présentons la multimodalité et les méthodes de fusion de score pour les deux approches combinaison et classification. Enfin les méthodes de fusion de scores sont comparées sur la base de données des visages XM2VTS et les scores de visages et voix de XM2VTS selon son protocole associé (protocole de Lausanne 1).

## Abstract

The identification and / or verification of faces have several advantages over other biometric technologies: it is natural, not intrusive and easy to use. Unimodal biometric systems can recognize a person using a single biometric modality, but cannot guarantee with certainty the proper identification. So the solution is the establishment of multimodal biometric systems obtained by merging several face recognition systems. In this work, we discuss several important issues concerning the multimodal biometrics. First, after taking a state of the art of face recognition and studied several methods for selecting the best face authentication systems. Then, we present multimodality and fusion methods score for both combination and classification approaches. Finally, the fusion methods are compared on the XM2VTS database of faces and scores of faces and voices according XM2VTS its associated protocol (Lausanne Protocol 1).

## ملخص

تحديد و/أو تحقق من الوجوه عدة مزايا أكثر من غيرها من التقنيات البيومترية (طبيعية وسهلة الاستعمال). يمكن لنظم التحقق من الهوية الأحادي الوسائط (xquadominu) التعرف على الشخص الذي يستخدم طريقة واحدة بيومترية, ولا كن هذا لا يضمن تحديد الهوية بشكل جيد. وبالتالي فلن الحل هو تنفيذ نظم التحقق من الهوية متعدد الوسائط (xquadomitlum) التي تم الحصول عليها عن طريق دمج العديد من أنظمة التعرف على الوجه. في هذا العمل, نهن بصدد مناقشة عدة قضايا هامة تتعلق بالقياسات البيومترية متعددة الوسائط. بداية بعد أن نقدم طرق التعرف على الوجه, دراسة العديد منها واختيار أفضلها. ثم نقدم نظم متعددة الوسائط وطرق دمج النتائج (serocs). أخيرا, نقوم بمقارنة طرق إدماج بواسطة قاعدة بيانات الوجوه و قاعدة بيانات النتائج الوجوه و الأصوات XM2VTS.

# Remerciements

Je tiens avant tout à remercier Madame **BENATIA BELAHCENEM** Mébarka qui m'a proposée ce sujet et sans laquelle je ne serais sûrement pas arrivé là ; certaines idées sont nées de conversations que nous avons eues ensemble. Sa rigueur, ses compétences techniques et sa grande pédagogie sont un exemple. Je la remercie également pour son temps et son investissement dans tous les aspects de mon travail et de ma rédaction.

Je remercie également Monsieur **BENAKCHA** Abdelhamid qui m'a encadré durant ces années de mémoire de magister. C'est grâce à nos discussions qu'une grande partie de nos travaux a abouti. Il a su faire preuve d'une pédagogie et d'une gentillesse remarquables et j'espère que nous pourrons continuer nos travaux dans le futur. Je le remercie pour ces grandes qualités humaines.

Je tiens également à remercier tous les membres du Jury : Mr **BAARIR** Zine Eddine, Mr **SBAA** Salim pour leur lecture et leurs commentaires très intéressants sur mon travail. Et en particulier, je remercie Mr **BENYOUCEF** Moussa, qui a accepté d'examiner mon travail et de l'évaluer et merci Monsieur pour votre déplacement pour assister à ma soutenance.

Un Grand Merci à Monsieur **AHMED TALEB** Abdelmalik qui nous a beaucoup aidé durant la réalisation de notre travail. Nous avons eu d'innombrables échanges scientifiques très constructifs. Ses conseils étaient très bénéfiques et son apport en idées et en documentation était très important pour nous.

Je tiens à remercier tout particulièrement ma famille et mes amis proches.

*Mon père* qui a toujours su être là dans tous les moments, les bons comme les difficiles ; son soutien et sa générosité constante ont été fondamentaux, sa droiture et sa conduite ont toujours été un modèle pour moi.

*Ma mère* qui a toujours su m'écouter; sa douceur, sa tendresse et sa volonté ont toujours mérité mon plus profond respect.

De manière générale, je remercie tous mes amis qui m'ont apporté un grand soutien et permis de passer de très bons moments de joie et de détente qui ont été nécessaires.

Je souhaite à présent adresser mes sincères remerciements à toutes les personnes avec qui j'ai eu la chance de travailler ou que j'ai eu l'honneur de côtoyer avant et pendant le déroulement de mon magister.

**Sommaire**

Sommaire.....	I
Listes des figures.....	I
Listes des tableaux.....	I
<b>Introduction générale.....</b>	<b>1</b>
<b>1. Reconnaissance de visage.....</b>	<b>5</b>
1.1 Introduction.....	5
1.2 État de l'art des techniques de reconnaissance de visage.....	6
1.2.1 Méthodes globales.....	6
1 Les techniques linéaires.....	7
2 Les techniques non linéaires.....	7
1.2.2 Méthodes locales.....	7
1.2.3 Méthodes hybrides.....	8
1.3 Présentation de modèle.....	10
1.3.1 Extraction des caractéristiques des visages.....	11
1 Extraction des caractéristiques par la transformée de Hough.....	11
2 Extraction des caractéristiques par les ondelettes de Gabor.....	14
1.3.2 Les algorithmes de réduction d'espace de données.....	17
1 Analyse en Composantes Principales (ACP).....	19
2 Analyse Discriminante Linéaire.....	21
3 Le modèle discriminant linéaire amélioré de Fisher (Enhanced Fisher linear discriminant Model (EFM)).....	24
4 La LDA directe (DLDA).....	26
5 L'Analyse Discriminante Linéaire Bidimensionnelle Orientée (ADL2Do).....	27
1.3.3 Comparaison.....	28
1.4 Mesure de la performance des systèmes d'authentification biométriques.....	29
1.4.1 Intervalle de confiance.....	31
1. Méthode non-paramétrique.....	31
2. Méthode paramétrique.....	31
1.5 Conclusion.....	32

<b>2. Authentification de visages et étude</b> .....	35
2.1 Introduction.....	35
2.2 La base de données XM2VTS.....	35
2.3 Comparaison des algorithmes de réduction l'espace de donnée.....	37
2.3.1 Prétraitement d'image.....	37
1 Découpage.....	37
2 Décimation.....	38
3 La composante S de l'espace de couleur TSV de l'image.....	38
4 Enchaînement des (lignes/ou colonnes).....	39
5 Photo normalisation.....	40
2.3.2 Les algorithmes de réduction d'espace.....	40
2.3.3 Projection des images.....	42
2.3.4 Comparaison.....	42
2.3.5 Résultats de simulation.....	43
2.4 La transformation de Hough.....	45
2.5 Les ondelettes de Gabor.....	47
2.5.1 Extraction des caractéristiques.....	47
2.5.2 Représentation du visage et choix des filtres de Gabor.....	50
1 Influence des caractéristiques de la famille des filtres de Gabor sur les performances de l'authentification de visages.....	50
2 Utilisation de la phase de Gabor pour l'authentification de visages.....	51
2.6 Conclusion.....	53
<b>3. La multimodalité</b> .....	54
3.1 Introduction.....	54
3.2 Les limitations des systèmes biométriques monomodaux.....	54
3.3 La multimodalité.....	56
3.3.1 Les différents multi- possibles.....	56
3.3.2 Les architectures.....	57
3.3.3 Les niveaux de fusion.....	58
1 La fusion pré-classification.....	59
2 La fusion post-classification.....	60

3.3.4	Choix des types de multimodalité et du niveau de fusion.....	61
1	Choix du type de multimodalité.....	61
2	Niveau de fusion.....	62
3.3.5	État de l'art de la fusion de scores.....	62
1	Approche par classification de scores.....	63
2	Approche par combinaison de scores.....	63
3.4	Conclusion.....	65
<b>4.</b>	<b>Les méthodes de fusion de scores.....</b>	<b>66</b>
4.1	Introduction.....	66
4.2	Normalisation de scores.....	67
4.2.1	Normalisation par la méthode Min-Max.....	68
4.2.2	Normalisation par la méthode Z-Score.....	68
4.2.3	Normalisation par la méthode tangente hyperbolique "Tanh".....	69
4.2.4	Normalisation par la médiane et l'écart absolu médian (MAD).....	69
4.2.5	Normalisation par une fonction Quadratique-Linéaire-Quadratique (QLQ).....	69
4.2.6	Normalisation par une fonction double sigmoïde.....	70
4.3	Approche par combinaison de scores.....	71
4.3.1	Méthode de combinaisons simples.....	71
4.3.2	Combinaison de scores par logique floue.....	72
1	Mesure floue.....	73
2	Intégrale floue.....	74
4.4	Approche par classification de scores.....	75
4.4.1	Fusion par méthode statistique de Fisher.....	75
4.4.2	Les machines à vecteurs de support (SVM).....	77
4.4.3	Réseaux de neurones artificiels.....	79
4.5	Conclusions.....	81
<b>5.</b>	<b>Conception du système multimodale.....</b>	<b>83</b>
5.1	Introduction.....	83
<u>Application I</u>		
5.2	La multi-modalité (multi-algorithmes).....	83
5.2.1	Les distributions et normalisation de scores.....	85
5.2.2	Comparaison des méthodes de normalisation.....	88

1 Méthodes de normalisation associées aux méthodes de combinaisons	
Simples.....	88
2 Comparaison des méthodes de normalisation par logique floue.....	90
5.2.2 Méthodes de normalisation associées aux méthodes de classifications.....	92
1 Classification de scores par méthode statistique de Fisher.....	92
2 Classification de scores avec les machines à vecteurs de support (SVM)....	93
3 Classification de scores avec réseaux de neurones artificiels.....	95
<u>Application II</u>	
5.3 La multi-modalité (multi-biométries).....	97
5.3.1 Les scores de visage et voix de la base de données XM2VTS.....	97
5.4. Conclusion.....	102
<b>Conclusion générale</b> .....	103
<b>Annexe</b> .....	108
<b>Bibliographie</b> .....	124

**List des figures**

Figure 1.1 : Scores de compatibilité pour différentes technologies biométriques dans un système MRTD.....	5
Figure 1.2 : Principe de fonctionnement d'un système de reconnaissance de visage.....	10
Figure 1.3: Transformée de Hough, (a) Plan cartésien (xy), (b) Plan des paramètres (ab)...	12
Figure 1.4 : Quantification du plan des paramètres (ab).....	13
Figure 1.5 : Paramétrage polaire d'une droite.....	13
Figure 1.6 : Quantification du plan des paramètres ( $\rho \theta$ ).....	14
Figure 1.7 : Transformée de Hough, (a) Plan cartésien (xy), (b) Plan des paramètres (ab)...	14
Figure 1.8 : Orientation de l'ondelette.....	15
Figure 1.9 : Fréquence centrale de l'ondelette.....	16
Figure 1.10 : Phase de l'ondelette.....	16
Figure 1.11 : Support temporel de l'ondelette.....	16
Figure 1.13 : Forme de l'enveloppe gaussienne.....	17
Figure 1.14: Les 80 masques d'ondelettes de Gabor.....	17
Figure 1.15 : Comparaison entre les projections de deux classes de points ("class 1" et "class 2") sur les axes principaux construits par ACP et par LDA.....	22
Figure 1.16 : Illustration du TFR et du TFA.....	30
Figure 1.17 : Courbe caractéristique de fonctionnement (ROC) d'un système de Vérification Biométrique.....	30
Figure 2.1: Exemples des images de la base de données XM2VTS.....	36
Figure 2.2: Exemples des images d'une personne de la base de données XM2VTS.....	36
Figure 2.3 : l'image de visage a)avant et b) après découpage.....	37
Figure 2.4 : image de visage a)avant b) après décimation.....	38
Figure 2.5: Différents systèmes colorimétriques : (a) plan du système de couleurs RVB, (b) cylindre du système de couleurs TSV.....	38
Figure 2.6: (a) Image découpée et décimer, (b) Image en système TSV, (c) La composons S de TSV.....	39
Figure 2.7 : Comparaison des TEE pour ACP, LDA, EFM et DLDA dans l'ensemble d'évaluation.....	43
Figure 2.8 : TEE dans l'ensemble d'évaluation pour différentes fonctions de mesure de distances pour (LDA2DoL).....	43

Figure 2.9 : TEE dans l'ensemble d'évaluation pour différentes fonctions de mesure de distances pour (LDA2DoC).....	44
Figure 2.10 : (a) la composons S de l'image découpée et décimée (b) après le filtrage par Roberts.....	45
Figure 2.11 : Organigramme de la Transformée de Hough.....	46
Figure 2.12 : TEE dans l'ensemble d'évaluation en fonction de nombre des caractéristiques	47
Figure 2.13: Résultats de la convolution d'une image de visage avec une famille de 40 filtres de Gabor (8 orientations (horizontales) et 5 résolutions (verticales)).(a) Image en espace de couleur TSV , l'ensemble (b) représente les amplitudes et (c) les phases de cette convolution.....	49
Figure 2.14 : Taux d'égale erreur dans l'ensemble d'évaluation en fonction du seuil $Th$ .....	51
Figure 3.1 : Les différents systèmes multimodaux.....	57
Figure 3.2 : Architecture de fusion en parallèle.....	58
Figure 3.4 : Les différents niveaux de fusion.....	59
Figure 4.1 : Schéma de la fusion de scores.....	66
Figure 4.2 : Normalisation QLQ.....	70
Figure 4.3 : Normalisation double sigmoïde.....	71
Figure 4.4 : Séparation linéaire dans un espace à deux dimensions.....	77
Figure 4.5 : Exemple de perceptron multicouches.....	80
Figure 5.1 : Organigramme du Système de fusion de scores des quatre systèmes d'authentification de visages.....	85
Figure 5.2 : Les distributions de scores des quatre systèmes d'authentification de visages sur la base de données XM2VTS. ....	86
Figure 5.3 : Normalisation de scores par la méthode du Min-Max.....	87
Figure 5.4 : Normalisation de scores.....	88
Figure 5.5 : la séparation non linéaire de SVM de distributions 2D des scores des six couples des systèmes pour les deux classes Client et Imposteur avec la normalisation de scores Z-norme.....	95
Figure 5.6 : Les distributions de 5 systèmes d'authentification de visages et 3 systèmes d'authentification de voix sur la base de données XM2VTS.....	99

## Liste des tableaux

Tableau 1.1 : Comparaison des performances des algorithmes proposés sur les bases de visages :Yale, AR, FERET, XM2VS, BANCA (G1) et ORL .....	9
Tableau 2.1 : Répartition des photos dans les différents ensembles.....	36
Tableau 2.2 : Performances des méthodes ACP, LDA, EFM, DLDA, LDA2DoL et LDA2DoC (intervalle de confiance paramétrique à 95%). .....	44
Tableau 2.3 : TF de Hough dans l'ensemble d'évaluation et test ( intervalle de confiance paramétrique à 95%).....	47
Tableau 2.4 : Taux d'égalité erreur TEE dans l'ensemble d'évaluation pour différent niveaux de résolutions et d'orientations pour les 40 filtres.....	50
Tableau 2.5 : Taux d'égalité erreur TEE dans l'ensemble d'évaluation pour différent résolution et d'orientations des filtres pour les 40 filtres.....	52
Tableau 2.6 : Résultats du système d'authentification de visages pour les quatre meilleurs systèmes retenus dans l'ensemble d'évaluation et test (intervalle de confiance paramétrique à 95%).....	52
Tableau 5.1: Performances des méthodes de normalisation qui ne modifient pas la forme des distributions associées aux méthodes de combinaison de scores simples (intervalle de confiance paramétrique à 95%).....	89
Tableau 5.2 : Performances des méthodes de normalisation qui modifient la forme des Distributions associées aux méthodes de combinaison de scores simples (intervalle de confiance paramétrique à 95%).....	90
Tableau 5.3 : Performances des méthodes de normalisation par logique floue ( intervalle de confiance paramétrique à 95%).....	92
Tableau 5.4 : Performances des méthodes de normalisation par fusion statistique de Fisher dans d'ensemble évaluation et test ( intervalle de confiance paramétrique à 95%).....	93
Tableau 5.5 : Performances des méthodes de normalisation associées au SVM dans d'ensemble évaluation et test (intervalle de confiance paramétrique à 95%).	94
Tableau 5.7: Performances des méthodes de normalisation associées à la classification MLP (intervalle de confiance paramétrique à 95%).....	96
Tableau 5.8: Performances des 5 systèmes de visages et 3 systèmes de voix (intervalle de confiance paramétrique à 95%).....	98
Tableau 5.9 : Performances des combinaisons de scores pour les 15 couples des systèmes (voix, visage) (intervalle de confiance paramétrique à 95%).....	99
Tableau 5.10: Performances des classification de scores pour les 15 couples des systèmes (voix, visage) ( intervalle de confiance paramétrique à 95%).....	100
Tableau 5.11 : Tableau de synthèse des résultats obtenus en fusion multimodale.....	101

### Introduction générale

La **Biométrie** est une science qui s'est développée depuis le début du siècle passé. Elle a été employée au début pour identifier les criminels. Cette technique initiée par **Alphonse Bertillon (1853-1914)**. Ce criminologue français fonda en **1870** le premier laboratoire de police scientifique d'identification criminelle et inventa l'**anthropométrie judiciaire** appelée "système Bertillon" ou "bertillonnage". Le système de reconnaissance des personnes, proposé par Bertillon, est basé sur la mesure de différents paramètres du corps humain [1]. Dans ses travaux, Bertillon affirme qu'en prenant **quatorze mensurations** (taille, pieds, mains, nez, oreilles, etc.) sur n'importe quel individu, il n'y a qu'une seule chance sur deux cent quatre-vingt-six millions pour qu'on retrouve les mêmes mesures chez une autre personne. On s'est intéressé ensuite de plus près à l'empreinte des doigts et à étudier sa forme et sa texture. La découverte de l'unicité pour chaque personne de celle-ci par **Johann Christoph Andreas Mayer**, est une étape importante dans la diffusion de ce procédé d'identification auprès des services de polices et de sécurité pour authentifier les signatures faites à l'époque par l'empreinte et pour le contrôle dans les prisons. Ensuite, on a exploré d'autres modalités biométriques : la voix, l'iris, la forme de la main, etc. La modalité la **plus naturelle** de se reconnaître entre Hommes est le **visage**. Elle est sujette à de nombreuses études pour comprendre ce mécanisme de reconnaissance et comment l'homme arrive à différencier les visages.

Les premières études théoriques sur la reconnaissance faciale remontent au moins au début des années **1970**. Le passage vers des produits commerciaux n'a reçu une impulsion décisive qu'à partir des années 1994-1996 en bonne partie grâce à la mise en œuvre d'un programme d'évaluation international **FERET** (Face Recognition Technology), organisé par le ministère de la Défense américaine (*Department of Defense, DoD*). On dispose alors d'une base d'images de référence pour les tests. La fin du programme en 1996 a marqué l'apparition sur le marché des premiers produits commerciaux. Le vif intérêt du marché potentiellement gigantesque a fait éclore un grand nombre d'algorithmes de reconnaissance de visages ou de variantes, dont la plupart n'étaient même pas présents lors des tests d'évaluation FERET. Le **DoD** décida alors de lancer des programmes d'évaluation, **FRVT** (*Facial Recognition Vendor Test*) successivement en 2000, 2002 et en 2006 dont l'objectif était d'évaluer les performances des produits commerciaux.

Comme tout autre sujet de reconnaissance de formes, les problèmes liés à la reconnaissance de visages peuvent être caractérisés par les deux difficultés suivantes : (a) la **variabilité intra-classes** et (b) la **similarité inter-classes**. La première difficulté est d'abord liée aux variations intrinsèques du visage produites par exemple par les changements d'expressions faciales, l'évolution de la physionomie du visage et les transformations du visage voulues ou subies. Cette **variabilité** englobe également des **variations environnementales** provoquées par exemple par les changements de conditions d'éclairage, l'orientation du visage de la personne ou l'occlusion d'un visage par des accessoires ou par d'autres objets. La seconde difficulté caractérise la **similarité** entre **des classes** d'individus différents. Ici, les apparences faciales de deux personnes différentes peuvent être similaires par exemple en raison des liens familiaux (jumeaux, jumelles, père/fils, mère/fille, etc.) ou plus généralement de l'appartenance à une même race.

Alors, le système biométrique de reconnaissance de visages et tous les systèmes biométriques en générale utilisant un **seul système** (appelés systèmes monomodaux) ont trois limitations principales : une limitation en termes de **performances**, une limitation en termes de **d'universalité** d'utilisation et une limitation en termes de **détection des fraudes**. La première limitation concerne les performances en reconnaissance parce que les biométries sont des caractéristiques physiques variables dans leur acquisition mais aussi dans leur nature. Le visage change selon les expressions, les conditions d'éclairage, la position et l'instant de capture. On ne pourra jamais trouver deux images du même visage strictement identiques alors qu'elles correspondent cependant à la même personne. Cette variabilité des caractéristiques biométriques et de leur acquisition entraîne des performances en reconnaissance d'identité plus ou moins bonnes selon les biométries utilisées. Par exemple on considérera qu'un système d'empreinte digitale fait de l'ordre 0.1% à 1% d'erreurs (il s'agit d'un ordre de grandeur comparatif, cette valeur peut être variable selon les capteurs, les personnes et les conditions d'acquisition), alors qu'un système de visage aura un taux d'erreur pouvant varier de 1% à 30% selon la variabilité des images.

La deuxième limitation est due à la non-universalité de certaines biométries. La non-universalité signifie que certaines biométries ne sont pas possédées par le sujet à authentifier ou ne sont pas assez informantes pour permettre la vérification de l'identité de certaines personnes. Par exemple, certaines personnes peuvent avoir les empreintes digitales inutilisables à cause d'un accident ou d'un travail manuel prolongé qui a endommagé la peau au niveau des empreintes digitales. Une personne muette ne peut utiliser la reconnaissance par la voix ou une personne handicapée ne peut signer. Pour toutes ces personnes, certains

systèmes biométriques ne sont pas accessibles et ceci risque alors de les exclure de certaines utilisations si aucune alternative ne leur est proposée. La troisième limitation concerne la détection des impostures. En effet, la fraude ou le vol d'identité est un problème qui a toujours existé et qui peut être réduit par la biométrie car il est plus facile de falsifier un mot de passe ou un papier d'identité que de reproduire un visage ou des empreintes digitales. Cependant ces impostures existent, en particulier pour les empreintes digitales qui est, de nos jours, la biométrie la plus utilisée et pour laquelle il est possible de "voler" (car les empreintes digitales laissent des traces) et de reproduire (à l'aide de silicone par exemple) les empreintes digitales d'une autre personne.

Toutes ces limitations peuvent être réduites voire supprimées par l'utilisation conjointe de **plusieurs systèmes biométriques** formant alors un **système biométrique multimodal**. Les systèmes multimodaux permettent d'améliorer les performances de reconnaissance. Ils permettent également de régler le problème de non-universalité de certaines biométries en proposant une alternative aux personnes ne pouvant utiliser certaines biométries. Et enfin, ils permettent de limiter les possibilités de fraude car il est plus difficile d'obtenir et de reproduire plusieurs caractéristiques qu'une seule. Pour toutes ces raisons, les systèmes biométriques multimodaux ont fait l'objet de beaucoup d'études. En effet, si les **premiers travaux sur la biométrie multimodale** ont été publiés au milieu et à la fin des années **1990** [2, 3, 4, 5], ces travaux précurseurs dans le domaine ont été essentiellement basés sur la combinaison de données audio-vidéo (voix et visage). Avec la diversification des biométries et le premier "workshop" **MMUA1 (MultiModal User Authentication)** consacré à la multimodalité en **2003**, plusieurs thèses sur le sujet on peut citer celles de Julian Fierrez-Aguilar [6], de Norman Poh [7], de Jonas Richiardi [8], de Karthik Nandakumar [9], Lorène Allano [10] et Nicolas Morizet [11].

Au moment du début de cette mémoire, beaucoup de questions et des problèmes concernant la biométrie multimodale et les techniques de fusion. Le premier problème a été celui des bases de données multimodales. Lorsque nous n'avons pas trouvé une base de données multimodale nous avons changé la direction des multi-biométries vers les multi-algorithmes de reconnaissance de visages. Alors le deuxième problème est de chercher des systèmes monomodaux de reconnaissances de visages pour cela nous avons consacré une grande partie de notre travail (chapitre1, chapitre2) pour chercher des systèmes efficaces de reconnaissance de visages. Le problème le plus important dans notre travail est la recherche de la méthode de fusion à utiliser parce que les techniques de fusion étaient encore ouvertes. En effet, les premiers travaux ont étudié la fusion de **décisions** par des méthodes de **vote** puis

sont arrivés des méthodes de fusion aux niveaux des scores. Pour la **fusion de scores**, des méthodes plus ou moins complexes de classification ont été proposées et la question de la différence "d'efficacité" entre ces méthodes était encore ouverte. Les études sur la multimodalité depuis l'année 2003 étaient essentiellement axées sur la comparaison des méthodes de fusion et la question : **quelle méthode de fusion et pour quelle situation** ? C'est pour cela que dans le chapitre 4 nous étudions plusieurs méthodes de fusion de scores et le **chapitre 5** présente la comparaison entre ces techniques sur la base de données des visages XM2VTS et les scores voix et visages de la base de scores XM2VTS.

Le reste de ce manuscrit est organisé de la façon suivante :

- **Le chapitre 1** présente l'état de l'art de reconnaissance de visage pour les trois grandes familles des méthodes de reconnaissance: les méthodes globales, locales et hybrides. Notre modèle de reconnaissance de visages est décrit en détail, et nous avons présenté quelques algorithmes utilisés dans les trois étapes d'un système de reconnaissance faciale. Enfin nous présenterons la mesure de la performance d'un système biométrique et l'intervalle de confiance.
- **Le chapitre 2** présente les résultats expérimentaux d'authentification monomodaux de visages et fait les comparaisons des méthodes détaillés dans le chapitre 1.
- **Le chapitre 3** présente les limitations des systèmes biométriques lorsqu'ils n'utilisent qu'une seule modalité biométrique. Et présente aussi la **biométrie multimodale** qui est l'objectif de notre étude
- **Le chapitre 4** récapitule l'ensemble des méthodes de fusion de scores pour les deux approches de fusion, l'approche de combinaison (précédées d'une phase de normalisation), qui traite les scores séparément avant de les combiner. Et l'approche de classification de scores qui cherche également à modéliser la séparatrice des classes dans l'espace multidimensionnel,
- Les résultats expérimentaux de la multimodalité sont décrits dans le **chapitre 5**,

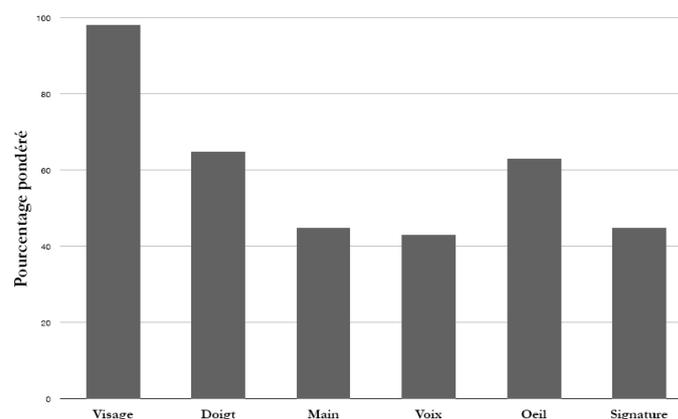
Enfin, nous dressons une conclusion globale sur ces travaux et nous présentons les perspectives futures.

## 1 Reconnaissance de visage

### 1.1 Introduction

L'identification et/ou la vérification (on en trouvera une explication approfondie en *Annexe A*) des visages ont attiré l'attention des chercheurs depuis quelques décennies, et reste encore et toujours un sujet de recherche attractif et très ouvert. Beaucoup de connaissances dans les domaines de la reconnaissance des formes, du traitement d'images, des statistiques ont été appliquées au domaine de la reconnaissance du visage. En plus, les capacités grandissantes des moyens informatiques et l'existence de bases de données de grandes tailles ont permis de mettre au point des algorithmes et des approches de plus en plus complexes et par conséquent, les performances de reconnaissance se sont trouvées améliorées.

La reconnaissance faciale possède plusieurs avantages sur les autres technologies biométriques : elle **est naturelle, non intrusive et facile à utiliser**. Parmi les six attributs biométriques considérés par *Hietmeyer* [12], les caractéristiques faciales marquent un score de compatibilité le plus élevé dans un système *MRTD* (“*Machine Readable Travel Documents*”) [13], ce score étant basé sur plusieurs facteurs d'évaluation tels que l'enrôlement, le renouvellement des données, les requis matériels et la perception des utilisateurs (figure 1.1).



**Figure 1.1 :** Scores de compatibilité pour différentes technologies biométriques dans un système MRTD.

Dans ce chapitre, nous donnerons tout d'abord l'état de l'art des méthodes de reconnaissance de visage 2D dans la section suivante, nous présenterons notre modèle de la

reconnaissance de visage. Enfin dans la dernière section, nous détaillerons la mesure de la performance d'un système biométrique.

## 1.2 État de l'art des techniques de reconnaissance de visage

Les méthodes de reconnaissance faciales peuvent être séparées en deux grandes familles, les méthodes **globales** (ou holistiques) et les méthodes **locales**, basées sur des modèles. Le choix a été fait de se concentrer sur ces deux types d'approches fondamentales et de n'aborder ni les *réseaux neuronaux*(NN) [14] (plus adaptés à la détection des visages), ni les *modèles cachés de Markov* (HMM) [15] (plus utilisés en reconnaissance de la parole) car ces deux dernières techniques rencontrent des problèmes lorsque le nombre d'individus augmente (les calculs deviennent très importants). Cependant, on pourra se référer aux travaux des équipes françaises du *LE2I* [16, 17] et du *GIPSA* [18, 19], qui les ont respectivement appliquées aux images de visage et aux signaux de parole [11].

Dans cette description de l'état de l'art et vu le **grand nombre de méthodes** appliquées à la reconnaissance du visage, la citation de la totalité des approches et des méthodes ne peut être complète. Nous nous sommes donc intéressé à décrire la plupart des grandes méthodes classiques et la majorité des approches qui ont apporté une vraie originalité et une avancée sensible dans le domaine. Dans la première section, nous commencerons par un résumé des méthodes **globales** liés aux méthodes de **réduction d'espace**. Dans la deuxième partie, nous résumerons les méthodes **locales** les plus populaires des algorithmes basés sur les points caractéristiques. Enfin nous finirons par les méthodes **hybrides**.

### 1.2.1 Méthodes globales

Les méthodes **globales** de classification par **réduction d'espace** restent parmi les techniques les plus utilisées et les plus populaires dans le domaine de la reconnaissance du visage 2D, sont basées sur des techniques **d'analyse statistique** bien connues. Il n'est pas nécessaire de repérer certains points caractéristiques du visage (comme les centres des yeux, les narines, le centre de la bouche, etc.) à part pour normaliser les images. **L'avantage** principal des méthodes globales est qu'elles sont relativement **rapides** à mettre en œuvre et que les calculs de base sont d'une **complexité moyenne**. En revanche, elles sont très sensibles aux variations **d'éclairément**, de **pose** et **d'expression faciale**. Ceci se comprend aisément puisque la moindre variation des conditions de l'environnement ambiant entraîne des changements inéluctables dans les valeurs des pixels qui sont traités directement.

Nous pouvons distinguer deux types de techniques parmi les méthodes globales : les *techniques linéaires* et les *techniques non linéaires*.

## 1 Les techniques linéaires

Les techniques linéaires **projetent linéairement** les données d'un espace de grande dimension (par exemple, l'espace de l'image originale) sur un sous-espace de dimension inférieure. Dans un sous-espace linéaire, les distances sont utilisées pour faire comparer des vecteurs de données. La technique linéaire la plus connue et sans aucun doute l'analyse en composantes principales (**PCA**), également appelée *transformée de Karhunen-Loeve* [20]. Une autre méthode qui vise à représenter le visage sans employer la notion de classe est la *factorisation non négative des matrices* (Non Negative Matrix Factorization, **NMF**) [21] Ou encore *l'analyse en composantes indépendantes* (**ICA**) [22].

Cependant, le PCA classique nécessite que les images de visage soient mises sous formes de vecteurs, ce qui a pour effet de détruire la structure géométrique de l'image. Pour ne pas perdre les informations de voisinage lors du passage de l'image en vecteur, une méthode *PCA bi-dimensionnelle* (**2-D PCA**) [23] a été étudiée. Cette méthode prend en entrée des images et non plus des vecteurs.

Il existe d'autres techniques également construites à partir de *décompositions linéaires* comme *l'analyse discriminante linéaire* (**LDA**) est parmi l'une des méthodes **les plus utilisées pour la reconnaissance de visage** [24], le *modèle discriminant Linéaire amélioré de Fisher* (**EFM**) [25], la directe LDA (**DLDA**) [26] et **RLDA** (Regression LDA) [27].

## 2 Les techniques non linéaires

Afin de pouvoir traiter ce problème de non-linéarité en reconnaissance faciale, de telles méthodes linéaires ont été étendues à des techniques non linéaires basées sur la notion mathématique de noyau ("*kernel*") comme le **Kernel PCA** [28] , le **Kernel LDA** [29] et **Kernel ICA** [30].

### 1.2.2 Méthodes locales

Les méthodes locales, basées sur des modèles, utilisent des connaissances a priori que l'on possède sur la morphologie du visage et s'appuient en général sur des points caractéristiques de celui-ci. Kanade [31] présenta un des premiers algorithmes de ce type. Les approches *Bayésiennes* (comme la méthode **BIC** [32]), les machines à vecteurs de support

(SVM) [33], la méthode des modèles actifs d'apparence (AAM) [34] ou encore la méthode "local binary pattern" (LBP) [35] ont été utilisées dans ce but.

Toutes ces méthodes ont l'avantage de pouvoir modéliser plus facilement les variations de pose, d'éclairage et d'expression par rapport aux méthodes globales. Toutefois, elles sont plus lourdes à utiliser puisqu'il faut souvent placer manuellement un assez grand nombre de points sur le visage, ce qui peut être fait automatiquement et de manière assez fiable par un algorithme de détection [36,11].

### 1.2.3 Méthodes 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. Les techniques hybrides, proches du fonctionnement du **système visuel humain**, elles permettent d'augmenter la stabilité de la performance de reconnaissance lors de changements de pose, d'éclairage et d'expressions faciales.

L'analyse de caractéristiques locales (*Local Feature Analysis : LFA*) [37] et les caractéristiques extraites par ondelettes de Gabor comme : *l'Elastic Graph Matching (EGM)* et *l'Elastic Bunch Graph Matching (EBGM)* [38] et l'algorithme *Log Gabor PCA (LG-PCA)* [39] sont des algorithmes hybrides typiques.

Le tableau 1.1 fournit une comparaison des performances des algorithmes proposés sur les bases de visages : Yale, AR, FERET, XM2VS, BANCA (G1) et ORL [40, 41, 42, 43].

Nous pouvons dire qu'au travers les résultats expérimentaux provenant de différentes sources (tableau 1.1) que la méthode globale des *fisherfaces* (**FLD** particulièrement **EFM**) est très performante, en comparaison avec les techniques citées dans l'état de l'art. Les *fisherfaces* présentent en plus l'avantage d'être facilement généralisables à des problèmes plus complexes. Pour toutes ces raisons notre choix s'est porté sur les méthodes globales à base des *fisherfaces* puis nous utilisons des méthodes d'extraction des caractéristiques associées à une méthode globale de projection linéaire à base des *fisherfaces* afin de construire notre méthode de reconnaissance et c'est ce que nous présentons dans la section suivante.

Bases	Méthodes	Année	Source	Taux de reconnaissance	Commentaires
Yale	Eigenfaces	1991	TP91[43]	71.5%	Stratégie d'évaluation <i>leave-one-out</i>
	Fisherfaces	1998	ZCK98[44]	91.5 %	
	ICA	2002	BMS02[45]	71.5%	
	SVM	2002	Yan02[46]	82 %	
	Eigenfaces à noyau	2002	[46]	75.8 %	
	Fisherfaces à noyau	2002	[46]	93.9 %	
AR	Eigenfaces	1991	[47]	70.7 %	Évalué en moyenne sur les 3 variations d'expressions de la base
	Fisherfaces	1997	BHK97[48]	81.7 %	
	ACL	2001	GSC01[49]	88 %	
FERET	Eigenfaces	1991	[47]	80 % et 20 %	<b>Base d'apprentissage</b> : 3323 images de 1196 personnes. <b>Base de test</b> : fb et fc
	Fisherfaces	1999	[48]	96 % et 59 %	
	EBGM	1998	[38]	95 % et 81 %	
	Eigenfaces	1991	[47]	77 %	<b>5 bases d'apprentissage</b> : Chacune incluant 140 images de 140 personnes. <b>5 bases de test</b> : Chacune avec 140 images ou plus.
	ACP à noyau	2002	[49]	87 %	
	ACP Bayésienne	2002	[48]	95 %	
	ICA	2002	[45]	77 %	
XM2VTS	Eigenfaces	2006	[41]	88.73 %	<b>Protocole 1</b>
	Fisherfaces	2006	[41]	96.34 %	
	<b>EFM</b>	<b>2006</b>	<b>[41]</b>	<b>96.78 %</b>	
	S de l'espace de couleur TSV+LDA	2007	[52]	96.74%	
BANCA (G1)	Eigenfaces	2009	[42]	83.19 %	<b>Protocole Mc</b>
	Direct LDA	2009	[42]	88.9 %	
ORL	Eigenfaces	1991	[47]	97.5 %	Stratégie d'évaluation <i>leave-one-out</i>
	Fisherfaces	1998	[44]	98.5 %	
	ICA	2002	[45]	93.75 %	
	SVM	2002	[49]	97 %	
	Eigenfaces à noyau	2002	[49]	98 %	
	Fisherfaces à noyau	2002	[49]	98.75 %	
	Fisherfaces	2008	[43]	91 %	<b>Partitions aléatoires en Base d'apprentissage</b> : 5 images /personne. <b>Base de test</b> : 5 images /personne.
	2D-LDA	2008	[43]	92.5 %	
	2D-WLDA	2008	[43]	93.5 %	
	SVM	2000	[51]	91.21 %	
	Rdn probabiliste	1997	[52]	96 %	

**Tableau 1.1** : Comparaison des performances des algorithmes proposés sur les bases de visages :Yale, AR, FERET, XM2VS, BANCA (G1) et ORL .

Avec :

**SVM** : Les machines à vecteurs de support (Support Vector Machine),

**ICA** : l'analyse en composantes indépendantes,

**ACL** : Analyse des Caractéristiques Locales,

**EBGM** : Elastic Bunch Graph Matching,

**EFM** : Le modèle discriminant Linéaire amélioré de Fisher,

**2D-LDA** : LDA bi-dimensionnel,

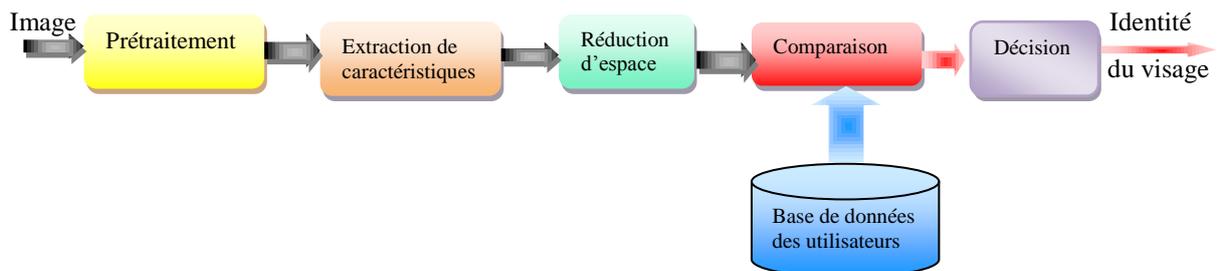
**2D-WLDA** : LDA pondérée bi-dimensionnel,

### 1.3. Présentation de modèle

La reconnaissance faciale est une tâche que les humains effectuent naturellement et sans effort dans leurs vies quotidiennes. La grande disponibilité d'ordinateurs puissants et peu onéreux ainsi que des systèmes informatiques embarqués ont suscité un énorme intérêt dans le traitement automatique des images au sein de nombreuses applications, incluant l'identification biométrique, la surveillance, l'interaction homme-machine et la gestion de données multimédia. La reconnaissance faciale, en tant qu'une des technologies biométriques de base, a pris une part de plus en plus importante dans le domaine de la recherche, ceci étant dû aux **avances rapides** dans des technologies.

Idéalement, un système de reconnaissance faciale doit pouvoir **identifier des visages** présents dans une image. Le système peut opérer dans les deux modes suivants : authentification ou identification. Le principe de fonctionnement de base de notre système de reconnaissance faciale (figure 1.2) peut être résumé en **quatre étapes** :

- Prétraitement,
- Extraction de caractéristiques,
- Réduction d'espace,
- Comparaison.



**Figure 1.2** : Principe de fonctionnement d'un système de reconnaissance de visage.

La première phase c'est la phase de prétraitement pour éliminer le bruit par des techniques de traitement et de restauration d'images et procéder à une localisation de visage dans l'image, cette opération est très complexe et importante puisque les tâches secondaires de vérification dépendent fortement du succès de la technique de localisation. Notre intérêt s'est porté sur les trois phases extraction des caractéristiques, réduction d'espace et la comparaison. Et à chaque étape nous avons essayé de retenir la meilleure méthode au sens de performance et stabilité du système. Cette conception ainsi obtenue représente la première partie de nos travaux. Car, par la suite et une fois le modèle de reconnaissance choisi nous l'appliquons à quatre pour en sortir avec quatre scores. Ces derniers sont normalisés par des techniques finement choisies et fusionnés par deux approches par combinaison et par classification. Finalement une décision est prise concernant le Client ou l'Imposteur.

### 1.3.1 Extraction des caractéristiques des visages

Dans un visage de la même personne, au point que deux images très semblables (à l'œil humain) pourraient être extrêmement différentes si elles sont **comparées pixel par pixel**. Il est donc nécessaire d'extraire les caractéristiques appropriées et discriminantes à partir des images du visage et de comparer ces caractéristiques au lieu de comparer des images approximatives. Naturellement, plus sont les caractéristiques discriminantes, plus la reconnaissance sera facile.

Dans le cadre de notre travail nous optons pour les deux descripteurs suivants pour caractériser les images de visages de la base de données utilisées la **XM2VTS** :

- La transformation de **Hough**.
- Les ondelettes de **Gabor**.

### 1 Extraction des caractéristiques par la transformée de Hough

La transformée de **Hough** est un outil de **détection de courbes paramétriques** dans l'image, elle a été proposée par **Hough** en **1960**[53]. Inaperçue pendant plusieurs années, cette dernière a été vulgarisée par les travaux de **Rosenfield** [54], **Duda** et **Hart** [55] au début des années **70** et fait l'objet par la communauté scientifique depuis cette date à ce jour d'une particulière attention. Depuis les années **80**, elle a quitté les laboratoires de recherche pour trouver des champs d'applications dans de nombreux domaines industriels tels que la vision par ordinateur et le traitement d'images. Elle est devenue une solution plus adaptée au problème de détection des lignes droites, cercles ou toute autre forme paramétrique dans l'image [56].

**Principe de la transformée de Hough**

Une droite est décrite dans le plan cartésien  $(x,y)$  par l'expression suivante :

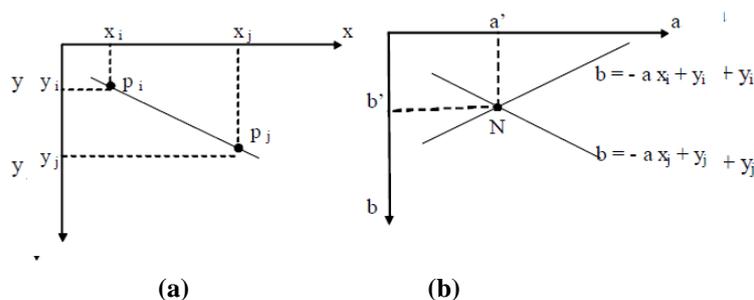
$$F(y, x, a, b) = y - a x - b = 0 \tag{1.1}$$

Sachant que **a** est la pente et **b** l'ordonnée à l'origine des abscisses. Etant donné un ensemble de contours d'objets représentés par un ensemble de points discrets **M<sub>i</sub>**, nous cherchons à déterminer si un ou plusieurs sous-ensembles de points **M<sub>i</sub>** font partie d'une courbe dont les paramètres **a** et **b** restent à définir. Si nous cherchons à tester les **n** points **M<sub>i</sub>** deux par deux, nous arriverons à un nombre exagéré d'itérations au moins supérieur à **n<sup>2</sup>**.

**Hough** puis **Rosenfeld** ont proposé une méthode pour **détecter les droites** à l'aide des points du plan  $(x, y)$ . Son principe est de calculer pour chaque point **M<sub>i</sub>** de coordonnées  $(x_i, y_i)$ , du contour d'un objet, l'ensemble des paramètres **a** qui vérifient l'équation :

$$f(y_i, x_i, a, b) = 0 \text{ avec } b \text{ fixé.}$$

Pour chaque point **M<sub>i</sub>**  $(x_i, y_i)$ , de l'image, il y a un ensemble de valeurs possibles pour les paramètres **a** et **b**. Cet ensemble forme une droite d'équation  $b = -ax + y$  dans l'espace des paramètres  $(a,b)$  appelé espace de **Hough**. Deux points **p<sub>i</sub>** et **p<sub>j</sub>** de coordonnées  $(x_i, y_i)$  et  $(x_j, y_j)$  respectivement, appartenant à la même droite, forment des droites dans l'espace des paramètres  $(a, b)$ , qui se coupent au point **N** de coordonnées  $(a', b')$ . De cette façon tous les points qui appartiennent à la même droite forment des droites dans le plan des paramètres  $(a,b)$  qui se coupent au même point. Ce concept est illustré dans les figures **1.3.a** et **1.3.b**.



**Figure 1.3:** Transformée de **Hough**, (a) Plan cartésien  $(xy)$ , (b) Plan des paramètres  $(ab)$ .

Le traitement **Hough** consiste en une quantification du plan des paramètres en cellules accumulatrices sur la **figure 1.4**. Où  $(a_{\min}, a_{\max})$ , et  $(b_{\min}, b_{\max})$  sont les valeurs limites de l'intervalle de la pente **a** et de l'ordonnée à l'origine des abscisses **b**.

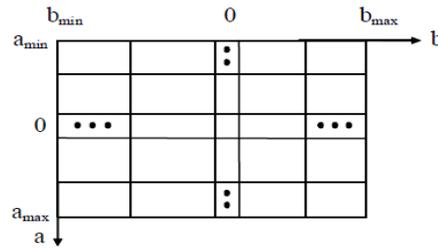


Figure 1.4 : Quantification du plan des paramètres (ab)

Chaque cellule de coordonnées  $(i, j)$  a une valeur accumulée  $A(i, j)$  et correspond à la cellule de coordonnées  $(a_i, b_j)$  dans le plan des paramètres  $(a, b)$ . Initialement, ces cellules sont mises à zéro. Pour chaque point de l'image de coordonnées  $(x_k, y_k)$  on calcule pour chaque valeur de  $a$  quantifié a la valeur  $p$  sur l'axe des  $a$ , son correspondant  $b$  en utilisant l'équation suivante :  $b = -ax_k + y_k$ . La valeur résultante  $b$  est arrondie à la valeur la plus proche de  $b$  quantifiées  $b_q$  sur l'axe des  $b$ . Si on obtient une valeur  $b_q$  suite à  $a_p$  choisie, on incrémente la valeur de la cellule correspondante:

$$A(p, q) = A(p, q) + 1 \tag{1.2}$$

A la fin de cette procédure, la valeur  $n$  de  $A(i, j)$  dans une cellule  $(i, j)$  correspond à  $n$  points dans le plan  $(x, y)$  qui vérifient l'équation  $y = a_i x + b_j$ , donc il existe  $n$  points qui appartient à la droite de pente  $a_i$  et de l'ordonnée à l'origine des abscisses  $b_j$ .

L'inconvénient majeur de cette procédure réside dans son incapacité de détecter les droites verticales. Pour remédier à ce problème, un paramétrage polaire  $(\rho, \theta)$  est plus satisfaisant. Ce paramétrage est illustré dans la figure 1.5.

Une droite est alors définie par l'équation suivante :

$$f(x, y, \rho, \theta) = \rho - x \cos \theta - y \sin \theta = 0 \tag{1.3}$$

Avec  $\rho$  la distance perpendiculaire à la droite de l'origine du plan  $(x, y)$  et  $\theta$  l'angle entre cette distance et l'axe de  $x$ .

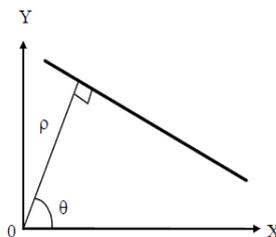


Figure 1.5 : Paramétrage polaire d'une droite

L'utilisation de cette représentation dans la construction du tableau accumulateur est identique à celle développée précédemment (représentation  $a, b$  sur la figure 1.6).

On précisera que le choix de quantification de l'espace des paramètres  $(\rho, \theta)$  doit porter sur les trois objectifs essentiels suivants:

- Garantir une précision de détection aussi bonne que possible.
- Diminuer la mémoire nécessaire au stockage des accumulateurs.
- Accélérer les calculs.

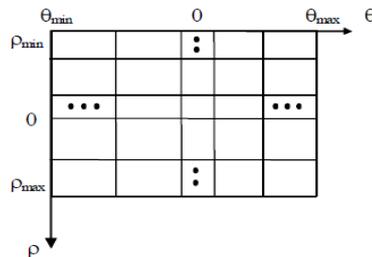


Figure 1.6 : Quantification du plan des paramètres  $(\rho, \theta)$

Chaque point  $M_i$  de coordonnées  $(x_i, y_j)$  d'une droite se transforme dans le plan des paramètres  $(\rho, \theta)$  en une sinusoïde d'équation :

$$\rho = x_i \cos\theta + y_j \sin\theta \tag{1.4}$$

Donc une droite sera représentée par un ensemble de sinusoïdes qui se coupent en un seul point de coordonnées polaires  $(\rho_0, \theta_0)$   $\rho, \theta$  caractéristique de cette droite dans le plan des paramètres (se référer aux figures 1.3-a et 1.3-b) [56].

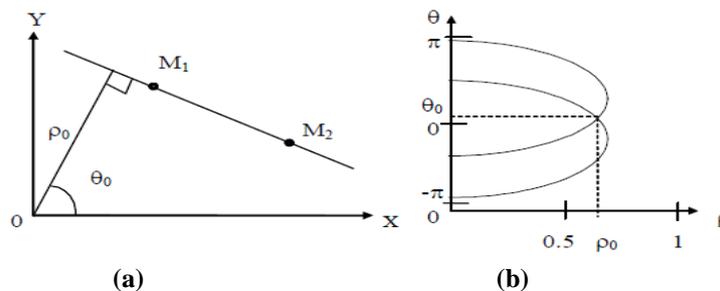


Figure 1.7 Transformée de Hough, (a) Plan cartésien  $(xy)$ , (b) Plan des paramètres  $(\rho, \theta)$

## 2 Extraction des caractéristiques par les ondelettes de Gabor

Une alternative à l'usage de la morphologie mathématique est l'application d'une **transformée d'ondelette**. Les ondelettes de Gabor sont connues comme le moyen d'analyse **espace-fréquence** qui minimise l'**incertitude de Heisenberg** à savoir si on est précis dans l'analyse de l'information dans l'espace du pixel on le sera moins dans l'espace fréquentiel et vice versa. Cette caractéristique fait des ondelettes de Gabor un moyen puissant pour **l'analyse de la texture**. Elles ont été développées au début pour l'analyse des signaux à une dimension. Leur expression générale est :

$$W(t, t_0, \omega) = e^{-\sigma(t-t_0)^2} e^{-i\omega(t-t_0)} \quad (1.5)$$

Où  $\omega$  est la fréquence et  $i^2 = -1$ . Cette ondelette est en fait une gaussienne modulée par une fonction sinusoidale. La généralisation de cette fonction en **2D** donne une fonction gaussienne à deux dimensions modulée par une harmonique. Les ondelettes de Gabor **2D** ont la caractéristique d'être modulables par différents paramètres qui contrôlent leurs **orientations**, **fréquences**, **phases** de la sinusoïde, **le support temporel** et **l'enveloppe de la gaussienne** [57]. L'équation générale d'une ondelette de **Gabor** en **2D** est :

$$W(x, y, \theta, \lambda, \varphi, \sigma, \gamma) = e^{-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}} e^{i(\frac{x'}{\lambda} + \varphi)} \quad (1.6)$$

Ou :

$$x' = x \cos \theta + y \sin \theta \quad \text{et} \quad y' = -x \sin \theta + y \cos \theta.$$

Il y a donc 5 paramètres qui contrôlent l'analyse par ondelette. Ce jeu de paramètres permet dès lors une analyse complète de la texture d'une région de l'image.

### L'orientation de l'ondelette ( $\theta$ )

Ce paramètre fait pivoter l'ondelette autour de son centre. L'orientation de l'ondelette (figure 1.8) détermine l'angle des contours **ou des lignes de l'image auxquelles l'ondelette va être sensible**.

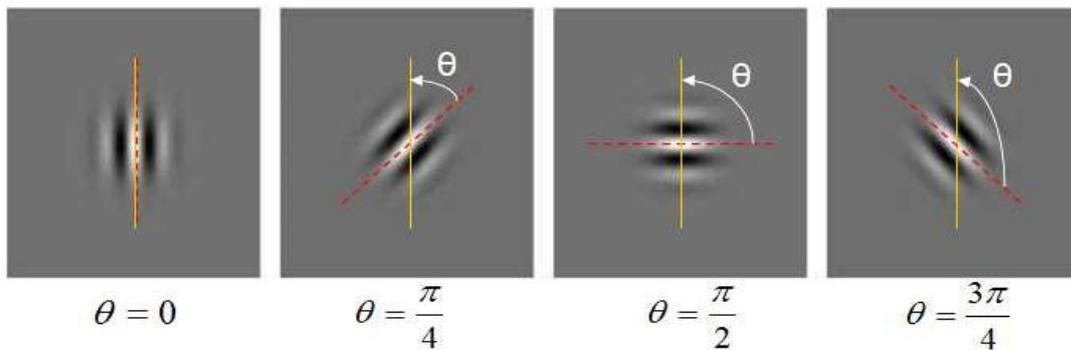


Figure 1.8 : Orientation de l'ondelette.

### La fréquence centrale de l'ondelette ( $\lambda$ )

Ce paramètre spécifie la **longueur d'onde du cosinus** ou inversement la fréquence centrale de l'ondelette (figure 1.9). Les ondelettes avec une grande longueur d'onde seront sensibles à des changements progressifs d'intensité dans une image. Les ondelettes avec une petite longueur d'onde seront sensibles à des contours et des lignes abruptes.

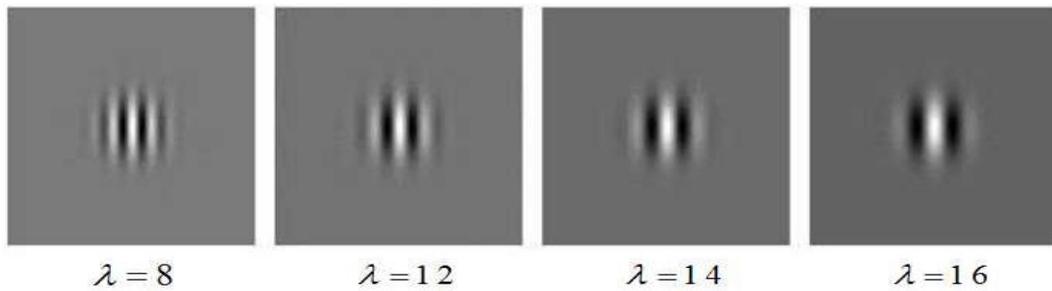


Figure 1.9 : Fréquence centrale de l'ondelette

### La phase de la sinusoïde ( $\varphi$ )

On utilise à la fois la partie réelle et la partie imaginaire de l'ondelette complexe de Gabor. Ce qui nous donne, en quelque sorte, deux ondelettes : une ondelette **paire** et une ondelette **impaire**. La convolution (incluant alors deux phases) donne un coefficient complexe basé sur deux ondelettes qui sont déphasées de  $\pi/2$  (figure 1.10).

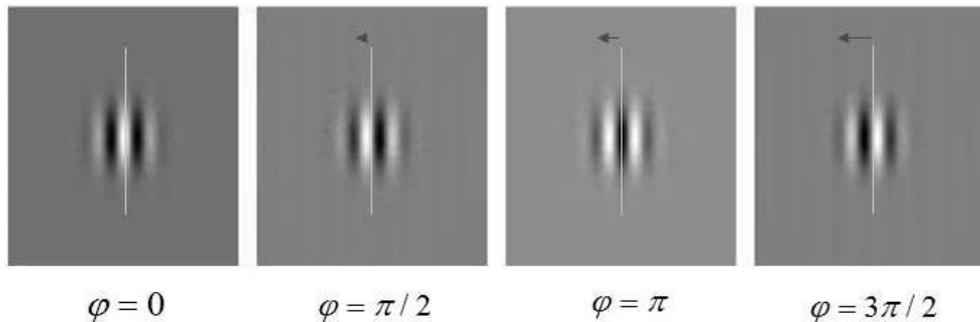


Figure 1.10 : Phase de l'ondelette.

### Le support temporel de l'ondelette ( $\sigma$ )

Ce paramètre spécifie le **rayon de la gaussienne** (figure 1.11). La **taille** de la gaussienne détermine la quantité de pixels de l'image qui vont être pris en compte dans la convolution.

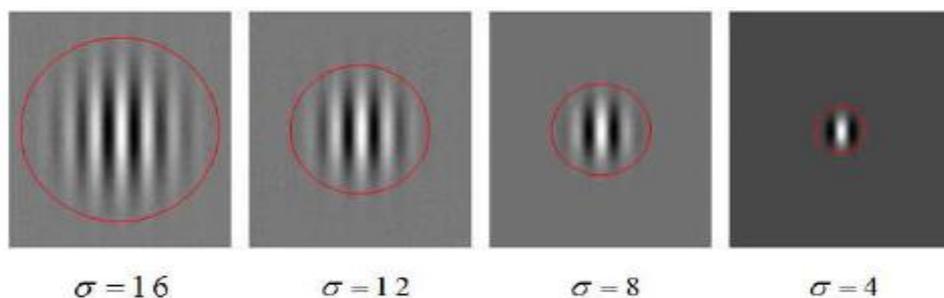


Figure 1.11 : Support temporel de l'ondelette.

L'enveloppe de la gaussienne ( $\gamma$ )

Ce paramètre agit sur la **forme de l'enveloppe gaussienne**, en l'étirant spatialement (figure 1.13). Ce paramètre a été inclus de manière à ce que les ondelettes puissent approximer certains modèles biologiques.

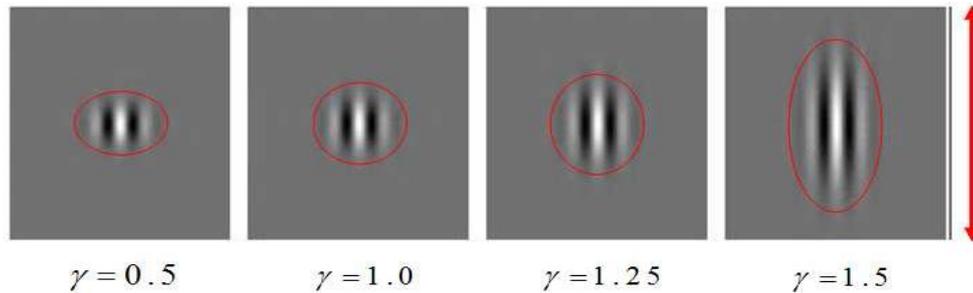


Figure 1.13 : Forme de l'enveloppe gaussienne.

La (figure 1.14) présentée quatre vingt masques d'ondelettes de Gabor différents avec cinq fréquences différentes, huit orientations différentes, et deux phases différentes [58, 11].

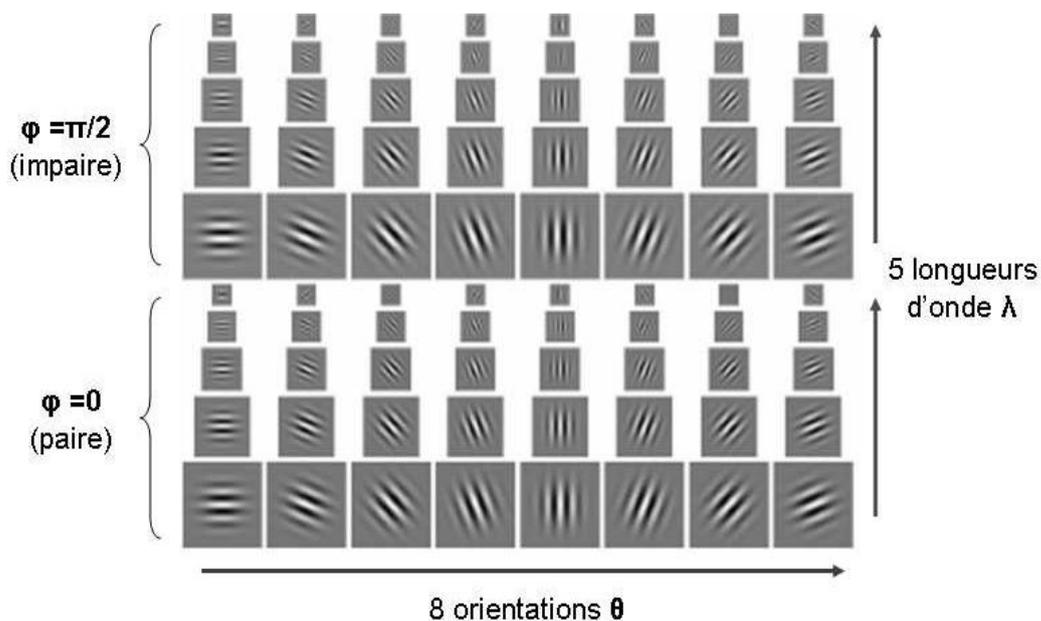


Figure 1.14: Les 80 masques d'ondelettes de Gabor.

1.3.2 Les algorithmes de réduction d'espace de données

Une image du visage est un **signal à 2 dimensions**, acquis par un capteur digital (caméra numérique, scanner...). Ce capteur codera la couleur ou l'intensité des différents points de l'image dans **une matrice de pixels** à deux dimensions.

Après une normalisation de l'image et une mise à l'échelle vers une taille fixe (ex :  $m*n$ ), l'image du visage peut être considérée comme un vecteur dans un espace multidimensionnel  $R^m$ . Ce vecteur est obtenu en mettant, simplement, dans un ordre fixe les

éléments de la matrice des pixels. Afin de définir une image (ou un point) dans l'espace des images, nous devons spécifier une valeur pour chaque pixel de cette image. Le nombre de points constituant cet espace devient rapidement très grand, même pour les images de petite dimension. Cette dimensionnalité pose un certain nombre de problèmes pour les algorithmes de reconnaissance, qui se basent sur cette représentation de l'image, à savoir :

- Dans un contexte de la reconnaissance, travailler dans un grand espace pose un problème de complexité de calcul.
- Pour les méthodes paramétriques, le nombre de paramètres à estimer peut rapidement dépasser le nombre d'échantillons d'apprentissage, ce qui pénalise l'estimation.
- Pour les méthodes non paramétriques, le nombre d'exemples nécessaires afin de représenter efficacement la distribution des données peut être insuffisant.

Le cas du visage est assez particulier. Le visage est formé par des surfaces lisses et une texture régulière. Il possède une forte symétrie (dans le cas des images frontales) et il est formé à partir de mêmes objets (yeux, nez et bouche), qui ont les mêmes positions pour tous les visages. Ces spécificités donnent lieu à plusieurs constatations :

- les différents pixels de l'image du visage sont fortement corrélés.
- les images des différents visages sont aussi corrélées.
- un important nombre de points dans l'espace des images ne représentent pas des visages.

En **1994**, **Ruderman** a démontré que les images naturelles possèdent une **grande redondance statistique** [59]. En **1996**, **Penev** [60] a démontré que dans le cas précis des images normalisées des visages, cette **redondance statistique est d'autant plus forte**.

L'appartenance à un même sous-espace et la redondance statistique qui caractérise les visages expliquent l'intérêt des chercheurs pour les méthodes de reconnaissance basées sur les approches par **réduction d'espace**, appelées également "**les approches globales**". Le fait que les visages soient confinés dans un sous-espace permet de définir une base (orthogonale, ou non orthogonale) dans cet espace. Cette base devrait représenter de façon plus pertinente n'importe quel point de l'espace des visages. L'avantage de la redondance statistique est qu'elle permet une extraction d'une structure simple des caractéristiques importantes et pertinentes de l'image du visage. Cette structure permettrait de représenter le visage tout en gardant l'information la plus importante, et par conséquent, de réduire la dimensionnalité de l'espace visage. Tout l'intérêt des approches globales est **la construction de cette base de projection** qui permettra de comparer, de reconnaître ou d'analyser l'information essentielle des visages [42].

## 1 Analyse en Composantes Principales (ACP)

L'analyse en composantes principales (**ACP**) est une méthode mathématique qui peut être utilisée pour **simplifier** un ensemble de données, en **réduisant** sa dimension et pour représenter efficacement les images de visages. Qui peuvent être approximativement reconstruites à partir d'un petit ensemble de poids et d'une image de visage standard (*Eigen Picture*). Ces poids sont obtenus en projetant l'image dans un espace de visage engendré par les visages propres (*Eigen faces*). Ainsi au lieu de stocker un vecteur image de taille résolution  $N$ , on stocke un vecteur de taille  $R \leq$  nombre d'image d'apprentissage  $M$ . Dans l'approche **PCA** la normalisation d'éclairage est toujours indispensable.

Les visages propres sont des images de la même taille que les nombres d'images d'apprentissage et qui montrent des visages ayant un aspect fantomatique. Mathématiquement, elles sont les composantes principales de la distribution des visages, ou les vecteurs propres de la matrice de covariance de l'ensemble des images de visage. Chaque image de visage de l'ensemble d'apprentissage peut être exactement représentée en termes de combinaison linéaire des visages propres et du visage moyen.

Le nombre possible des visages propres est égal au nombre d'images de visage dans l'ensemble d'apprentissage. Néanmoins les visages peuvent être approximés, en utilisant seulement les meilleurs visages propres (ayant les plus larges valeurs propres qui représentent en fait, la plupart de variance dans l'ensemble d'images de visage), permettant ainsi de diminuer les calculs.

La **ACP** est une technique rapide, simple et populaire dans l'identification de modèle, c'est l'une des meilleures techniques. Les projections de la **ACP** sont optimales pour la reconstruction d'une base de dimension réduite. Cependant, la **ACP** n'est pas optimisée pour la séparabilité (discrimination) de classes. Une alternative qui est *l'analyse discriminante linéaire LDA* tient compte de ceci [77].

### Le processus de reconnaissance (Karhunen-Loève)

Au début de la phase d'apprentissage, on fait l'acquisition, la lecture et la normalisation des images d'apprentissage (de taille  $N$ ). Puis on calcule le visage moyen de ces images. On soustrait ensuite le visage moyen de chaque image d'apprentissage normalisée (ces nouvelles images représentent les colonnes d'une matrice  $\bar{X}$ ). On fait le calcul de la matrice de covariance.

Chaque image est représentée par un vecteur de taille  $N$ .

$$x^i = [x_1^i, x_2^i, \dots, x_N^i]^T \quad (1.7)$$

Où  $N$  est le nombre de pixels dans l'image.

Toutes les images sont centrées en soustrayant l'image moyenne de chaque vecteur image,

$$\bar{x}^i = x^i - m \quad \text{où} \quad m = \frac{1}{P} \sum_{i=1}^P x^i \quad (1.8)$$

$m$  : La moyenne de toutes les images d'apprentissage.

Ces vecteurs sont combinés, côte à côte, pour créer une matrice de données d'apprentissage de taille  $N \times P$  (où  $P$  est le nombre d'images d'apprentissage,  $N$  la taille de la vecteur image).

$$\bar{X} = [\bar{x}^1 \ \bar{x}^2 \ \bar{x}^3 \ \dots \ \bar{x}^P] \quad (1.9)$$

Le calcul de la matrice de covariance qui est définie par la formule :

$$\Omega = \bar{X}\bar{X}^T \quad (1.10)$$

On calcule ensuite les vecteurs propres  $V$  et les valeurs propres  $D$  de la matrice  $\Omega$ .

Les valeurs propres de la matrice  $\Omega$  représentent le taux de variation le long de l'axe du vecteur propre associé. Dès lors, on pourra éliminer les vecteurs propres les moins importants sans perte conséquente d'information.

Puis on passe à l'étape de calcul des visages propres selon la formule

$$U = \bar{X} * V * (abs(D))^{-0.5} \quad (1.11)$$

Et finalement on calcule les poids des visages de la base (de taille  $M$ ) en les projetant dans le sous-espace engendré par les visages propres "Face Space".

Pendant la phase de vérification on fait l'acquisition, la lecture et la normalisation de l'image de vérification (de taille  $N$ ). Puis on soustrait le visage moyen (calculé auparavant lors de la phase d'apprentissage) de l'image de vérification normalisée. Ensuite on calcule le poids de l'image (de taille  $M$ ) en utilisant les visages propres comme une base de projection. Et finalement on utilise les distances de similitude [42].

### Le choix de la dimension $M$ de l'espace de visage [41]

En général le choix de la dimension  $M$  ou le nombre des vecteurs propres n'est pas facile. En effet, plusieurs techniques ont été utilisées pour déterminer cette dimension intrinsèque  $M$  de l'espace des visages.

- L'utilisation d'un **critère d'énergie** [61]: M est choisi tels que les valeurs propres maintenues contiennent 95% de toute l'énergie.

$$\sum_{j=1}^m \lambda_j = 0.95 \sum_{j=1}^{\min(l,n)} \lambda_j \quad (1.12)$$

- La **sélection** est basée sur les caractéristiques des **vecteurs propres** [62] : les auteurs notent qu'au-dessus du **rang 300**, les vecteurs propres encodent principalement le bruit et devraient être jetés pour des tâches de classification de visages.
- Les **résultats empiriques** basé sur une large gamme de bases de données de visages [63]: il est suggéré de choisir un nombre M qui représente le **40%** du nombre de tous les vecteurs propres.
- Une **sélection** de type **bayésienne** est utilisée dans [64] pour trouver une dimension correcte. Une approche pragmatique consiste en choisissant M de telle sorte de minimiser les taux d'erreur de classification sur l'ensemble de données.
- Le choix dépend de l'étape du **prétraitement des images**. Par exemple, si les visages sont normalisés par une transformation souple de sorte que l'information de la forme soit uniforme comme dans [65] le sous-espace de visages est de dimension basse car seule la variation de l'information de niveau de gris doit être représentée par le modèle.
- Dans [83] la partie du visage utile pour la classification est déterminée en étudiant les **rangs statistiques** (rank statistics) entre les pixels dans l'image. La bonne localisation de visage donne une région étroite du visage en éliminant le front et le fond et donc permet de réduire la dimension M.

#### ❖ Inconvénient

L'**ACP** ne tient pas compte de l'aspect (discrimination) des classes et on ne devrait pas s'attendre ainsi une performance optimale pour des tâches telle que l'authentification de visage en utilisant la **ACP** pour le codage.

#### ❖ Solution

Une solution qui a été proposée pour résoudre ce nouveau problème est d'employer l'analyse discriminante linéaire de Fisher (**FLD**) pour le but même de réaliser une séparabilité élevée entre les différentes classes dans lesquelles la classification est intéressante.

## 2 Analyse Discriminante Linéaire

L'analyse discriminante linéaire (**LDA**) et les discriminants linéaires de Fisher (**FLD**) relatifs, sont utilisés pour trouver la combinaison linéaire des caractéristiques qui séparent le mieux les classes d'objet ou d'événement. Les combinaisons résultantes peuvent être

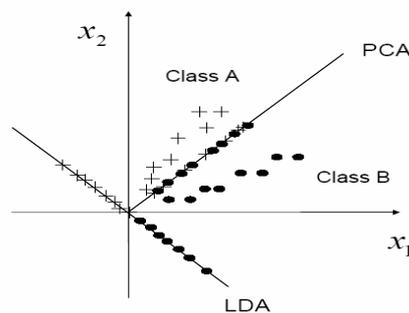
employées comme **classificateur linéaire**, ou généralement dans la réduction de caractéristiques.

La **LDA** est étroitement liée à la **ACP**, du fait que toutes les deux recherchent des combinaisons linéaires des variables qui représentent au mieux les données. La **LDA** essaye explicitement de modéliser la différence entre les classes des données. La **ACP** quand à elle, ne tient pas compte des différences entre les classes.

Chaque visage, qui se compose d'un grand nombre de pixels, est réduit à un plus petit ensemble de combinaisons linéaires avant la classification.

Chacune des nouvelles dimensions est une combinaison linéaire des valeurs de pixels. Les combinaisons linéaires obtenues en utilisant **FLD** s'appellent les Fisherfaces, en analogie avec les visages propres (Eigenfaces) [67].

La **LDA** est une technique qui cherche les directions qui sont efficaces pour la discrimination entre les données. La figure 1.15 représente un exemple de classification de deux nuages de points. L'axe principal de la LDA est l'axe de projection qui maximise la séparation entre les deux classes. Il est clair que cette projection est optimale pour la séparation des deux classes par rapport à la projection sur l'axe principal calculé par **ACP**.



**Figure 1.15 :** Comparaison entre les projections de deux classes de points ("class 1" et "class 2") sur les axes principaux construits par ACP et par LDA.

Les étapes de l'analyse discriminante linéaire sont :

**1) Calcul des moyennes**

Nous calculons la moyenne des images dans chaque classe ( $m_i$ ) et la moyenne de toute les images  $m$

$$m_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_j^i ; \quad i = 1, 2, \dots, C \tag{1.13}$$

$$m = \frac{1}{P} \sum_{i=1}^P x_i ; \tag{1.14}$$

$P$  : Le nombre d'images d'apprentissage.

$n_i$  : Le nombre d'images dans chaque classe.

$C$  : Le nombre de classes.

**2) Centrer les images dans chaque classe**

Soustraire la moyenne de chaque classe de l'image de cette classe

$$\forall x \in X_i, X_i \in X, \hat{x} = x - m_i \quad (1.15)$$

**3) Centrer les moyennes de chaque classe**

Soustraire la moyenne totale des moyennes de classes.

$$\hat{m}_i = m_i - m \quad (1.16)$$

**4) Calculer la matrice de dispersion intra-classes  $S_W$  (within class scatter matrix)**

La matrice de dispersion intra-classes ( $S_W$ ) est la somme des matrices de dispersion des classes. Pour la  $i^{\text{ème}}$  classe la matrice de dispersion  $S_i$  est calculée par la somme des matrices de covariance des images centrées.

$$S_i = \sum_{x \in X_i} \hat{x} \hat{x}^T \quad (1.17)$$

La matrice de dispersion intra-classes ( $S_W$ ) est la somme de toutes les matrices de dispersion.

$$S_W = \sum_{i=0}^C S_i \quad (1.18)$$

Où  $C$  est le nombre de classe.

**5) Calculer de la matrice de dispersion inter-classes  $S_B$  (Between class scatter matrix)**

La matrice de dispersion inter-classes ( $S_B$ ) est la somme de dispersion entre classes.

$$S_B = \sum_{i=1}^C n_i \hat{m}_i \hat{m}_i^T \quad (1.19)$$

Où  $n_i$  : Le nombre d'images dans la classe.

$\hat{m}_i$  : La moyenne des classes.

Le but est de maximiser les distances inter-classes tout en minimisant les distances intra-classes, ce qui revient à retrouver la matrice de transformation  $W$  qui maximise le critère :

$$J(W) = \frac{W^T S_b W}{W^T S_w W}$$

Donc  $W$  est optimale pour :

$$W_{opt} = \underset{W}{arg \max} \left( \frac{|W^T S_b W|}{|W^T S_w W|} \right) = [w_1, w_2, \dots, w_m] \quad (1.20)$$

**6) Résoudre le problème de valeurs propres généralisé**

- Résoudre le problème généralisé de vecteurs propres ( $V$ ) et des valeurs propres ( $\Lambda$ ) de la matrice de dispersion  $S_W$  et la matrice de dispersion  $S_B$ .

$$S_B V = \Lambda S_W V \quad (1.21)$$

La solution est rendue par calcul des vecteurs propres et des valeurs propres de la matrice  $S_W^{-1} * S_B$ .

- Ordonner les vecteurs propres par ordre décroissant des valeurs propres correspondantes. La matrice de transformation de la **LDA** est constituée par les premiers vecteurs propres.

❖ **Inconvénient**

**FLD** exige un grand nombre d'échantillons de l'ensemble d'apprentissage pour la bonne généralisation. Quand un tel besoin n'est pas répondu, la **FLD** crée un problème de sur-ajustement aux données d'apprentissage et ceci s'apprête mal aux nouvelles données de test [41, 67, 68].

❖ **Solution**

Le modèle discriminant Linéaire Amélioré de Fisher (**Enhanced Fisher Model EFM**).

**3 Le modèle discriminant linéaire amélioré de Fisher (Enhanced Fisher linear discriminant Model (EFM))**

Ce modèle discriminant linéaire de **Fisher** améliore la capacité de généralisation de la **FLD** en décomposant la procédure **FLD** en diagonalisation simultanée des deux matrices de dispersion intra-classe et inter-classe [70]. La diagonalisation simultanée est une étape sagement équivalente à deux opérations comme l'a souligné **Fukunaga** [91]. Blanchiment de la matrice de dispersion intra-classe et l'application de la ACP sur la matrice de dispersion intrre-classe en utilisant les données transformées. Durant l'opération du blanchiment de la matrice de dispersion intra-classe apparaisse dans le dénominateur de la séparabilité des petites valeurs propres qui tendent à capturer du bruit [70]. Pour atteindre des performances améliorées **l'EFM** préserve un équilibre approprié entre la sélection des valeurs propres (correspondant à la composante principale de l'espace de l'image originale) qui tiennent compte de la plupart de l'énergie spectrale des données brutes, c.-à-d., représentation adéquate

et l'exigence que les valeurs propres de la matrice de dispersion intra-classe (de l'espace ACP réduit) ne sont pas trop petites, c.-à-d., meilleure généralisation.

Le choix de rang des composantes principales ( $m$ ) pour la réduction de la dimension, prend en compte de l'ordre de grandeur de l'énergie spectrale. Les valeurs propres de la matrice de covariance fournit un bon indicateur pour répondre au critère de l'énergie. Il faut ensuite calculer les valeurs propres de la matrices de dispersion intra-classe dans l'espace ACP réduit pour faciliter le choix du rang des composantes principales de sorte que l'exigence de grandeur est respectée. A cette fin, on effectue la FLD par des étapes comme décrit ci-dessous. En particulier, ces étapes FLD permettent de trouver les valeurs propres et les vecteurs propres de  $S_W^{-1}S_b$  comme résultat de la diagonalisation simultanée de  $S_W$  et  $S_b$ .

Alors l'étapes de EFM sont présentées comme suit :

- Blanchissons d'abord la matrice de dispersion intra-classe :

$$S_W V = V A \quad \text{et} \quad V^T V = I \quad (1.22)$$

$$A^{-1/2} V^T S_W V A^{-1/2} = I \quad (1.23)$$

Où  $V, A \in \mathbf{R}^{m \times m}$  sont la matrice des vecteurs propres et la matrice diagonale des valeurs propres de  $S_W$  respectivement.

Les valeurs propres de la matrice de dispersion intra-classe dans l'espace ACP réduit peut être obtenu en (équation 1.22).

Donc, EFM diagonalise en premier lieu la matrice de dispersion intra-classe  $S_W$  (1.22) et (1.23). Notez que  $V$  et  $A$  sont les matrices des vecteurs propres et des valeurs propres correspondants aux vecteurs caractéristiques.

- En second lieu EFM procède à calculer la nouvelle matrice de dispersion inter-classe comme suit:

$$A^{-1/2} V^T S_b V A^{-1/2} = K_b \quad (1.24)$$

- Diagonalisons maintenant la nouvelle matrice de dispersion inter-classe  $K_b$  :

$$K_b V_b = V_b A_b \quad \text{et} \quad V_b^T V_b = I \quad (1.25)$$

Où  $V_b, A_b \in \mathbf{R}^{m \times m}$  sont la matrice des vecteurs propres et la matrice diagonale des valeurs propres de  $K_b$  respectivement.

- La matrice de transformation globale de l'EFM est définie comme suit [91]:

$$U = V A^{-1/2} V_b \quad (1.26)$$

#### 4 La LDA directe (DLDA)

Une nouvelle approche, appelée la **DLDA**, a été proposée dans [73] afin de maximiser le critère de **Fisher**. La première étape de cette approche est la diagonalisation de la matrice  $S_b$  par le calcul de la matrice  $V$  :

$$V^T S_b V = \Lambda \quad (1.27)$$

Avec :  $V^T V = I$ . Le problème est alors ramené à la résolution d'un problème de valeurs propres. Chaque vecteur de la matrice  $V$  est un vecteur propre de la matrice  $S_b$  et  $\Lambda$  contient toutes les valeurs propres de  $S_b$ . Comme la matrice  $S_b$  peut être singulière,  $\Lambda$  peut contenir des valeurs propres nulles ou très faibles. Ces valeurs ainsi que les vecteurs propres associés doivent être écartés car une projection dans la direction de ces vecteurs n'apporte aucune information discriminante entre les classes.

Soit  $Y$  la sous-matrice composée des  $m$  premières colonnes de  $V$  ( $Y$  étant de dimension  $m \times n$  et  $n$  étant la taille de l'échantillon d'entrée) :

$$Y^T S_b Y = D_b > 0 \quad (1.28)$$

Avec :  $D_b$  la sous-matrice de  $\Lambda$ , de dimension  $m \times m$ .

En multipliant par  $D_b^{-1/2T}$  et  $D_b^{-1/2}$  de part et d'autre on aboutit à :

$$(Y D_b^{-1/2})^T S_b (Y D_b^{-1/2}) = D_b$$

En posant  $Z = Y D_b^{-1/2}$  ;

Nous obtenons :

$$Z^T S_b Z = D_b \quad (1.29)$$

On constate que  $Z$  permet de réduire la dimension de  $S_b$  de  $n \times m$  à  $m \times m$ .

Considérons la diagonalisation de la matrice  $Z^T S_w Z$  par la résolution du problème des valeurs propres :

$$U^T Z^T S_w Z U = D_w \quad (1.30)$$

Avec :  $U^T U = I$ .  $D_w$  peut contenir des valeurs nulles sur la diagonale.

L'objectif est de minimiser l'éparpillement des inter-classes. Il est donc important de garder les vecteurs de projection associés aux valeurs propres les plus faibles, spécialement les valeurs nulles, et d'écarter ceux associés aux valeurs propres les plus fortes.

En posant la matrice  $A = U^T Z^T$ ,  $A$  permet la diagonalisation du numérateur et du dénominateur du critère de **Fisher** :

$$A S_w A^T = D_w \quad ; \quad A S_w A^T = I \quad (1.31)$$

En posant  $\phi = D_w^{-1/2} A$ , nous obtenons la matrice de projection qui répond au critère de Fisher [42].

## 5 L'Analyse Discriminante Linéaire Bidimensionnelle Orientée (ADL2Do)

Les techniques de projection statistique des vecteurs-images ont été très utilisées dans le contexte de la reconnaissance de visages (ACP, LDA, DLDA, EFM...). Ces analyses de données en très grandes dimensions sont coûteuses en temps de calcul, et souffrent d'instabilité numérique. Pour pallier ces inconvénients, **Yang et al.** ont introduit dans [74] la technique d'Analyse en Composantes Principales bidimensionnelle (**ACP2D**), plus performante et robuste aux différentes sources de variabilité que la méthode des *eigenfaces* [75]. Nous avons donc présenté l'Analyse Discriminante Linéaire Bidimensionnelle Orientée (**ADL2Do**).

On dispose d'une base de données  $\Omega$  contenant  $n$  images de visages, avec plus d'une vue par personne. L'ensemble des images correspondant à une même personne est désigné par le terme de *classe*. La base de données contient  $C$  classes. Chaque image est stockée sous la forme d'une matrice de pixels de taille  $h \times w$ , et sa classe d'appartenance  $\Omega_c$  est connue.

L'**ADL2Do** se décline en deux versions : l'*ADL2D orientée en lignes* (**ADL2DoL**) et l'*ADL2D orientée en colonnes* (**ADL2DoC**). Présentons dans un premier temps l'**ADL2DoL**. Considérons une matrice de projection  $P$ , de taille  $w \times k$ , et la projection suivante :

$$X_i^p = X_i P \quad (1.32)$$

Où  $X_i^p$  est la matrice de taille  $h \times k$  correspondant à la projection de la matrice-image  $X_i$  sur  $P$ , et constitue la matrice de la signature associée à l'image  $X_i$  par l'**ADL2DoL**.

Nous recherchons la matrice  $P$  qui maximise après projection la distance Euclidienne entre classes différentes tout en minimisant la distance entre les éléments provenant d'une même classe. Sous l'hypothèse que les lignes de pixels des images constituent des vecteurs gaussiens et que les lignes provenant de classes différentes ont la même covariance intra classes, on peut montrer que la matrice  $P^*$  maximisant le critère de Fisher généralisé suivant :

$$P^* = \underset{P \in \mathbb{R}^{w \times k}}{\text{Argmax}} \frac{|P^T S_b P|}{|P^T S_w P|} \quad (1.33)$$

Où :  $S_w$  et  $S_b$  sont respectivement les *matrices de covariance intra et inter classes généralisées* de  $\Omega$  :

$$S_w = \sum_{c=1}^C \sum_{X_i \in \Omega_c} (X_i - \bar{X}_c)^T (X_i - \bar{X}_c) \quad (1.34)$$

$$S_b = \sum_{c=1}^C n_c (\bar{X}_c - \bar{X})^T (\bar{X}_c - \bar{X}) \quad (1.35)$$

Où :  $\bar{X}_c$  est la matrice moyenne des  $n_c$  visages appartenant à la classe  $\Omega_c$  et  $\bar{X}$  est la matrice moyenne de tous les visages de  $\Omega$ . Sous l'hypothèse que  $S_w$  est inversible (on peut aisément montrer que cette hypothèse est généralement vérifiée), les  $k$  colonnes de la matrice  $P^*$  sont les  $k$  premiers vecteurs propres de la matrice  $S_w^{-1}S_b$  (c.-à-d. ceux associés aux plus grandes valeurs propres).

De la même manière, pour l'ADL2DoC, la projection considérée est la suivante :

$$X_i^Q = Q^T X_i \quad (1.36)$$

Où  $Q$  est une matrice de projection de taille  $h \times k$ ,  $Q^T$  est sa transposée et la matrice  $X_i^Q$  de taille  $k \times w$ , constitue la signature associée au visage  $X_i$  par l'ADL2DoC. Sous les hypothèses de multinormalité et d'homoscédasticité des colonnes des matrices-images, nous pouvons considérer que la matrice  $Q^*$ , optimale pour une taille  $h \times k$  fixée, maximise le critère de Fisher généralisé suivant :

$$Q^* = \underset{Q \in \mathbb{R}^{h \times k}}{\text{Argmax}} \frac{|Q^T \Sigma_b Q|}{|Q^T \Sigma_w Q|} \quad (1.37)$$

Où :  $\Sigma_w$  et  $\Sigma_b$  désignent respectivement les matrices de covariance intra et inter classes généralisées de la base constituée des  $(X_i^T)_{i \in \{1 \dots n\}}$

$$\Sigma_w = \sum_{c=1}^C \sum_{X_i \in \Omega_c} (X_i - \bar{X}_c)(X_i - \bar{X}_c)^T \quad (1.38)$$

$$\Sigma_b = \sum_{c=1}^C n_c (\bar{X}_c - \bar{X})(\bar{X}_c - \bar{X})^T \quad (1.39)$$

On peut montrer que les  $k$  colonnes de la matrice  $Q^*$  sont les vecteurs propres de  $\Sigma_w^{-1} \Sigma_b$  associés aux plus grandes valeurs propres [40].

### 1.3.3 Comparaison

Lorsqu'on souhaite comparer deux vecteurs de caractéristiques réduit issus du module de réduction de données d'un système biométrique, on peut soit effectuer une mesure de similarité (ressemblance), soit une mesure de distance (divergence).

Bien que la distance euclidienne soit optimale dans la théorie, les diverses expériences ont trouvé que la distance euclidienne est surpassée par d'autres distances. Une d'elles est la **corrélation normalisée** (normalized correlation) [76] qui est définie par :

$$S(A, B) = \frac{A^T B}{\|A\| \|B\|} \quad (1.40)$$

Cette fonction calcule simplement le cosinus de l'angle entre les deux vecteurs caractéristiques A et B. Une valeur élevée de corrélation normalisée correspond à une bonne similarité entre les deux vecteurs.

Et lorsqu'on souhaite comparer deux matrices de caractéristiques réduites. Dans [77] **Yang** a proposé une nouvelle mesure de similarité au plus proche voisin pour la reconnaissance de visages. La distance de **Yang** se base sur la classification de matrices caractéristiques obtenues par l'**ACP2D**. Cette distance a été, également, adoptée par **Visani et al** [78] et **Bengherabi** [79]. Elle est définie pour deux matrices caractéristiques réduites

$$Y_j = [y_1^{(j)} y_2^{(j)} \dots y_{d_1}^{(j)}] \text{ et } Y_i = [y_1^{(i)} y_2^{(i)} \dots y_{d_1}^{(i)}] \text{ comme suit :}$$

$$d(Y_i, Y_j) = \sum_{k=1}^{d_1} \|y_k^{(i)} - y_k^{(j)}\|_2 \quad (1.41)$$

Où  $\|y_k^{(i)} - y_k^{(j)}\|_2$  désigne la distance Euclidienne entre les deux vecteurs  $y_k^{(i)}$  et  $y_k^{(j)}$  d'où:

$$d(Y_i, Y_j) = \sum_{k=1}^{d_1} \left( \sum_{h=1}^n (y_{hk}^{(i)} - y_{hk}^{(j)})^2 \right)^{1/2} \quad (1.42)$$

$$\text{Avec } y_k^{(i)} = [y_{1k}^{(i)} y_{21}^{(i)} \dots y_{d_1 k}^{(i)}] \text{ et } y_k^{(j)} = [y_{1k}^{(j)} y_{21}^{(j)} \dots y_{d_1 k}^{(j)}]$$

Dans [80], **Zuo** a proposé « l'Assembled Matrix Distance: **AMD** » définie par:

$$d(Y_i, Y_j) = \left( \sum_{k=1}^{d_1} \left( \sum_{h=1}^n (y_{hk}^{(i)} - y_{hk}^{(j)})^2 \right)^{p/2} \right)^{1/p} \text{ avec } p > 0 \quad (1.43)$$

Pour laquelle la distance de **Yang** est obtenue pour  $p = 1$  et la distance de Frobenius pour  $p = 2$ .

## 1.4 Mesure de la performance des systèmes d'authentification biométriques

Tout d'abord, afin de comprendre comment déterminer la performance d'un système biométrique, il nous faut définir clairement **quatre critères principaux** :

- Le premier critère s'appelle le **taux de faux rejet** ou **TFR** ("False Reject Rate" ou **FRR**). Ce taux représente le pourcentage de personnes censées être reconnues mais qui sont rejetées par le système,
- Le deuxième critère est le **taux de fausse acceptation** ou **TFA** ("False Accept Rate" ou **FAR**). Ce taux représente le pourcentage de personnes censées ne pas être reconnues mais qui sont tout de même acceptées par le système,

- Le troisième critère est connu sous le nom de taux d'égale erreur ou **TEE** ("Equal Error Rate" ou **EER**). Ce taux est calculé à partir des deux premiers critères et constitue un point de mesure de performance courant. Ce point correspond à l'endroit où **TFR = TFA**, c'est-à-dire le meilleur compromis entre les faux rejets et les fausses acceptations.
- Le quatrième critère est le **taux de réussite** ou **TR égale** ( $100 - (TFR + TFA)$ ).

La **figure 1.16** illustre le **TFR** et le **TFA** à partir de distributions des scores clients et imposteurs tandis que le **TEE** est représenté sur la **figure 1.17**

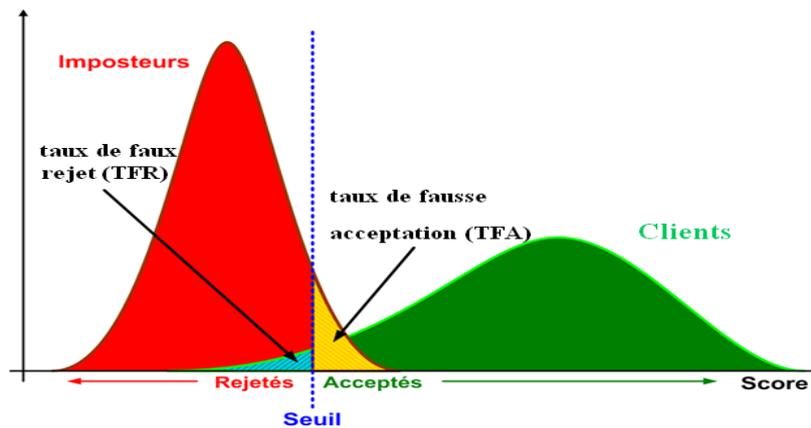


Figure 1.16 : Illustration du TFR et du TFA.

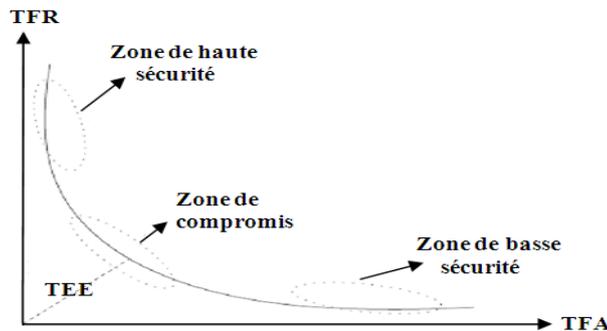


Figure 1.17 : Courbe caractéristique de fonctionnement (ROC) d'un système de Vérification Biométrique.

La performance globale d'un système de vérification d'identité est mieux caractérisée par sa courbe caractéristique de fonctionnement (**Receiver Operating Characteristic** ou **ROC**), La courbe ROC (**figure 1.17**) trace le taux de faux rejet en fonction du taux de fausse acceptation [81]. Plus cette courbe tend à épouser la forme du repère, plus le système est performant, c'est-à-dire possédant un taux de reconnaissance global élevé.

Le seuil de décision doit donc être ajusté en fonction de l'application ciblée : haute sécurité, basse sécurité ou compromis entre les deux.

### 1.4.1. Intervalle de confiance

Le choix des critères d'évaluation des performances des systèmes biométriques, nous devons faire face à une autre difficulté de l'évaluation d'un système biométrique. En général, on ne dispose que de **peu de données** de chaque personne et **de peu de personnes** dans les bases. Certaines bases sont composées de quelques dizaines de personnes seulement et on a peu de données par personne, même si récemment des bases de l'ordre de centaines de personnes ont fait leur apparition. Il est alors impératif de mesurer en plus des taux d'erreurs calculés, une mesure de confiance associée.

Deux types de mesure des intervalles de confiance sont introduits par **Mellakh** [42] et **Allano** [10] : l'une basée sur une méthode non-paramétrique appelée "**bootstrap**" et l'autre basée sur une **méthode paramétrique**.

#### 1 Méthode non-paramétrique

Pour utiliser la méthode d'estimation d'intervalles de confiance par le "**bootstrap**", il faut réaliser un ré-échantillonnage de l'ensemble de test. Ce ré-échantillonnage consiste en un tirage aléatoire avec remplacement de **M** exemples. Dans notre cas, on prendra **M** égal au nombre d'exemples. Pour l'estimation du **TFR**, on prendra donc **M** =  $N_c$  avec  $N_c$  le nombre de scores "client" et pour le **TFA**, on prendra **M** =  $N_i$ ,  $N_i$  étant le nombre de scores "imposteur".

Pour chaque tirage, on va donc choisir **M** exemples avec remplacement, c'est à dire que les exemples peuvent être utilisés plusieurs fois ou jamais. Pour chaque tirage, on estime ensuite les valeurs du **TFR** et du **TFA** pour une valeur du **seuil**  $\theta$ . On répète ensuite cette procédure **T** fois avec **T** tirages (en général, 1000 tirages sont suffisants). A partir de ces **T** valeurs de **TFA**( $\theta$ ) et de ces **T** valeurs de **TFR**( $\theta$ ), on peut estimer alors l'intervalle de confiance à 90% ou à 95% autour de la médiane. Le problème de cette méthode est changée les protocoles d'évaluation biométrique, les bases d'apprentissage et les listes des tests.

#### 2 Méthode paramétrique

L'hypothèse principale qui est faite lorsqu'on utilise cette méthode d'estimation de l'intervalle de confiance est que les scores sont issus de variables indépendantes et identiquement distribuées même pour les scores d'une personne, c'est-à-dire que tous les scores sont issus de la **même loi** et sont mutuellement indépendants. Cette hypothèse semble très stricte mais est nécessaire pour l'application des théorèmes. Ceci dit, elle n'est généralement pas vérifiée en pratique.

Il s'agit ensuite d'estimer les taux d'erreur  $\text{TFA}(\theta)$  et  $\text{TFR}(\theta)$  pour une ou plusieurs valeurs du seuil  $\theta$ . Pour estimer ces erreurs, l'estimation des fonctions de répartition, appelées "fonctions de répartition empiriques", pour les deux classes "client" et "imposteur" doit être réalisée par une **loi normale**. L'estimation des intervalles de confiance des taux d'erreur  $\text{TFA}$  et  $\text{TFR}$  à un seuil  $\theta$  est obtenue par :

$$\text{TFA}(\theta) \in \left[ \widehat{\text{TFA}}(\theta) - \lambda \times \sqrt{\frac{\widehat{\text{TFA}}(\theta)(1-\widehat{\text{TFA}}(\theta))}{N_i}}, \widehat{\text{TFA}}(\theta) + \lambda \times \sqrt{\frac{\widehat{\text{TFA}}(\theta)(1-\widehat{\text{TFA}}(\theta))}{N_i}} \right], \quad (1.44)$$

$$\text{TFR}(\theta) \in \left[ \widehat{\text{TFR}}(\theta) - \lambda \times \sqrt{\frac{\widehat{\text{TFR}}(\theta)(1-\widehat{\text{TFR}}(\theta))}{N_c}}, \widehat{\text{TFR}}(\theta) + \lambda \times \sqrt{\frac{\widehat{\text{TFR}}(\theta)(1-\widehat{\text{TFR}}(\theta))}{N_c}} \right], \quad (1.45)$$

Avec  $\widehat{\text{TFA}}(\theta)$ ,  $\widehat{\text{TFR}}(\theta)$  les valeurs calculées pour un test donné et  $N_c$  les scores "client" et  $N_i$  les scores "imposteur" du test. Pour des intervalles de confiance à **90%**, on a  $\lambda = 1.645$  et pour des intervalles de confiance de **95%**, on a  $\lambda = 1.96$ .

Cette méthode non paramétrique sera utilisée dans la suite afin de calculer les intervalles de confiance. Ce choix est justifié pour éviter le changement dans les protocoles d'évaluation de notre base de données et cette méthode est la plus utilisée [10, 11, 42, 57, 82].

## 1.5 Conclusion

Dans ce chapitre, nous avons présenté **l'état de l'art** et **notre modèle** de reconnaissance de visage, ce chapitre n'a pas pour finalité de décrire tous les algorithmes de reconnaissance faciale mais nous nous sommes contentés de présenter les algorithmes qui ont introduit une nette évolution dans le domaine de cette biométrie, permettant ainsi une véritable amélioration des performances. Dans l'état de l'art nous avons décomposé les méthodes en trois grandes familles : les méthodes **globales**, pour lesquelles les caractéristiques sont extraites directement depuis l'ensemble des valeurs de pixels des images et les méthodes **locales**, basées sur l'extraction de signatures extraites localement du visage. Les techniques dites hybrides utilisent conjointement ces deux types de modélisation.

Parmi les méthodes globales, on compte notamment les méthodes basées sur la **projection statistique**, aussi appelées techniques de **réduction de dimension**. Elles visent à définir un espace de projection dans lequel les données sont projetées puis classées. Deux types de critères peuvent être utilisés pour déterminer ce sous-espace : un critère de

représentation des données (on cherche à préserver la distribution des données) ou de séparabilité en fonction de la classe d'appartenance. Pour optimiser le premier critère, on utilise essentiellement l'ACP ou l'ACI tandis que, pour le second critère, l'LDA est généralement préférée. Ces méthodes sont caractérisées par un apprentissage rapide, un faible nombre de paramètres à ajuster et de très bonnes performances. Nous avons vu à travers de résultats expérimentaux provenant de différentes sources (tableau 1.1) que la méthode globale des *fisherfaces* (LDA) est très performante, en comparaison avec les autres techniques de l'état de l'art. Les *fisherfaces* présentent de plus l'avantage d'être facilement généralisables à des problèmes plus complexes alors notre choix est porter sur les méthodes globale à base des *fisherfaces*.

Dans la section de présentation de modèle, nous avons essayé de présenter quelques algorithmes utilisés dans les trois étapes d'un système de reconnaissance faciale. Nos efforts sont concentrés la dessus :

- pour l'étape d'extraction des caractéristiques nous utilisons deux types de descripteurs de visages la transformée de Hough et les ondelettes de Gabor et un choix du meilleur filtre s'impose.
- Mais pour l'étape de réduction l'espace de données nous présentons plusieurs d'algorithmes de réduction globale à une dimension (Analyse en Composantes Principales (ACP), Analyse Discriminante Linéaire (LDA), le Modèle Discriminant Linéaire Amélioré de Fisher (Enhanced Fisher Linear Discriminant Model (EFM) et LDA directe (DLDA) et la LDA à deux dimensions (L'Analyse Discriminante Linéaire Bidimensionnelle Orientée (ADL2Do).
- Et finalement nous présentons des algorithmes de comparaisons entre deux vecteurs et deux matrices à base de plusieurs métriques. La distance qui donne les meilleures performances est retenue.

Enfin nous détaillons la mesure de la performance et l'intervalle de confiance d'un système biométrique.

## 2 Authentification de visages et étude expérimentale

### 2.1 Introduction

Jusqu'ici nous avons approché la reconnaissance automatique des visages d'un point de vue théorique. Comme l'authentification est un problème pratique, il est maintenant d'examiner les différentes alternatives de solutions et d'essayer de tirer quelques conclusions.

Les résultats expérimentaux que nous allons présenter sont divisés en deux parties. Nous donnerons tout d'abord les résultats des comparaisons des méthodes de réduction d'espace. Nous développons ensuite la partie d'extraction de paramètres par deux méthodes : la transformation de **Hough** et les filtres de **Gabor**. Finalement, le modèle ainsi étudié est appliqué sur les images de visages de la base de données **XM2VTS** et les résultats obtenus sont exposés et les meilleures systèmes d'authentification de visages est retenue dans la conclusion.

### 2.2 La base de données XM2VTS

La base de données XM2VTS (Extended Multi Modal Verification for Teleservices and Security applications) [83] est une base multimodale développée au sein du projet européen ACTS offre des enregistrements synchronisés des photos de visages prises de face et de profil et des paroles de 295 personnes des deux sexes hommes et femmes de différents âges, elle est utilisée pour la vérification d'identité. Pour chaque personne huit prises ont été effectuées en quatre sessions distribuées pendant cinq mois afin de prendre en compte le changement d'apparence des personnes selon plusieurs facteurs (présence et/ou absence de lunettes, barbe, changement de coiffure, pose, etc.). Par contre, seules les expressions neutres ont été considérées, et chaque session est composée de deux enregistrements, une pour les séquences de parole et l'autre pour les séquences vidéo de la tête. Les vidéos et photos sont en couleur de haute résolution (format ppm), la taille est de 256 x 256 pixels pour les images et de très bonne qualité codé sur 24 bits dans l'espace RGB. Cela permet de travailler en niveaux de gris ou en couleur.

Le choix principal de XM2VTS est sa taille grande, avec 295 personnes et 2360 images en total et sa popularité puisqu'elle est devenue une norme dans la communauté biométrique audio et visuelle de vérification multimodale d'identité.

## Chapitre 2 Authentification de visages et étude expérimentale

Pour la tâche de vérification, un protocole standard d'estimation de performances a été développé. Appelé «*Lausanne protocolsplitsrandomly*», il existe deux configurations différentes, la **configuration I** et la **configuration II** (LPI et LPII) et définis dans [83]. Nous n'utiliserons la **configuration I** puisqu'elle est la plus dure, il consiste à séparer les individus en deux classes, client et imposteur. Le groupe client contient 200 sujets, alors que le groupe imposteur est divisé en 25 imposteurs pour l'évaluation et 70 imposteurs pour le test. La répartition des photos dans les différents ensembles sont présentée dans le tableau 2.1, et les figures 2.1, 2.2 présentent des images de face de la base de données XM2VTS.



Figure 2.1: Exemples des images de la base de données XM2VTS.



Figure 2.2: Exemples des images d'une personne de la base de données XM2VTS

Ensemble	Clients	Imposteurs
Apprentissage	600 images (3 par personne)	0 images
Evaluation	600 images (3 par personne)	200 images (8 par personne)
Test	400 images (2 par personne)	560 images (8 par personne)

Tableau 2.1 Répartition des photos dans les différents ensembles

### 2.3 Comparaison des algorithmes de réduction l'espace de donnée

Dans un visage de la même personne, au point que deux images très semblables (à l'œil humain) pourraient être extrêmement différentes si elles sont comparées pixel par pixel. Il est donc nécessaire d'extraire les caractéristiques appropriées et discriminantes à partir des images du visage et de comparer ces caractéristiques au lieu de comparer des images approximatives. Naturellement, plus les caractéristiques discriminantes, plus l'authentification sera facile.

#### 2.3.1 Prétraitement d'image

Le prétraitement est une phase importante dans le processus globale d'authentification. C'est une méthode simple qui augmente en général les performances du système. Elle permet souvent une première réduction des données et elle atténue les effets d'une différence de conditions lors des prises de vues. Pour notre travail nous avons besoin de trois phases de prétraitement : le découpage de l'image puis la décimation conversion de la couleur en l'espace de couleur TSV.

##### 1 Découpage

Le découpage de l'image consiste à conserver les maximums des variations intrinsèques du visage, et de supprimer les autres informations comme l'arrière plan, les cheveux, les cols de chemise, les oreilles...etc. afin d'améliorer les performances du système d'authentification.

Une fenêtre rectangulaire de taille (161x121) est centrée autour des caractéristiques les plus stables liées aux yeux, aux sourcils, au nez et à la bouche a été utilisée. La Figure 2.3 montre l'utilisation de cette fenêtre.



Figure 2.3 : l'image de visage a) avant et b) après découpage

##### 2 Décimation

Consiste à ne prendre qu'un pixel sur quatre. Cela réduit bien entendu la résolution des images. L'image de visage passera ainsi d'une dimension  $256 \times 256 = 65536$  pixels vers une dimension de  $81 \times 61 = 4941$  pixels après le découpage et la décimation, comme il est illustré sur la Figure 2.4.



(a) (b)

Figure 2.4 : image de visage a) avant b) après décimation.

### 3 La composante S de l'espace de couleur TSV de l'image

Les systèmes d'authentification de visages utilisent souvent la représentation en niveaux de gris de l'image de visage comme caractéristique d'entrée. Mais lorsque les images sont en couleurs, il est avantageux d'utiliser cette information supplémentaire pour améliorer les performances de ce système d'authentification. Alors nous avons utilisé l'espace de couleur TSV (Teinte, Saturation, Valeur) ou HSV (Hue, Saturation, Value) car est le système le plus utilisé dans la littérature [42]. La quantification de la couleur de notre base de données XM2VTS est faite dans le système RVB (Rouge, Vert et Bleu). Le passage du système de couleurs RVB au système TSV se fait par une fonction de transfert non linéaire. La figure 2.5 présente les deux systèmes colorimétriques RVB et TSV.

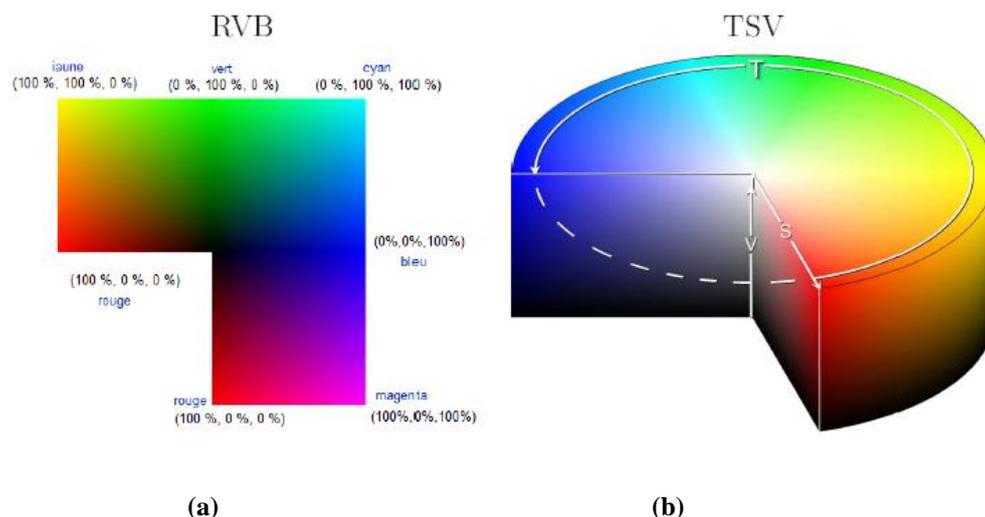


Figure 2.5: Différents systèmes colorimétriques : (a) plan du système de couleurs RVB, (b) cylindre du système de couleurs TSV.

Soient  $min$  et  $max$ , respectivement, les valeurs minimales et maximales des composantes R, V et B ( $min = minimum(R, V, B)$  et  $max = maximum(R, V, B)$ ) et soient  $x$  et  $y$  les coordonnées d'un pixel de l'image.

Les transformations des espaces RVB vers TSV sont données par les équations suivantes[42].

$$T(x, y) = \begin{cases} \text{non défini, si } min = max \\ 60 * \frac{V - B}{max - min}, \text{ si } max = R \text{ et } V \geq B \\ 60 * \frac{max - min}{V - B} + 360, \text{ si } max = R \text{ et } V < B \\ 60 * \frac{max - min}{B - R} + 120, \text{ si } max = V \\ 60 * \frac{max - min}{R - V} + 240, \text{ si } max = B \end{cases}$$

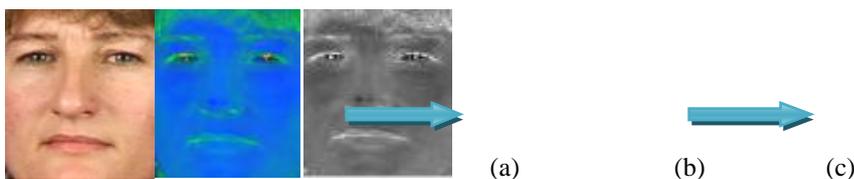
$$(2.2) S(x, y) = \begin{cases} 0, \text{ si } max = 0 \\ 1 - \frac{min}{max}, \text{ sinon} \end{cases}$$

$$V(x, y) = max(2.3)$$

Pour plus de détails sur cette transformation, voir [84].

Dans [85] est étudiée les espaces de couleurs pour la base de donnée XM2VTS pour les deux protocoles de Lausanne est trouvée que la composante S est donnée le meilleur résultat pour la tâche de vérification d'identité de visages, à partir de ce résultat nous avons utilisé la composante S de l'espace de couleur TSV comme caractéristique d'entrée de l'algorithme de réduction d'espace de donnée.

La figure 2.6 présente la conversion de couleur et l'extraction de composante S de visage .



**Figure 2.6:** (a) Image découpée et décimée, (b) Image en système TSV, (c) La composante S de TSV.

### 4Enchaînement des (lignes/ou colonnes)

C'est la conversion de la matrice image vers un vecteur. Extrayons de l'image la matrice de composant couleur S. Puis nous faisons la conversion de la matrice à un vecteur par l'enchaînement des lignes ou colonnes de la matrice correspondante.

### 5Photo normalisation

La photo normalisation a un double effet : d'une part elle supprime pour tout vecteur image un éventuel décalage par rapport à l'origine et ensuite elle supprime tout effet d'amplification (multiplication par un scalaire). Pour chaque image on effectue l'opération suivante :

$$photo\ normalisation(x) = \frac{x - mean(x)}{std(x)} \quad (2.4)$$

mean(x) : le moyen de x

std(x) : l'écart type de x.

### 2.3.2 Les algorithmes de réduction d'espace

Nous avons étudié 6 algorithmes de réduction d'espace de donnée pour choisir le meilleur algorithme, les étapes de chaque algorithme est présenter comme suite :

#### 1 L'analyse en composantes principales(PCA)

1. Prétraitement.
2. Création de la matrice des données X (L'ordonnancement des vecteur images des clients coté à coté).
3. Centrer la matrice des données  $\bar{X}$
4. Calcul de la matrice de covariance des données  $X X^T$ .
5. Calcul des valeurs et vecteurs propres de la matrice de covariance.
6. Ordonner les vecteurs propres par ordre décroissant des valeurs propres correspondantes.
7. La matrice de transformation de la PCA est les m premiers vecteurs propres ( $U_{PCA}$ ).

#### 2 Analyse Discriminante Linéaire (LDA)

1. Prétraitement.
2. Calcul la matrice de dispersion intra-classes  $S_W$  (équation 1.17).
3. Calcul de la matrice de dispersion inter-classes  $S_B$  (équation 1.18).
4. Calcul des valeurs et vecteurs propres de la matrice  $S_W^{-1} * S_B$ .
5. Ordonner les vecteurs propres par ordre décroissant des valeurs propres correspondantes.
6. La matrice de transformation de la LDA est les m premiers vecteurs propres ( $U_{LDA}$ ).

#### 3 Le modèle discriminant Linéaire amélioré de Fisher (EFM)

1. Prétraitement.
2. Calcul la matrice de dispersion intra-classes  $S_W$ ( équation1.17).
3. Calcul de la matrice de dispersion inter-classes  $S_B$ ( équation1.18).
4. Calcul des valeurs ( $\lambda$ ) et vecteurs ( $V$ ) propres de la matrice  $S_W$ .
5. Calculer la nouvelle matrice de dispersion inter-classe  $K_b = A^{-1/2} V^T S_b V A^{-1/2}$
6. Calcul des valeurs ( $\lambda_b$ ) et vecteurs ( $V_b$ ) propres de la matrice  $K_b$ .
7. Calcul de la matrice  $U = V A^{-1/2} V_b$ .
8. La matrice de transformation de l'EFM est les  $m$  premiers vecteurs de  $U$  ( $U_{EFM}$ ).

### 4La LDA directe (DLDA)

1. Prétraitement.
2. Calcul la matrice de dispersion intra-classes  $S_W$ ( équation1.17).
3. Calcul de la matrice de dispersion inter-classes  $S_B$ ( équation1.18).
4. Calcul des valeurs ( $D_b$ ) et vecteurs ( $Y$ ) propres de la matrice  $S_B$  avec  $Y^T S_b Y = D_b > 0$ .
5. Calcul la matrice  $K_W$  avec  $K_W = D^{-1/2} Y^T S_W Y D^{-1/2}$ .
6. Calcul des valeurs ( $U_W$ ) et vecteurs ( $D_W$ ) propres de la matrice  $K_W$ .
7. La matrice de transformation de la DLDA est ( $U_{DLDA}$ ) avec  $U_{DLDA} = Y D_b^{-1/2} U_W D_W^{-1/2}$ .

### 5L'Analyse Discriminante Linéaire Bidimensionnelle Orientée en lignes (ADL2DoL)

1. Prétraitement.
2. Calcul la matrice de dispersion intra-classes  $S_W$ ( équation1.34).
3. Calcul de la matrice de dispersion inter-classes  $S_B$ ( équation1.35).
4. Calcul des valeurs et vecteurs propres de la matrice  $S_W^{-1} * S_B$ .
5. Ordonner les vecteurs propres par ordre décroissant des valeurs propres correspondantes.
6. La matrice de transformation de l'ADL2DoL est les  $m$  premiers vecteurs propres ( $U_{ADL2DoL}$ ).

### 6 L'Analyse Discriminante Linéaire Bidimensionnelle Orientée en colonnes(ADL2DoC)

1. Prétraitement.
2. Calcul la matrice de dispersion intra-classes  $S_W$ ( équation1.38).
3. Calcul de la matrice de dispersion inter-classes  $S_B$ ( équation1.39).
4. Calcul des valeurs et vecteurs propres de la matrice  $S_W^{-1} * S_B$ .

5. Ordonner les vecteurs propres par ordre décroissant des valeurs propres correspondantes.
6. La matrice de transformation de l'ADL2DoC est les  $m$  premiers vecteurs propres ( $U_{ADL2DoC}$ ).

### 2.3.3 Projection des images

- Pour l'algorithme de réduction d'espace ACP

Le vecteur image  $X_i$  se transforme en vecteur  $Y_i$  par la relation :

$$Y_i = U_{ACP}^T * X_i \quad (2.5)$$

- Pour les algorithmes de réduction d'espace LDA, EFM est DLDA nous avons effectuée deux opérations de projection la première par ACP pour réduire les dimensions de vecteur image et la deuxième projection par LDA, EFM est DLDA pour faire la réduction et la séparation entre les classes. Le vecteur image  $X_i$  se transforme en vecteur  $Y_i$  par la relation :

$$Y_i = U_{ACP}^T * U_{LDA, EFM, DLDA}^T X_i \quad (2.6)$$

- Pour l'algorithme de réduction d'espace ADL2DoL

La matrice image  $X_i$  se transforme en matrice  $Y_i$  par la relation :

$$Y_i = U_{ADL2DoL} * X_i \quad (2.7)$$

- Pour l'algorithme de réduction d'espace ADL2DoC

La matrice image  $X_i$  se transforme en matrice  $Y_i$  par la relation :

$$Y_i = U_{ADL2DoC}^T * X_i \quad (2.8)$$

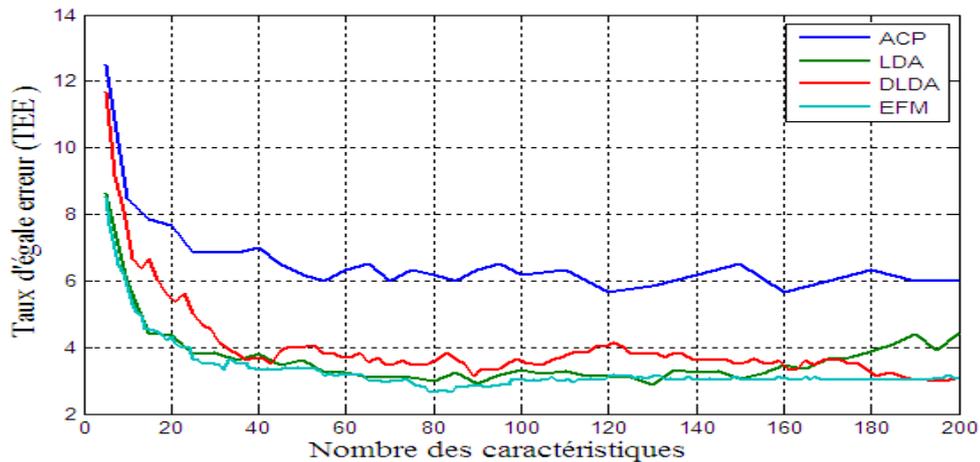
### 2.3.4 Comparaison

Nous utilisons pour la comparaison la distance de corrélation normalisée donnée par l'équation 1.40 pour comparer deux vecteurs caractéristiques réduits. Mais pour la comparaison de deux matrices caractéristiques réduites nous utilisons la distance AMD (Assembled Matrix Distance) définie par l'équation 1.43.

### 2.3.5 Résultats de simulation

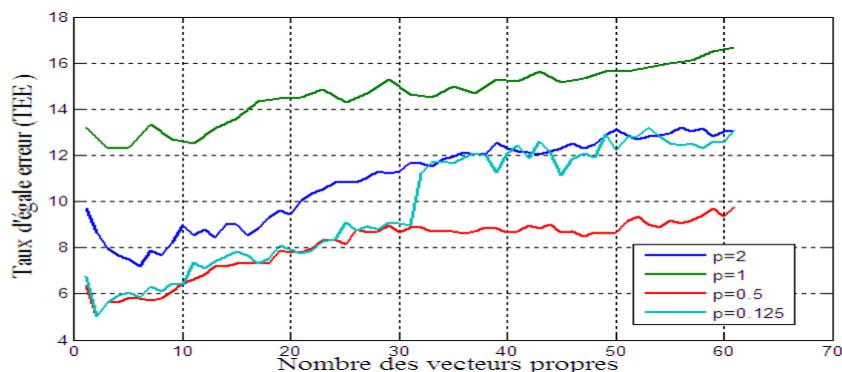
Nous utilisons plusieurs algorithmes de réduction d'espace, on peut légitimement se poser la question : quel algorithme choisir ? Pour répondre à cette question nous avons effectué nos expériences sur tous ces algorithmes, pour le choix du meilleur.

La **figure 2.7** nous montre que la méthode EFM est mieux performante que la LDA suivie de la DLDA suivie de L'ACP en employant un nombre restreint de caractéristiques (vecteurs propres) dans notre base de données XM2VTS.

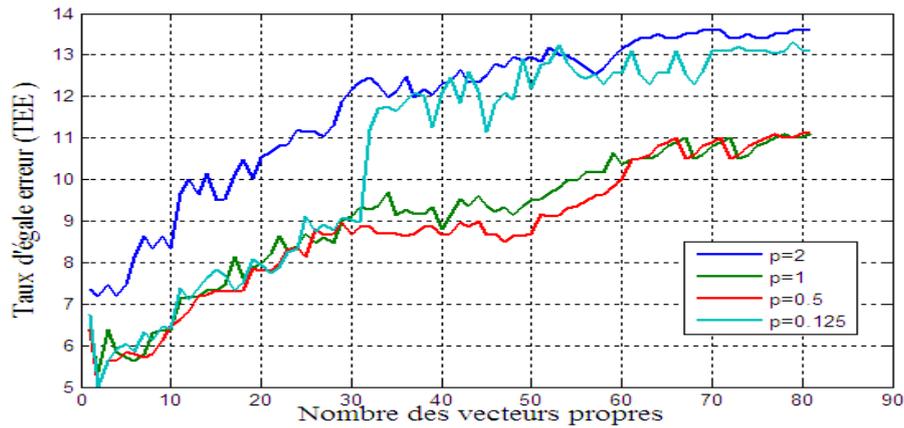


**Figure 2.7 :** Comparaison des TEE pour ACP, LDA, EFM et DLDA dans l'ensemble d'évaluation.

Les **figures 2.8** et **2.9** illustrent les taux d'égale erreur TEE obtenus sur l'ensemble d'évaluation de deux méthodes (ADL2DoL, ADL2DoC) en appliquant la AMD pour la mesure de distance avec  $p = 0.125, 0.5, 1, 2$ . Nous remarquons que le meilleur résultat est obtenu par  $p = 0.125$  pour les deux méthodes d'extraction de caractéristiques par un nombre restreint des vecteurs propres. Le **tableau 2.2** montre les meilleurs résultats obtenus dans l'ensemble d'évaluation et l'ensemble de test pour tous les algorithmes de réduction d'espace étudiés.



**Figure 2.8 :** TEE dans l'ensemble d'évaluation pour différentes fonctions de mesure de distances pour (LDA2DoL).



**Figure 2.9 :** TEE dans l'ensemble d'évaluation pour différentes fonctions de mesure de distances pour (LDA2DoC).

Méthodes	Ensemble	Ensemble de test			Nombre de caractéristiques
	évaluation	TFR %	TFA %	TR %	
ACP	5.7 ±1.04	2.7 ±1.58	8.1 ±0.17	89.2 ±1.75	200
ACP+LDA	2.83 ±0.74	2 ±0.08	4.12 ±1.94	93.88 ±2.03	90
ACP+EFM	<b>2.66 ±0.13</b>	<b>2 ±1.37</b>	<b>3.66 ±0.12</b>	<b>94.33 ±1.49</b>	<b>80</b>
ACP+DLDA	3 ±0.72	2.1 ±1.4	4.3 ±0.12	93.6 ±1.53	195
LDA2DOL	5.1 ±0.98	2 ±1.37	6.3 ±0.15	91.7 ±1.52	2*81
LDA2DOC	5.01 ±0.97	2 ±1.37	6.19 ±0.15	91.81 ±1.52	2*61

**Tableau 2.2 :** Performances des méthodes ACP, LDA, EFM, DLDA, LDA2DoL et LDA2DoC (intervalle de confiance paramétrique à 95%).

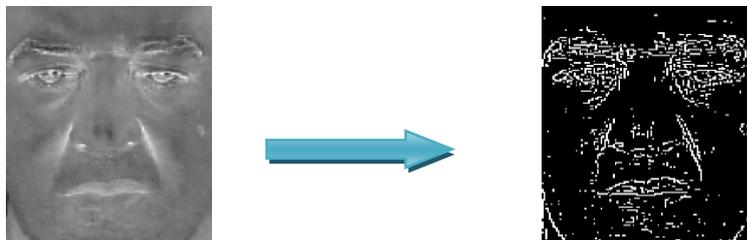
Le **tableau 2.2** nous montre bien que le meilleur résultat est obtenu avec **EFM** avec un **TEE=2.66 ±0.13%** et un **TR=94.33 ±1.49%** en utilisant seulement un nombre **80** de paramètres caractéristiques, et nous remarquons aussi que les **TR** et **TEE** obtenus par la **DLDA** et la **LDA** sont assez proches. Alors que le mauvais résultat est obtenu par **PCA** appliquée uniquement et c'est un résultat bien évident puisque cette méthode est gourmande en quantité de données et n'assure pas une discrimination optimale. Les deux méthodes **LDA2DOL** et **LDA2DOC** donnent presque le même résultat avec un nombre de caractéristiques  $2*81=162$  pour **LDA2DOL** et  $2*61=122$  pour **LDA2DOC**. Donc, ces méthodes à deux dimensions n'améliorent pas la performance d'authentification de visage par rapport aux **LDA**, **EFM** et **DLDA** car la base de données **XM2VTS** a été réalisée sur une longue période ce qui a permis d'obtenir plusieurs images d'une même personne et donc une grande variabilité d'apparence (changement de coiffure, présence et/ou absence de lunettes,

etc...). Parcontre, seules les expressions neutres ont été considérées, donc nous proposons l'utilisation de ces méthodes pour les bases de données qui contiennent plusieurs expressions.

L'algorithme de réduction d'espace de donnée retenu reste le modèle discriminant Linéaire Amélioré de Fisher (EFM).

### 2.4 La transformation de Hough

Cette transformation nécessite un filtrage passe haut et un seuillage de la composante S de l'espace de couleur TSL de l'image découpée et décimée, nous permettant de déterminer les contours de l'image par le filtre de Roberts, la figure 2.10 montre cette opération.



(a)(b)

**Figure 2.10 :** (a) la composante S de l'image découpée et décimée (b) après le filtrage par Roberts.

La transformée de **Hough (TH)** opère sur des données binaires des points contours de l'image. La quantification du plan des paramètres  $\rho$   $\theta$  revient à quantifier l'intervalle  $0 \leq \theta < \pi$  pour la dimension de  $\theta$  et l'intervalle  $-R \leq \rho \leq R$  pour la dimension de  $\rho$ , avec  $R$  la moitié de la diagonale de l'image.

$\rho_k$  et  $\theta_k$  pas de quantification des dimensions de  $\rho$  et  $\theta$ .

$n_\theta$  et  $n_\rho$  le nombre de valeurs discrètes dans les intervalles de  $\rho$  et  $\theta$ . Les valeurs de  $\rho$  et  $\theta$  discrétisées s'écrivent comme suit :

$$\theta = t * \theta_k \quad 0 \leq t \leq n_\theta$$

$$\rho = -R + r * \rho_k \quad 0 \leq r \leq n_\rho \quad \text{Avec : } n_\theta = \pi / \theta_k \quad \text{et} \quad n_\rho = 2R / \rho_k$$

Pour chaque point contour de coordonnées  $(x, y)$  de l'image, nous calculons pour chaque valeur discrète de  $\theta$  la valeur discrète de  $\rho$  suivant l'équation suivante :

$$\rho = x * \cos\theta + y * \sin\theta$$

Ensuite, nous incrémentons la cellule correspondante dans le tableau accumulateur. Nous faisons la même chose pour tous les points contours de l'image. L'organigramme ci-dessous de la **figure 2.11** décrit ce calcul [56].

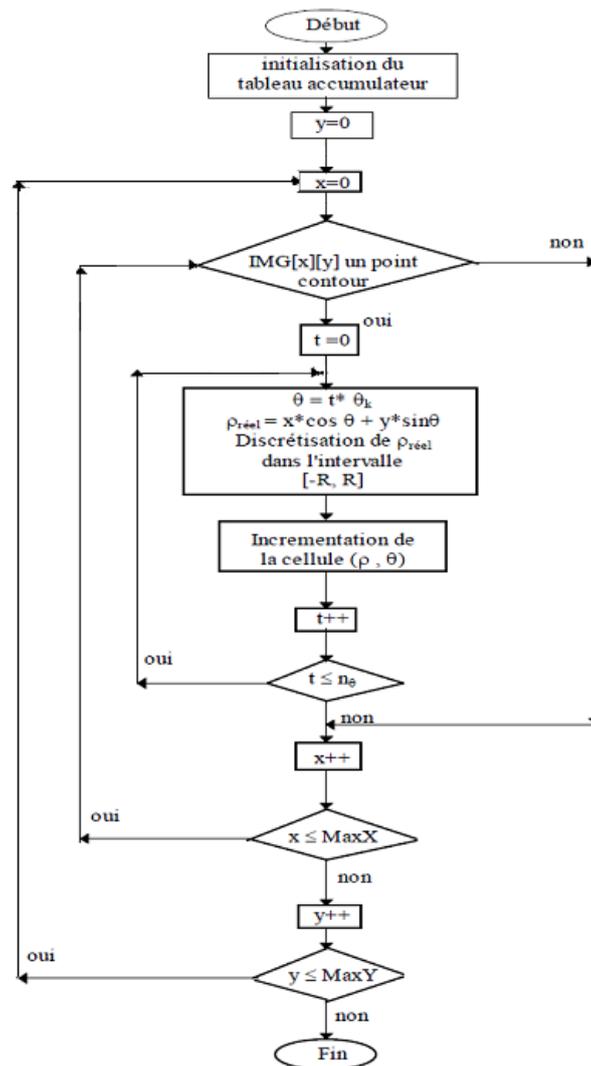


Figure 2.11 : Organigramme de la Transformée de Hough.

Finalement, le vecteur caractéristique de l'image de visage est construit par l'enchaînement des lignes ou colonnes de le **tableau accumulateur** puis la **Photo Normalisation**. Nous utilisons **ACP + EFM** comme méthode de réduction d'espace et la **corrélation** pour la mesure de similarité. La courbe de taux d'égalité erreur **TEE** obtenu dans l'ensemble d'évaluation de notre approche du système d'authentification de visages en utilisant la transformation de **Hough** comme méthode d'extraction des caractéristiques de visage est montrée par la **figure 2.12**.

Cette figure nous montre que le meilleur résultat est obtenu avec un nombre de caractéristiques égale  $N_p=140$ . Le **tableau 2.3** donne le meilleur **TEE** obtenu dans l'ensemble d'évaluation et l'ensemble de **test** pour la transformation de **Hough**.

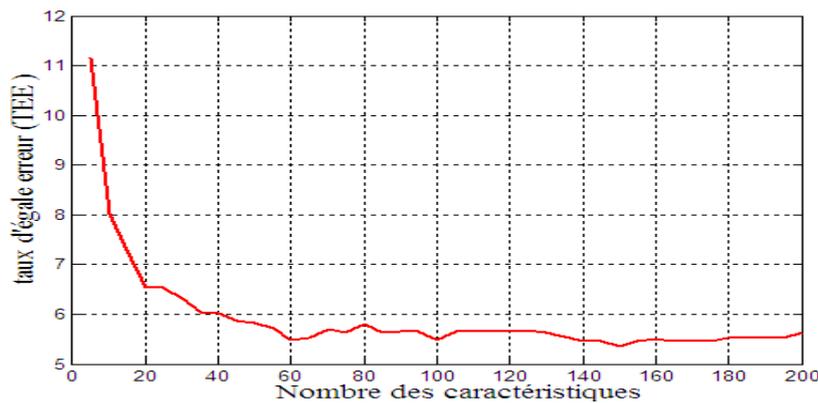


Figure 2.12 : TEE dans l'ensemble d'évaluation en fonction de nombre des caractéristiques

Méthode	Ensemble évaluation	Ensemble de test			Nombre des caractéristiques
	TEE %	TFR %	TFA %	TR %	
TF Hough	5.01 ±0.97	2 ±1.37	6.19 ±0.15	91.81 ±1.52	140

Tableau 2.3 : TF de Hough dans l'ensemble d'évaluation et test (intervalle de confiance paramétrique à 95%).

Le meilleur TEE obtenu dans le cas de cette transformation est égale **5.01 ±0.97 %** dans l'ensemble d'évaluation et un TR égale **91.81 ±1.52** dans l'ensemble de test qui demeure un résultat nécessitant des améliorations pour la vérification de visages.

## 2.5 Les ondelettes de Gabor

### 2.5.1 Extraction des caractéristiques

La méthode de reconnaissance que nous étudions dans cette partie repose principalement sur l'extraction des caractéristiques du visage par filtrage de Gabor 2D. Les filtres de Gabor sont connus comme un moyen d'analyse espace-fréquence très robuste. Cette spécificité a fait des filtres de Gabor un moyen puissant d'analyse de textures et de classification. Les filtres de Gabor analysent la texture d'un objet suivant différentes résolutions et différents angles. Dans le domaine spatial, un filtre de Gabor 2D est une fonction à noyau gaussien modulée par une onde sinusoïdale plane complexe. L'équation générale d'une ondelette de Gabor en 2D est :

$$W(x, y, \theta, \lambda, \varphi, \sigma, \gamma) = e^{-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}} e^{i(\frac{x'}{\lambda} + \varphi)} \quad (2.9)$$

Où :

$$x' = x \cos \theta + y \sin \theta \text{ et } y' = -x \sin \theta + y \cos \theta.$$

Il y a donc cinq paramètres qui contrôlent l'analyse par ondelettes, à partir [57,86]. Nous choisissons les cinq paramètres comme suit :

- $\theta$  spécifie l'orientation du filtre. Nous utilisons dans ce cas huit orientations :  
 $\theta = \{0, \pi/8, \pi/4, 3\pi/8, \pi/2, 5\pi/8, 3\pi/4, 7\pi/8\}$ .
- $\lambda$  spécifie la longueur d'onde et donc la fréquence de la sinusoïde. Cette variable peut prendre cinq valeurs comprises entre 4 et 16 :  $\lambda = \{4, 4\sqrt{2}, 8, 8\sqrt{2}, 16\}$ .
- $\phi$  spécifie la phase de la sinusoïde. Elle vaut 0 ou  $\pi/2$  selon que l'on veut la partie réelle ou imaginaire.
- $\sigma$  spécifie la variance de la gaussienne (sa taille ou résolution). Elle est proportionnelle à la longueur d'onde de la sinusoïde. Dans notre cas  $\sigma = \lambda$ .
- $\gamma$  spécifie l'aspect de la gaussienne. Ici les gaussiennes sont de forme circulaire :  
 $\gamma = 1$ .

La représentation de Gabor d'une image de visage est obtenue par la convolution de l'image avec la famille des filtres de Gabor, définie par  $IG(r, \theta) = I * G(r, \theta)$  où  $IG(r, \theta)$  est le résultat de la convolution de l'image par le filtre de Gabor à une certaine résolution  $r$  et à une orientation  $\theta$ .

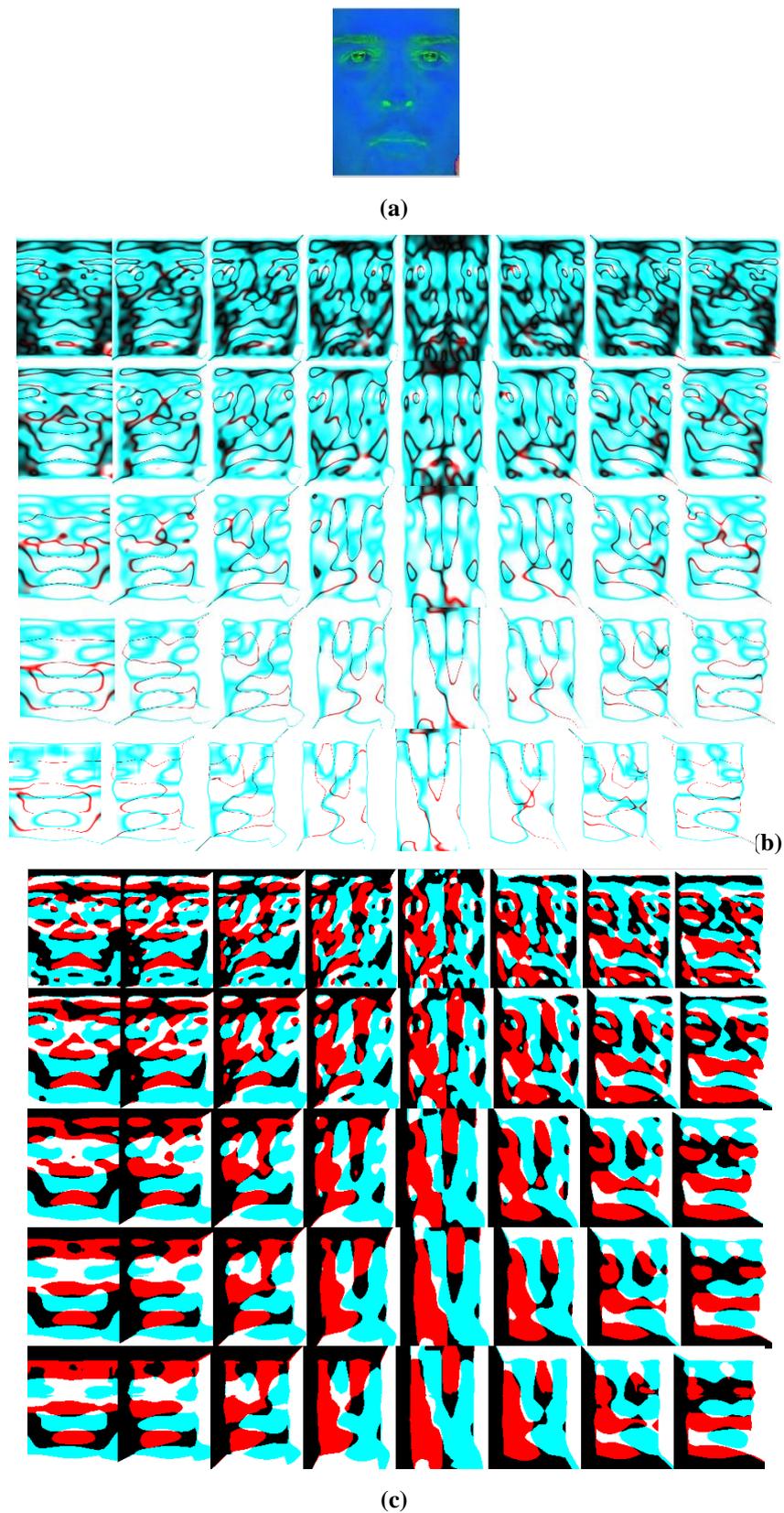
La famille de filtres de Gabor est alors caractérisée par un certain nombre de résolutions ou fréquences et orientations, qui seront appelées "caractéristiques" par la suite. Comme on peut le remarquer dans l'équation 2.9, les filtres de Gabor ont une forme complexe qu'il est possible d'exploiter. En effet, étant en quadrature de phase, il est important d'utiliser les informations données par la partie réelle et la partie imaginaire des coefficients de Gabor. Deux choix triviaux s'offrent à nous : l'étude de l'amplitude et l'étude de la phase de Gabor. Dans [87], Oppenheim a démontré que la phase des analyses multi-résolutions est plus informative que les amplitudes qui peuvent souffrir des problèmes d'illumination.

L'amplitude et la phase seront notées respectivement par :

$$M(IG_{r,\theta}(x, y)) = \sqrt{\text{Im}(IG_{r,\theta}(x, y))^2 + \text{Real}(IG_{r,\theta}(x, y))^2} \quad (2.10)$$

$$P(IG_{r,\theta}(x, y)) = \arctan\left(\frac{\text{Im}(IG_{r,\theta}(x, y))}{\text{Real}(IG_{r,\theta}(x, y))}\right) \quad (2.11)$$

La **figure 2.13** présente les amplitudes et les phases de la convolution de l'image de visage en l'espace de couleur **TSV** par quarante filtres de Gabor.



**Figure 2.13:** Résultats de la convolution d'une image de visage avec une famille de 40 filtres de Gabor (8 orientations (horizontales) et 5 résolutions (verticales)).(a) Image en espace de couleur TSV , l'ensemble (b) représente les amplitudes et (c) les phases de cette convolution.

### 2.5.2 Représentation du visage et choix des filtres de Gabor

La majorité des approches se basant sur le filtrage de Gabor pour la reconnaissance de visages (approches globales par réduction d'espace ou approches locales par points caractéristiques) utilisent uniquement l'amplitude de la réponse ou la fusion de l'amplitude et de la partie réelle [88, 89, 90]. L'utilisation de ces représentations est rarement argumentée et la non-utilisation de la phase de Gabor est dans la plupart des cas, justifiée par une instabilité temporelle de la réponse [57].

#### 1 Influence des caractéristiques de la famille des filtres de Gabor sur les performances de l'authentification de visages

Nous commencerons par étudier l'influence des caractéristiques des familles des filtres de Gabor sur les performances de l'authentification pour en déduire le choix optimal. La représentation de l'image considérée est l'**amplitude** des réponses des filtres de Gabor. L'image d'entrée est une image de visage coupée et **décimée**. Nous considérons la composante **S** de l'image couleur dans l'espace **TSV**. L'algorithme de réduction d'espace utilisé est **ACP + EFM** appliquée à l'amplitude des réponses des **filtres de Gabor** et la mesure de similarité utilisée est la **corrélation**. Le **tableau 2.4** présente les résultats de l'authentification en terme taux d'égale erreur **TEE** dans l'ensemble d'évaluation pour différents niveaux de résolution et d'orientations de quarante filtres de Gabor et un nombre de caractéristiques dans la phase de réduction d'espace par **EFM** égale  $N_p=100$ .

résolution ( $\sigma = \lambda$ )	Orientations des filtres ( $\theta$ )							
	0	$\pi/8$	$\pi/4$	$3\pi/8$	$\pi/2$	$5\pi/8$	$3\pi/4$	$7\pi/8$
4	9.28	10.1	8.13	8.02	8.01	8.04	8.63	7.3
$4\sqrt{2}$	8.33	9.01	7.54	9.61	5.35	7.95	7.2	8.5
8	9.31	7.34	8.7	5.7	7.85	<b>5.13</b>	8.02	8.17
$8\sqrt{2}$	9.54	8.64	7.31	9.36	10.19	8.3	7.07	7.54
16	9.17	8.48	8.65	9.18	9.18	8.64	7.84	7.62

**Tableau 2.4 :** Taux d'égale erreur **TEE** dans l'ensemble d'évaluation pour différents niveaux de résolutions et d'orientations pour les 40 filtres.

Dans ce tableau nous constatons que le meilleur **TEE= 5.13%** avec résolution ( $\sigma=8$ ) et orientation ( $\theta= 5\pi/8$ ) et ce n'est pas encore un bon résultat. Pour cela nous nous intéressons aux phases des filtres de Gabor.

**2Utilisation de la phase de Gabor pour l’authentification de visages**

Lorsque nous considérons une image du visage, certaines parties du visage n'ont aucune texture informative qui pourrait être analysée par les basses résolutions des filtres de Gabor. Pour ces régions, l'analyse par filtrage de Gabor donne  $Real(IG_{s,o}) \cong 0$  et  $Im(IG_{s,o}) \cong 0$ .

Même si ces valeurs sont très près de 0, l'amplitude de la convolution n'est pas affectée par ce problème, alors que la phase devient une forme indéterminée pour ces régions spécifiques.

Pour éviter les formes indéterminées, nous sélectionnons les régions informatives par seuillage de l'amplitude à chaque point d'analyse.

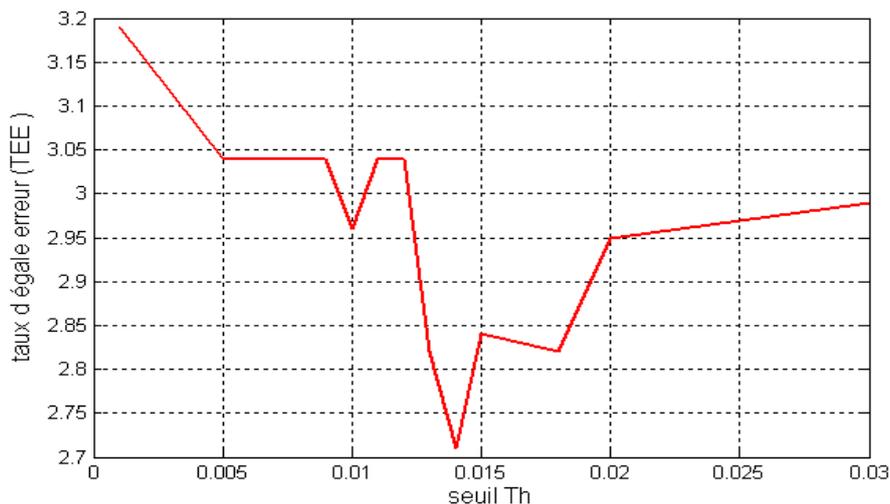
$$P(IG_{s,o}(x, y)) = \begin{cases} \arctan\left(\frac{Im(IG_{s,o}(x, y))}{Real(IG_{s,o}(x, y))}\right) & \text{Si } M(IG_{s,o})(x,y) > Th \end{cases} \quad (2.12)$$

$$0 \quad \text{Si } M(IG_{s,o})(x,y) < Th$$

Où  $(x; y)$  sont les coordonnées du point analysé et  $Th$  est le seuil de sélection de la phase.

**Optimisation du seuil pour la sélection de la phase**

Afin d'étudier l'influence du seuillage de la phase en fonction des performances. La **figure 2.14** montre l'évolution des taux d'égale erreur **TEE** en fonction du seuil **Th** par un **filtre de Gabor** avec résolution  $\sigma = \lambda=4$  et orientations  $\theta = /2$ .



**Figure 2.14 :** Taux d'égale erreur dans l'ensemble d'évaluation en fonction du seuil **Th**.

La courbe de la **figure 2.14** montre que la variation du taux de **TEE** en utilisant la **phase de Gabor** est bien **liée aux seuils du filtrage**.

## Chapitre 2 Authentification de visages et étude expérimentale

Notre choix s'est porté sur le seuil du filtrage  $Th = 0.014$  qui comme le montre la courbe présente le meilleur **TEE**. Notre seuil est presque égale à celui obtenu par Mellakh [42]. La même étude a été publiée dans [91] mais sur la base FRGC.

Dans cette deuxième phase, nous choisissons les filtres de Gabor optimaux pour la phase. Le **tableau 2.5** présente les résultats de l'authentification de visages en terme taux d'égale erreur **TEE** dans l'ensemble d'évaluation pour différents niveaux de résolutions et d'orientations des quarantefiltres de Gabor et un nombre de caractéristiques dans la phase de réduction d'espace par **EFM<sub>Np</sub>=100**.

résolution( $\sigma = \lambda$ )	Orientations des filtres ( $\theta$ )							
	0	$\pi/8$	$\pi/4$	$3\pi/8$	$\pi/2$	$5\pi/8$	$3\pi/4$	$7\pi/8$
4	4.79	5.14	4.12	4.96	<b>2.69</b>	<b>3.21</b>	<b>3.79</b>	4.64
$4\sqrt{2}$	4.8	5.29	6	5.28	4.15	4.88	4.87	5.3
8	6.03	6.53	7.16	6.79	6.04	6.85	6.29	7.04
$8\sqrt{2}$	6.64	7.47	7.29	8.21	8.52	8.14	7.66	7.8
16	6.5	7.01	8.16	8.45	9.01	8.61	7.99	7.84

**Tableau 2.5 :** Taux d'égale erreur **TEE** dans l'ensemble d'évaluation pour différent résolution et d'orientations des filtres pour les 40 filtres.

Nous remarquons que la première résolution et l'orientations :  $\theta = \pi/2, 5\pi/8, 5\pi/8$ , donnent le meilleur **TEE**. Les résultats obtenus par les phases de Gabor sont satisfaisants et encourageants. Nous utiliserons donc dans ce qui suit et pour la conception de notre système multi-algorithmes ces trois phases des filtres de Gabor.

Les meilleurs systèmes d'authentification de visages sont présentés dans le **tableau 2.6**.

Méthodes	Ensemble évaluation	Ensemble de test		
	<b>TEE %</b>	<b>TFR %</b>	<b>TFA %</b>	<b>TR %</b>
Système 1	$2.66 \pm 0.72$	$2 \pm 1.37$	$3.66 \pm 0.11$	$94.33 \pm 1.48$
Système 2	$2.69 \pm 0.72$	$0.5 \pm 0.69$	$4.07 \pm 0.12$	$95.43 \pm 0.81$
Système 3	$3.3 \pm 0.8$	$2 \pm 1.37$	$4.41 \pm 0.12$	$93.59 \pm 1.49$
Système 4	$3.79 \pm 0.85$	$0.5 \pm 0.69$	$4.47 \pm 0.12$	$95.03 \pm 0.82$

**Tableau 2.6** Résultats du système d'authentification de visages pour les quatre meilleurs systèmes retenus dans l'ensemble d'évaluation et test (intervalle de confiance paramétrique à 95%).

Avec :

- Le **système 1** utilise l'étape de prétraitement (voir paragraphe 2.3.1) puis

**ACP + EFM** comme étape de réduction d'espace et une comparaison par la métrique **corrélation**.

- Le **système 2** utilise la **phase filtrée** de la convolution de la composante **S** de l'espace de couleur **TSV** de l'image découpée et décimée et filtrée par le filtre de Gabor de la première résolution ( $\sigma = \lambda = 4$ ) et l'orientation ( $\theta = \pi/2$ ) puis **ACP + EFM** comme étape de réduction d'espace et enfin la **corrélation** pour la comparaison.
- Le **système 3** est identique au système 2 avec ( $\sigma = \lambda = 4$ ) et l'orientation ( $\theta = 5\pi/8$ ).
- Le **système 4** est aussi identique aux systèmes 2 et 3 avec ( $\sigma = \lambda = 4$ ) et l'orientation ( $\theta = 3\pi/4$ ).

### 2.6 Conclusion

Dans ce chapitre nous avons présenté les résultats obtenus en effectuant une étude comparative tout d'abord sur les méthodes de réduction d'espace. Puis une comparaison entre la transformation de Hough et les filtres de Gabor où l'information apportée par les phases des filtres est retenue et un seuil étudié argumente bien notre choix. Ces résultats montrent :

- L'efficacité de l'approche **ACP+EFM** par rapport **ACP, LDA, DLDA, ADLB2Doc** et **ADLB2Dol**.
- La transformation de Hough ne donne aucune amélioration au sens performance.
- Les phases filtrées de la convolution des images de visages par les filtres de Gabor contiennent beaucoup d'information et donnent de bons résultats surtout pour des orientations particulières.

Malgré des résultats satisfaisants pour les systèmes d'authentification de visages obtenus dans ce chapitre les taux d'erreur associés à des systèmes biométriques uni-modaux restent relativement élevés. Ce qui les rend inacceptables pour un déploiement d'applications critiques de sécurité. Pour pallier à ces inconvénients, une solution est l'utilisation de **plusieurs modalités biométriques** au sein d'un même système. On parle alors de **système biométrique multimodal** qui est étudié dans le reste de notre mémoire.

### 3 La multimodalité

#### 3.1 Introduction

Comme il a été annoncé dans le résultat obtenu dans le chapitre 2 concernant les systèmes uni-modaux d'authentification de visages. Chaque système en soi ne peut pas toujours être utilisé de manière fiable pour effectuer la reconnaissance. Cependant, la consolidation d'informations présentées par les différentes modalités peut permettre une authentification précise de l'identité. Alors la biométrie multimodale est la combinaison de plusieurs modalités biométriques différentes, ce qui implique l'utilisation d'informations complémentaires pour une personne donnée. Dans ce chapitre, nous allons d'abord présenter les limitations des systèmes biométriques lorsqu'ils n'utilisent qu'une seule modalité biométrique. Ensuite, nous présenterons la **biométrie multimodale** qui est l'objectif de notre étude et :

- Les différents multi- possibles,
- Les architectures,
- Les niveaux de fusion,
- Choix des types de multimodalité et du niveau de fusion,
- État de l'art de la fusion de scores,

#### 3.2 Les limitations des systèmes biométriques monomodaux

Bien que les techniques de reconnaissance biométrique promettent d'être très performantes, on ne peut garantir actuellement un excellent taux de reconnaissance avec des *systèmes biométriques unimodaux*, basés sur une unique signature biométrique. De plus, ces systèmes sont souvent affectés par les problèmes suivants [92] :

- **Bruit introduit par le capteur**: le bruit peut être présent dans les données biométriques acquises, ceci étant principalement dû à un capteur défaillant ou mal entretenu. Par exemple, un mauvais focus de caméra entraînant du flou dans des images de visage. Le taux de reconnaissance d'un système biométrique est très sensible à la qualité de l'échantillon biométrique et des données bruitées peuvent sérieusement compromettre la précision du système [93],
- **Non-universalité** : si chaque individu d'une population ciblée est capable de présenter une modalité biométrique pour un système donné, alors cette modalité est dite

*universelle*. Ce principe d'universalité constitue une des conditions nécessaires de base pour un module de reconnaissance biométrique. Cependant, toutes les modalités biométriques ne sont pas vraiment universelles. Le *National Institute of Standards and Technologies* (NIST) a rapporté qu'il n'était pas possible d'obtenir une bonne qualité d'empreinte digitale pour environ **2%** de la population [94].

- **Manque d'individualité** : les caractéristiques extraites à partir de données biométriques d'individus différents peuvent être relativement similaires. Par exemple, une certaine partie de la population peut avoir une apparence faciale pratiquement identique due à des facteurs génétiques (père et fils, vrais jumeaux, etc...). Ce manque d'unicité augmente le taux de fausse acceptation **TFA** d'un système biométrique.
- **Manque de représentation invariante**: les données biométriques acquises à partir d'un utilisateur lors de la phase de reconnaissance ne sont pas identiques aux données qui ont été utilisées pour générer le modèle de ce même utilisateur lors de la phase d'enrôlement. Ceci est connu sous le nom de "*variations intra-classe*". Ces variations peuvent être dues à une mauvaise interaction de l'utilisateur avec le capteur (par exemple, changements de pose et d'expression faciale lorsque l'utilisateur se tient devant une caméra), à l'utilisation de capteurs différents lors de l'enrôlement et de la vérification, à des changements de conditions de l'environnement ambiant (par exemple, changements en éclairage pour un système de reconnaissance faciale). Cependant, dans la plupart des systèmes biométriques, ces caractéristiques ne sont pas invariantes et, par conséquent, des algorithmes complexes sont requis pour prendre en compte ces variations. De grandes *variations intra-classe* augmentent généralement le taux de faux rejet **TFR** d'un système biométrique.
- **Sensibilité aux attaques**: bien qu'il semble très difficile de voler les modalités biométriques d'une personne, il est toujours possible de contourner un système biométrique en utilisant des modalités biométriques usurpées. Des études [95] ont montré qu'il était possible de fabriquer de fausses empreintes digitales en gomme et de les utiliser pour contourner un système biométrique. Les modalités biométriques comportementales telles que la signature et la voix sont plus sensibles à ce genre d'attaque que les modalités biométriques physiologiques [11].

### 3.3 La multimodalité

**La multimodalité est l'utilisation de plusieurs systèmes biométriques.** La combinaison de plusieurs systèmes a pour objectif d'en diminuer les problèmes vus au paragraphe 3.2. En effet, l'utilisation de plusieurs systèmes a pour but premier d'améliorer les performances de reconnaissance. En augmentant la quantité d'informations discriminante de chaque personne, on souhaite augmenter le pouvoir de reconnaissance du système. De plus, le fait d'utiliser plusieurs modalités biométriques réduit le risque d'impossibilité d'enregistrement ainsi que la robustesse aux fraudes.

### 3.3.1 Les différents multi- possibles

Les systèmes biométriques multimodaux diminuent les contraintes des systèmes biométriques monomodaux en combinant plusieurs systèmes. On peut différencier **5 types** de systèmes multimodaux selon les systèmes qu'ils combinent (**figure 3.1**). On les appelle [10]:

- **multi-algorithmes** : lorsque plusieurs algorithmes traitent la même image acquise, cette multiplicité des algorithmes peut intervenir dans le module d'extraction en considérant plusieurs ensembles de caractéristiques et/ou dans le module de comparaison en utilisant plusieurs algorithmes de comparaison.
- **multi-instances** : lorsqu'ils associent plusieurs instances de la même biométrie, par exemple l'acquisition de plusieurs images de visage avec des changements de pose, d'expression ou d'illumination.
- **multi-capteurs** : lorsqu'ils associent plusieurs capteurs pour acquérir la même modalité, par exemple un capteur optique et un capteur capacitif pour l'acquisition de l'empreinte digitale.
- **multi-échantillons** : lorsqu'ils associent plusieurs échantillons différents de la même modalité, par exemple deux empreintes digitales de doigts différents ou les deux iris. Dans ce cas les données sont traitées par le même algorithme mais nécessitent des références différentes à l'enregistrement contrairement aux systèmes multi-instances qui ne nécessitent qu'une seule référence.
- **multi-biométries** : lorsque l'on considère plusieurs biométries différentes, par exemple visage et iris.

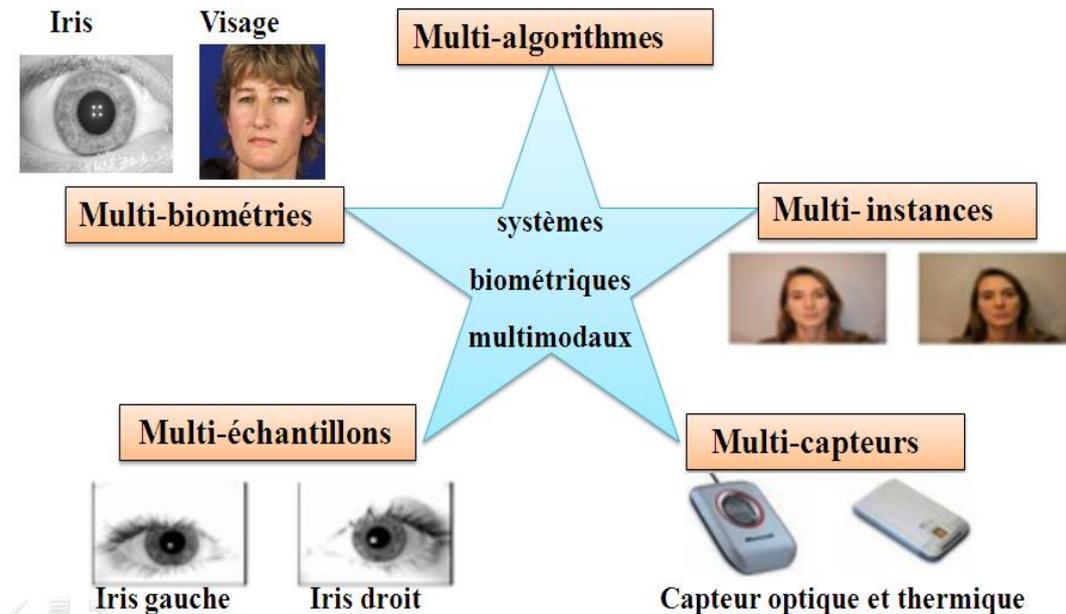


Figure 3.1 : Les différents systèmes multimodaux

### 3.3.2 Les architectures

Les systèmes multimodaux associent plusieurs systèmes biométriques et nécessitent donc l'acquisition et le traitement de plusieurs données. L'acquisition et le traitement peuvent se faire successivement, on parle alors **d'architecture en série**, ou simultanément, on parle alors **d'architecture en parallèle**.

L'architecture est en réalité surtout liée au traitement. En effet, l'acquisition des données biométriques est en général séquentielle pour des raisons pratiques. Il est difficile d'acquérir en même temps une empreinte digitale et une image d'iris dans de bonnes conditions. Il existe cependant certains cas où les acquisitions peuvent être faites simultanément lorsque les différentes données utilisent le même capteur par exemple les capteurs d'empreintes multi-doigts qui permettent d'acquérir plusieurs doigts simultanément ou même les empreintes palmaires.

L'architecture est donc en général liée au traitement et en particulier à la décision. En effet la différence entre un système multimodal en série et un système multimodal en parallèle réside dans le fait d'obtenir un score de similarité à l'issue de chaque acquisition (fusion en série) ou de procéder à l'ensemble des acquisitions avant de prendre une décision (fusion en parallèle).

L'architecture en parallèle (**figure 3.2**) est la plus utilisée car elle permet d'utiliser toutes les informations disponibles et donc d'améliorer les performances du système. En revanche, l'acquisition et le traitement d'un grand nombre de données biométriques est coûteux

en temps et en matériel, et réduite le confort d'utilisation. C'est pour cela que l'architecture en série (**figure 3.3**) peut être privilégiée dans certaines applications ; par exemple si la multimodalité est utilisée pour donner une alternative pour les personnes ne pouvant pas utiliser l'empreinte digitale. Pour la majorité des individus seule l'empreinte est acquise et traitée mais pour ceux qui ne peuvent pas être ainsi authentifiés on utilise un système à base d'iris alternativement [10].

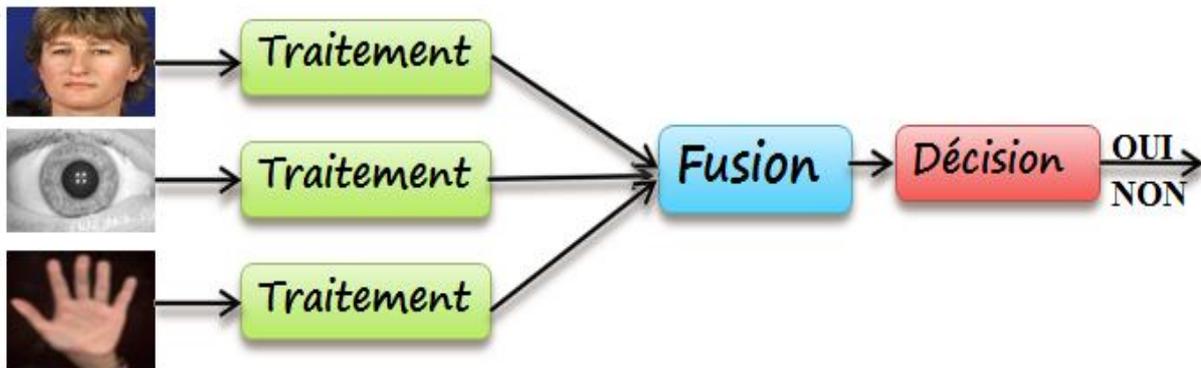


Figure 3.2 : Architecture de fusion en parallèle

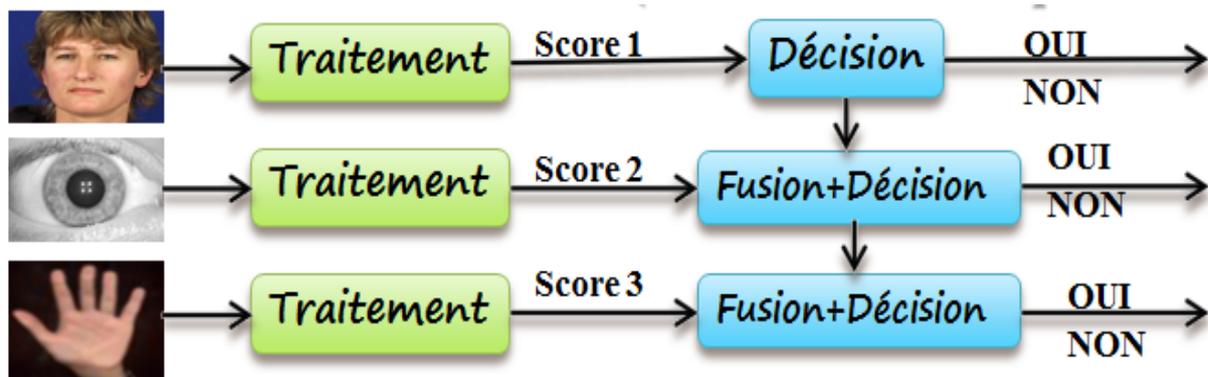


Figure 3.3 : Architecture de fusion en série (incrémentale ou séquentielle)

### 3.3.3 Les niveaux de fusion [10,11]

La combinaison de plusieurs systèmes biométriques peut se faire à quatre niveaux différents : au niveau des **données**, au niveau des **caractéristiques** extraites, au niveau des **scores** issus du module de comparaison ou au niveau des **décisions** du module de décision (**figure 3.4**).

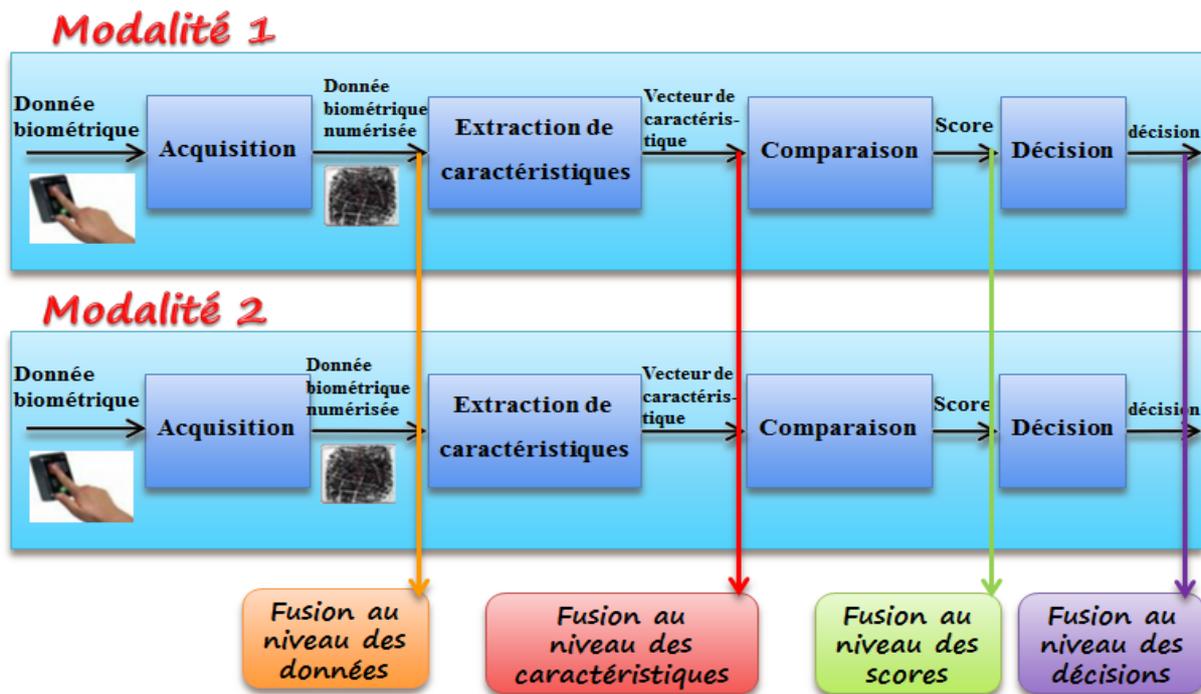


Figure 3.4 : Les différents niveaux de fusion

Ces quatre niveaux de fusion peuvent être classés en deux sous-ensembles :

- la fusion pré-classification (avant comparaison),
- la fusion post-classification (après la comparaison).

### 1 La fusion pré-classification

La fusion pré-classification correspond à la fusion des informations issues de plusieurs données biométriques au niveau du capteur (images brutes) ou au niveau des caractéristiques extraites par le module d'extraction de caractéristiques.

#### Niveau du capteur (*Sensor Level*)

La fusion au niveau **capteur** est relativement peu utilisée car se faire uniquement si les diverses captures sont des instances du même trait biométrique obtenu à partir de **plusieurs capteurs** compatibles entre eux ou **plusieurs instances** du même trait biométrique obtenu à partir d'un seul capteur. De plus, les captures doivent être compatibles entre elles et la correspondance entre les points dans les données brutes doit être connue par avance. Par exemple, les images de visage obtenues à partir de plusieurs caméras peuvent être combinées pour former un modèle 3D du visage. Un autre exemple de fusion au niveau capteur consiste à mettre en mosaïque plusieurs images d'empreintes digitales afin de former une image d'empreinte digitale finale plus complexe. La fusion au niveau capteur n'est généralement pas

possible si les instances des données sont incompatibles (par exemple, il est peut être difficile de fusionner des images de visages provenant de caméras ayant des résolutions différentes).

### Niveau Caractéristiques (*Feature Level*)

La fusion au niveau des caractéristiques est moins limitée par la nature des données biométriques. Cependant une certaine homogénéité est nécessaire pour la plupart des méthodes de fusion au niveau des caractéristiques comme par exemple la moyenne de plusieurs "templates" d'empreintes ou de visages. Un exemple de fusion au niveau des caractéristiques qui ne nécessite pas vraiment d'homogénéité est la concaténation de plusieurs vecteurs de caractéristiques avant le traitement par l'algorithme de comparaison. Par exemple, dans [96], Jing et al proposent une méthode de fusion de caractéristiques pour la fusion de visage et d'empreinte palmaire. La fusion est effectuée par concaténation d'images obtenues par transformée de Gabor sur les images de visage et d'empreinte de la main. Mais la concaténation pose le problème de la dimension de l'espace de classification qui lorsqu'il augmente, rend plus difficile la tâche de classification.

Les méthodes de fusion pré-classification **sont assez peu utilisées** car elles posent un certain nombre de contraintes qui ne peuvent être remplies que dans certaines applications très spécifiques. En revanche, la fusion post-classification est très étudiée par les chercheurs.

### 2 La fusion post-classification

La fusion post-classification peut se faire au niveau des scores issus des modules de **comparaison** ou au niveau des **décisions**. Dans les deux cas, la fusion est en fait un problème bien connu de la littérature sous le nom de "Multiple Classifier systems".

### Niveau Décision (*Decision Level*)

La fusion au niveau des décisions est souvent utilisée pour sa simplicité. En effet, chaque système fournit une décision binaire sous la forme **OUI** ou **NON** que l'on peut représenter par **0** et **1**, et le système de fusion de décisions consiste à prendre une décision finale en fonction de cette série de **0** et de **1**. Les méthodes les plus utilisées sont des méthodes à base de **votes** telles que le **OR** (si un système a décidé 1 alors OUI), le **AND** (si tous les systèmes ont décidé 1 alors OUI) ou le vote à la **majorité** (si la majorité des systèmes ont décidé 1 alors OUI). On peut également utiliser des méthodes plus complexes qui pondèrent les décisions de chaque sous-système ou qui utilisent des classifieurs dans l'espace de décisions telles que **BKS** (Behaviour Knowledge Space) [97]. Dans [98], **Verlinde** présente un

grand nombre de méthodes de fusion de décision. Ces méthodes de fusion au niveau des décisions sont très simples mais utilisent très peu d'information (Oou 1).

### Niveau Score (*Score Level*)

La fusion au niveau des scores est le type de fusion le plus utilisé car elle peut être appliquée à tous les types de systèmes (contrairement à la fusion pré-classification), dans un espace de dimension limité (un vecteur de scores dont la dimension est égale au nombre de sous-systèmes), avec des méthodes relativement simples et efficaces mais traitant plus d'information que la fusion de décisions. La fusion de scores consiste donc à la classification : **OUI** ou **NON** pour la décision finale, d'un vecteur de nombres réels dont la dimension est égale au nombre de sous-systèmes. Il existe un grand nombre de méthodes de fusion de scores qui seront présentées dans le paragraphe 3.3.5 et seront détaillées dans le chapitre 4.

### 3.3.4 Choix des types de multimodalité et du niveau de fusion

Comme nous l'avons vu dans les sections 3.3.1 et 3.3.3 il existe une multitude de systèmes biométriques multimodaux qui se distinguent par leurs types de modalités et de niveau de fusion. Après notre étude et pour des raisons de disponibilité de base de données visage (Face) notre choix est la fusion multi-algorithmique qui se présente comme suit :

#### 1 Choix du type de multimodalité

Nous essayons de combiner plusieurs algorithmes d'authentification de visages (multi-algorithmes) pour une raison de manque d'autres BDDs la seule que nous avons est les visages 2D de la base de données XM2VTS et ensuite parce que :

- Les caractéristiques faciales marquent un score de compatibilité le plus élevé dans un système *MRTD* [13].
- La modalité du visage est non-intrusive, cela veut dire qu'elle n'atteint pas l'intimité de l'individu ; elle constitue l'une des voies les plus naturelles pour reconnaître une personne, elle permet d'effectuer un travail à la volée et son coût de déploiement est relativement faible : une simple caméra reliée à un ordinateur peut suffire.
- L'acquisition et le traitement d'un grand nombre de données biométriques est coûteux en temps et en matériel, et réduit le confort d'utilisation.
- Les systèmes combinant plusieurs informations issues de la même image de visage permettent d'améliorer les performances en reconnaissance.

### 2 Niveau de fusion

Pour le choix du niveau de fusion nous sommes basés sur les travaux de référence [10, 11]. Nous utilisons dans notre travail la fusion des scores pour les raisons suivantes :

- La relation entre les espaces de caractéristiques des différents systèmes biométriques n'est pas forcément connue. Dans le cas où la relation est connue, on doit prendre soin d'éliminer les caractéristiques qui sont fortement corrélées. Cela requiert l'application d'algorithmes de sélection de caractéristiques avant l'étape de classification,
- La concaténation de deux vecteurs de caractéristiques peut engendrer un vecteur de caractéristiques ayant une grande dimension, menant au fameux problème de la "*malédiction de la dimensionnalité*". Bien que ce soit un problème général dans la plupart des applications de reconnaissance de forme, cela est encore plus marquant dans les applications biométriques à cause du temps, de l'effort et du coût impliqués dans la collecte de grandes quantités de données biométriques,
- La plupart des systèmes biométriques commerciaux ne fournissent pas l'accès aux vecteurs de caractéristiques qui sont utilisés dans leurs produits. Ainsi, très peu de chercheurs ont étudié la fusion aux niveaux caractéristiques et la plupart d'entre eux se tournent généralement vers les schémas de fusion de scores,
- La fusion au niveau des scores est le type de fusion le plus utilisé car elle peut être appliquée à tous les types de systèmes (contrairement à la fusion pré-classification).
- Après les vecteurs de caractéristiques, les scores donnés en sortie par les systèmes contiennent l'information la plus riche à propos du modèle d'entrée.
- La fusion au niveau score donne le meilleur compromis entre la richesse d'information et la facilité d'implémentation. Aussi, il est relativement facile d'accéder et de combiner les scores générés par les différents systèmes.
- L'intégration d'information au niveau score est l'approche la plus courante dans les systèmes biométriques multimodaux.

#### 3.3.5 État de l'art de la fusion de scores

Au cours des dix dernières années les systèmes multi-algorithmes ou multi-classifieurs (multiple classifier systems *MCS*) sont devenus une méthode établie pour concevoir des systèmes de classification. Un grand nombre d'éléments de preuve à la fois théoriques et expérimentales montre que les systèmes multi-classifieurs peuvent surpasser un classifieur unique dans plusieurs applications réelles, en termes de précision de la classification (voir, par

exemple[99,100]). En particulier, plusieurs auteurs ont montré que *MCS* peut permettre d'améliorer l'authentification biométrique de visages [101, 102, 103].

Il existe deux approches pour fusionner les scores obtenus par différents classifieurs :

### 1 Approche par classification de scores

Plusieurs classifieurs ont été utilisés pour combiner les scores de correspondance afin d'arriver à une décision.

*Wang et al*[104] considèrent les scores provenant de modules de reconnaissance faciale et de reconnaissance par l'iris comme un vecteur de caractéristiques à deux dimensions. Une analyse discriminante linéaire de Fisher (**LDA**) et un classifieur par **réseau de neurones** combiné à une fonction de base **radiale (RBF)** sont ensuite utilisés pour la classification. *Verlinde et Chollet*[105] combinent les scores provenant de deux modules de reconnaissance faciale et d'un module de reconnaissance de la parole avec l'aide de trois classifieurs: un premier classifieur utilisant la méthode des "**k plus proches voisins**" ("k-nearest neighbor algorithm", "**k-NN**") avec une quantification vectorielle, un deuxième classifieur basé sur un **arbre décisionnel** et un dernier classifieur basé sur un modèle de **régression logistique**.

*Chatzis et al*[106] utilisent une méthode de regroupement ("clustering") appelée "**fuzzy k means**" et une **quantification vectorielle floue** ("fuzzy vector quantization"), couplée à un classifieur de réseau de neurones de **RBF** médiane pour fusionner les scores obtenus à partir de systèmes biométriques basés sur des caractéristiques visuelles (visage) et acoustiques (voix).

*Sanderson et al*[107] utilisent un classifieur basé sur une Machine à Vecteurs de Support (**SVM**) pour combiner les scores issus d'un module de reconnaissance faciale et d'un module de reconnaissance de la parole. Ils montrent que la performance d'un tel classifieur se détériore en la présence de conditions de bruit en entrée. Afin de surmonter ce problème, ils implémentent des classifieurs résistants au bruit structurel comme un classifieur linéaire défini par morceaux ("Piece-Wise Linear Classifier") et un Classifieur Bayésien Modifié.

*Ross et Jain*[108] utilisent un **arbre décisionnel** et des **classifieurs discriminants linéaires** pour combiner les scores des modalités du visage, d'empreinte digitale et de géométrie de la main.

### 2 Approche par combinaison de scores

*Kittler et al*[109] ont développé un cadre théorique pour combiner les informations d'identification obtenues à partir de plusieurs classifieurs en utilisant des schémas tels que la règle somme ("*sumrule*"), la règle "produit" ("*productrule*"), la règle maximum ("*max rule*"), la règle minimum ("*min rule*") et la règle médiane ("*medianrule*"). Afin d'employer ces schémas, les scores doivent être convertis en probabilités a posteriori conformément à un utilisateur client ou imposteur. Ils considèrent le problème de classer un modèle d'entrée  $\mathbf{X}$  dans l'une des  $m$  classes possibles (dans un système de vérification,  $m = 2$ ) d'après l'information d'identification fournie par  $R$  classifieurs ou matchers différents.

*Kwak et al* [110] réalisent l'intégration floue de **Choquet** pour deux systèmes multi-classifieurs. Le premier système combine des images traitées du visage (yeux, nez et la bouche) et images globales de visage. Le deuxième système est basé sur des classifieurs opérant sur quatre ensembles de sous-image générée par décomposition en ondelettes.

*Rasheed et al* [111] utilisent l'intégrale floue de **Sugeno** pour la combinaison de scores pour un système multi-classifieurs pour la décomposition d'un signal électromyographique (EMG).

*Chia et al* [112] utilisent une méthode hybride de calcul du minimum, maximum ou somme pour la combinaison des scores de deux systèmes d'authentification de visages et voix.

*Yong Li et al*[113] utilisent la somme pondérée pour la fusion par combinaison des scores.

*Shukla et al*[114] proposent un calcul adaptatif et hybride utilisant la combinaison de scores par **logique floue** à base de l'**intégrale de Sugeno** ou de **Choquet**.

### Travaux récents

**Morizet** et **Allano** ont essayé d'utiliser les deux approches et arrivent à des taux de réussite très intéressants.

**Allano**[10] utilise les deux approches basées sur le **SVM** pour l'approche par classification de scores et le **moyenne**, le **produit**, le **minimum**, le **maximum**, la **médiane** et **somme pondérée** pour l'approche par combinaison de scores.

**Morizet**[11] utilise l'approche basée sur la combinaison de scores avec les méthodes simples (**moyenne**, le **produit**, le **minimum**, le **maximum**, la **médiane** et **somme pondérée**) et une nouvelle technique de fusion adaptative appelée "**WaveletScore Denoising Statistical Fusion**" (**WSDSF**),

Nous nous sommes inspirés de leurs travaux et nous utilisons les deux approches l'étude des stratégies de fusion basée sur des méthodes de ces approches de fusion dont l'objectif est de réduire le coût et le temps d'utilisation des systèmes multimodaux et améliorer

les performances du systèmes biométrique. Les approches de fusion utilisées sont les suivantes et sont bien développées dans le chapitre 4:

- la **première approche** est de traiter le sujet comme un problème de *combinaison* par les méthodes simples : **moyenne**, le **produit**, le **minimum**, le **maximum**, la **médiane** et **somme pondérée** [113,115] et la combinaison de scores par **logique floue** à base de **l'intégrale de Sugeno** ou de **Choquet** [111,114,116].
- la **deuxième approche** est de voir cela comme un problème de *classification* comme méthode **statistique de Fisher** [117], les machines à vecteurs de support **SVM** [102,10,118,115] et les **réseaux de neurones artificiels (MLP)** [119].

Une autre question importante traitée dans ce travail **est la normalisation des scores** comme présentée par [120] qui est une phase nécessaire avant de faire la combinaison.

### 3.4 Conclusion

Dans ce chapitre nous avons commencé par la présentation des limitations des systèmes biométriques lorsqu'ils n'utilisent qu'une seule modalité biométrique. Nous avons ensuite présenté une façon de réduire les limitations des systèmes biométriques monomodaux en combinant plusieurs systèmes menant ainsi à la biométrie multimodale. Les systèmes multimodaux peuvent être de différente nature. Nous avons donc présenté dans ce chapitre les différents types de combinaisons de modalités possibles, mais aussi les architectures et les niveaux de fusion qui peuvent être utilisés dans un système multimodal. Nous avons présenté encore le type de modalités et le niveau de fusion et l'état de l'art de la fusion de scores. Dans notre travail nous n'étudierons pas tous les aspects des systèmes multimodaux. Nous nous concentrerons sur les systèmes faisant la fusion au niveau des scores qui sera développée dans le chapitre 4. Nous traitons les systèmes multi-algorithmes ou multi-classifieurs. Ces systèmes de fusion de scores seront traités par des architectures en parallèle pour les systèmes multi-algorithmes appliqués sur la base de données XM2VTS d'une part et d'autre part les multi-biométries appliqués sur les scores de visage et voix de la base de données XM2VTS dans le chapitre 5.



### 4 Les méthodes de fusion de scores

#### 4.1 Introduction

Après avoir introduit et défini la biométrie multimodale en général au chapitre 3. Nous allons maintenant nous intéresser aux méthodes de **fusion de scores**. Les méthodes de fusion de scores combinent les informations au niveau des scores issus des modules de comparaison comme indiqué sur la **figure 4.1**.

Un système de fusion est constitué de deux modules, un module de **fusion** et un module de **décision** (voir **figure 4.1**). Le problème devient donc un problème de classification à 2 classes (**OUI** ou **NON**, Client ou Imposteur) à partir d'un vecteur de nombre réels dont la dimension est égale au nombre de sous-systèmes du système **multi-algorithmes**.

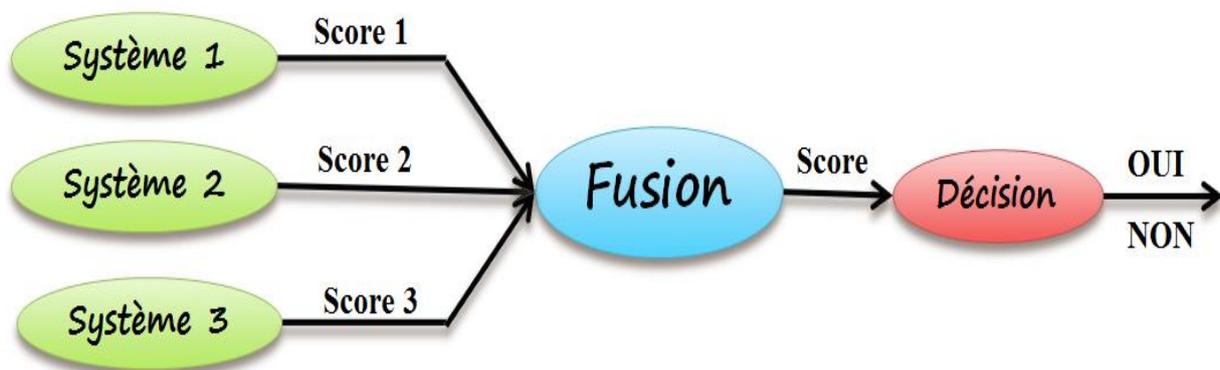


Figure 4.1 : Schéma de la fusion de scores.

Il existe deux approches pour combiner les scores obtenus par différents systèmes :

- i) La première approche est de traiter le sujet comme un problème de **combinaison**,
- ii) tandis que l'autre approche est de voir cela comme un problème de **classification**.

Il est important de noter que *Jain et al* ont montrés que les approches par combinaison sont plus performantes que la plupart des méthodes de classification [121, 122].

Dans **l'approche par combinaison**, les scores individuels sont combinés de manière à former un unique score qui est ensuite utilisé pour prendre la décision finale. Afin de s'assurer

que la combinaison de scores provenant de différents systèmes soit cohérente, les scores doivent d'abord être transformés dans un domaine commun : on parle alors de **normalisation de score** [123].

Dans l'**approche par classification**, un vecteur de caractéristiques est construit en utilisant les scores donnés en sortie par les systèmes individuels ; ce vecteur est ensuite attribué à une des deux classes : Client ou Imposteur. En général, le classifieur utilisé pour cette opération est capable d'apprendre la frontière de décision sans tenir compte de la manière dont le vecteur de caractéristiques a été généré. Ainsi, les scores en sortie de différents systèmes peuvent être non-homogènes (mesure de distance ou de similarité, différents intervalles de valeurs prises, etc....) et aucun traitement n'est requis avant de les envoyer dans le classifieur.

### 4.2. Normalisation de scores

Les méthodes de normalisation de scores ont pour objectif de transformer individuellement chacun des scores issus des systèmes pour les rendre **homogènes** avant de les combiner. En effet, les scores issus de chaque système peuvent être de nature différente.

Certains systèmes produisent des scores de **similarité** (plus le score est grand, plus la référence ressemble au test, donc l'utilisateur est un Client), d'autres produisent des **distances** (plus la distance est faible, plus la référence et le test sont proches, plus l'utilisateur est un Client). De plus chaque système peut avoir des intervalles de variations des scores différents, par exemple pour un système les scores varient entre 0 et 1 et pour un autre les scores varient entre 0 et 1000.

On comprend bien la nécessité de normaliser les scores avant de les combiner. Les méthodes de normalisation présentées dans la suite, traitent des scores qui varient déjà tous dans le même sens (en général on considère tous les scores sous forme de similarité). Pour transformer des distances en similarité il existe deux solutions : l'inverse ou l'opposé. Dans toute la suite, nous considérerons que tous les scores à fusionner ont été transformés en scores de similarité (**scores Client > scores Imposteur**) [10].

Les différentes techniques de normalisation de scores utilisés sont :

- **Normalisation par la méthode Min-Max**
- **Normalisation par la méthode Z-Score**
- **Normalisation par la méthode tangente hyperbolique "Tanh"**
- **Normalisation par la médiane et l'écart absolu médian (MAD)**
- **Normalisation par une fonction quadratique-linéaire-quadratique (QLQ)**

### 4.2.1 Normalisation par la méthode Min-Max [10, 11, 124]

La technique de normalisation la plus simple est la **normalisation Min-Max**. Elle est la plus adaptée dans le cas où les bornes (valeurs minimales et maximales) des scores produits par des systèmes sont connues. Dans ce cas, on peut facilement translater les scores minimums et maximums respectivement vers 0 et 1. Cependant, même si les scores ne sont pas bornés, on peut estimer les valeurs minimales et maximales pour un jeu de scores donné et appliquer ensuite la normalisation Min-Max. Soit  $s_{ij}$  le  $j^{\text{ème}}$  score de sortie du  $i^{\text{ème}}$  système, où  $i = 1, 2, \dots, N$  et  $j = 1, 2, \dots, M$  ( $N$  est le nombre de systèmes et  $M$  le nombre de scores disponibles dans l'ensemble de données d'entraînement). Le score normalisé Min-Max pour

le score de test  $s_{ik}$  est donné par (4.1):

$$s'_{ik} = \frac{s_{ik} - \min(\{s_i\})}{\max(\{s_i\}) - \min(\{s_i\})} \quad (4.1)$$

Où  $\{s_i\} = \{s_{i1}, s_{i2}, \dots, s_{iM}\}$ . Quand les valeurs minimales et maximales sont estimées à partir du jeu d'entraînement de scores donné, cette méthode n'est pas robuste (c'est à dire que cette méthode est fortement sensible aux valeurs aberrantes dans les données utilisées pour l'estimation). La normalisation Min-Max conserve la distribution de scores originale à un facteur d'échelle près et transforme tous les scores dans l'intervalle  $[0, 1]$ .

### 4.2.2 Normalisation par la méthode Z-Score [10, 11, 124]

La technique de normalisation de score la plus employée est certainement la **Z-Score** qui utilise la **moyenne** arithmétique et l'**écart-type** des données. On peut s'attendre à ce que cette méthode fonctionne bien si on a une connaissance à priori du score moyen et des variations de score d'un système. Si on n'a pas de connaissance à priori sur la nature de l'algorithme de reconnaissance, nous devons alors estimer la moyenne et l'écart-type des scores à partir d'un jeu de scores donné. Les scores normalisés sont donnés par (4.2) :

$$(4.2) \quad s'_{ik} = \frac{s_{ik} - \mu}{\sigma},$$

Où  $\mu$  est la moyenne arithmétique et  $\sigma$  l'écart-type des données. Cependant, la moyenne et l'écart-type sont tous les deux sensibles aux valeurs aberrantes et donc cette méthode n'est pas robuste. De plus, la normalisation *Z-Score* ne garantit pas un intervalle commun pour les scores normalisés provenant de différents systèmes. Si la distribution des scores n'est pas gaussienne, la normalisation *Z-Score* ne conserve pas la

distribution d'entrée en sortie. Cela est simplement dû au fait que la moyenne et l'écart-type sont les paramètres de position et d'échelle optimaux seulement pour une distribution gaussienne.

Pour une distribution arbitraire, la moyenne et l'écart-type sont respectivement des estimateurs raisonnables de position et d'échelle, mais ne sont pas optimaux.

### 4.2.3 Normalisation par la méthode tangente hyperbolique "Tanh" [10, 11, 125]

Les scores normalisés sont donnés par (4.3):

$$s'_{ik} = \frac{1}{2} \left\{ \tanh\left(0.001 \frac{s_{ik} - \mu}{\sigma}\right) + 1 \right\}, \quad (4.3)$$

Où  $\mu$  est la moyenne arithmétique,  $\sigma$  l'écart-type des données et  $\tanh$  la tangente hyperbolique. La méthode tangente hyperbolique met chaque score normalisé dans l'intervalle  $[0, 1]$ .

### 4.2.4 Normalisation par la médiane et l'écart absolu médian (MAD) [11, 124]

Ces deux méthodes sont insensibles aux valeurs aberrantes et aux points aux extrémités d'une distribution. Ainsi, une méthode de normalisation utilisant la médiane et la MAD (l'écart absolu médian) serait robuste et est donnée par (4.4):

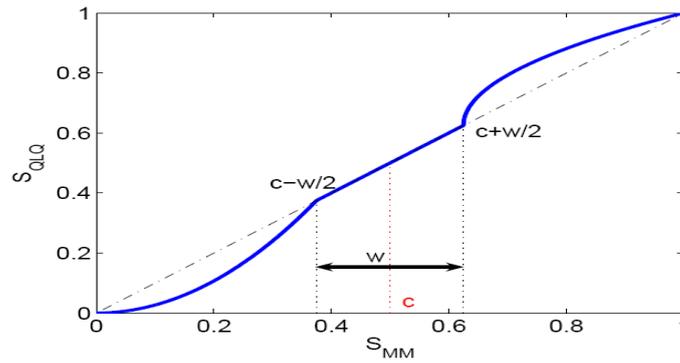
$$s'_{ik} = \frac{s_{ik} - \text{median}(\{s_i\})}{MAD},$$

Où:  $MAD = \text{median}(\{|s_i - \text{median}(\{s_i\})| \})$

Cependant, les estimateurs issus de la médiane et de la MAD ont une faible efficacité comparée aux estimateurs issus de la moyenne et de l'écart-type, c'est-à-dire que lorsque la distribution de score n'est pas gaussienne, la médiane et la MAD sont de pauvres estimateurs des paramètres de position et d'échelle. Ainsi, cette technique de normalisation ne conserve pas la distribution d'entrée et ne transforme pas les scores dans un intervalle commun.

### 4.2.5 Normalisation par une fonction Quadratique-Linéaire-Quadratique (QLQ) [11, 126, 127, 128]

*Snelick et al* [126] utilisent une fonction Quadratique-Linéaire-Quadratique (QLQ) pour normaliser des scores au préalable transformés dans l'intervalle  $[0, 1]$  ( $s_{MM}$ ) grâce à une normalisation *MinMax*( $s_{MM}$ ). Cette normalisation QLQ prend comme paramètres le centre  $c$  et la largeur  $w$  de la zone de recouvrement des distributions des scores Imposteurs et Client (**figure 4.2**).



**Figure 4.2 :** Normalisation QLQ.

La zone de recouvrement reste inchangée tandis que les autres régions sont transformées à l'aide de deux fonctions quadratiques par segment. Le score normalisé est donné par (4.5):

$$S_{QLQ} \begin{cases} \frac{1}{c - \frac{\omega}{2}} s_{MM}^2, & \text{si } s_{MM} \leq \left(c - \frac{\omega}{2}\right) \\ s_{MM}, & \text{si } \left(c - \frac{\omega}{2}\right) < s_{MM} \leq \left(c + \frac{\omega}{2}\right) \\ \left(c + \frac{\omega}{2}\right) + \sqrt{\left(1 - c - \frac{\omega}{2}\right) \left(s_{MM} - c - \frac{\omega}{2}\right)}, & \text{sinon} \end{cases} \quad (4.5)$$

#### 4.2.6 Normalisation par une fonction double sigmoïde [11,129]

Cappelli et al [127] ont utilisé une **fonction double sigmoïde** pour la normalisation de score dans un système biométrique multimodal qui combine différents systèmes d'empreinte digitale. Le score normalisé est donné par (4.6) :

$$s'_{ik} = \begin{cases} \frac{1}{1 + \exp\left(-2\left(\frac{s_{ik} - t}{r_1}\right)\right)} & \text{si } s_{ik} < t, \\ \frac{1}{1 + \exp\left(-2\left(\frac{s_{ik} - t}{r_2}\right)\right)} & \text{sinon} \end{cases} \quad (4.6)$$

Où  $t$  est le point de fonctionnement de référence et  $r_1$  et  $r_2$  sont respectivement les bords gauche et droit, la figure 4.3 montre un exemple d'une normalisation double sigmoïde, où les scores étant dans l'intervalle  $[0, 300]$  sont transformés dans l'intervalle  $[0, 1]$ , avec  $t = 200$ ,  $r_1 = 20$  et  $r_2 = 30$ .

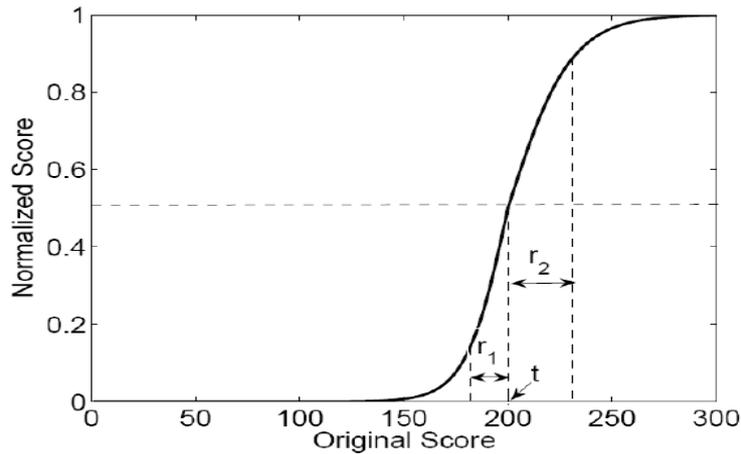


Figure 4.3 : Normalisation double sigmoïde.

Ce schéma transforme les scores dans l'intervalle  $[0,1]$ . Mais il demande de paramétrer prudemment  $t$ ,  $r_1$  et  $r_2$  afin d'obtenir une bonne efficacité. En général,  $t$  est choisie de façon à être une valeur appartenant à la région de recouvrement entre les distributions des scores client et imposteurs, et  $r_1$  et  $r_2$  sont fixés de telle sorte qu'ils correspondent respectivement aux extrémités de recouvrement entre les deux distributions vers la gauche et vers la droite de  $t$  (on peut également voir cela comme les valeurs correspondant respectivement au minimum des scores de similarité des clients et au maximum des scores de similarité des imposteurs). Ce schéma de normalisation fournit donc une transformation linéaire des scores dans la région de recouvrement, tandis que les scores à l'extérieur de cette région sont transformés de manière non linéaire.

### 4.3 Approche par combinaison de scores

#### 4.3.1 Méthode de combinaisons simples

Les méthodes de combinaisons de scores simples sont des méthodes très simples dont l'objectif est d'obtenir un score final  $s$  à partir des  $N$  scores disponibles  $s_i$  pour  $i = 1$  à  $N$  issus de  $N$  systèmes.

Les méthodes les plus utilisées sont la moyenne, le produit, le minimum, le maximum ou la médiane [10,113].

- Combiner les scores par la moyenne consiste à calculer  $s$  tel que

$$s = \frac{1}{N} \sum_{i=1}^N s_i \quad (4.7)$$

- Combiner les scores par le produit consiste à calculer  $s$  tel que

$$s = \prod_{i=1}^N s_i$$

- Combiner les scores par le minimum consiste à calculer  $s$  tel que

$$s = \min (s_i) \quad (4.9)$$

- Combiner les scores par le maximum consiste à calculer  $s$  tel que

$$s = \max (s_i) \quad (4.10)$$

- Combiner les scores par la médiane consiste à calculer  $s$  tel que

$$s = \text{med} (s_i) \quad (4.11)$$

- La somme pondérée c'est une méthode un peu plus évoluée qui nécessite une adaptation par le réglage de paramètres.

$$s = \sum_{i=1}^N \omega_i s_i$$

La somme pondérée permet de donner des poids différents  $\omega_i$  à chacun des systèmes en fonction de leur performance individuelle ou de leur intérêt dans le système multi-algorithmes.

### 4.3.2 Combinaison de scores par logique floue

La théorie de la **logique floue** (des sous-ensembles flous) a été introduite par Zadeh en 1965 [130] comme une extension de la logique binaire d'une part et une amélioration de la logique multivaluée (admettant plusieurs valeurs de vérité) d'autre part. L'importance de la logique floue réside dans le fait qu'elle s'approche du **raisonnement humain** par l'intégration et le traitement du caractère approximatif, vague, imprécis ou flou de la connaissance humaine. Les termes linguistiques tels que « **environ** », « **moyenne** », « **approximativement** » sont de nature à donner un caractère flou aux phrases énoncées. Par exemple, la règle « si le prix est inférieur à 6000 DA, j'achète » sera intuitivement utilisable si le prix est de 6002 DA, mais elle ne pourrait être exploitée en logique classique puisque le prix indiqué ne satisferait pas la prémisse.

Dans [110] on réalise l'**intégration floue de Choquet** pour deux approches de fusion :

- La **première approche** est de combiner des images des traits du visage (yeux, nez et la bouche) et images globales de visage.
- La **deuxième approche** est à l'agrégation des classificateurs opérant sur quatre ensembles de sous-image générée par décomposition en ondelettes.

### 1 Mesure floue

Un jeu de fonction  $g : P(Y) \rightarrow [0, 1]$  est appelée une **mesure floue** si les conditions suivantes sont remplies :

1. conditions aux limites:  $g(\emptyset) = 0, g(Y) = 1$
2. monotonie :  $g(A) \leq g(B)$ , si  $A \subset B$  et  $A, B \in P(Y)$
3. continuité :  $\lim_{i \rightarrow \infty} g(A_i) = g(\lim_{i \rightarrow \infty} A_i)$ , si  $\{A_i\}_i^\infty$  est une suite croissante d'ensembles mesurables.

A partir de cette définition, Sugeno [131] a introduit un soi-disant  $g_\lambda$  **mesure floue** qui est livré avec une propriété supplémentaire

$$g(A \cup B) = g(A) + g(B) + \lambda g(A)g(B) \quad (4.13)$$

Pour tous les  $A, B \subset Y$  et  $A \cap B = \emptyset$ , et pour certains  $\lambda > -1$ .

Évidemment quand  $\lambda = 0$ , le  $g_\lambda$  mesure floue devient une mesure de probabilité standard.

En général, la valeur de  $\lambda$  peut être déterminée en raison de l'état limite de la mesure floue  $g_\lambda$ . Cette condition pour  $g(Y) = 1$ . Par conséquent, la valeur de  $\lambda$  est déterminée par la résolution de ce qui suit:

$$g_\lambda(Y) = \frac{1}{\lambda} \left( \prod_{i=1}^n (1 + \lambda g^i) - 1 \right), \lambda \neq 0 \quad (4.14)$$

est l'équivalent de :

$$\lambda + 1 = \prod_{i=1}^n (1 + \lambda g^i) \quad (4.15)$$

Où  $\lambda \in (-1, +\infty)$ ,  $\lambda \neq 0$ , et  $g^i$  est la valeur de **la fonction de densité floue**. La solution peut être facilement obtenue; évidemment on s'intéresse à la racine supérieure  $a-1$  [132].

### 2 Intégrale floue

L'intégrale floue de la fonction  $h$  calculée sur  $Y$  par rapport à une mesure floue  $g$  est définie sous la forme

$$\int_Y h(y) \circ g(\cdot) = \sup_{\alpha \in [0,1]} [\min[\alpha, g(\{y | h(y) \geq \alpha\})]] \quad (4.16)$$

**Intégrale floue de Sugeno**[110,111,114,133]

Lorsque les valeurs de  $h(\cdot)$  sont classés dans l'ordre décroissant,

$$h(y_1) \geq h(y_2) \geq \dots \geq h(y_n)$$

L'intégrale floue de Sugeno est calculé comme suit:

$$\int_Y h(y) \circ g(\cdot) = \max_{i=1:n} [\min(h(y_i), g(A_i))] \quad (4.17)$$

Où  $A_i = \{y_1, y_2, \dots, y_i\}$  désigne un sous-ensemble d'éléments. Les valeurs de  $g(A_i)$  pris en charge par la mesure floue sur les sous-ensembles correspondant d'éléments peut être déterminée de manière récursive sous la forme :

$$g(A_1) = g(y_1) = g^1 \quad (4.18)$$

$$g(A_i) = g^i + g(A_{i-1}) + \lambda g^i g(A_{i-1}), \quad (4.19)$$

Le calcul de la fonction de densité floue  $g^i$  sur la base des données sont assurées par la manière suivante

$$g^i = \beta p_i, \quad \begin{cases} i = 1 \\ i = 2, 3, 4 \end{cases} \quad (4.20)$$

Où  $p_i$  est le taux de classification dans l'intervalle  $[0, 1]$  pour chaque système.

$B \in [0, 1]$  est un facteur qui met en place un certain équilibre entre les résultats de la classification.

**Intégrante floue de Choquet**[110,111,114,133]

Il a été démontré que (4.17) n'est pas une extension correcte de l'intégration de **Lebesgue** habituelle. En d'autres termes, lorsque la mesure est additive l'expression ci-dessus ne retourne pas l'intégrale au sens de **Lebesgue**. Afin de remédier à cet inconvénient, Murofushi et Sugeno [134] ont proposés un soi-disant intégrante floue de Choquet calculé de la manière suivante:

$$\int_Y h(y) d g(\cdot) = \sum_{i=1}^n [h(y_i) - h(y_{i+1})] g(A_i) h(y_{n+1}) = 0 \quad (4.21)$$

**4.4 Approche par classification de scores**

L'approche par classification de scores [135] cherche à prendre une décision à partir du vecteur de scores de dimension  $N$ . Le but des méthodes de fusion basées sur des classificateurs de chercher une séparatrice entre les 2 classes, Client et Imposteur, dans l'espace à  $N$  dimensions des scores. La séparatrice peut être plus ou moins complexe selon le classificateur utilisé. Toutes les méthodes classiques de reconnaissance des formes et d'apprentissage automatique peuvent être utilisées pour la fusion de scores qui est un problème de classification comme un autre.

Dans le domaine des systèmes multi-algorithmes, dont l'objectif est d'utiliser un classificateur pour combiner plusieurs systèmes sources (ici ce sont des systèmes biométriques que l'on veut fusionner) afin de consolider la décision de classification, un grand nombre de méthodes ont été utilisées. Les méthodes utilisées dans notre travail sont :

- la méthode **statistique de Fisher**,
- les machines à vecteurs de support (**SVM** : Support Vector Machine)
- les réseaux de neurones **RN**.

### 4.4.1 Fusion par méthode statistique de Fisher

La méthode **statistique** introduite ici, se base sur les travaux de **Fischer** [68] et fait usage d'une frontière de décision linéaire pour séparer deux populations données, à savoir les clients et les imposteurs dans notre cas. Dans le contexte particulier de l'authentification de personnes, Duc et al [136] utilisent eux aussi une règle de décision linéaire, mais l'optimisent en se fixant un critère dérivé de la théorie de Bayes [3].

Envisageons à présent la règle de décision développée par Fisher. Elle se base sur le rapport de vraisemblance repris ci-dessous:

$$\frac{T(z|c)}{T(z|i)} > k \quad (4.22)$$

Où  $k$  représente un seuil d'acceptation dont la valeur dépend du compromis **TFA/TFR** que l'on veut obtenir. Dans le problème qui nous préoccupe,  $T(z|c)$  et  $T(z|i)$  sont inconnus et doivent être estimés à partir des données d'apprentissage. Une hypothèse courante consiste à approcher les distributions réelles par des distributions normales à  $p$  variables  $N_p(\mu_A, \Sigma)$ , où  $A = \{c, i\}$  représente la classe d'individus,  $\mu_A$  le vecteur des scores moyens et  $\Sigma$  la matrice de covariance entre experts. En un premier temps, on supposera la matrice  $\Sigma$  indépendante de la classe d'individus. Sous de telles hypothèses, les fonctions de densité de probabilité s'écrivent sous la forme:

$$f_A(z) = (2\pi)^{-p/2} |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} (z - \mu_A)' \Sigma^{-1} (z - \mu_A) \right\}$$

Les paramètres  $\mu_c$ ,  $\mu_i$  et  $\Sigma$  sont inconnus, mais peuvent être estimés à partir des données d'apprentissage, soit  $x$  les  $n_c$  données relatives aux accès clientset  $y$ , les  $n_i$  données relatives aux accès imposteurs (simulés). On a :

$$\hat{\mu}_c = \sum_{q=1}^{n_x} x_q / n_c$$

$$(4.2) \hat{\mu}_i = \sum_{q=1}^{n_u} y_q / n_i$$

$$(4.2) \hat{\Sigma}_c = \sum_{q=1}^{n_x} (x_q - \hat{\mu}_c)(x_q - \hat{\mu}_c)' / (n_c - 1)$$

$$(4.2) \hat{\Sigma}_i = \sum_{q=1}^{n_y} (y_q - \hat{\mu}_i)(y_q - \hat{\mu}_i)' / (n_i - 1)$$

$$(4.2) \hat{\Sigma} = [(n_c - 1)\hat{\Sigma}_c + (n_i - 1)\hat{\Sigma}_i] / (n_c + n_i - 2)$$

Notons que l'on tient compte ici, par l'intermédiaire de  $\Sigma$ , de la dépendance qui peut exister entre experts.

En combinant les équations (4.22) à (4.28), on peut réécrire  $\ln(\hat{f}_c(z)/\hat{f}_i(z)) \geq k \Rightarrow D_L(z) \geq \ln(k) = k^*$  où

$$(4.2) D_L(z) = (z - \frac{1}{2}(\hat{\mu}_c + \hat{\mu}_i))' \hat{\Sigma}^{-1} (\hat{\mu}_c - \hat{\mu}_i)$$

Fisher fut le premier à utiliser cette fonction à des fins de classification.

Comme  $D_L(z)$  est linéaire en  $z$ , elle fut communément appelée **fonction linéaire discriminante** ou **LDF** (Linear Discriminant Function). Ainsi, la procédure à suivre pour vérifier l'identité d'un candidat, consiste à calculer à l'aide de  $\hat{\mu}_c$ ,  $\hat{\mu}_i$  et  $\hat{\Sigma}$  données d'entraînement (ce qui est effectué une fois pour toutes), puis  $D_L(z)$  et comparer celui-ci au seuil  $k^*$  donné. Si  $D_L(z) \geq k^*$ , le candidat est accepté comme client.

Dans le cas de distributions de scores clients et imposteurs ne satisfaisant pas l'hypothèse d'une covariance  $\Sigma$  unique, la règle de décision (4.29) peut se réécrire sous la forme  $D_Q(z) \geq 2k^*$  où

$$(4.30) \quad \begin{aligned} D_Q(z) = & (z - \hat{\mu}_i)' \hat{\Sigma}_i^{-1} (z - \hat{\mu}_i) \\ & - (z - \hat{\mu}_c)' \hat{\Sigma}_c^{-1} (z - \hat{\mu}_c) \\ & + \ln(|\hat{\Sigma}_i|/|\hat{\Sigma}_c|) \end{aligned}$$

$D_Q(z)$  est appelée **fonction quadratique discriminante** ou **QDF** (Quadratic Discriminant Function). Cette règle permet d'obtenir de meilleurs résultats que ceux obtenus dans le cadre d'une fusion linéaire, mais requiert un ensemble d'apprentissage suffisamment étendu pour pouvoir évaluer  $\Sigma_c$  et  $\Sigma_i$  avec précision [117].

### 4.4.2 Les machines à vecteurs de support (SVM)

Les machines à vecteurs de support (Support Vector Machine : SVM), également appelés **Séparateurs à Vaste Marge**, introduits par **Vapnik** [137] est utilisée dans la fusion de données dans les références [10, 85, 129, 138, 139, 140, 141]. Ils ont pour but de définir un hyperplan de séparation entre les deux classes, hyperplan qui minimise l'erreur de classification sur un ensemble d'apprentissage. Cette méthode d'apprentissage supervisée peut apprendre une séparatrice plus ou moins complexe selon la nature du noyau choisi. Le noyau le plus simple est le noyau linéaire qui correspond à chercher une séparatrice linéaire dans l'espace à  $N$  dimensions des scores. Le but des fonctions noyau est de transformer l'espace initial (des scores à  $N$  dimensions) en un espace de dimension **plus grande** dans lequel les données pourraient être linéairement séparables. La séparatrice est donc toujours linéaire dans l'espace transformé par la fonction noyau, mais n'est plus dans l'espace des scores.

Le but du SVM est donc de trouver une séparatrice qui minimise l'erreur de classification sur l'ensemble d'apprentissage mais qui sera également performante en généralisation sur des données non utilisées en apprentissage. Pour cela le concept utilisé est celui de marge (d'où le nom de séparateurs à vaste marge). La marge est la distance quadratique moyenne entre la séparatrice et les éléments d'apprentissage les plus proches de celle-ci appelés vecteurs de support (figure 4.4). Ces éléments sont appelés vecteurs de support car c'est uniquement sur ces éléments de l'ensemble d'apprentissage qu'est optimisée la séparatrice.

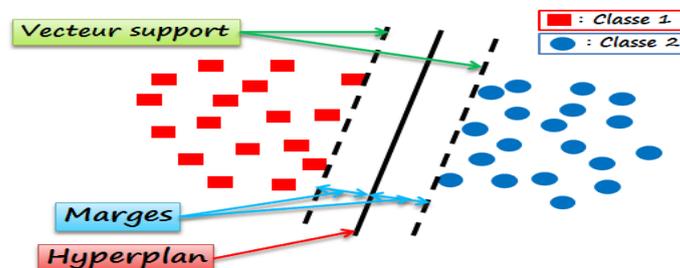


Figure 4.4 : Séparation linéaire dans un espace à deux dimensions.

Tout classifieur a pour but de classifier un élément  $x$ , ici  $x = (s_1, \dots, s_N)$  est un vecteur de scores de dimension  $N$ , dans l'une des classes possibles. Dans notre problème il y a deux

classes, Client ou Imposteur, dont l'étiquette sera notée avec  $y = -1$  et  $1$ ,  $-1$  correspondant à la classe des Imposteurs et  $1$  à la classe des Clients. Le classifieur a donc pour but de déterminer  $f$  telle que :

$$y = f(x) \quad (4.31)$$

Le **SVM** a pour but de trouver la meilleure séparatrice linéaire (en terme de marge maximale, c'est à dire la meilleure généralisation) dans l'espace transformé par la fonction de noyau  $K$ , c'est à dire de déterminer le vecteur  $w$  et la constante  $b$  tels que la séparatrice ait pour équation :

$$w \cdot k(x) + b = 0 \quad (4.32)$$

La distance entre un point de l'espace  $x_i$  et l'hyperplan d'équation  $w \cdot K(x) + b = 0$  est égale à :

$$(4.33) \quad h(x_i) = \frac{w \cdot K(x_i) + b}{\|w\|}$$

Pour maximiser la marge, il faut donc minimiser  $\|w\|$  en maximisant  $w \cdot K(x_i) + b$  pour les  $x_i$  définis comme vecteurs de support. Ces vecteurs de supports sont les  $x_i$  pour  $i = 1 : m$  de la base d'apprentissage tels que  $w \cdot K(x_i) + b = \pm 1$ .

La résolution de ce problème d'optimisation est faite par l'utilisation des multiplicateurs de **Lagrange** où le Lagrangien est donné par :

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^m \alpha_i (y_i (w \cdot K(x_i) + b) - 1)$$

Avec les coefficients  $\alpha_i$  appelés multiplicateurs de **Lagrange**. Pour résoudre ce problème d'optimisation, il faut minimiser le **Lagrangien** par rapport à  $w$  et  $b$  et le maximiser par rapport à  $\alpha$ .

Dans la pratique, il est souvent impossible de trouver une séparatrice linéaire (même dans l'espace transformé par la fonction noyau) car il y a toujours des erreurs de classification. Il a donc été introduit par **Vapnik** [137] la technique de marge souple. Ce principe de marge souple tolère les mauvais classements par l'introduction de variables ressorts  $\xi_i$  qui permettent de relâcher les contraintes sur les éléments d'apprentissage qui ne doivent plus être à une distance supérieure ou égale à 1 de la marge (l'égalité correspondant aux vecteurs de support), mais à une distance supérieure ou égale à  $1 - \xi_i$ , c'est à dire :

$$(4.35) \quad y_i (w \cdot K(x_i) + b) \geq 1 - \xi_i$$

Avec  $\xi_i \geq 0$  pour  $i = 1 : M$ ,  $M$  étant le nombre d'éléments de la base d'apprentissage.

Le problème d'optimisation est donc modifié et le **Lagrangien** devient :

$$(4.3) L(w, b, \alpha) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^M \xi_i - \sum_{i=1}^M \alpha_i (y_i (w \cdot K(x_i) + b) - 1)$$

Où  $C$  est une constante strictement positive qui permet de régler le compromis entre le nombre d'erreurs de classification et la largeur de la marge [10].

### 4.4.3. Réseaux de neurones artificiels

Le principe général des **Réseaux de Neurones Artificiels (RNA)** est à l'origine inspiré de certaines fonctions de base des neurones naturels du cerveau.

Un réseau de neurones artificiel est généralement organisé en **plusieurs couches** :

- une couche d'entrée,
- une couche de sortie,
- des couches intermédiaires appelées couches cachées.

La présence de couches cachées permet de **discriminer** des classes d'objets non linéairement séparables.

En général, un réseau de neurones est fondamentalement un classifieur, il réalise un travail de classification pendant la phase d'apprentissage, et de classement lors de la reconnaissance. Mais on peut l'utiliser pour réaliser la fusion de données [119, 142] pour séparer deux populations données, à savoir les clients et les imposteurs dans notre cas.

Les grands avantages des réseaux de neurones résident dans leur capacité **d'apprentissage automatique** (approximation universelle (**Cybenko, Hornik**)), ce qui permet de résoudre des problèmes sans nécessiter l'écriture de règles complexes, tout en étant tolérant aux erreurs. Ils résident aussi dans leur capacité à prendre une décision à partir de critères **non formalisables explicitement** [143].

### Le réseau perceptron multicouche MLP (**Multi-Layer perceptron**)

#### 1) Structure du réseau

Les réseaux monocouches ont rapidement montré des limites car ils ne permettaient de traiter ni le cas des classes non linéairement séparables, ni le cas multi-classes [144]. Le perceptron multicouche s'est imposé par la suite comme une réponse à ces problèmes, grâce à sa propriété d'approximateur universel [145], qui le rend capable d'approcher une fonction quelconque avec une précision arbitraire. Comme son nom l'indique, sa structure est organisée en couches (voir figure 4.5) :

## Chapitre 4 Les méthodes de fusion de scores

Chaque neurone élémentaire est connecté à tous les neurones de la couche qui suit celle à laquelle il appartient. Les fonctions d'activation peuvent être quelconques, et le nombre de neurones en sortie dépend généralement du codage adopté. Par exemple, dans un problème de discrimination, chacun d'eux est généralement dédié à une classe donnée.

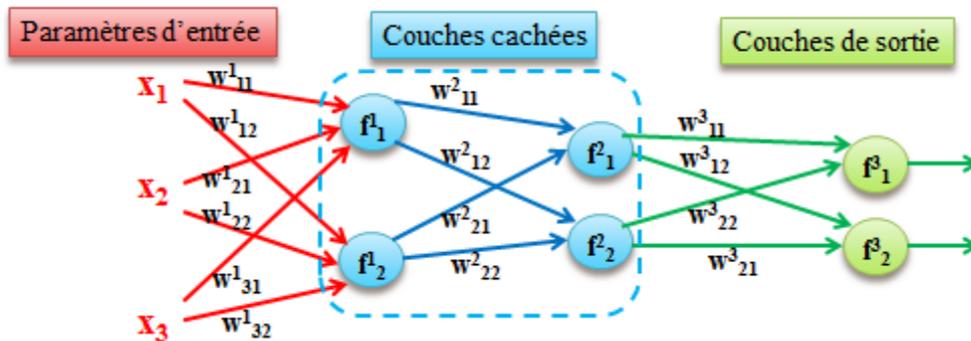


Figure 4.5 : Exemple de perceptron multicouches

La figure 4.5 illustre un exemple de perceptron multicouches avec 3 entrées, 2 couches cachées comportant 2 neurones chacune et une couche de sortie comportant 2 neurones.

### 2) Contrôle de la complexité du réseau

Le problème des méthodes d'apprentissage est qu'elles risquent de donner lieu à un réseau trop adapté à l'ensemble d'apprentissage. Dans ce cas les performances du réseau seront **excellentes sur l'ensemble d'apprentissage**, mais **médiocres sur des exemples inconnus**. Ceci peut arriver lorsque l'algorithme d'apprentissage n'est pas arrêté à temps.

Pour éviter ce phénomène de surapprentissage, une technique consiste à utiliser un ensemble de validation. Les données, qui doivent être en quantité suffisante, sont alors séparées en deux sous-ensembles :

- l'ensemble d'apprentissage
- l'ensemble de validation.

L'ensemble d'apprentissage est utilisé pour déterminer les poids par l'algorithme de rétropropagation du gradient.

L'ensemble de validation, lui, permet de vérifier l'aptitude du réseau sur des échantillons qu'il ne connaît pas. Si le risque empirique diminue sur l'ensemble d'apprentissage, alors qu'il augmente de manière significative sur l'ensemble de validation, on arrête alors le processus d'apprentissage [146].

### 3) Sélection de modèle

La mise en œuvre pratique des perceptrons multicouches soulève une autre question délicate : combien de couches et de neurones cachés utiliser ? Dans un problème de régression, cela dépend de la fonction que l'on cherche à approcher. Dans un problème de classification, c'est la complexité de la surface de séparation entre les classes qui importe. Plus le nombre de neurones est élevé, plus la taille de l'espace des fonctions atteignables est grande, et plus les surfaces de séparation sont complexes. Cependant, encore une fois, si le réseau a une structure trop complexe, il aura tendance à être trop bien adapté à l'ensemble d'apprentissage (cas de surparamétrisation), et il risque de donner de l'importance à des exemples marginaux. Au contraire, s'il ne l'est pas suffisamment, la modélisation sera trop sommaire. Dans les deux cas, les capacités de généralisation du réseau seront faibles, ce qui aura des conséquences négatives sur ses performances.

Deux approches permettent de déterminer une taille convenable pour les perceptrons multicouches. La première consiste à déterminer l'architecture optimale du réseau [146]:

- soit par une méthode constructive : on construit une estimée grossière à l'aide d'un réseau de petite taille, puis on ajoute successivement des cellules pour affiner le modèle [147],
- soit par élagage : on part d'un réseau surdimensionné et on annule les poids participant peu à la détermination de la sortie [148].

La seconde approche consiste à travailler à architecture fixée et surestimée, puis à pénaliser certains poids [149].

Pour plus de détails sur les réseaux de neurones avec d'exemples programmer par MATLAB dans [50].

### 4.5 Conclusions

Ce chapitre nous a permis de connaître un certain nombre de méthodes de fusion de scores. Nous avons présenté les deux approches essentielles de fusion des scores : la première est l'approche par combinaison. Cette dernière traite les scores séparément avant de les combiner et elle est basée sur les méthodes les plus utilisées telles que la moyenne, le produit, le minimum, le maximum ou la médiane.

Nous avons également montré qu'il y a différents niveaux de "complexité" pour les méthodes de combinaison.

La deuxième approche étudiée est basée sur la classification de scores qui cherche également à modéliser la séparatrice des classes dans l'espace multidimensionnel. Une phase

## **Chapitre 4** Les méthodes de fusion de scores

---

de normalisation préalable aux deux approches est réalisée c'est ce qui fait l'originalité de notre travail. Car dans les travaux antérieurs la normalisation se limite à la première approche seulement. Cependant, il y a toujours de nombreuses questions ouvertes qui sont laissées sans réponse.

Toutes ces méthodes ont des avantages et des inconvénients selon la complexité, le besoin de données d'apprentissage ou les paramètres à optimiser. Le choix d'une méthode va donc dépendre de l'application visée selon le compromis entre performance, simplicité, confort et quantité de données d'apprentissage disponible. Toutes ces méthodes seront comparées sur la base de données XM2VTS nous verrons cela dans le chapitre 5.

## **5 Conception du système multimodale**

### **5.1 Introduction**

Dans ce chapitre nous tentons de présenter les résultats expérimentaux de la multi-modalité qui sont divisés en deux catégories.

Nous nous intéressons tout d'abord aux résultats sur la fusion multi-algorithmes en exploitant deux approches de fusion en étudiant la normalisation de scores pour chaque cas. Dans la deuxième partie nous testons les deux approches de fusion de scores des visages et des voix de la base de données XM2VTS et ceci dans le but de construire un système multi-biométrique.

### **Application I**

### **5.2 La multi-modalité (multi-algorithmes)**

Le système multimodale étudié dans cette section est un système multi-algorithmes ou multi-classifieurs (**Multiple Classifier Systems**). Celui-ci est construit par les quatre systèmes retenus. En effet lorsque l'on cherche à fusionner plusieurs systèmes d'authentification de visages, comment choisit la méthode la plus appropriée ? Le choix de cette méthode dépend de plusieurs critères qui sont :

- les systèmes à combiner et surtout les distributions de leurs scores,
- la difficulté des données, c'est-à-dire la séparabilité des classes **Client** et **Imposteur**,
- la quantité de données disponibles pour l'apprentissage
- l'application visée.

Selon la nature des scores de comparaison issus de chaque système, une étape de normalisation des scores peut être nécessaire. Cette étape ne pose pas vraiment de problème à partir du moment où l'on connaît la loi des scores ou que l'on possède un échantillon d'apprentissage pour l'estimer.

La difficulté des données est associée à la possibilité de séparer les classes **Client** et **Imposteur**. Cette caractéristique est directement liée à la performance de reconnaissance. La difficulté des données doit être considérée séparément pour chaque système biométrique à combiner pour une connaissance préalable à l'étape de fusion des différentes performances

individuelles des systèmes. En effet, on ne combinera pas de la même façon des systèmes avec des performances à peu près équivalentes et des systèmes avec des performances très différentes, car dans ce dernier cas, on voudra que le "bon" système ait plus de poids dans la décision que le "mauvais". Ainsi, l'étude des performances individuelles peut nous aider à choisir la méthode de fusion adéquate. Mais c'est surtout la séparabilité des données multidimensionnelles (c'est-à-dire en considérant les vecteurs de scores) qui va nous donner une indication sur la complexité de la méthode de fusion à utiliser. En effet, des données très bien séparées dans l'espace des scores à  $N$  dimensions ne nécessiteront pas l'utilisation d'une méthode très complexe, une séparatrice linéaire sera suffisante. En revanche, dans le cas de données non séparables linéairement, on pourra essayer d'utiliser une méthode de fusion plus complexe avec une séparatrice ou des transformations de score non linéaires.

L'étude de la séparabilité des données ou l'utilisation de méthodes de fusion complexes nécessitent la disponibilité de données étiquetées pour l'apprentissage. Selon la quantité et la représentativité des données disponibles.

L'application visée a également une part importante dans le choix de la méthode de fusion. En revanche, les exigences de performance sont impactées par le choix de la méthode de fusion. En effet, selon les applications (de sécurité ou de confort), les performances attendues ne sont pas les mêmes et surtout les deux types d'erreurs de classification que sont le taux de fausse acceptation et le taux de faux rejet ne sont pas considérés de la même façon. Selon le point de fonctionnement visé, on ne choisira pas les mêmes méthodes de fusion ; par exemple, un classifieur qui optimise sa séparatrice en considérant de la même façon les deux types d'erreurs n'est donc peut être pas approprié pour une application de sécurité qui cherche à minimiser les fausses acceptations.

Cette section a donc pour but de comparer toutes les méthodes de normalisation des scores et toutes les méthodes de fusion de scores pour les deux approches combinaison et classification présentées au Chapitre 4 pour les quatre systèmes d'authentification de visages appliquée sur la base de données XM2VTS. La figure 5.1 présente notre système multi-algorithmes.

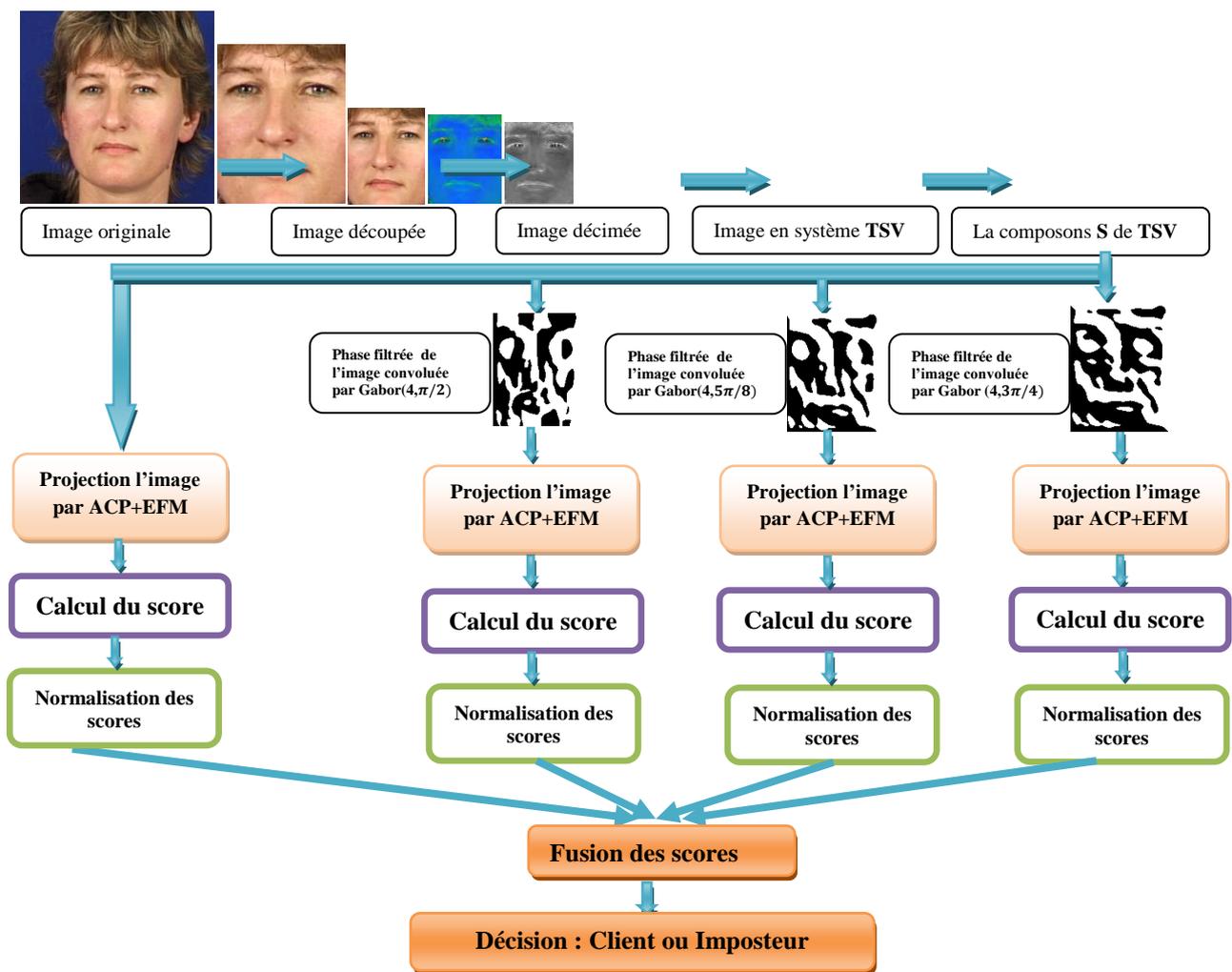


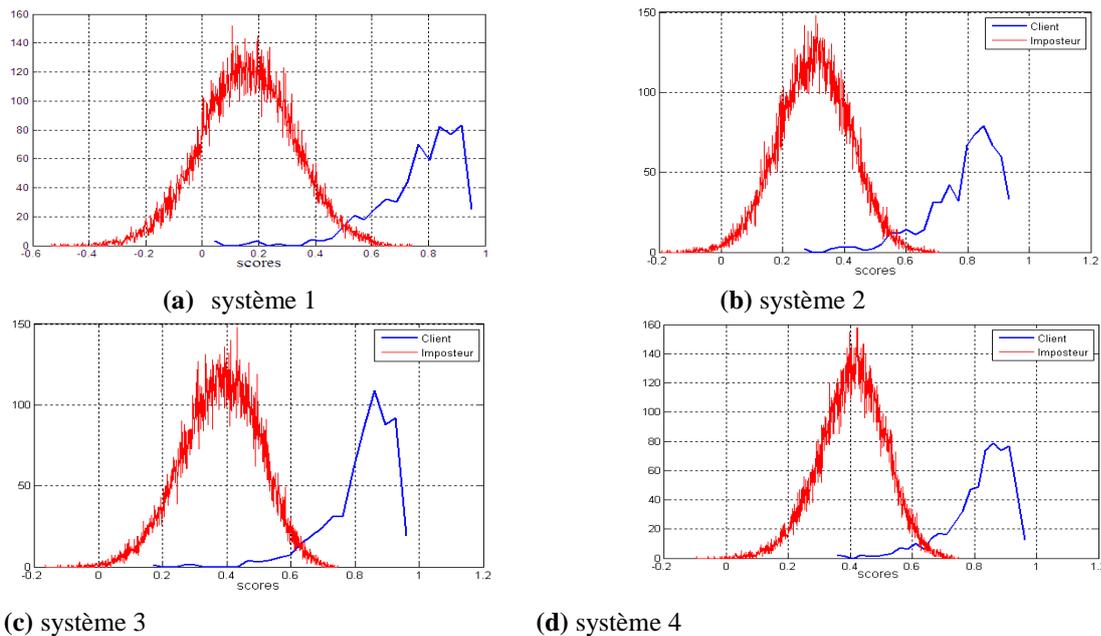
Figure 5.1 : Organigramme du Système de fusion de scores des quatre systèmes d'authentification de visages.

### 5.2.1. Les distributions et normalisation de scores

Les distributions de scores pour les quatre systèmes d'authentification de visages représentées sur la figure 5.2.

Nous remarquons que les quatre systèmes donnent des distributions Client et Imposteur différentes. Les distributions sont différentes en termes de plage de variation, ce qui rendra nécessaire l'étape de normalisation des scores détaillée au Chapitre 4. Elles sont également différentes dans leurs formes et leurs recouvrements entre les deux classes (Client et Imposteur). Même si toutes les distributions ont un seul mode (mais pas forcément symétrique).

Voir par exemple l'effet de la normalisation MinMax pour les quatre systèmes sur la figure 5.3, nous remarquons que toutes les distributions normalisées sont dans un intervalle conjoint et la variation des distributions se trouve entre 0 et 1.



**Figure 5.2 :** Les distributions de scores des quatre systèmes d’authentification de visages sur la base de données XM2VTS.

**Remarques :**

- **Nombre de scores dans l'ensemble d'évaluation :**
  - Pour les clients : Il est égale au Nombre des Clients \* Nombre des Poses de Clients dans l'ensemble d'évaluation, et c'est égale à  $200 * 3 = 600$ .
  - Pour les imposteurs : Il est égale au Nombre des Clients \* Nombre des Imposteurs dans l'ensemble d'évaluation \* Nombre de Poses des Imposteurs dans l'ensemble d'évaluation, et c'est égale à  $200 * 25 * 8 = 40000$ .
- **Nombre de scores dans l'ensemble de test:**
  - Pour les clients : Il est égale au Nombre des Clients \* Nombre des Poses des Clients dans l'ensemble de test, et c'est égale à  $200 * 2 = 400$ .
  - Pour les imposteurs : Il est égale au Nombre des Clients \* Nombre des Imposteurs dans l'ensemble de test \* Nombre des Poses des Imposteurs dans l'ensemble de test, et égale  $200 * 75 * 8 = 120000$ .

Nous pouvons dire que le nombre de scores Imposteurs est beaucoup plus important que le nombre de scores des Clients. C'est ce qui explique la différence de forme pour les deux distributions de Client et Imposteur.

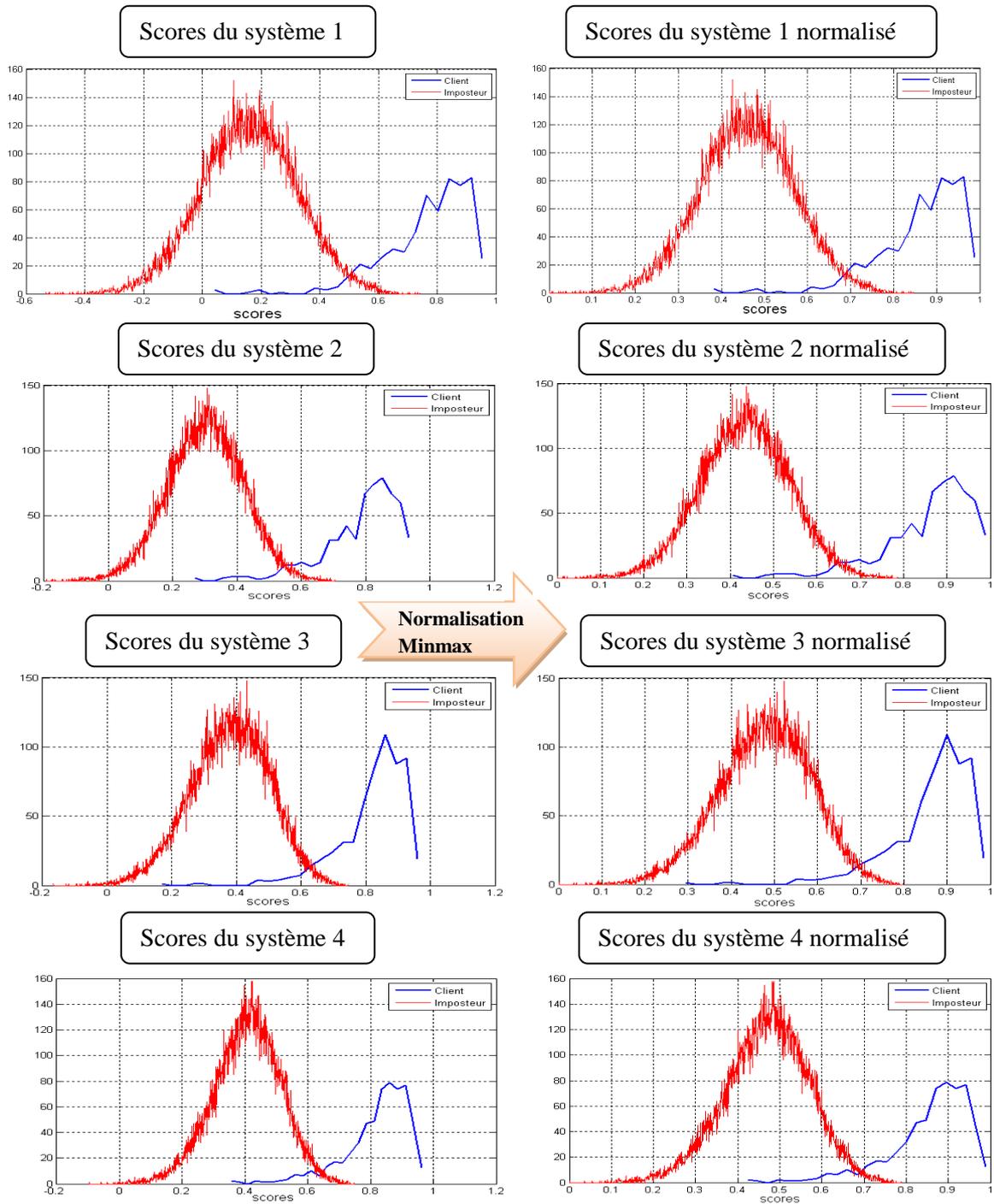
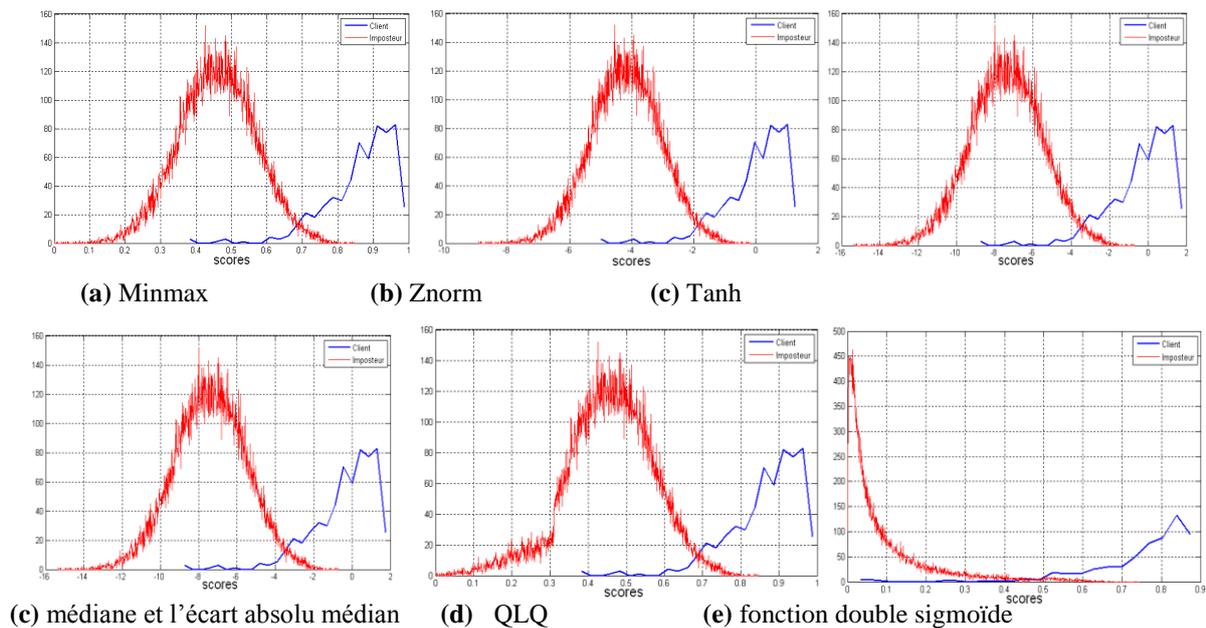


Figure 5.3 : Normalisation de scores par la méthode du Min-Max.

Dans la **figure 5.4** la transformation des scores du premier système est présentée pour les six normalisations.



**Figure 5.4 :** Normalisation de scores

Nous remarquons que les méthodes de normalisation **Min-Max**, **Znorm**, **Tanh**, la **médiane** et l'**écart absolu médian** ne modifient pas la forme des distributions mais leur différence se situe dans la façon dont chaque distribution sera répartie dans l'intervalle et donc dans la façon dont les systèmes à combiner se juxtaposeront dans l'intervalle défini (avec **Znorm** la moyenne de distribution des Clients de chaque système sera en 0. Alors qu'avec **Min-Max** c'est les extrémités qui sont définies et la moyenne sera positionnée différemment selon les systèmes. De plus, les méthodes de normalisation par fonction **Quadratique-Linéaire-Quadratique (QLQ)** et fonction double sigmoïde modifient la forme des distributions.

## 5.2.2 Comparaison des méthodes de normalisation

### 1 Méthodes de normalisation associées aux méthodes de combinaisons simples

Nous avons vu dans le paragraphe précédent que les scores ont des distributions et intervalles de variation différents et qu'il est donc nécessaire de les normaliser avant de les fusionner. Comme nous l'avons déjà cité, il existe deux types :

- des méthodes de normalisations de scores qui ne modifient pas la forme des distributions,
- les méthodes qui modifient la forme des distributions.

Ces méthodes de normalisation sont ensuite associées aux méthodes de combinaison de scores simples que sont la somme, le produit, le minimum, le maximum, la médiane et la somme pondérée. En effet, pour réaliser la fusion il faut combiner les quatre scores normalisés pour obtenir un score final qui permet de calculer les taux d'erreurs. Les performances de ces différentes associations de méthodes de normalisation avec les méthodes de combinaison sont présentées dans les tableaux 5.1 et 5.2.

Méthodes de normalisation	Taux de performance	Règles de Fusion					
		moyenne	produit	minimum	maximum	médiane	Somme pondérée
Min-Max	TEE	1.95±0.62	2.03±0.63	3.11±0.77	2.04±0.63	2.18±0.65	<b>1.95±0.62</b>
	TFR	0.5±0.69	0.5±0.69	0.5±0.69	1±0.97	0.5±0.69	<b>0.5±0.69</b>
	TFA	2.8±0.1	2.99±0.1	4.49±0.12	2.93±0.1	15.36±0.22	<b>2.8±0.1</b>
	TR	96.7±0.79	96.51±0.79	95.01±0.82	96.07±1.08	84.14±0.91	<b>96.7±0.79</b>
Znorm	TEE	1.95±0.62	2.03±0.63	3±0.76	2.03±0.63	2.21±0.66	<b>1.95±0.62</b>
	TFR	0.5±0.69	0.5±0.69	1±0.97	0.5±0.69	0.5±0.69	<b>0.5±0.69</b>
	TFA	2.79±0.1	3.01±1	4.31±0.12	2.92±0.1	15.12±0.22	<b>2.77±0.1</b>
	TR	96.71±0.79	96.49±0.79	93.69±1.1	96.58±0.79	84.38±0.91	<b>96.73±0.79</b>
Tangente Hyperbolique "Tanh"	TEE	1.95±0.62	1.96±0.62	3±0.76	2.03±0.63	2.21±0.66	<b>1.95±0.62</b>
	TFR	0.5±0.69	0.5±0.69	1±0.97	0.5±0.69	0.5±0.69	<b>0.5±0.69</b>
	TFA	2.79±0.1	2.82±0.1	4.31±0.12	2.92±0.1	15.12±0.22	<b>2.77±0.1</b>
	TR	96.71±0.79	96.68±0.79	95.59±1.1	96.58±0.79	84.38±0.91	<b>96.73±0.77</b>
La médiane et l'écart absolu médian (MAD)	TEE	2±0.62	2.04±0.63	3.16±0.78	1.99±0.62	2.31±0.67	<b>1.95±0.62</b>
	TFR	0.5±0.69	0.5±0.69	1±0.97	0.5±0.69	0.5±0.69	<b>0.5±0.69</b>
	TFA	2.92±0.1	3.04±0.1	4.41±0.12	2.78±0.1	16.26±0.22	<b>2.8±0.1</b>
	TR	96.58±0.79	96.46±0.79	94.59±1.1	96.72±0.79	83.24±0.91	<b>96.7±0.19</b>

**Tableau 5.1:** Performances des méthodes de normalisation qui ne modifient pas la forme des distributions associées aux méthodes de combinaison de scores simples (intervalle de confiance paramétrique à 95%).

Dans le **tableau 5.1** nous constatons qu'il n'est pas de grandes différences de performances selon les normalisations et les méthodes de combinaison. En particulier la méthode médiane qui donne mauvais résultat pour les données de test.

La méthode somme donne de bons résultats pour les quatre normalisations et fait toujours mieux que le meilleur système. La somme pondérée est la meilleure méthode de combinaison. Elle consiste en la définition de poids associés à chaque système, poids calculés de façon à minimiser le taux d'erreur sur la base d'évaluation. Nous constatons que la somme pondérée améliore les performances comparée à la deuxième meilleure méthode de combinaison qui est la somme simple (avec des poids identiques pour toutes les systèmes, ici ¼ sauf pour la normalisation Min-Max pour laquelle les performances sont équivalentes).

Méthodes de normalisation	Taux de performance	Règles de Fusion					
		moyenne	produit	minimum	maximum	médiane	Somme pondérée
Fonction Quadratique-Linéaire-Quadratique (QLQ)	TEE	2.18±0.65	2.68±0.72	3.21±0.79	2.95±0.75	2.13±0.64	2.13±0.64
	TFR	0.5±0.69	0.5±0.69	0.5±0.69	1.5±1.29	0.5±0.69	0.5±0.69
	TFA	3.05±0.1	3.88±0.12	4.49±0.12	3.21±0.09	25.27±0.27	2.9±0.1
	TR	96.45±0.79	95.62±0.81	95.01±0.81	95.29±1.3	74.23±0.96	96.6±0.79
Fonction Double Sigmoide	TEE	2.16±0.65	2.33±0.67	2.81±0.74	1.96±0.62	2.15±0.65	2.17±0.65
	TFR	0.5±0.69	0.5±0.69	1±0.97	0.5±0.69	0.5±0.69	0.5±0.69
	TFA	3.15±0.1	3.5±0.11	4.07±0.13	2.74±0.1	14.03±0.21	3.14±0.1
	TR	96.35±0.79	96±0.8	94.93±1.1	96.76±0.79	85.47±0.9	96.36±0.79

**Tableau 5.2:** Performances des méthodes de normalisation qui modifient la forme des distributions associées aux méthodes de combinaison de scores simples (intervalle de confiance paramétrique à 95%).

Dans le **tableau 5.2** nous constatons que les méthodes de normalisation modifiant la forme des distributions n'améliorent pas la performance de vérification associées aux méthodes de combinaison par rapport aux méthodes de normalisation qui ne modifient pas la forme des distributions. Mais nous aspirons que ces méthodes donnent de bons résultats par la combinaison de scores par la logique floue.

Enfin nous pouvons affirmer que la somme et la somme pondérée sont les meilleures méthodes de combinaison simples. La somme pondérée étant légèrement meilleure mais nécessitant un apprentissage de poids.

## 2 Comparaison des méthodes de normalisation par logique floue

Dans ce paragraphe, nous avons sélectionné trois méthodes de normalisation de score :

- une méthode qui ne modifie pas la forme des distributions : c'est la **Znorm** (ayant donné le meilleur résultat par rapport Min-Max, la médiane et l'écart absolu médian et c'est une méthode simple par rapport à tangente hyperbolique),
- deux méthodes modifiant la forme des distributions : qui sont la fonction Quadratique-Linéaire-Quadratique (**QLQ**), fonction double sigmoïde.

Pour comparer les méthodes de combinaisons par logique floue à base de l'intégrale de **Sugeno** et de **Choquet**.

L'algorithme suivant présente comment on fait la combinaison par les deux intégrales :

1. Calcul la fonction de densité floue  $g^i$

$$g^i = \beta p_i, \begin{cases} i = 1 \\ i = 2,3,4 \end{cases} \quad (5.1)$$

Avec:

$p_i$  est le taux de classification dans l'intervalle  $[0, 1]$  pour chaque système.

$B \in [0, 1]$  est un facteur qui met en place un certain équilibre entre les résultats de la classification.

$i$  : l'indice de chaque système.

2. Calcul de  $\lambda$  par :  $\lambda + 1 = \prod_{i=1}^n (1 + \lambda g^i)$  (5.2)

Où  $\lambda \in (-1, +\infty)$ ,  $\lambda \neq 0$ .

3. Calcul de  $g(A_i)$  la mesure floue sur les sous-ensembles par :

$$g(A_1) = g(y_1) = g^1 \quad (5.3)$$

$$g(A_i) = g^i + g(A_{i-1}) + \lambda g^i g(A_{i-1}), \quad (5.4)$$

4. Calcul de l'intégrale floue de **Sugeno** par :

$$\int_Y h(y) \circ g(\cdot) = \max_{i=1:n} [\min(h(y_i), g(A_i))] \quad (5.5)$$

$h(y_i)$  sont les scores et sont classés dans l'ordre décroissant,  $n = 4$ .

Ou calcule l'intégrale floue de **Choquet** par :

$$\int_Y h(y) d g(\cdot) = \sum_{i=1}^n [h(y_i) - h(y_{i+1})] g(A_i) h(y_{n+1}) = 0 \quad (4.21)$$

Le **tableau 5.3** présente la comparaison des méthodes de normalisation associées aux méthodes de combinaisons par logique floue.

Dans le **tableau 5.3** nous remarquons que la méthode de normalisation par la fonction Quadratique-Linéaire-Quadratique (**QLQ**) est intéressante pour les deux intégrales **Sugeno** et **Choquet**. Nous obtenons une bonne performance par la fonction **double sigmoïde** associée au l'intégrale de **Sugeno**.

Nous pouvons dire que les méthodes de normalisation de score qui modifient la forme des distributions (La fonction double sigmoïde) nous donne le meilleur résultat pour l'intégrale floue de **Sugeno** avec un **TEE=1.14±0.47%** et **TR=98.36 ± 0.75%**.

Méthodes de normalisation	Taux de performance	Règles de Fusion			
		Intégrale flouede Sugeno	B	Intégrale floue de Choquet	B
Znorm	TEE	2.08 ± 0.64	0.3	2.16 ± 0.65	0.1
	TFR	0.5 ± 0.69		0.5 ± 0.69	
	TFA	3.22 ± 0.11		3.19 ± 0.1	
	TR	96.28 ± 0.8		96.31 ± 0.79	
Fonction Quadratique-Linéaire-Quadratique (QLQ)	TEE	1.22 ± 0.49	0.523	1.89 ± 0.61	0.13
	TFR	0.5 ± 0.69		0.5 ± 0.69	
	TFA	1.45 ± 0.07		2.42 ± 0.09	
	TR	98.05 ± 0.76		97.08 ± 0.78	
Fonction Double Sigmoide	TEE	<b>1.14 ± 0.47</b>	0.509	2.4 ± 0.68	0.12
	TFR	<b>0.5 ± 0.69</b>		0.5 ± 0.69	
	TFA	<b>1.14 ± 0.06</b>		3.82 ± 0.11	
	TR	<b>98.36 ± 0.75</b>		95.68 ± 0.81	

Tableau 5.3 : Performances des méthodes de normalisation par logique floue( intervalle de confiance paramétrique à 95%).

### 5.2.3 Méthodes de normalisation associées aux méthodes de classifications

Dans cette section nous étudions la fusion de scores par l’approche de classification de scores (voir paragraphe 4.4).

Normalement les méthodes de fusion par classification de scores cherchent à prendre une décision à partir du vecteur de scores de dimension  $N$  sans normalisation de scores. Dans notre cas nous proposons d’utiliser les méthodes de normalisation de scores pour transformer individuellement chacun des scores issus des systèmes afin de les rendre homogènes avant de fusionner et ceci pour augmenter la performance d’authentification. Pour cela nous utilisons trois méthodes de normalisation: Z-Score, fonction Quadratique-Linéaire-Quadratique (QLQ) et normalisation par une fonction double sigmoïde (sont sélectionnées dans le paragraphe 5.2.2).

#### 1 Classification de scores par méthode statistique de Fisher

Les différents taux d’erreur et de succès dans l’ensemble d’évaluation et de test en utilisant la fusion par fonction linéaire discriminante (LDF) et fonction quadratique

discriminante (**QDF**)(voir paragraphe 4.4.1) sans et avec les méthodes de normalisation sont présentés dans le **tableau 5.4**.

Méthodes de normalisation	Taux d'erreur	Règles de Fusion	
		LDF	QDF
sans normalisation	TEE	2.14 ± 0.64	2.15 ± 0.65
	TFR	0.5 ± 0.69	0.5 ± 0.69
	TFA	3.18 ± 0.1	3.12 ± 0.1
	TR	96.32 ± 0.79	96.38 ± 0.79
Znorm	TEE	1.98 ± 0.62	2 ± 0.62
	TFR	0.5 ± 0.69	0.5 ± 0.69
	TFA	2.84 ± 0.1	2.89 ± 0.1
	TR	96.66 ± 0.79	96.61 ± 0.79
fonction quadratique-linéaire-quadratique(QLQ)	TEE	<b>1.97 ± 0.62</b>	2.2 ± 0.68
	TFR	<b>0.5 ± 0.69</b>	0.5 ± 0.69
	TFA	<b>2.68 ± 0.1</b>	3.06 ± 0.1
	TR	<b>96.82 ± 0.79</b>	96.44 ± 0.79
La une fonction double sigmoïde	TEE	2.17 ± 0.65	2.18 ± 0.65
	TFR	0.5 ± 0.69	0.5 ± 0.69
	TFA	3.15 ± 0.1	3.19 ± 0.1
	TR	96.35 ± 0.79	96.31 ± 0.79

**Tableau 5.4 :** Performances des méthodes de normalisation par fusion statistique de Fisher dans d'ensemble évaluation et test( intervalle de confiance paramétrique à 95%).

Le **tableau 5.4** montre clairement que les méthodes de normalisation améliorent la performance de l'authentification, et on constate qu'il n'y'a pas de grandes différences de **TEE** pour les deux fonctions de fusion linéaire discriminante (**LDF**) et quadratique discriminante (**QDF**). Le meilleur résultat est obtenu avec un **TEE = 1.97± 0.62%** et **TR= 96.82±0.79%** par la méthode de normalisation de score la fonction Quadratique-Linéaire-Quadratique(**QLQ**) qui modifiela forme des distributions et la fonction de fusion **LDF**.

## 2 Classification de scores avec les machines à vecteurs de support (SVM)

Nous utilisonsune machine à vecteurs de support (**SVM**) avec un noyau **RBF** (Radial BasisFunction ou fonction à base radiale) (voir équation4.32).

Le **SVM** a été implémenté en utilisant la librairie **LibSVM** disponible gratuitementau site (<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>).Le noyau **RBF** utilisé est de la forme :

$$K_{RBF}(u, v) = e^{-\gamma\|u-v\|^2} \quad (5.7)$$

Où  $\gamma$  est un paramètre qui règle la largeur de la marge. Le SVM nécessite également le réglage du paramètre C qui règle le compromis entre la largeur de la marge et le nombre d'erreurs admissibles (comme décrit dans équation 4.36).

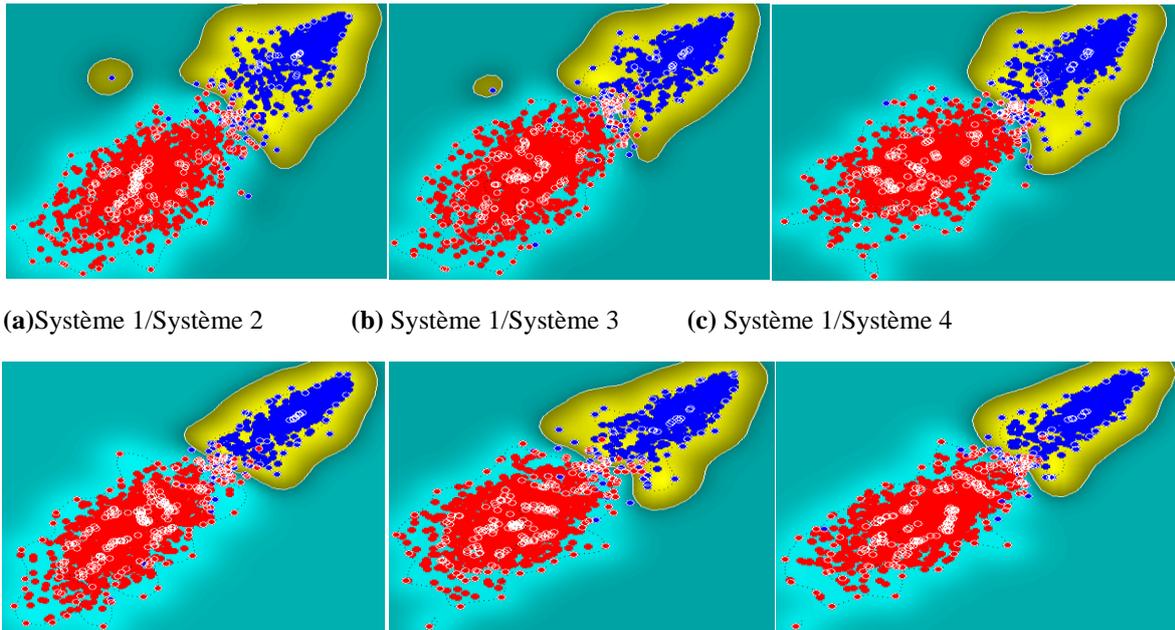
Les différents taux d'erreur et de succès dans l'ensemble d'évaluation et test en utilisant la fusion par une machine à vecteurs de support (SVM) avec un noyau RBF sans et avec les méthodes de normalisation sont montrés dans le **tableau 5.5**.

Règles de Fusion	Ensemble	Ensemble de test				$\gamma$	C
	d'évaluation	TEE %	TFR %	TFA %	TR %		
Sans normalisation	2±0.62	0.5±0.69	2.93±0.1	96.53±0.79	1.5	0.6	
	<b>1.5±0.54</b>	<b>0.5±0.69</b>	<b>2.06±0.08</b>	<b>97.44±0.77</b>	<b>1.5</b>	<b>0.55</b>	
Znorm	1.96±0.62	0.5±0.69	2.14±0.09	97.36±0.78	2	0.7	
Fonction Quadratique- Linéaire - Quadratique(QLO)	1.66±0.0.62	0.5±0.69	2.39±0.9	97.11±0.78	1.6	3	

**Tableau 5.5 :** Performances des méthodes de normalisation associées au SVM dans d'ensemble évaluation et test (intervalle de confiance paramétrique à 95%).

Le **tableau 5.5** montre clairement que les méthodes de normalisation améliorent la performance de l'authentification de visages. Notons également que le meilleur résultat avec un **TEE=1.5±0.54 %** et un **TR=97.44± 0.77%** est obtenu par la méthode de normalisation Znorm qui ne modifie pas la forme des distributions.

La **figure 5.5** présente la séparation de SVM de distributions 2D des scores des six couples des systèmes pour les deux classes Client et Imposteur car est impossible de représenter les distributions dans l'espace à 4 dimensions.



(d)Système 2/Système 3(e) Système 2/Système 4(f) Système 3/Système 4

**Figure 5.5** : la séparation non linéaire de SVM de distributions 2D des scores des six couples des systèmes pour les deux classes Client et Imposteur avec la normalisation de scores Z-norme

### 3 Classification de scores avec réseaux de neurones artificiels

Nous essayons finalement d'appliquer la fusion par classification de scores avec un réseaux de neurones simple de type **MLP** (**M**ulti **L**ayer **P**erceptron). Sa structure est organisée en couches (voir figure 4.5) :

- des paramètres d'entrée,
- une ou plusieurs couches cachées,
- une couche de sortie.

Chaque neurone élémentaire est connecté à tous les neurones de la couche qui suit celle à laquelle il appartient. Les fonctions d'activation peuvent être quelconques, et le nombre de neurones en sortie dépend généralement du codage adopté [146].

Les paramètres choisis par notre expérience sont :

- Quatre entrées sont les scores pour chaque système.
- Une couche cachée avec dix neurones et la fonction d'activation sigmoïde.
- Deux neurones dans la couche de sortie et la fonction d'activation tangente hyperbolique.

Les différents taux d'erreur et de succès dans deux ensembles évaluation et test en utilisant un classifieur **MLP** sans et avec les méthodes de normalisation se trouvent dans le **tableau 5.7**.

Règles de Fusion	Ensemble d'évaluation	Ensemble de test		
	TEE %	TFR %	TFA %	TR %
Sans normalisation	1.83±0.6	1±0.97	2.75±0.1	96.25±1.07
Znorm	1.69±0.57	1.5±1.19	1.81±0.08	96.69±1.27
Fonction Quadratique- Linéaire - Quadratique(QLQ)	1.66±0.57	0.5±0.69	1.93±0.08	<b>97.84±0.77</b>
Fonction double sigmoïde	<b>1.33±0.51</b>	0.5±0.69	2.01±0.08	97.49±0.77

**Tableau 5.7:** Performances des méthodes de normalisation associées à la classification MLP (intervalle de confiance paramétrique à 95%).

Le **tableau 5.7** montre bien là aussi que les méthodes de normalisation améliorent la performance de l'authentification. Nous constatons qu'il n'y'a pas de grandes différences de TEE selon les méthodes de normalisation. Les méthodes de normalisation qui modifient la forme des distributions donnent de bons résultats avec **TEE=1.66±0.57%** et **TR = 97.84±0.77%** par fonction Quadratique-Linéaire -Quadratique(QLQ) et **TEE= 1.33±0.51 %** et **TR = 97.49±0.77%** par la fonction double sigmoïde.

Nous pouvons dire que la méthode proposée qui utilise la normalisation de scores avant la classification de scores améliore en générale la performance d'authentification de visages. Dans la majorité des cas les méthodes de normalisation qui modifient la forme des distributions donnent de meilleurs résultats que les méthodes qui ne modifient pas la forme des distributions. Les deux méthodes de classification de scores SVM (TEE=1.5±0.54 % et un TR=97.44± 0.77%) et MLP (TEE= 1.33±0.51 % et TR = 97.49±0.77%) donnent presque le même résultat.

La méthode retenue reste la fusion par combinaison de scores par intégrale floue de Sugeno (TEE=1.14 ±0.47 % et TR= 98.36±0.75 %) qui donne le meilleur système multi-algorithmes d'authentification de visages avec un temps de calcul très réduit **t=0.94 seconde** (Langage de programmation utilisé est MATLAB R2009b et ordinateur : Intel Pentium Dual CPU 2.2 GHz , 1.49 GHz de RAM) . Le meilleur système de fusion est donc celui qui est construit par la **normalisation** de scores par la fonction **double sigmoïde** et la **fusion** par combinaison par l'intégrale floue de **Sugeno**.

## Application II

### 5.3 La multi-modalité (multi-biométries)

Dans cette section nous utilisons la base des scores de visage et voix de la base de données XM2VTS pour construire un système multi-biométries par les deux modalités visage et voix pour deux buts :

- Faire la comparaison pour les deux approches de fusion de scores combinaison et classification.
- Faire la comparaison entre les deux systèmes multimodale multi-algorithmes et multi-biométries.

#### 5.3.1 Les scores de visage et voix de la base de données XM2VTS

Pour les deux protocoles de Lausanne (LP1 et LP2), Norman Poh a mis à disposition un ensemble de scores issus de 5 classifieurs de visage et 3 classifieurs de voix présentés dans [151]. Ces scores donnent 21 configurations différentes pour 21 couples de classifieurs voix et visage (15 couples pour le protocole I et 6 couples pour le protocole II). Chaque configuration correspond à un couple des systèmes : un système de voix et un système de visage. Chaque système est composé d'un module d'extraction de caractéristiques et d'un classifieur. Un seul classifieur est utilisé pour la voix : le **GMM** (mélange de gaussiennes). Deux classifieurs sont utilisés pour le visage : le **GMM** (mélange de gaussiennes) et le **MLP** (réseau de neurones appelé "Multi Layer Perceptron"). Trois modules d'extraction de caractéristiques sont disponibles pour le visage (**FH** : Face Histogram, **DCTs** : Discrete Cosinus Transform (small size) et **DCTb** : Discrete Cosinus Transform (big size)) et trois modules d'extractions de caractéristiques pour la voix (**LFCC** : Linear Filter-bank Cepstral Coefficient, **PAC-MFCC** : Phase Auto-Correlation Mel Filter-bank Cepstral Coefficient et **SSC** : Spectral Subband Centroid). Tous les détails des classifieurs et des modules d'extraction de caractéristiques sont donnés dans [2].

Nous allons maintenant nous intéresser uniquement au Protocole I (LP1) pour lequel nous allons comparer les méthodes de fusion.

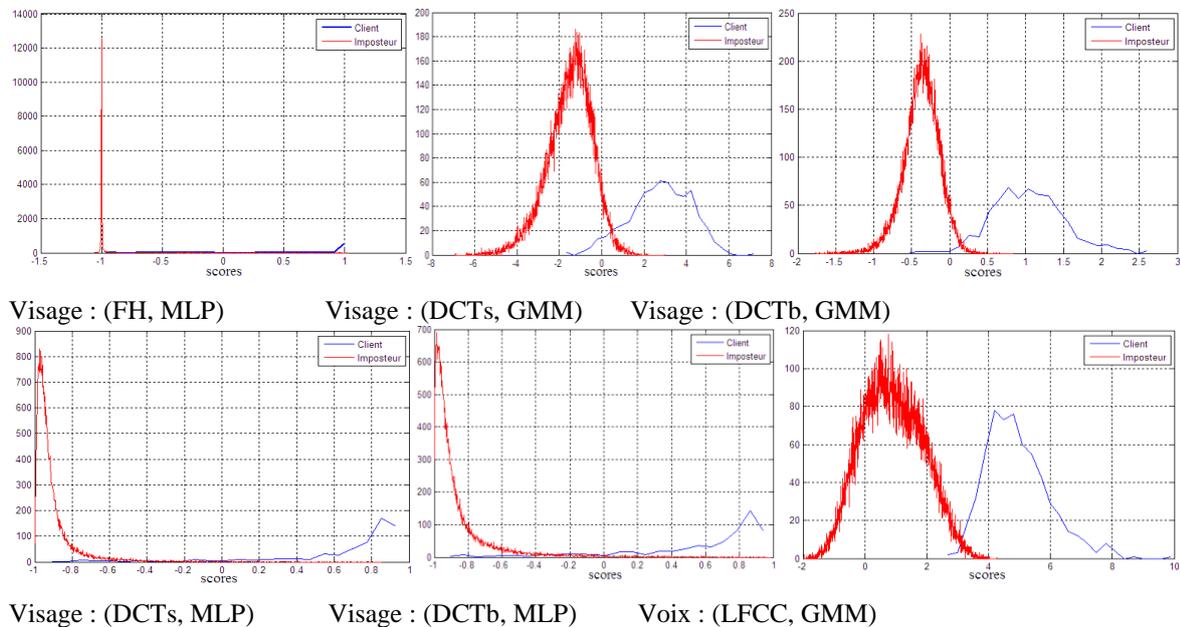
Le **tableau 5.8** présente la performance des 5 systèmes de visages et les trois systèmes de voix dans l'ensemble d'évaluation et de test. Nous remarquons dans ce tableau que le système de **Voix : (LFCC, GMM)** donne un bon résultat avec un **TEE=0.84±0.4%** et un **TR=97.7±1.15%**.

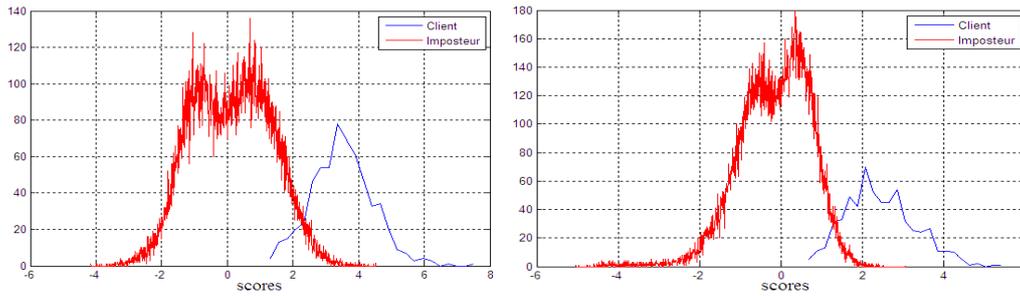
Méthodes	Ensemble d'évaluation		Ensemble de test	
	TEE %	TFR %	TFA %	TR %
Visage : (FH, MLP)	1.66±0.57	2±1.37	1.74±0.08	96.25±1.45
Visage : (DCTs, GMM)	4.37±0.91	4±0.91	4.44±0.12	91.55± 2.04
Visage : (DCTb, GMM)	<b>1.53±0.55</b>	<b>1.5±1.19</b>	<b>1.87±0.08</b>	<b>96.62±1.27</b>
Visage : (DCTs, MLP)	2.96±0.76	3.5±1.80	3.17±0.1	93.32±1.91
Visage : (DCTb, MLP)	4.82±0.96	8±2.65	4.41±0.12	87.58± 2.78
Voix : (LFCC, GMM)	<b>0.84±0.4</b>	<b>1.25±1.08</b>	<b>1.04±0.06</b>	<b>97.7± 1.15</b>
Voix : (PAC, GMM)	5.82±1.05	7.5±2.58	4.85±0.13	87.64±2.71
Voix : (SSC, GMM)	4.36±0.19	4.5±2.03	4.42±0.12	91.07±2.15

Tableau 5.8: Performances des 5 systèmes de visages et 3 systèmes de voix(intervalle de confiance paramétrique à 95%).

Les distributions de scores pour les 5 systèmes d'authentification de visages et les 3 systèmes d'authentification de voix représentées sur la figure 5.6.

Dans la figure 5.6 nous remarquons que les systèmes donnent des distributions Client et Imposteur différentes. Les distributions sont différentes en termes de plage de variation et différentes dans leurs formes et leurs recouvrements entre les deux classes, ce qui rendra nécessaire l'étape de normalisation de scores.





Voix : (PAC, GMM)

Voix : (SSC, GMM)

**Figure 5.6 :** Les distributions de 5 systèmes d’authentification de visages et 3 systèmes d’authentification de voix sur la base de données XM2VTS.

Le **tableau 5.9** présente la performance en terme de taux d’égale erreur **TEE** dans l’ensemble d’évaluation et en terme taux de réussite **TR** dans l’ensemble de test pour les 15 couples de systèmes voix et visage (voix, visage). La méthode de normalisation utilisée est la fonction double sigmoïde (voir l’annexe C les trois méthodes de normalisation Minmax, fonction Quadratique-Linéaire-Quadratique (QLQ) et la fonction double sigmoïde) ainsi que l’approche de fusion appliquée la combinaison de scores par la somme pondérée (Celle qui a été retenue voir paragraphe 5.2.2.), l’intégrales de floue de **Sugeno** et **Choquet**.

Méthodes Voix ,Visage	somme pondérée		Intégrale floue de Sugeno		Intégrale floue de Choquet	
	TEE%	TR %	TEE%	TR %	TEE%	TR %
(LFCC, GMM),(FH, MLP)	0.16±0.17	99.7±0.26	<b>0.002±0.02</b>	<b>99.99±0.04</b>	0.28±0.23	99.62±0.3
(LFCC, GMM),(DCTs, GMM)	0.2±0.2	99.39±0.38	<b>0.001±0.01</b>	<b>99.99±0.04</b>	0.20±0.2	99.09±0.46
(LFCC, GMM),(DCTb, GMM)	0.16±0.17	99.21±0.43	<b>0.007±0.03</b>	<b>99.93±0.12</b>	0.14±0.16	99.5±0.34
(LFCC, GMM), (DCTs, MLP)	0.04±0.01	99.31± 0.4	<b>0.03±0.07</b>	<b>99.91±0.14</b>	0.13±0.16	99.58±0.31
(LFCC, GMM), (DCTb, MLP)	0.34±0.08	98.95±0.5	<b>0.04±0.08</b>	<b>99.88±0.17</b>	0.54±0.32	99.04±0.48
(PAC, GMM),(FH, MLP)	1.46±0.53	98.6±0.57	<b>0.17±0.18</b>	<b>99.49±0.35</b>	1.69±0.57	99.00±0.49
(PAC, GMM),(DCTs, GMM)	0.68±0.36	98.79±0.53	<b>0.24±0.21</b>	<b>98.69±0.56</b>	1.36±0.52	97.28±0.8
(PAC, GMM),(DCTb, GMM)	1.35±0.51	96.07±0.96	<b>0.03±0.07</b>	<b>99.39±0.28</b>	0.82±0.4	96.75±0.88
(PAC, GMM), (DCTs, MLP)	0.69±0.37	97.78± 0.72	<b>0.04±0.08</b>	<b>99.66±0.28</b>	0.54±0.32	98.67±0.56
(PAC, GMM), (DCTb, MLP)	0.61±0.34	98.55±0.58	<b>0.63±0.53</b>	<b>99.21±0.43</b>	1.33±0.51	96.87±0.86
(SSC, GMM),(FH, MLP)	0.64±0.35	98.77±0.54	<b>0.09±0.13</b>	<b>99.72±0.25</b>	0.99±0.44	98.41±0.61
(SSC, GMM),(DCTs, GMM)	0.98±0.44	97.51±0.77	<b>0.28±0.23</b>	<b>99.48±0.35</b>	1.2±0.48	96.57±0.90
(SSC, GMM),(DCTb, GMM)	0.68±0.36	98.57±0.58	<b>0.03±0.07</b>	<b>99.64±0.29</b>	0.97±0.44	97.45±0.78
(SSC, GMM), (DCTs, MLP)	0.67±0.36	98.38±0.62	<b>0.2±0.2</b>	<b>99.40±0.37</b>	0.79±0.39	97.41±0.78
(SSC, GMM), (DCTb, MLP)	1.36±0.26	96.55±0.90	<b>0.45±0.3</b>	<b>97.65±0.74</b>	1.49±0.54	96.00±0.97

**Tableau 5.9 :** Performances des combinaisons de scores pour les 15 couples des systèmes (voix, visage) (intervalle de confiance paramétrique à 95%).

Le **tableau 5.9** montre que la combinaison de scores par l’intégrale floue de **Sugeno** donne le meilleur résultat par un système multi-biométries avec le couple voix visage (**LFCC**,

GMM),(DCTs, GMM) et une performance d'authentification presque parfaite avec un **TEE=0.001±0.01%** et un **TR=99.99±0.04%** .

Le **tableau 5.10** présente la performance de fusion de scores par l'approche de classification (méthode statistique de **Fisher, SVM** et **MLP**)pour les 15 couples de systèmes voix et visage (voix ,visage). La méthode de normalisation utilisée est la fonction double sigmoïde (voir l'annexe nous avons utilisé trois méthode de normalisation Minmax, fonction Quadratique-Linéaire-Quadratique (**QLQ**) et la fonction double sigmoïde).

Méthodes Parole ,Visage	Fonction linéaire discriminante (LDF)		Fonction quadratique discriminante(QDF)		Machines à vecteurs de support (SVM)		Le perceptron multicouche (MLP)	
	TEE%	TR %	TEE%	TR %	TEE%	TR %	TEE%	TR %
(LFCC, GMM),(FH, MLP)	0.54±0.32	98.45±0.60	0.19±0.19	99.65±0.29	<b>0.17±0.18</b>	<b>99.67±0.28</b>	0.23±0.21	99.57±0.32
(LFCC, GMM),(DCTs, GMM)	<b>0.18±0.19</b>	<b>99.21±0.43</b>	0.82±0.4	99.06±0.47	0.19±0.19	99.14±0.45	0.19±0.19	99.12±0.45
(LFCC, GMM),(DCTb, GMM)	0.16±0.17	99.20±0.43	0.51±0.21	99.21±0.43	<b>0.14±0.16</b>	<b>99.51±0.34</b>	0.22±0.21	99.04±0.48
(LFCC, GMM), (DCTs, MLP)	0.17±0.18	99.15±0.45	0.45±0.3	99.20±0.43	<b>0.03±0.07</b>	<b>99.30±0.41</b>	0.05±0.1	99.28±0.41
(LFCC, GMM), (DCTb, MLP)	0.35±0.26	99.19±0.44	1.95±0.62	97.99±0.69	<b>0.33±0.25</b>	<b>99.24±0.42</b>	0.25±0.22	98.97±0.49
(PAC, GMM),(FH, MLP)	4.5±0.93	89.42±1.57	2.37±0.68	93.25±1.26	0.83±0.4	99.06±0.47	<b>0.81±0.4</b>	<b>98.10±0.67</b>
(PAC, GMM),(DCTs, GMM)	1.36±0.52	95.82±0.99	1.45±0.53	97.26±0.80	1.49±0.54	97.36±0.79	<b>1.11±0.47</b>	<b>97.37±0.79</b>
(PAC, GMM),(DCTb, GMM)	1.69±0.57	93.04±1.28	1.13±0.47	94.50±1.14	0.81±0.59	98.78±0.54	<b>0.69±0.37</b>	<b>98.39±0.62</b>
(PAC, GMM), (DCTs, MLP)	2.68±0.72	92.01±1.37	0.82±0.4	94.18±1.17	0.61±0.34	98.43±0.61	<b>0.38±0.27</b>	<b>98.36±0.62</b>
(PAC, GMM), (DCTb, MLP)	1.63±0.56	94.29±1.16	1.45±0.53	96.07±0.96	1.85±0.6	96.25±0.94	<b>1.13±0.47</b>	<b>96.30±0.93</b>
(SSC, GMM),(FH, MLP)	4.21±0.92	91.73±1.39	3.96±0.87	39.06±3.49	0.46±0.3	98.72±0.55	<b>0.43±0.29</b>	<b>98.86±0.52</b>
(SSC, GMM),(DCTs, GMM)	2.13±0.64	94.27±1.16	2.85±0.74	45.60±3.34	0.95±0.43	97.73±0.73	<b>0.74±0.38</b>	<b>97.20±0.81</b>
(SSC, GMM),(DCTb, GMM)	3.01±0.76	93.20±1.26	3.2±0.79	43.73±3.38	0.54±0.32	98.36±0.62	<b>0.41±0.28</b>	<b>97.62±0.75</b>
(SSC, GMM), (DCTs, MLP)	3.37±0.18	92.79±1.30	3.14±0.78	43.09±3.40	0.67±0.36	97.84±0.71	<b>0.59±0.34</b>	<b>98.31±0.63</b>
v(SSC, GMM), (DCTb, MLP)	2.61±0.71	93.87±1.20	3.12±0.78	42.89±3.40	1.2±0.48	96.70±0.88	<b>0.99±0.44</b>	<b>97.18±0.82</b>

**Tableau 5.10:** Performances des classification de scores pour les 15 couples des systèmes (voix, visage) (intervalle de confiance paramétrique à 95%).

Dans le **tableau 5.10** nous constatons qu'il n'y'a pas de différences de performances selon la classification de scores par **SVM** et **MLP**. Cependant le **SVM** donne la meilleure performance par le couple de systèmes (LFCC, GMM),(DCTs, MLP) avec un **TEE=0.03±0.07%** et un **TR=99.30±0.41%**.

Enfin de cette série de résultats, un tableau de synthèse (**tableau 5.11**) permet de comparer notre système d'authentification multi-algorithmes et multi-biométries (visage,voix)par quelque travaux connus enfusion multimodale. Ainsi nous pouvons faire la comparaison entre les systèmes multi-algorithmes et multi-biométries sur différentes bases de données telles la base de données XM2VTS et la base de scores XM2VTS selon le protocole de Lausanne I, la base FERET, CASIA et la base de scores de BioSecure (DS3).

Date	Auteur	Multimodalités	Niveau de fusion	Type de fusion	Méthodes de fusion	Base de données	Meilleurs résultats
2007	Sadeghi et Khoshrou [85]	multi-classifieursvisage	Score	Combinaison	Confidence Based Gating of Scores	XM2VTS	TFA=0.48%, TFR= 0.5% TR=98.45%
2009	Villegas et Paredes[152]	multi-biométries visage,voix	Score	Combinaison	Maximizing the Area under the ROC Curve	Base de scores XM2VTS	TEE=0.28%
2009	Li et al [153]	multi-biométries visage,voix	Score	Combinaison	Nouvelle méthode basé sur TFR et TFA	Base de scores XM2VTS	TEE=0.27%
2009	EJarque et al [139]	multi-biométries visage,voix	Caractéristiques	Classification	SVM	XM2VTS	TEE=0.008%
2009	Chia et al [112]	multi-biométries visage,voix	Score	Combinaison	Hybride par min ou max ou somme	Base de scores XM2VTS	TEE=0.5%
2009	Morizet[11]	multi-biométries visage, iris	Score	Combinaison	WSDSF	RERET / CASIA	TEE=0.011%
2009	Allano [10]	multi-biométriesvoix, visage et signature	Score	Classification	SVM (noyau RBF)	Base de scores de BioSecure DS3	TEE=5.54%
2010	Lobrano et al [154]	multi-biométries visage,voix	Score	Combinaison	Dynamic Linear Combination(DLC)	Base de scores XM2VTS	TEE=0.0045 %
2011	Nous	multi-classifieursvisage	Score	Combinaison	Intégrales de floue de Sugeno	XM2VTS	TEE=1.14% TR=98.36%
2011	Nous	multi-classifieursvisage	Score	Classification	MLP	XM2VTS	TEE=1.66% TR=97.84%
2011	Nous	multi-classifieursvisage	Score	Classification	SVM (noyau RBF)	XM2VTS	TEE=1.66% TR=97.84%
2011	Nous	multi-biométries visage ,voix	Score	Combinaison	Intégrales de floue de Sugeno	Base de scores XM2VTS	TEE=0.001% TR=99.99%
2011	Nous	multi-biométries visage ,voix	Score	Classification	MLP	Base de scores XM2VTS	TEE=0.23% TR=99.57%
2011	Nous	multi-biométries visage ,voix	Score	Classification	SVM (noyau RBF)	XM2VTS	TEE=0.17% TR=99.67%

Tableau 5.11 : Tableau de synthèse des résultats obtenus en fusion multimodale.

Nous pouvons dire que la performance de notre système multi-algorithme d'authentification de visages et presque équivalent à celui obtenu en 2007 par Sadeghi et Khoshrou avec un TEE =1.14% et un TR=98.36. Notre système multi-biométries d'authentification de visage et voix ainsi conçu est le meilleur comparé aux travaux 2009 et 2010 avec un TEE=0.001% et un TR=99.99% qui est justifié par l'efficacité de la méthode de normalisation de scores la fonction double sigmoïde et la fusion de scores par intégrale floue de Sugeno.

Pour la comparaison entre les deux systèmes multi-algorithmes et multi-biométries le système multi-biométries donne une performance presque parfaite avec un TR=99.99%, mais avec quelques inconvénients: temps de capture de visage et voix très grand [17], le coût du matériel, le confort d'utilisateurs etc... Ceci dit la performance satisfaisante est

obtenue par un système multi-algorithmes d'authentification de visages avec un TR= 0.98 \* TR de système multi-biométries et avec l'avantage de temps de calcul réduit (parce que le temps de capture est très supérieure que le temps de traitement [10]), le confort d'utilisateurs etc....

## 5.4. Conclusion

Dans ce chapitre les résultats sont présentés sous forme de deux applications comme suit :

Dans la première partie nous avons construit un système multi-algorithmes par la fusion des quatre systèmes retenus à travers l'étude menée dans le chapitre 2. Une étape de normalisation de scores avant la fusion est nécessaire. Pour cela nous comparons six méthodes de normalisation de scores. Nous étudions deux approches de fusion de scores qui sont la combinaison et la classification. Enfin le meilleur système multi-algorithmes est obtenu avec un **TEE=1.14±0.47%** et un **TR=98.36 ± 0.75%** avec la **normalisation** de scores par la **fonction double sigmoïde** pour l'**intégrale floue de Sugeno**.

Dans la dernière partie nous testons les deux approches de fusion de scores des visages et des voix de la base de données XM2VTS pour construire un système multi-biométries. Le meilleur résultat dans ce cas est réalisé avec un **TEE=0.001±0.01%** et un **TR=99.99±0.04%** pour l'**intégrale floue de Sugeno**. Enfin nous confirmons l'efficacité de notre système multimodale par la comparaison de notre modèle avec quelques travaux connus en fusion multimodale.

### Conclusion générale

Dans ce présent travail nous avons tout d'abord donné un aperçu sur l'état de l'art des méthodes de reconnaissance de visage 2D, ensuite nous présentons notre **modèle de la reconnaissance** de visage. Puis, nous détaillons la **mesure de la performance** d'un système biométrique. Dans l'état de l'art nous avons décomposé les méthodes en trois grandes familles : les **méthodes globales**, pour lesquelles les caractéristiques sont extraites directement depuis l'ensemble des valeurs de pixels des images et les **méthodes locales**, basées sur l'extraction de signatures extraites localement du visage. Les techniques dites **hybrides** utilisent conjointement ces deux types de modélisation. Parmi les méthodes globales, on compte notamment les méthodes basées sur la **projection statistique**, aussi appelées techniques de **réduction de dimension**. Nous avons vu au travers des résultats expérimentaux provenant de différentes sources (**tableau 1.1**) que la méthode globale des **fisherfaces (LDA)** est très performante, en comparaison avec les autres techniques de l'état de l'art. Donc notre choix s'est porté sur les méthodes globales à base de **fisherfaces**. Dans notre **modèle de la reconnaissance** de visage, nous avons essayé de présenter quelques algorithmes utilisés dans les **trois étapes** d'un système de reconnaissance faciale (l'étape d'extraction des caractéristiques, l'étape de réduction l'espace de données et la comparaison entre deux vecteurs et deux matrices). Enfin pour la mesure de la performance nous avons défini quatre critères principaux et une **mesure de confiance associée** (taux de faux rejet (**TFR**), taux de fausse acceptation (**TFA**), taux d'égale erreur (**TEE**) et taux de réussite (**TR**)). L'étude expérimentale de l'authentification de visages sur la base de données XM2VTS selon son protocole associé «protocole de Lausanne I» est présentée dans le **chapitre 2**. Le choix principal de cette base de données est sa grande taille et sa popularité, puisqu'elle est devenue une norme dans la communauté biométrique audio et visuelle de vérification multimodale d'identité. D'abord nous commençons par la comparaison des algorithmes de réduction d'espace, puis nous avons étudié deux méthodes d'extraction des caractéristiques de visage.

Pour la comparaison de méthodes de réduction d'espace de donnée nous avons utilisé deux types des méthodes de réduction le premier type sont les méthodes en une dimension (Analyse en Composantes Principales (**ACP**), Analyse Discriminante Linéaire (**LDA**), Le modèle discriminant Linéaire amélioré de Fisher (**EFM**) et LDA directe (**DLDA**)) et le deuxième type sont les méthodes en deux dimension (L'Analyse Discriminante Linéaire

Bidimensionnelle Orientée (**ADL2Do**). Dans cette étude nous déduisons que la méthode une dimension **EFM** est la meilleure et les méthodes en deux dimension n'améliore pas la performance d'authentification de visage par rapport aux LDA, EFM et DLDA car la base de données XM2VTS a été réalisée sur une longue période ce qui a permis d'obtenir plusieurs images d'une même personne et donc une grande variabilité d'apparence (changement de coiffure, présence et/ou absence de lunettes, etc.). Parcontre, seules les expressions neutres ont été considérées, donc nous proposons utiliser ces méthodes 2D pour les bases de données qui contiennent plusieurs expressions.

Dans notre étude deux méthodes d'extraction de caractéristiques de visage ont été réalisées. D'abord nous avons utilisé la transformation de **Hough** comme méthode de détection de contours de l'image de visage et nous avons remarqué que cette méthode n'améliore pas la performance de vérification de visage en termes de précision. Puis nous avons utilisé l'information **spatio-fréquentielle** du visage pour obtenir une bonne représentation. Cette information est extraite de l'image des pixels par application de filtres de **Gabor** à différentes **résolutions** et à diverses **orientations**. Cette représentation a largement été utilisée et étudiée dans la littérature. Dans la majorité de ces études, seule **l'amplitude** de la réponse des filtres de Gabor a été employée alors que la **phase** est omise. **Nous avons expliqué les raisons de la limitation de l'utilisation de cette phase** et nous avons fourni une solution simple pour pallier cette limitation par **une seuillage de la phase**. Nous avons également montré que les phases filtrées de la convolution des images de visages par les filtres de Gabor ont données les meilleurs résultats par rapport aux amplitudes. Et finalement nous sélectionnons les **meilleurs systèmes** d'authentification de visages pour construire un système multi-algorithmes dans la deuxième partie de notre mémoire.

Dans la deuxième partie de notre travail les limitations des systèmes biométriques lorsqu'ils n'utilisent qu'une seule modalité biométrique est passée en revue. Ensuite nous présentons une façon de réduire les limitations des systèmes biométriques monomodaux en combinant plusieurs systèmes menant ainsi à la biométrie multimodale. Les systèmes multimodaux peuvent être de différentes natures. Pour cela nous étions amenés à étudier les différents types de combinaisons de modalités possibles, mais aussi les architectures et les niveaux de fusion qui peuvent être utilisés dans un système multimodal. Nous présentons donc le type de multimodalité et le niveau de fusion. Parmi les divers niveaux de fusion existant, nous avons choisi de travailler au niveau score car il offre **le meilleur compromis entre la richesse d'information et la facilité d'implémentation**. De plus, ce niveau de fusion

est **complètement indépendant des algorithmes qui sont utilisés** pour générer des scores. Ceci offre donc une grande souplesse d'intégration dans des systèmes existants qui peuvent être mis à jour sans interférer avec la technique de fusion à mettre en place. Un état de l'art de la fusion de scores est présenté dans la fin du **chapitre 3** et la théorie des méthodes de fusion de scores pour les deux approches **combinaison** et **classification** de scores ainsi que six méthodes de **normalisation de scores** sont présentées dans le **chapitre 4**. Les résultats expérimentaux de la multimodalité sont divisés en deux parties dans le **chapitre 5** comme suit :

Dans la première partie nous étudions un système **multi-algorithmes** ou **multi-classifieurs** (Multiple Classifier Systems) d'authentification de visages. Celui-ci est construit par les quatre systèmes retenus dans l'étude expérimentale de l'authentification de visages dans le chapitre 2. Tout d'abord nous présentons les **distributions de scores** et nous remarquons que les distributions sont différentes en termes de plages de variations et différentes dans leurs formes et leur recouvrements entre les deux classes (Client et Imposteur). Ce qui rend **nécessaire** l'étape de **normalisation** des scores. Ainsi nous appliquons les six méthodes de normalisation de scores sur le **système 1** d'authentification de visages et nous étudions les méthodes de normalisation :

- **Minmax, Znorm, Tanh et la médiane et l'écart absolu médian** qui sont des méthodes qui ne modifient pas la forme des distributions. Mais leurs différences se situent dans la façon dont chaque distribution sera répartie dans l'intervalle commun,
- et les méthodes de normalisation par **fonction quadratique-linéaire-quadratique (QLQ)** et **fonction double sigmoïde** des méthodes qui modifient la forme des distributions.

Ensuite nous comparons les méthodes de normalisation associées aux méthodes de **combinaisons de scores**, pour les méthodes de **combinaisons de scores simples** : la somme, le produit, le minimum, le maximum, la médiane et la somme pondérée. Nous constatons qu'il n'y a pas de grandes différences de performances selon les normalisations. Mais surtout selon les méthodes de combinaison, en particulier la médiane.

Nous pouvons dire que la **somme** et la **somme pondérée** sont les meilleures, la somme pondérée étant légèrement meilleure mais nécessitant un apprentissage des poids.

Pour la combinaison de scores par **logique floue** à base de l'intégrale floue de **Sugeno** et de **Choquet** nous obtenons une **bonne performance** d'authentification de visage par l'intégrale floue de **Sugeno** et la méthode de normalisation de scores **la fonction double sigmoïde** avec un **TEE=1.14±0.47%** et **TR=98.36 ± 0.75%**.

Dans la section de comparaison des méthodes de **classification de scores** nous **proposons** d'utiliser les méthodes de normalisation de scores avant la classification c'est ce qui fait l'originalité de notre travail. Les méthodes de classification de scores utilisées sont : **méthode statistique de Fisher**, les **machines à vecteurs de support (SVM)** et un **réseau de neurones** simple de type **MLP** (Multi layer perceptron), **sans et avec** les méthodes de fusion suivant : **Z-Score**, **Fonction Quadratique-Linéaire-Quadratique (QLQ)** et normalisation par une fonction **double sigmoïde**. D'après toute cette étude et les résultats obtenus nous affirmons que les méthodes de normalisation de scores améliorent la performance en générale de l'authentification de visages pour toutes les méthodes de classification de scores utilisées. Et les deux méthodes de classification de scores non linéaire **SVM** et **MLP** donnent des résultats équivalents, ceci dit le **SVM** présente une simplicité d'application. Les deux fonctions de classification de **Fisher** : la fonction linéaire discriminante (**LDF**) et fonction quadratique discriminante (**QDF**) ont des performances moindres que celles obtenues par le **SVM** et le **MLP**.

Dans la deuxième partie nous appliquons nos deux approches de fusion à la base des scores des visages et voix de la base de données des scores **XM2VTS** pour construire un système **multi-biométries** par les deux modalités **visage** et **voix** avec **15 couples** des systèmes voix, visage pour deux buts :

- faire la **comparaison** pour les **deux approches** de fusion de scores **combinaison** et **classification**,
- faire la **comparaison** entre notre système multimodale **multi-algorithmes** et **multi-biométries**.

La **combinaison de scores** par **l'intégrale floue de Sugeno** donne le meilleur résultat pour un système multi-biométries avec le couple (**voix,visage**)(**LFCC**, **GMM**),(**DCTs**, **GMM**) et une **performance d'authentification presque parfaite** avec un **TEE=0.001±0.01%** et un **TR=99.99±0.04%**.

Pour la **comparaison** entre les deux systèmes **multi-algorithme** et **multi-biométries** :

- Le système **multi-biométries** donne une **performance presque parfaite** avec un **TR=99.99%**, mais avec beaucoup d'inconvénients : temps de capture de visage et voix assez lent, coût du matériel, confort de l'utilisateur ...etc.
- Nous obtenons une **performance très satisfaisante** pour un système **multi-algorithmes** d'authentification de visages avec un **TR= 0.98 x TR** du système **multi-**

**biométries** et avec pas mal d'avantages : temps de calcul réduit parce que le temps de capture est très supérieure que le temps de traitement, le confort d'utilisateurs ...etc.

### Perspectives

Nous avons le long de notre travaillé uniquement sur les scores issus des systèmes monomodaux pour réaliser la fusion. Malgré des résultats très intéressants, nous nous sommes rendu compte que nous atteignons des limites en ne considérant que les scores. En effet, les scores ne donnent qu'une quantité d'information limitée. Si l'on souhaite améliorer davantage les performances en fusion, il faudrait se placer dans un espace contenant plus d'information, comme par exemple celui des caractéristiques.

Comme suite à notre travail et comme travaux futurs nous pouvons proposer certaines idées. Les perspectives d'évolution de ce travail sont alors multiples :

- Application de nos systèmes multimodaux sur les bases de données génériques qui contiennent de très grand nombre d'individus, et les bases de personnes virtuelles.
- Chercher des systèmes d'authentification monomodaux plus performants.
- Etudier d'autres systèmes multi-biométries comme le visage et l'empreinte.
- Faire de la fusion au niveau des caractéristiques. Dans un espace de dimension plus grande par les **ondelettes**, un classifieur de type **SVM** ou la **modélisation multidimensionnelle**, parce que le score augmente de beaucoup la taille de l'espace de travail. C'est ce que l'on recherche pour augmenter l'information disponible, mais cela engendre également des contraintes et des difficultés en termes de complexité et de temps de calcul.
- Se mettre en **3D** et **4D** pour la modalité visage.

## Annexe A

### Authentification et identification

Il existe plusieurs types d'application qui nécessitent la reconnaissance des utilisateurs. Ces applications peuvent être séparées en 2 catégories que sont la **vérification d'identité** ou **l'identification** (figure A.1).

#### A.1. La vérification ou authentification

La vérification est une comparaison **un à un**, dans lequel le système valide l'identité d'une personne en comparant les données biométriques saisie avec le modèle biométrique de cette personne stockée dans la base de données du système. Dans un tel mode, le système doit alors répondre à la question suivante: «*Suis-je réellement la personne que je suis en train de proclamer ?*». Actuellement la vérification est réalisée via un numéro d'identification personnel, un nom d'utilisateur, ou bien une carte à puce.

#### A.2. L'identification

L'identification est une comparaison **un à N**, dans lequel le système reconnaît un individu en l'appariant avec un des modèles de la base de données. La personne peut ne pas être dans la base de données. Ce mode consiste à associer une identité à une personne. En d'autres termes, il répond à des questions du type: «*Qui suis-je ?*».

Les différents modules qui composent un système biométrique sont représentés sur la **figure A.1**. Leur fonctionnement peut être résumé comme suit :

- **Le processus d'enrôlement** : C'est une **phase d'apprentissage** qui a pour but de recueillir des informations biométriques sur les personnes à identifier. Plusieurs campagnes d'acquisitions de données peuvent être réalisées afin d'assurer une certaine robustesse au système de reconnaissance aux variations temporelles des données. Pendant cette phase, les caractéristiques biométriques des individus sont saisies par un capteur biométrique, puis représentées sous forme numérique (signatures), et enfin stockées dans la base de données. Le traitement lié à l'enrôlement n'a pas de contrainte de temps, puisqu'il s'effectue « hors-ligne ».

- **Module capteur biométrique** : correspond à la **lecture** de certaines caractéristiques physiologiques, comportementales ou biologiques d'une personne, au moyen d'un terminal de capture biométrique,
- **Module extraction des données** : **extraie** les informations pertinentes à partir des données biométriques brutes, par exemple des images de visage ou des régions caractéristiques de visage,
- **Module création d'une signature** : crée un modèle numérique afin de **représenter** la donnée biométrique acquise. Ce modèle, appelé aussi **signature** ou **caractéristique**, sera conservé sur un support portable (puce ou autre) ou dans une base de données,
- **Module comparaison** : compare les caractéristiques biométriques d'une personne soumise à contrôle avec les « signatures » mémorisées. Ce module fonctionne soit en mode vérification (pour une identité proclamée) ou bien en mode identification (pour une identité recherchée).
- **Module base de données** : stocke les modèles biométriques des utilisateurs enrôlés.

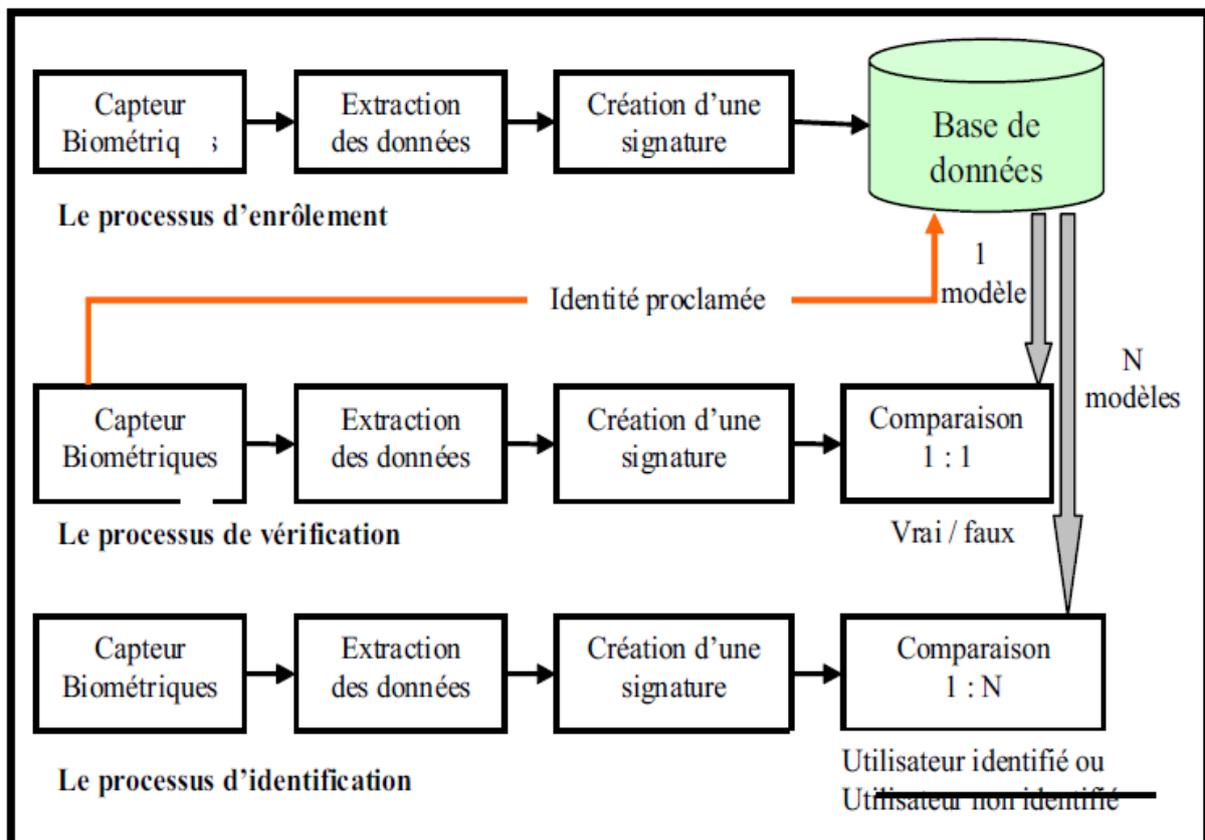


Figure A.1 : Principaux modules d'authentification et d'identification d'un système biométrique.

## Annexe B

### Introduction à Matlab pour le traitement d'image

#### B.1 Les images numériques

Une image est un signal bidimensionnel. Une image analogique est par exemple celle formée sur la rétine de l'œil ou les images obtenue par la photographie argentique classique.

Une image numérique est un signal numérique composé d'unités élémentaires (appelées pixels) qui représentent chacun une portion de l'image. Contrairement au cas unidimensionnel, nous étudierons uniquement les images numériques (discrètes).

Une image numérique est définie par :

- le nombre de pixels qui la composent en largeur et en hauteur (qui peut varier presque à l'infini),
- l'étendue des teintes de gris ou des couleurs que peut prendre chaque pixel (on parle de dynamique de l'image).

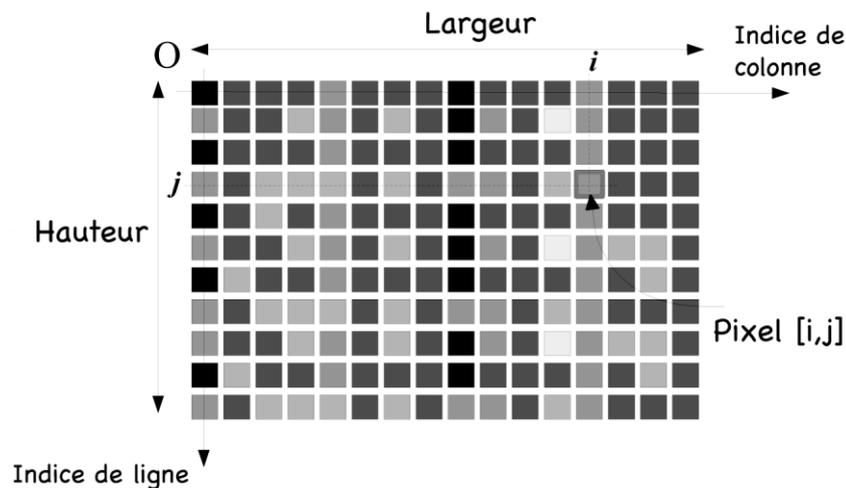


Figure B.1. Pixels et niveaux de gris.

$I$  est l'image numérique :  $I[i, j]=n$  est la valeur du niveau de gris. Lorsque  $n \in [N_{min}, N_{max}]$ ,  $N_{max} - N_{min}$  est le nombre de niveaux de gris. La dynamique de l'image est donnée par  $\text{Log}_2(N_{max} - N_{min})$ .

##### B.1.1 Les images binaires (noir ou blanc)

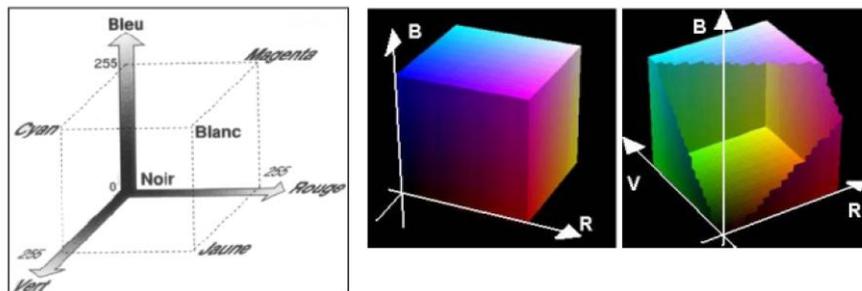
Exemple, images les plus simples, un pixel peut prendre uniquement les valeurs noir ou blanc. C'est typiquement le type d'image que l'on utilise pour scanner du texte quand celui-ci est composé d'une seule couleur.

### B.1.2 Les images en niveaux de gris

En général, les images en niveaux de gris renferment 256 teintes de gris. Par convention la valeur zéro représente le noir (intensité lumineuse nulle) et la valeur 255 le blanc (intensité lumineuse maximale). Le nombre 256 est lié à la quantification de l'image. En effet chaque entier représentant un niveau de gris est codé sur 8 bits. Il est donc compris entre 0 et  $2^8 - 1 = 255$ . C'est la quantification la plus courante. On peut coder une image en niveaux de gris sur 16 bits.

### B.1.3 Les images couleurs

L'espace couleur est basé sur la synthèse additive des couleurs, c'est à dire que le mélange de trois composantes (par exemple (R, V, B)) donne une couleur. On garde l'information couleur, ou intensité lumineuse et chromaticité. Un pixel est codé par trois valeurs numériques. La signification de ces valeurs dépend du type de codage choisi. Le plus utilisé pour le maniement des images numériques est l'espace couleur Rouge, Vert, Bleu (R,V, B) (RGB en anglais). La restitution des couleurs sur écran utilise cette représentation (synthèse additive).



**Figure B.2.** L'espace couleur Rouge Vert Bleu (RGB)

Il en existe beaucoup d'autres : CMJ (ou CMY en anglais), TSV (ou HSV en anglais), YUV, YIQ, Lab, XYZ.

## B.2 Introduction à Matlab

### B.2.1 Q'est-ce que MATLAB

- MATLAB est l'abréviation de MATrix LABoratory. C'est un logiciel pour le calcul numérique et la visualisation optimisé pour le calcul matriciel.
- MATLAB peut être considéré comme un langage de programmation, il dispose d'une syntaxe spécifique mais elle est simple et intuitive.

- MATLAB est un interpréteur : les instructions sont interprétées et exécutées ligne par ligne.
- MATLAB propose deux types de fonctionnement :
  - un fonctionnement en ligne de commande où MATLAB exécute les instructions au fur et à mesure qu'elles sont données par l'utilisateur.
  - un fonctionnement via des scripts, dans ce cas MATLAB exécute ligne par ligne un fichier texte que l'on désigne par programme.
- MATLAB propose également une aide très complète illustré d'exemple, étant donné le très grand nombre d'instructions utilisables, il est important de bien savoir l'utiliser.

### B.2.2 Les variables

Les variables manipulés sont en priorité des matrices à éléments réels ou complexes.

Un scalaire est une matrice 1x1. Un exemple de matrice :

» a=[1,2,3 ;4,5,6]

Lorsque les composantes d'un vecteur forment une suite régulièrement espacée, on peut utiliser une boucle dite (implicite) de la forme :

» a= (0 :2 :10) ce qui est équivalent à a=[0,2,4,6,8,10]

On accède au premier élément d'une matrice par a(1,1) et à la première ligne de la matrice par a(1, :).

Remarque : Il faut respecter les minuscules majuscules dans l'appellation des variables.

### B.2.3 Les opérations matricielle

- u \* v multiplication matricielle
- u + v addition matricielle
- u - v soustraction matricielle
- u^v exponentiation
- u. \* v multiplication terme à terme u et v
- u./v division du tableau u par le tableau v terme à terme.
- u.^v exponentiation terme à terme
- A\b résolution du système linéaire Ax=b
- u' prend le transposé de u
- inv(A) inverse d'une matrice
- det(A) déterminant d'une matrice
- rank(A) rang d'une matrice

### B.2.4 Les variables prédéfinies

- `ones(L,C)` : matrice de L lignes C colonnes contenant des 1.
- `zeros(L,C)` : matrice de L lignes C colonnes contenant des 0.
- `randn(L,C)` : matrice de L lignes et C colonnes contenant un échantillon de distribution gaussienne centrée de variance 1.
- `rand(L,C)` : matrice de L lignes et C colonnes contenant un échantillon de distribution uniforme sur l'intervalle (0,1).
- `eye(L,C)` : matrice de L lignes C colonnes contenant des 1 sur la diagonale principale et des 0 ailleurs (matrice identité).
- `diag(V)` : matrice carrée avec le vecteur V sur la diagonale et des 0 ailleurs.
- La constante pi vaut 3.14159265358979.
- Les constantes i, j sont aussi prédéfinies et  $i^2 = j^2 = -1$ .

### B.2.5 Les fonctions

- Les fonctions mathématiques de base :  
abs, sqrt, real, imag, conj, round, fix, floor, ceil, sign, rem, exp, log.
- Les fonctions trigonométriques :  
sin, cos, tan, asin, acos, atan, sinh, cosh, tanh, asinh, acosh, atanh.
- Autres fonctions utiles :
  - La fonction eig permet d'obtenir les valeurs propres et vecteurs propres d'une matrice.
  - La fonction poly donne le polynôme caractéristique associé à la matrice.
  - La fonction roots fournit les racines d'un polynôme.

### B.2.6 Les entrées sorties

Affichage des résultats :

- MATLAB affiche automatiquement le résultat après l'exécution d'une ligne de commande. Il faut placer un « ; » après une commande si l'on veut empêcher cet affichage.

Les entrées sorties :

- `input` et `ginput` pour les acquisitions au clavier.
- `disp` et `sprintf` pour les affichages sur l'écran.
- `gtext`, `plot`, `grid`, `title` pour les affichages graphiques.

- load et save pour charger ou sauvegarder dans un fichier format MATLAB.
- fopen, fread, fwrite pour charger sauvegarder dans des fichiers sous un autre format.

### B.2.7 Structure des programmes

Un programme MATLAB est un fichier texte contenant des suites d'instructions MATLAB dont le nom a comme extension .m Par exemple« programme. m ». Dans la fenêtre commande, si l'on tape « programme. m », les instructions contenues dans le fichier programme. m seront exécutées une par une.

### B.2.8 Les toolboxes

En plus de fonctions de bases pour le calcul matriciel, MATLAB dispose de nombreuses bibliothèques de fonctions spécialisées appelées « toolbox » dans différents domaines. On peut citer notamment les toolboxes : control system, data acquisition, fuzzy logic, neural network, optimization, statistics, signal processing, wavelet et la toolbox image processing qui nous intéresse plus particulièrement.

## B.3 Traitement d'image sous MATLAB

### B.3.1 Image Processing Toolbox

Les fonctions de la bibliothèque « traitement d'image » de MATLAB :

- Lecture, écriture et affichage d'une couleur ou niveau de gris,
- Transformations spatiales et transformations fréquentielles,
- Filtrage linéaire et non linéaire,
- Binarisation et morphologie mathématique,
- Analyse, et restauration d'image,
- Changement d'espace couleur,
- ...

### B.3.2 Codage d'une image, représentation spatiale

Une image est considérée comme un ensemble de points ou pixels (picture element), associé au quadrillage rectangulaire de l'image d'origine. La représentation d'une image se fait donc par l'intermédiaire d'une matrice d'entier codé entre 0 et 255. Les images en niveau de gris sont représentées par des matrices 2D, les images couleurs représentés par 3 composantes (Rouge, Vert, Bleu) sont représentées par des matrices 3D. On accède à un pixel grâce a son

indice de ligne et son indice de colonne. Le premier pixel d'une image est le pixel en haut à gauche. Cette représentation est appelé représentation spatiale de l'image.

### B.3.3 Lecture, écriture et affichage d'une image

Matlab est capable de lire et de décoder les fichiers images JPEG, TIFF, BMP, PNG, HDF, PCX ou XWD. Une image sous Matlab peut être représentée sous plusieurs formes, mais toujours sous forme d'une matrice. Avant de traiter une image dans Matlab, il faut la lire et décoder son format afin de la transformer en une matrice de valeurs. L'exemple ci-dessous permet de lire une image au format TIFF, de la décoder dans la variable `img` et de l'afficher à l'écran dans une figure. La commande `axis('image')` rend l'image affichée carrée pour garder les proportions. L'appel à `axis on` permet l'affichage des graduations des axes. Enfin, `colorbar` affiche la barre des couleurs de l'image.

```
-----
» img=imread('saturn.tif');
» figure;imshow(img);
» axis('image');
» axis on
» colorbar
-----
```

L'accès à un élément particulier d'une image est indexé par le nom et la position de cet élément. Par exemple, si on conserve l'image `img` ci-dessus, on peut récupérer les valeurs ou les modifier aisément.

```
-----
» img(3,2)
» img(1:10,30:40)
» img(1:3,31:39) = 0;
» figure;imshow(img);
-----
```

Matlab offre une possibilité intéressante, en effet, il est possible d'afficher plusieurs images dans la même figure. Pour ce faire, il faut utiliser la commande `subplot`. Elle s'utilise avec comme arguments le nombre de ligne, le nombre de colonnes et le numéro de l'image dans la figure. Dans l'exemple ci-dessous on souhaite afficher deux images sur la même ligne dans une seule figure.

```
-----
» img=imread('blood1.tif');
» img2=imread('alumgrns.tif');
» figure;subplot(1,2,1);imshow(img);
» subplot(1,2,2);imshow(img2);
-----
```

Afin de fermer une figure sous Matlab, on tape *close* s'il s'agit de la dernière figure ouverte ou bien *close* avec en paramètre le numéro de figure pour fermer la figure donnée en paramètres. Pour fermer toutes les figures, on demande la fermeture avec le mot-clé *all*.

```
-----
» img=imread('rice.tif');
» img2=imread('alumgrns.tif');
» figure;imshow(img);
» figure;imshow(img2);
» close(1)
» close all
-----
```

Pour sélectionner une figure, on demande l'affichage de celle-ci à l'aide de la commande *figure* qui prend en paramètre le numéro de la figure. Un exemple illustre ce cas ci-dessous.

```
-----
» figure(1)
-----
```

Matlab autorise l'exportation d'images sous divers formats: BMP, TIFF, EPS, PS... La commande qui permet de sauvegarder une figure est *print -dFORMAT fichier*. Un exemple est donné ci-dessous. Dans cet exemple, on affiche une image dans une figure et grâce à la commande *print*, on exporte le résultat dans le format JPEG avec pour nom de fichier *result.jpg*.

```
-----
» img=imread('rice.tif');
» figure;imshow(img);
» print -djpeg result.jpg
-----
```

Les valeurs des images lues sous Matlab sont entières, mais dans certaines circonstances, on a besoin de travailler sur des valeurs réelles. La transformation pour passer d'entier à réel utilise la fonction *im2double*.

```
-----
» img=imread('rice.tif');
» figure;imshow(img);
» imgdbl=im2double(img);
» figure;imshow(imgdbl);
» imgint=im2uint8(imgdbl);
» figure;imshow(imgint);
» imwrite(imgint,'test.jpg','jpeg');
» whos
-----
```

### B.3.4 Opérations géométriques

Les opérations géométriques classiques sont permises avec la boîte à outils de traitement d'images: rotation, changement de taille, découpage...

L'exemple ci-dessous illustre la rotation d'une image avec Matlab. Dans le premier cas, `imgrot1` est plus grande que `img`. Dans le second cas, le paramètre `'crop'` impose un découpage de l'image et la taille de l'image `imgrot2` est la même que celle de l'image `img`.

```
-----
» img=imread('bacteria.tif');
» figure;imshow(img);
» imgrot1=imrotate(img,3,'bilinear');
» figure;imshow(imgrot1);
» imgrot2=imrotate(img,3,'bilinear','crop');
» figure;imshow(imgrot2);
-----
```

Le zoom permet d'agrandir une partie de l'image à l'aide de la souris, pour cela, il faut entrer la commande dans l'éditeur de Matlab puis sélectionner une zone de l'image à agrandir. Le bouton droit de la souris permet de revenir à la taille normale.

```
-----
» zoom
-----
```

Un autre outil géométrique très utile pour détecter les niveaux de gris dans une image est la fonction de profil. On exécute cette fonction, on choisit à l'aide de la souris une ligne de l'image et on obtient le profil du niveau de gris le long de cette ligne.

```
-----
» img=imread('bacteria.tif');
» figure;imshow(img);
» improfile
-----
```

### B.3.5. Détection de contours

La détection de contours permet de repérer les différents objets qui constituent la scène de l'image. Il existe de nombreuses méthodes pour trouver les contours des objets, la plupart sont basées sur les dérivées premières et secondes de l'image.

```
-----
» img = imread('rice.tif');
» cont1 = edge(img,'prewitt');
» cont2 = edge(img,'canny');
» figure; imshow(img);
» figure; imshow(cont1);
» figure; imshow(cont2);
-----
```

### B.3.6. Histogramme - seuillage

L'histogramme d'une image donne la répartition de ses niveaux de gris. Ainsi pour une image qui possède 256 niveaux de gris, l'histogramme représente le niveau de gris en fonction du nombre de pixels à ce niveau de gris dans l'image.

```
-----
» img = imread('rice.tif');
» histo = imhist(img,256);
» figure;plot(histo);
-----
```

On sait que les niveaux de gris à zéro correspondent au noir et que les niveaux de gris à 1 indiquent le blanc. L'histogramme donne donc une excellente idée de la séparation entre quelque chose qui est clair et quelque chose qui est foncé dans l'image. Typiquement, une utilisation de ce fait est le seuillage d'une image, ce terme désigne la définition d'un seuil au-dessus ou en-dessous duquel on va garder certaines valeurs de niveaux de gris.

```
-----
» img=imread('saturn.tif');
» figure;imshow(img);
» img=im2double(img);
» figure;subplot(1,2,1);imshow(img);
» result=(img>0.5).*img;
» subplot(1,2,2);imshow(result);
-----
```

### B.3.7. Espace de couleurs

La couleur est une donnée importante pour une image, elle modifie la perception que l'on a de l'image. L'espace de représentation standard décompose une image en trois plans de couleur: le rouge, le vert et le bleu (Red/Green/Blue RGB en anglais). Les couleurs finales sont obtenues par synthèse additive de ces trois couleurs primaires. Il existe cependant des problèmes qui peuvent nécessiter de changer d'espace de couleur pour percevoir différemment l'image. Il y a des images où la couleur importe peu, par exemple des photographies de cellules vivantes (pseudo-transparentes), des images radar, des images satellites... Dans ce cas, l'espace RGB n'est plus utilisé. On lui préfère d'autres espaces comme HSV Hue/Saturation/Value ou YCbCr Luminance/Chrominance bleue/Chrominance rouge.

```
-----
» img=imread('blood1.tif');
» figure;imshow(img)
» colorbar
» colormap(hot)
» colormap(hsv)
» colormap(gray)
-----
```

```
» colormap(bone)
» colormap(copper)
» colormap(pink)
» colormap(white)
...
» help graph3d
```

-----

La boîte à outils images de Matlab gère les espaces de couleur RGB, HSV, YCbCr, NTSC. Un exemple ci-dessous permet de se rendre compte de l'utilisation d'un changement d'espace de couleur. On lit une image colorée, on la passe dans l'espace HSV.

```
-----
» img=imread('flowers.tif');
» figure;imshow(img);
» imgHSV=rgb2hsv(img);
» figure;imshow(imgHSV);
-----
```

Une fonction utile à connaître est la fonction de passage d'une image couleur à une image en niveaux de gris. Cette fonction **rgb2gray** permet d'obtenir une image en niveaux de gris en partant d'une image couleur. Dans certains cas, cette transformation est très utile.

```
-----
» img=imread('blood1.tif');
» figure;imshow(img)
» nivgris=rgb2gray(img);figure;imshow(nivgris);
-----
```

---

## Annexe C

### Les deux approches de fusion de scores pour trois méthodes de normalisation de scores pour les scores de la base de données XM2VTS

Les tableaux C.1, C.2, C.3, C.4, C.5, C.6 et C.7 présentent les performances en terme de taux d'égale erreur TEE dans l'ensemble d'évaluation et en terme taux de réussite TR dans l'ensemble de test pour les 15 coupe des systèmes voix et visage (voix , visage ) de la base de scores XM2VTS. Les méthodes de normalisation de scores utilisée sont : Minmax, fonction quadratique-linéaire-quadratique (QLQ) et la fonction double sigmoïde, et les méthodes de fusion sont : la somme pondérée, l'intégrale floue de Sugeno et de Choquet, la méthode de Fisher (LDF), la méthode de Fisher (QDF), la méthode à base de classifieur (SVM) et méthode à base de classifieur (MLP).

Méthodes  Parole ,Visage	Normalisation Z-Score		Normalisation par une fonction QLQ		Normalisation par une fonction double sigmoïde		La meilleur système unimodale	
	TEE%	TR %	TEE%	TR %	TEE%	TR %	TEE%	TR %
(LFCC, GMM),(FH, MLP)	0.18	98.91	0.15	98.99	<b>0.16</b>	<b>99.7</b>	0.84	97.7
(LFCC, GMM),(DCTs, GMM)	0.17	98.99	0.19	98.66	<b>0.2</b>	<b>99.39</b>	0.84	97.7
(LFCC, GMM),(DCTb, GMM)	0.16	99.21	0.12	99.07	<b>0.16</b>	<b>99.21</b>	0.84	97.7
(LFCC, GMM), (DCTs, MLP)	0.11	98.84	0.14	98.76	<b>0.04</b>	<b>99.31</b>	0.84	97.7
(LFCC, GMM), (DCTb, MLP)	<b>0.32</b>	<b>99.03</b>	0.32	98.76	0.34	98.95	0.84	97.7
(PAC, GMM),(FH, MLP)	0.68	98.03	<b>0.67</b>	98.01	1.46	<b>98.6</b>	1.66	96.25
(PAC, GMM),(DCTs, GMM)	1.35	96.29	1.28	97.28	<b>0.68</b>	<b>98.79</b>	4.37	91.55
(PAC, GMM),(DCTb, GMM)	<b>0.82</b>	<b>96.69</b>	1.65	95.1	1.35	96.07	1.53	96.62
(PAC, GMM), (DCTs, MLP)	<b>0.61</b>	<b>98.23</b>	0.62	98.05	0.69	97.78	2.96	93.32
(PAC, GMM), (DCTb, MLP)	1.32	93.75	1.29	94.28	<b>0.61</b>	<b>98.55</b>	4.82	87.58
(SSC, GMM),(FH, MLP)	0.48	98.65	<b>0.47</b>	98.45	0.64	<b>98.77</b>	1.66	96.25
(SSC, GMM),(DCTs, GMM)	0.87	<b>97.8</b>	<b>0.84</b>	97.59	0.98	97.51	4.37	91.55
(SSC, GMM),(DCTb, GMM)	<b>0.64</b>	97.59	1.36	96.07	0.68	<b>98.57</b>	1.53	96.62
(SSC, GMM), (DCTs, MLP)	0.68	97.87	<b>0.62</b>	97.94	0.67	<b>98.38</b>	2.96	93.32
(SSC, GMM), (DCTb, MLP)	1.18	96.25	<b>1.15</b>	96.30	1.36	<b>96.55</b>	4.82	87.58

Tableau C.1 : Performances de combinaison de scores par la somme pondérée pour les 15 couples des systèmes (voix, visage) dans d'ensemble d'évaluation et test.

Méthodes Parole ,Visage	Normalisation Z-Score		Normalisation par une fonction QLQ		Normalisation par une fonction double sigmoïde		La meilleur système unimodale	
	TEE%	TR %	TEE%	TR %	TEE%	TR %	TEE%	TR %
(LFCC, GMM),(FH, MLP)	0.45	98.92	<b>0.001</b>	99.72	0.002	<b>99.99</b>	0.84	97.7
(LFCC, GMM),(DCTs, GMM)	0.77	98.02	0.09	99.71	<b>0.001</b>	<b>99.99</b>	0.84	97.7
(LFCC, GMM),(DCTb, GMM)	0.27	98.87	0.08	99.73	<b>0.007</b>	<b>99.93</b>	0.84	97.7
(LFCC, GMM), (DCTs, MLP)	0.65	98.07	<b>0.01</b>	<b>99.96</b>	0.03	99.91	0.84	97.7
(LFCC, GMM), (DCTb, MLP)	1.39	96.65	<b>0.02</b>	<b>99.88</b>	0.04	99.88	0.84	97.7
(PAC, GMM),(FH, MLP)	1.4	96.6	<b>0.05</b>	99.16	0.17	<b>99.49</b>	1.66	96.25
(PAC, GMM),(DCTs, GMM)	2.02	94.66	<b>0.14</b>	98.64	0.24	<b>98.69</b>	4.37	91.55
(PAC, GMM),(DCTb, GMM)	1.72	94.88	0.08	99.12	<b>0.03</b>	<b>99.39</b>	1.53	96.62
(PAC, GMM), (DCTs, MLP)	1.65	94.88	<b>0.02</b>	99.45	0.04	<b>99.66</b>	2.96	93.32
(PAC, GMM), (DCTb, MLP)	2.57	92.41	<b>0.27</b>	97.95	0.63	<b>99.21</b>	4.82	87.58
(SSC, GMM),(FH, MLP)	1.13	97.70	0.17	99.49	<b>0.09</b>	<b>99.72</b>	1.66	96.25
(SSC, GMM),(DCTs, GMM)	1.60	95.28	0.33	99.28	<b>0.28</b>	<b>99.48</b>	4.37	91.55
(SSC, GMM),(DCTb, GMM)	1.07	95.69	0.04	99.60	<b>0.03</b>	<b>99.64</b>	1.53	96.62
(SSC, GMM), (DCTs, MLP)	1.42	95.63	<b>0.13</b>	<b>99.60</b>	0.2	99.40	2.96	93.32
(SSC, GMM), (DCTb, MLP)	2.26	93.21	<b>0.27</b>	<b>98.76</b>	0.45	97.65	4.82	87.58

Tableau C.2 : Performances de combinaison de scores par la mesure floues et intégrale floue de Sugeno pour les 15 couples des systèmes (voix, visage) dans d'ensemble d'évaluation et test.

Méthodes Parole ,Visage	Normalisation Z-Score		Normalisation par une fonction QLQ		Normalisation par une fonction double sigmoïde		La meilleur système unimodale	
	TEE%	TR %	TEE%	TR %	TEE%	TR %	TEE%	TR %
(LFCC, GMM),(FH, MLP)	0.33	99.12	<b>0.14</b>	99.31	0.28	<b>99.62</b>	0.84	97.7
(LFCC, GMM),(DCTs, GMM)	0.29	<b>99.22</b>	0.2	98.66	<b>0.20</b>	99.09	0.84	97.7
(LFCC, GMM),(DCTb, GMM)	0.19	99.39	0.46	98.60	<b>0.14</b>	<b>99.5</b>	0.84	97.7
(LFCC, GMM), (DCTs, MLP)	0.29	98.91	0.59	98.70	<b>0.13</b>	<b>99.58</b>	0.84	97.7
(LFCC, GMM), (DCTb, MLP)	0.84	97.93	0.61	97.81	<b>0.54</b>	<b>99.04</b>	0.84	97.7
(PAC, GMM),(FH, MLP)	0.790	98.19	<b>0.63</b>	98.53	1.69	<b>99.00</b>	1.66	96.25
(PAC, GMM),(DCTs, GMM)	<b>1.18</b>	97.08	1.33	95.22	1.36	<b>97.28</b>	4.37	91.55
(PAC, GMM),(DCTb, GMM)	0.86	96.70	1.66	94.35	<b>0.82</b>	<b>96.75</b>	1.53	96.62
(PAC, GMM), (DCTs, MLP)	0.64	98.24	0.54	98.19	<b>0.54</b>	<b>98.67</b>	2.96	93.32
(PAC, GMM), (DCTb, MLP)	1.48	95.06	1.34	94.31	<b>1.33</b>	<b>96.87</b>	4.82	87.58
(SSC, GMM),(FH, MLP)	<b>0.47</b>	<b>98.73</b>	0.48	98.48	0.99	98.41	1.66	96.25
(SSC, GMM),(DCTs, GMM)	0.95	97.79	<b>0.83</b>	<b>97.84</b>	1.2	96.57	4.37	91.55
(SSC, GMM),(DCTb, GMM)	<b>0.62</b>	<b>97.91</b>	1.34	96.18	0.97	97.45	1.53	96.62
(SSC, GMM), (DCTs, MLP)	0.7	97.85	<b>0.70</b>	<b>98.52</b>	0.79	97.41	2.96	93.32
(SSC, GMM), (DCTb, MLP)	1.38	95.64	<b>1.32</b>	<b>96.35</b>	1.49	96.00	4.82	87.58

Tableau C.3 : Performances de combinaison de scores par la mesure floues et intégrale floue de Choquet pour les 15 couples des systèmes (voix, visage) dans d'ensemble d'évaluation et test.

Méthodes Parole ,Visage	Normalisation Z-Score		Normalisation par une fonction QLQ		Normalisation par une fonction double sigmoïde		La meilleur système unimodale	
	TEE%	TR %	TEE%	TR %	TEE%	TR %	TEE%	TR %
(LFCC, GMM),(FH, MLP)	<b>0.2</b>	<b>99.11</b>	0.7	97.79	0.54	98.45	0.84	97.7
(LFCC, GMM),(DCTs, GMM)	<b>0.15</b>	98.75	0.45	98.41	0.18	<b>99.21</b>	0.84	97.7
(LFCC, GMM),(DCTb, GMM)	0.28	98.85	0.69	98.06	<b>0.16</b>	<b>99.20</b>	0.84	97.7
(LFCC, GMM), (DCTs, MLP)	<b>0.13</b>	98.79	0.70	97.79	0.17	<b>99.15</b>	0.84	97.7
(LFCC, GMM), (DCTb, MLP)	<b>0.28</b>	98.84	0.78	97.79	0.35	<b>99.19</b>	0.84	97.7
(PAC, GMM),(FH, MLP)	2.34	92.81	5.63	87.83	4.5	89.42	<b>1.66</b>	<b>96.25</b>
(PAC, GMM),(DCTs, GMM)	2.19	92.39	3.2	91.09	<b>1.36</b>	<b>95.82</b>	4.37	91.55
(PAC, GMM),(DCTb, GMM)	3.47	90.36	5.37	88.29	1.69	93.04	<b>1.53</b>	<b>96.62</b>
(PAC, GMM), (DCTs, MLP)	3.16	91.17	5.62	87.85	<b>2.68</b>	92.01	2.96	<b>93.32</b>
(PAC, GMM), (DCTb, MLP)	3.52	89.81	5.46	88.01	<b>1.63</b>	<b>94.29</b>	4.82	87.58
(SSC, GMM),(FH, MLP)	0.66	98.53	<b>0.36</b>	91.06	4.21	91.73	1.66	<b>96.25</b>
(SSC, GMM),(DCTs, GMM)	<b>1.01</b>	<b>96.67</b>	3.46	92.53	2.13	94.27	4.37	91.55
(SSC, GMM),(DCTb, GMM)	<b>1.51</b>	95.83	4.31	91.18	3.01	93.20	1.53	<b>96.62</b>
(SSC, GMM), (DCTs, MLP)	<b>1.03</b>	<b>96.81</b>	4.35	91.08	3.37	92.79	2.96	93.32
(SSC, GMM), (DCTb, MLP)	<b>1.49</b>	<b>95.92</b>	4.35	91.08	2.61	93.87	4.82	87.58

Tableau C.4 : Performances de classification de scores par la méthode de Fisher (LDF) pour les 15 couples des systèmes (voix, visage) dans d'ensemble d'évaluation et test.

Méthodes Parole ,Visage	Normalisation Z-Score		Normalisation par une fonction QLQ		Normalisation par une fonction double sigmoïde		La meilleur système unimodale	
	TEE%	TR %	TEE%	TR %	TEE%	TR %	TEE%	TR %
(LFCC, GMM),(FH, MLP)	0.64	98.54	10.83	82.48	<b>0.19</b>	<b>99.65</b>	0.84	97.7
(LFCC, GMM),(DCTs, GMM)	<b>0.19</b>	98.69	10.70	81.67	0.82	<b>99.06</b>	0.84	97.7
(LFCC, GMM),(DCTb, GMM)	<b>0.33</b>	98.97	10.84	82.71	0.51	<b>99.21</b>	0.84	97.7
(LFCC, GMM), (DCTs, MLP)	<b>0.21</b>	98.92	10.53	82.50	0.45	<b>99.20</b>	0.84	97.7
(LFCC, GMM), (DCTb, MLP)	<b>0.51</b>	<b>98.81</b>	10.83	82.50	1.95	97.99	0.84	97.7
(PAC, GMM),(FH, MLP)	<b>0.83</b>	<b>98.04</b>	16.12	65.25	2.37	93.25	1.66	96.25
(PAC, GMM),(DCTs, GMM)	1.53	94.18	15.87	65.39	<b>1.45</b>	<b>97.26</b>	4.37	91.55
(PAC, GMM),(DCTb, GMM)	1.65	93.08	15.99	64.80	<b>1.13</b>	94.50	1.53	<b>96.62</b>
(PAC, GMM), (DCTs, MLP)	<b>0.54</b>	<b>97.07</b>	15.85	65.63	0.82	94.18	2.96	93.32
(PAC, GMM), (DCTb, MLP)	<b>1.36</b>	93.16	15.79	65.47	1.45	<b>96.07</b>	4.82	87.58
(SSC, GMM),(FH, MLP)	<b>0.99</b>	<b>98.19</b>	14.51	74.28	3.96	39.06	1.66	96.25
(SSC, GMM),(DCTs, GMM)	<b>0.87</b>	<b>97.77</b>	14.45	75.15	2.85	45.60	4.37	91.55
(SSC, GMM),(DCTb, GMM)	<b>0.71</b>	<b>97.47</b>	14.52	74.25	3.2	43.73	1.53	96.62
(SSC, GMM), (DCTs, MLP)	<b>0.87</b>	<b>98.11</b>	14.49	74.31	3.14	43.09	2.96	93.32
(SSC, GMM), (DCTb, MLP)	<b>1.15</b>	<b>96.27</b>	14.47	74.34	3.12	42.89	4.82	87.58

Tableau C.5 : Performances de classification de scores par la méthode de Fisher (QDF) pour les 15 couples des systèmes (voix, visage) dans d'ensemble d'évaluation et test.

Méthodes Parole ,Visage	Normalisation Z-Score		Normalisation par une fonction QLQ		Normalisation par une fonction double sigmoïde		La meilleur système unimodale	
	TEE%	TR %	TEE%	TR %	TEE%	TR %	TEE%	TR %
(LFCC, GMM),(FH, MLP)	0.18	99.18	0.30	99.10	<b>0.17</b>	<b>99.67</b>	0.84	97.7
(LFCC, GMM),(DCTs, GMM)	<b>0.15</b>	98.76	0.18	98.70	0.19	<b>99.14</b>	0.84	97.7
(LFCC, GMM),(DCTb, GMM)	0.15	98.92	0.19	98.85	<b>0.14</b>	<b>99.51</b>	0.84	97.7
(LFCC, GMM), (DCTs, MLP)	0.04	98.83	0.17	98.97	<b>0.03</b>	<b>99.30</b>	0.84	97.7
(LFCC, GMM), (DCTb, MLP)	0.33	99.01	0.53	98.76	<b>0.33</b>	<b>99.24</b>	0.84	97.7
(PAC, GMM),(FH, MLP)	<b>0.78</b>	98.01	0.81	98.06	0.83	<b>99.06</b>	1.66	96.25
(PAC, GMM),(DCTs, GMM)	<b>1.31</b>	97.08	2.15	94.87	1.49	<b>97.36</b>	4.37	91.55
(PAC, GMM),(DCTb, GMM)	<b>0.78</b>	97.82	1.12	96.97	0.81	<b>98.78</b>	1.53	96.62
(PAC, GMM), (DCTs, MLP)	<b>0.51</b>	<b>98.50</b>	0.61	98.42	0.61	98.43	2.96	93.32
(PAC, GMM), (DCTb, MLP)	<b>1.28</b>	95.44	1.32	95.37	1.85	<b>96.25</b>	4.82	87.58
(SSC, GMM),(FH, MLP)	0.46	98.72	<b>0.45</b>	<b>98.74</b>	0.46	98.72	1.66	96.25
(SSC, GMM),(DCTs, GMM)	0.95	97.72	<b>0.87</b>	<b>97.77</b>	0.95	97.73	4.37	91.55
(SSC, GMM),(DCTb, GMM)	0.61	98.29	0.96	97.23	<b>0.54</b>	<b>98.36</b>	1.53	96.62
(SSC, GMM), (DCTs, MLP)	<b>0.52</b>	<b>98.02</b>	0.63	97.93	0.67	97.84	2.96	93.32
(SSC, GMM), (DCTb, MLP)	1.31	96.46	<b>1.12</b>	96.32	1.2	<b>96.70</b>	4.82	87.58

Tableau C.6 : Performances de classification de scores par SVM pour les 15 couples des systèmes (voix, visage) dans d'ensemble d'évaluation et test.

Méthodes Parole ,Visage	Normalisation Z-Score		Normalisation par une fonction QLQ		Normalisation par une fonction double sigmoïde		La meilleur système unimodale	
	TEE%	TR %	TEE%	TR %	TEE%	TR %	TEE%	TR %
(LFCC, GMM),(FH, MLP)	<b>0.08</b>	99.50	0.25	99.01	0.23	<b>99.57</b>	0.84	97.7
(LFCC, GMM),(DCTs, GMM)	0.19	98.64	0.19	98.63	<b>0.19</b>	<b>99.12</b>	0.84	97.7
(LFCC, GMM),(DCTb, GMM)	<b>0.18</b>	<b>99.14</b>	0.40	98.51	0.22	99.04	0.84	97.7
(LFCC, GMM), (DCTs, MLP)	0.05	98.79	0.06	98.76	<b>0.05</b>	<b>99.28</b>	0.84	97.7
(LFCC, GMM), (DCTb, MLP)	0.36	98.92	0.55	98.72	<b>0.25</b>	<b>98.97</b>	0.84	97.7
(PAC, GMM),(FH, MLP)	0.58	98.35	<b>0.47</b>	<b>98.40</b>	0.81	98.10	1.66	96.25
(PAC, GMM),(DCTs, GMM)	1.22	96.27	1.42	97.20	<b>1.11</b>	<b>97.37</b>	4.37	91.55
(PAC, GMM),(DCTb, GMM)	<b>0.64</b>	97.95	1.18	96.38	0.69	<b>98.39</b>	1.53	96.62
(PAC, GMM), (DCTs, MLP)	0.51	97.51	0.56	96.82	<b>0.38</b>	<b>98.36</b>	2.96	93.32
(PAC, GMM), (DCTb, MLP)	1.22	94.16	1.19	94.22	<b>1.13</b>	<b>96.30</b>	4.82	87.58
(SSC, GMM),(FH, MLP)	0.48	98.67	0.51	98.38	<b>0.43</b>	<b>98.86</b>	1.66	96.25
(SSC, GMM),(DCTs, GMM)	0.86	96.98	0.88	<b>97.76</b>	<b>0.74</b>	97.20	4.37	91.55
(SSC, GMM),(DCTb, GMM)	0.52	<b>98.11</b>	0.88	97.27	<b>0.41</b>	97.62	1.53	96.62
(SSC, GMM), (DCTs, MLP)	0.64	97.73	<b>0.42</b>	97.81	0.59	<b>98.31</b>	2.96	93.32
(SSC, GMM), (DCTb, MLP)	1.01	96.36	1.4	96.52	<b>0.99</b>	<b>97.18</b>	4.82	87.58

Tableau C.6 : Performances de classification de scores par MLP pour les 15 couples des systèmes (voix, visage) dans d'ensemble d'évaluation et test.

**Bibliographie**

- [1] Alphonse Bertillon. *Ethnographie moderne : les races sauvages*. 1883.
- [2] Patrick Verlinde, Gérard Chollet, and Marc Acheroy. Multi-modal identity verification using expert fusion. *Information Fusion*, pp:17-33, 2000.
- [3] E. S. Bigün, Josef Bigün, Benoît Duc, and S. Fischer. Expert conciliation for multi modal person authentication systems by bayesian statistics. In *AVBPA 97 : Proceedings of the First International Conference on Audio and Video-Based Biometric Person Authentication*, pp: 291-300, London, UK, 1997. Springer-Verlag.
- [4] Josef Kittler, Mohamad Hatef, Robert P. W. Duin, and Jiri Matas. On combining classifiers. *IEEE Trans. Pattern Anal. Mach. Intell.*, pp: 226-239, 1998.
- [5] S. Ben-Yacoub, Yousri Abdeljaoued, and E. Mayoraz. Fusion of face and speech data for person identity verification. *IEEE Transactions on Neural Networks*, pp: 1065-1074, 1999.
- [6] Julian Fierrez-Aguilar. *Adapted Fusion Schemes for Multimodal Biometric Authentication*. PhD thesis, Universidad Politecnica de Madrid, 2006.
- [7] Norman Poh. *Multi-system biometric authentication*. PhD thesis, EPFL, Lausanne, 2006.
- [8] Jonas Richiardi. *Probabilistic models for multi-classifier biometric authentication using quality measures*. PhD thesis, EPFL, Lausanne, 2007.
- [9] Karthik Nandakumar. *Multibiometric Systems : Fusion Strategies and Template Security*. PhD thesis, Michigan State University, 2008.
- [10] Lorène ALLANO. *La Biométrie multimodale : stratégies de fusion de scores et mesures de dépendance appliquées aux bases de personnes virtuelles*. Thèse présentée pour l'obtention du grade de Docteur de l'Institut National des Telecommunications, soutenue le 12 janvier 2009.
- [11] Nicolas MORIZET. *Reconnaissance Biométrique par Fusion Multimodale du Visage et de l'Iris*. Thèse présentée pour obtenir le grade de Docteur de l'Ecole Nationale Supérieure des Télécommunications Spécialité : Signal et Images , Soutenue le 18 Mars 2009.
- [12] R. Hietmeyer. Biometric identification promises fast and secure processing of airline passengers. *The International Civil Aviation Organization Journal*, 17(9), pp: 10-11, 2000.
- [13] Machine Readable Travel Documents (MRTD). 2008. <http://www2.icao.int/en/mrtd/Pages/default.aspx>
- [14] S. Lin, S. Kung, and L. Lin. Face Recognition/Detection by Probabilistic Decision-Based Neural-Network. *IEEE Transactions on Neural Networks*, 8(1), pp: 114-132, January 1997.
- [15] A. Nefian and M. Hayes. Face Detection and Recognition Using Hidden Markov Models. In *International Conference on Image Processing (ICIP)*, pp: 141-145, 1998.
- [16] F. Smach, M. Atri, J. Miteran, and M. Abid. Design of a Neural Networks Classifier for Face Detection. *Journal of Computer Science, Sciences Publications*, 2(3), pp: 261-268, 2006.

- [17] F. Yang and M. Paindavoine. Implementation of a RBF neural network on embedded systems : Real time face tracking and identity verification. *IEEE Transactions on Neural Networks*, 14(5), pp:1162-1175, September 2003.
- [18] G. Bailly, F. Elisel, P. Badin, and C. Savariaux. Degrees of freedom of facial movements in face-to-face conversational speech. In : *International Workshop on Multimodal Corpora*, pp: 33-36, 2006.
- [19] L. Girin. Joint matrix quantization of face parameters and LPC coefficients for low bit rate audiovisual speech coding. *IEEE Transactions on Speech and Audio Processing*, Vol. 12, No. 3, pp: 265-276, 2004.
- [20] M. Kirby and L. Sirovich. Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 1, 1990.
- [21] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, pp: 788-791, October 1999.
- [22] M. Bartlett, J. Movellan, and T. Sejnowski. Face recognition by independent component analysis. *IEEE Transactions on Neural Networks*, 13(6), pp: 1450-1464, 2002.
- [23] Young-Gil Kim , Young-Jun Song , Un-Dong Chang , Dong-Woo Kim , Tae-Sung Yun , Jae-Hyeong Ahn, Face recognition using a fusion method based on bidirectional 2DPCA, 2008 Published by Elsevier, *Applied Mathematics and Computation* ,pp: 601–607, 2008.
- [24] P. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs. Fisherfaces : Recognition Using Class Specific Linear Projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19, pp: 711-720, 1997
- [25] X.F. Gao, S.Z. Li, R. Liu, and P.R. Zhang. Standardization of face image sample quality. In *ICB07*, pp: 242-251, 2007.
- [26] Hua Yu and Jie Yang. A direct lda algorithm for high-dimensional data with application to face recognition. *Pattern Recognition*, pp :2067- 2070, 2001.
- [27] J. Lu, K.N. Plataniotis, and A.N. Venetsanopoulos. Regularization studies of linear discriminant analysis in small sample size scenarios with application to face recognition. *PRL*, 26(2), pp :181-191, January 2005.
- [28] B. Schölkopf, A. Smola, and K.-R. Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, Vol. 10, No. 5, pp: 1299- 1319, 1998.
- [29] S. Mika, G. Ratsch, J. Weston, B. Schölkopf, and K.-R. Müller. Fisher Discriminant Analysis With Kernels. In : *Neural Networks for Signal Processing IX*, pp:41-48, 1999.
- [30] F.R. Bach and M.I. Jordan. Kernel independent component analysis. *Journal of Machine Learning Research*, pp: 1-48, 2002.
- [31] T. Kanade. Picture Processing System by Computer Complex and Recognition of Human Faces. In : *Doctoral dissertation, Kyoto University*, November 1973.
- [32] C. Liu and H. Wechsler. A Unified Bayesian Framework for Face Recognition. In : *Proceedings of the IEEE International Conference on Image Processing*, pp: 151-155, 1998.

- [33]G. Guo, S. Li, and K. Chan. Face Recognition by Support Vector Machines.In :Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, pp: 196-201, 2000.
- [34]T. Cootes, G. Edwards, and C. Taylor. Active Appearance Models.IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 6,pp: 681-685, June 2001.
- [35]T. Ahonen, A. Hadid, and M. Pietikainen. Face Recognition with Local Binary Patterns. 2004.
- [36] S. Arca, P. Campadelli, and R. Lanzarotti. A Face Recognition System Based On Automatically Determined Facial Fiducial Points. Pattern Recognition, Vol. 39, No. 3, pp: 432-443, 2006.
- [37]P. Penev and J. Atick. Local features analysis : A general statistical theory for object representation. Neural Systems, Vol. 7, No. 3, pp: 477-500, 1996.
- [38]L. Wiskott, J. M. Fellous, N. Kruger et C. von der Malsburg. Face Recognition by Elastic Bunch Graph Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7), pp :775-779,1997.
- [39]V. Perlibakas. Face recognition using Principal Component Analysis and Log-Gabor Filters. March 2005.
- [40]Muriel VISANI. Vers de nouvelles approches discriminantes pour la reconnaissance de visages. pour obtenir le grade de Docteur de l'Institut National des Sciences Appliquées de Lyon Spécialité Informatique ;soutenue le 25 Novembre 2005.
- [41] Djamel Saigaa, Contribution à l'authentification d'individus par reconnaissance de visages, thèse de Doctorat d'état en automatique, Université Mohamed khider Biskra Algérie, Novembre 2006.
- [42]Anour Mellakh. Reconnaissance des visages en conditions dégradées. THÈSE pour l'obtention du grade de Docteur de l'Institut National des Télécommunications Spécialité : Traitement d'image et de signal, soutenue 07 Avril 2009.
- [43]Ruicong Zhi , Qiuqi Ruan. Two-dimensional direct and weighted linear discriminant analysis for face recognition. ELSEVIER Neurocomputing.(71),pp: 3607-3611,2008.
- [44]W. Zhao ,R. Chellappa et A. Krishnaswamy. Discriminant of Principal Components for Face Recognition .Dans Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition (FGR) , pp: 336-341,1998.
- [45]M.S. Bartlett, J.R. Movellan et T.J. Sejnowski. Face Recognition by Independent Component Analysis . IEEE Transactions on Neural Networks , 13(6), pp :1450-1464,2002
- [46]M.H. Yang. Kernel Eigenfaces vs. Kernel Fisherfaces : Face recognition Using Automatic Face and Gesture Recognition (FGR),pp: 215-220,2002.
- [47]M.A. Turk et A.D. Pentland. Eigenfaces for Recognition .Journal of Cognitive Neuroscience ,3,pp :71-86,1991.
- [48]W. Zhao. Robust Image Based 3 D Face Recognition . PhD thesis , University of Maryland, 1999 .

- [49] M.H. Yang. Kernel Eigenfaces vs. Kernel Fisherfaces : Face recognition Using Automatic Face and Gesture Recognition (FGR), pp: 215-220, 2002.
- [50] B. Moghaddam. Principal Manifolds and Probabilistic Subspaces for Visual Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(6), pp: 780-788, 2002.
- [51] k. Jonsson J. Kittler, Y. Li et J. Matas. Learning Support International Conference on Automatic Face and Gesture Recognition (FGR) , pp: 208-213, 2000.
- [52] S. H Lin, S. Y. Kung et L. J. Lin. Face Recognition /Detection by Probabilistic Decision based Neural Networks. IEEE Transactions on Neural Networks, 8, pp: 114-132, 1997.
- [53] Hough P.V.C. Methods and Means for Recognizing Complex Patterns .No.3, December 1962.
- [54] Rosenfield A. Picture Processing by Computer. Academic, New York, 1969
- [55] Duda R.O, Hart P.E. Use of the Hough transformation to detect lines and curves in pictures. Communication of ACM, 15, pp: 11-15, January 1972.
- [56] BOUTARFA Abdelhalim, Reconnaissance de formes 3D par approche neuronale associant la transformée de Hough en robotique mobile application à la productique. THESE Présentée au Département d'Electronique Faculté des Sciences de l'Ingénieur Université de Batna Pour l'obtention du diplôme de Doctorat Es-Sciences en Electronique Industrielle Option : Contrôle soutenue 2006.
- [57] Walid Hizem . Capteur Intelligent pour la Reconnaissance de Visage. Thèse Doctorat délivré conjointement par l'Institut National des Télécommunications et l'Université Pierre et Marie Curie - Paris 6 Spécialité : Electronique/Informatique , Soutenue le 2009.
- [58] David S. Bolme , Elastic Bunch Graph Matching. Master of Science Colorado State University Fort Collins, Colorado Summer 2003.
- [59] Daniel L. Ruderman. The statistics of natural images. Network : Computation in Neural Systems, pp: 517-548, 1994.
- [60] P.S. Penev and J.J. Atick. Local feature analysis : A general statistical theory for object representation. Network : Computation in Neural Systems, 7(3), pp : 477-500, 1996.
- [61] D. Swets and J. Weng. Using discriminant eigenfeatures for image retrieval. IEEE Trans. On Pattern Analysis and Machine intelligence, 18(8), 1996.
- [62] W. Zhao, R. Chellappa, and A. Krishnaswamy. Discriminant analysis of principal components for face recognition. In Proc. of Int. Conf. on Automatic Face and Gesture Recognition, pp: 336- 341, 1998.
- [63] H. Moon and P. J. Phillips. Analysis of PCA-based face recognition algorithms. In Empirical Evaluation Techniques in Computer Vision, 1998.
- [64] T. P. Minka. Automatic choice of dimensionality of pca. Technical Report TR 514, MIT Media Laboratory Perceptual computing section, 2000.
- [65] T. Cootes, G. Edwards, and C. Taylor. Active appearance models. In Proceedings of the European Conference on Computer Vision, 1998.

- [66] D. Gorodnichy. Facial recognition from video. In Proc. of Int. Conf. on Audio-and Videobased Person Authentication, 2003.
- [67] L.F. Chen, H.Y.M. Liao, M.T. Ko, J.C. Lin, and G.J. Yu. A new lda-based face recognition system which can solve the small sample size problem. Pattern Recognition, 33(10), pp:1713-1726, October 2000.
- [68] R.A Fisher. The use of multiple measures in taxonomic problems. Ann. Eugenics, 7, pp :179-188, 1936.
- [70] C. Liu and H.Wechsler, Robust coding schemes for indexing and retrieval from large face databases, IEEE Trans. on Image Processing, 9(1), pp: 132–137, 2000.
- [71] K. Fukunaga, Introduction to Statistical Pattern Recognition, Academic Press, second edition, 1991
- [72] Chengjun Liu and Harry Wechsler , Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition, IEEE Trans. Image Processing, 11(4), pp:467-476, 2002.
- [73] Hua Yu and Jie Yang. A direct lda algorithm for high-dimensional data with application to face recognition. Pattern Recognition, 34(10), pp:2067- 2070, 2001.
- [74] J. Yang, D. Zhang et A.F. Frangi. Two-Dimensional PCA : A New Approach to AppearanceBased Face Representation and Recognition. IEEE Trans. OnPAMI, 26(1), pp:131-137, 2004.
- [75] M. Visani, C. Garcia et C. Laurent. Comparing Robustness of Two-Dimensional PCA and Eigenfaces for Face Recognition.DansProc. of the Int. Conf.on Image Analysis and Recognition, SpringerLNCS 3211, A. Campilho, M. Kamel,pp: 717–724, Octobre 2004.
- [76] Wonjun Hwang, Haitao Wang, Hyunwoo Kim, Member, IEEE, Seok-CheolKee, and Junmo Kim, Member, IEEE.Face Recognition System Using Multiple Face Model of Hybrid Fourier Feature under Uncontrolled Illumination Variation. IEEE TRANSACTIONS ON IMAGE PROCESSING 2011.
- [77] Turk M.A. and Pentland A.P. Eigenfaces for Recognition, Journal of Cognitive Neuroscience, 3(1):71-86, 1991.
- [78] Visani M. Garcia C. Jolion J.M. Two Dimensional-Oriented Discriminant Analysis for Face Recognition.In Proc.of the Int. Conf. On Computer Vision and Graphics,pp: 1008-1017, 2004.
- [79] Bengherabi M., Mezai L., Harizi F., Cheriet M., Guessoum A. Face recognition based on 2DPCA, DIAPCA and DIA2DPCA in DCT domain.The 5th International Multi-Conference on Systems, Signals and Devices, IEEE SSD, Amman, Jordan, July, pp: 20-23, 2008.
- [80] Zuo W., Wang K., Zhang D. Assembled Matrix Distance Metric for 2DPCA-based Face and Palmprint Recognition.Proceeding of 4th International Conference on MachineLearning and Cybernetics, Guangzhou,pp: 18-21, August 2005.
- [81] F. Perronnin and J.-L. Dugelay.Introduction à la biométrie Authentification des individus par traitement audio-vidéo. Traitement du signal, 19(4), 2002.

- [82] SOUHILA GUERFI ABABSA. Authentification d'individus par reconnaissance de caractéristiques biométriques liées aux visages 2D/3D. THÈSE pour obtenir le titre de Docteur de l'Université Evry Val d'Essonne Spécialité: Sciences de l'Ingénieur. Soutenue le 03 octobre 2008.
- [83] K. Messer, J. Matas, J. Kittler, J. Luetin, and G. Maitre. XM2VTSDB: The Extended M2VTS Database. In Proceedings, International Conference on Audio- and Video-Based Person Authentication, pp: 72-77, 1999.
- [84] Rafael C. Gonzalez and Richard E. Woods. Digital Image Processing (2<sup>nd</sup> Edition). Prentice Hall, January 2002.
- [85] Mohammad T. Sadeghi, Samaneh Khoshrou, and Josef Kittler. Confidence Based Gating of Colour Features for Face Authentication. Springer-Verlag Berlin Heidelberg, pp: 121-130, 2007.
- [86] Mohamad Hoseyn Sigari, Mahmood Fathy. Best Wavelength Selection for Gabor Wavelet using GA in EBGM Algorithm. Proceedings of the International MultiConference of Engineers and Computer Scientists Hong Kong, pp: 19-21 March, 2008.
- [87] A. V. Oppenheim and J. S. Lim. The importance of phase in signals. Proceedings of the IEEE, 69(5), pp: 529-541, May 1981.
- [88] Chengjun Liu. Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance. Pattern Analysis and Machine Intelligence, 28, pp: 725-737, May 2006.
- [89] L.L. Shen and L. Bai. Gabor feature based face recognition using kernel methods. In AFGR04, pp: 170-175, 2004.
- [90] Mian Zhou and Hong Wei. Face verification using gabor wavelets and adaboost. In IEEE Computer Society Proceedings of the 18th International Conference on Pattern Recognition, Washington, DC, USA, pp: 404-407, 2006.
- [91] Dijana Petrovska-Delacrétaz, Gérard Chollet, and Bernadette Dorizzi. Guide to Biometric Reference Systems and Performance Evaluation. Springer, 2009.
- [92] A. K. Jain and A. Ross. Multibiometric systems. Communications of the ACM, special issue on multimodal interfaces, 47(1), pp: 34-40, January 2004.
- [93] Y. Chen, S. Dass, and A. Jain. Fingerprint Quality Indices for Predicting Authentication Performance. In : Proceedings of Fifth International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), New York, NY, USA., pp: 160- 170, July 2005.
- [94] NIST report to the United States Congress. Summary of NIST Standards for Biometric Accuracy, Tamper Resistance, and Interoperability. November 2002. Available at [ftp://sequoyah.nist.gov/pub/nist\\_internal\\_reports/NISTAPP\\_Nov02.pdf](ftp://sequoyah.nist.gov/pub/nist_internal_reports/NISTAPP_Nov02.pdf).
- [95] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino. Impact of Artificial "Gummy" Fingers on Fingerprint Systems. In : Proceedings of SPIE : Optical Security and Counterfeit Deterrence Techniques IV, pp: 275-289, January 2002.

- [96] Xiao-Yuan Jing, Yong-Fang Yao, David Zhang, Jing-Yu Yang, and Miao Li. Face and palmprint pixel level fusion and kernel DCV-RBF classifier for small sample biometric recognition. *Pattern Recognition*, 40(11),pp :3209-3224, 2007.
- [97] Y.S. Huang and C.Y. Suen. A method of combining multiple experts for the recognition of unconstrained handwritten numerals. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(1),pp :90-94, Jan 1995.
- [98] Patrick Verlinde. A contribution to multi-modal identity verification using decision fusion.PhDthesis, Ecole Nationale des Télécommunications, France, 1999.
- [99] Haindl, M., Kittler, J., Roli, F. (eds.):MCS 2007. LNCS, vol. 4472. Springer, Heidelberg (2007)
- [100]Kittler, J., Hatef,M., Duin, R.P.,Matas, J.: On combining classifiers. *IEEE Trans. PatternAnalysis and Mach. Intell.* 20(3), pp: 226-239 ,1998.
- [101] Zhan Yu, Mi Young Nam, SumanSedai, and PhillKyu Rhee. Evolutionary Fusion of a Multi-Classifier System for Efficient Face Recognition.ICROS, KIEE and Springer. *International Journal of Control, Automation, and Systems* 7(1), pp:33-40. 2009.
- [102] Xingming Zhang, Dian Liu, and Jianfu Chen. An Illumination Independent Face Verification Based on Gabor Wavelet and Supported Vector Machine. Springer-Verlag Berlin Heidelberg, pp: 153–160, 2008.
- [103]Zeng-Shng Zhao, Ji- Zhen Wang,Qing-JiTian, Jian-Ming Zhou. Classifier Fusion of Improved Laplacianfaces. *IEEEProceedings of the Ninth International Conference on Machine Learning and Cybernetics*, Qingdao, pp:11-14 July 2010.
- [104] Y.Wang, T. Tan, and A. Jain.Combining face and iris biometrics for identityverification.In :Proceedings of Fourth International Conference on Audio- andVideo-Based Authentication (AVBPA), pp:805-813., June 2003.
- [105] P. Verlinde, P. Druyts, G. Cholet, and M. Acheroy. Applying Bayes basedclassifiers for decision fusion in a multi-modal identity verification system.In :Proceedings of International Symposium on Pattern Recognition “InMemoriamPierre Devijver”, Brussels, Belgium, February 1999.
- [106] V. Chatzis, A. Bors, and I. Pitas.Multimodal decision-level fusion for personauthentication.IEEE Transactions on Systems, Man and Cybernetics, Part A :Systems and Humans, Vol. 29, No. 6, pp. 674–681, November 1999.
- [107] C. Sanderson and K. Paliwal.Information fusion and person verificationusing speech and face information. Tech. Rep. IDIAP-RR 02-33, IDAIP, September2002.
- [108] A. Ross and A. Jain.Information fusion in biometrics.*Pattern RecognitionLetters*, Vol. 24, No. 13, pp. 2115–2125, 2003.
- [109] Kittler, J., Hatef,M., Duin, R.P.,Matas, J. On combining classifiers.IEEE Trans. *PatternAnalysis and Mach. Intell.* 20(3), 226–239 (1998)
- [110]Keun-Chang Kwak and WitoldPedrycz, Face recognition: A study in information fusion using fuzzy integral , Elsevier, 2004.

- [111] SarbastRasheed , Daniel W. Stashuk, Mohamed S. Kamel, Diversity-based combination of non-parametric classifiers for EMG signal decomposition, Springer-Verlag London Limited , pp: 385–408,2008.
- [112] Chaw Chia, Nasser Sherkat, and Lars Nolle. Confidence Partition and Hybrid Fusion in Multimodal Biometric Verification System. J. Fierrez et al. (Eds.): BioID\_MultiComm, LNCS 5707, pp: 212–219, Springer-Verlag Berlin Heidelberg 2009.
- [113] Yong Li, Jianping Yin, En Zhu, Chunfeng Hu, and Hui Chen, Studies of Fingerprint Template Selection and Update, T.-h. Kim et al. (Eds.): FGCN 2008 Workshops and Symposia, CCIS 28, pp:150–163, Springer-Verlag Berlin Heidelberg 2009.
- [114] Chapter 18 : Multimodal Biometric Systems, A. Shukla et al, Towards Hybrid and Adaptive Computing, SCI 307, pp: 401–418, Springer-Verlag Berlin Heidelberg 2010.
- [115]HaihongShen, Liqun Ma, Qishan Zhang. Multi-Modal Face Recognition.Proceedings of the 8<sup>th</sup> World Congress on Intelligent Control and Automation July 6-9 2010, Jinan, China .
- [116] Girija and Chetty, Member IEEE.Biometric Liveness Checking Using Multimodal Fuzzy Fusion.IEEE,2010.
- [117] Stéphane Pigeon, Authentification Mulimodale D'Identité. Thèse présentée en vue de l'obtention du grade de Docteur en Sciences Appliquées. Laboratoire de Télécommunications et Télédétection Place du Levant, 2 1348 Louvain-la-Neuve – Belgique , Février 1999
- [118] ZhiguoNiu and XuehongQiu. Facial Expression Recognition based on weighted principal component analysis and support vector machines. International Conference on Advanced Computer Theory and Engineering (ICACTE), 2010.
- [119]SitalakshmiVenkataraman, SiddhivinayakKulkarni, Risk-Based Neuro-Grid Architecture for Multimodal Biometrics, T. Sobh, K. Elleithy (eds.), Innovations in Computing Sciences and Software Engineering, Springer Science+Business Media B.V. 2010.
- [120]F. WANG and J. HAN, Multimodal biometric authentication based on score level fusion using support vector machine, Opto–Electron. 17(1), Springer 2009.
- [121] A. Jain, K. Nandakumar, and A. Ross. Score normalization in multimodalbiometric systems. Pattern Recognition, 38(12), pp: 2270–2285, December2005.
- [122] A. Ross and A. Jain.Information fusion in biometrics.Pattern RecognitionLetters, 24(13), pp: 2115–2125, 2003.
- [123] R. Brunelli and D. Falavigna.Person identification using multiple cues.IEEE Transactions on Pattern Analysis andMachine Intelligence, 17, pp: 955–966, 1995.
- [124] Sergey Tulyakov, Stefan Jaeger, and VenuGovindaraju, and David Doermann , Review of Classifier Combination Methods, Studies in Computational Intelligence (SCI) 90, 361–386, Springer-Verlag Berlin Heidelberg 2008 .
- [125]Arun A. Ross, KarthikNandakumar, and Anil K. Jain.Handbook of Multibiometrics.Springer, 2006.

- [126] R. Snelick and M. Indovina. Large-Scale Evaluation of Multimodal Biometric Authentication Using State-of-the-Art Systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 3, pp. 450–455, 2005.
- [127] R. Cappelli, D. Maio, and D. Maltoni. Combining Fingerprint Classifiers. In *Proceedings of the First International Workshop on Multiple Classifier Systems*, pp: 351–361, Springer-Verlag, London, UK, 2000.
- [128] F. WANG and J. HAN, Multimodal biometric authentication based on score level fusion using support vector machine, *Opto–Electron.* 17(1), Springer 2009.
- [129] F. WANG and J. HAN, Multimodal biometric authentication based on score level fusion using support vector machine, *Opto–Electron.* 17(1), Springer 2009.
- [130] L.A. Zadeh. Fuzzy sets, *Information Control*, Vol. 8, pp: 338–353, 1965.
- [131] M. Sugeno, Fuzzy measures and fuzzy integrals A survey, in *Fuzzy Automata and Decision Processes*, M. M. Gupta, G. N. Saridis, and B.R. Gaines, Eds. Amsterdam, The Netherlands: North Holland, pp: 89–102, 1977.
- [132] Keun-Chang Kwak, Member, IEEE, and Witold Pedrycz, Fellow, IEEE ; Face Recognition Using Fuzzy Integral and Wavelet Decomposition Method; *IEEE Transactions on Systems , Part B: Cybernetics*, 34(4), 2004
- [133] Anne M.P. Canuto and Marjory C.C. Abreu, Using Fuzzy, Neural and Fuzzy-Neural Combination Methods in Ensembles with Different Levels of Diversity, Springer-Verlag Berlin Heidelberg, Part I, LNCS 4668, pp: 349–359. 2007.
- [134] T. Murofushi and M. Sugeno, An interpretation of fuzzy measures and the Choquet integral as an integral with respect to a fuzzy measure. *Fuzzy Sets Syst.* vol. 29, pp: 201–227, 1988.
- [135] Patrick Verlinde. A contribution to multi-modal identity verification using decision fusion. PhD thesis, Ecole Nationale des Télécommunications, France, 1999.
- [136] B. Duc, E. S. Bigün, J. Bigün, G. Maître, and S. Fischer. Fusion of audio and video information for multi modal person authentication. *Pattern Recognition Letters*, 18(9), pp: 835–843, September 1997.
- [137] Vladimir N. Vapnik. *The nature of statistical learning theory*. Springer Verlag New York, Inc., New York, NY, USA, 1995.
- [138] Dzati Athiar Ramli, Salina Abdul Samad, and Aini Hussain. Score Information Decision Fusion Using Support Vector Machine for a Correlation Filter Based Speaker Authentication System. Springer-Verlag Berlin Heidelberg, pp: 235–242, 2009.
- [139] Pascual Ejarque, Javier Hernado, David Hernando, and David Gómez, Eigenfeatures and Supervectors in Feature and Score Fusion for SVM Face and Speaker Verification, *BioID\_MultiComm2009*, LNCS 5707, pp: 81–88, Springer-Verlag Berlin Heidelberg 2009.
- [140] Dakshina Ranjan Kisku, Phalguni Gupta, and Jamuna Kanta Sing, Fusion of Multiple Matchers Using SVM for Offline Signature Identification. Springer-Verlag Berlin Heidelberg, pp. 201–208, 2009.

- [141] Mireia Farrús, Pascual Ejarque, Andrey Temko, and Javier Hernando. Histogram Equalization in SVM Multimodal Person Verification. Springer-Verlag Berlin Heidelberg, pp: 819–827, 2007.
- [142] Rachid Benmokhtar and Benoit Huet. Classifier Fusion: Combination Methods For Semantic Indexing in Video Content, Springer-Verlag Berlin Heidelberg pp: 65–74, 2006.
- [143] Muhammad ARIF. Fusion de Données : Ultime Etape de Reconnaissance de Formes, Applications à l'Identification et à l'Authentification. Thèse pour obtenir le grade de Docteur de l'université de Tours Soutenue Le 25 novembre 2005.
- [144] M. Minsky et S. Papert. Perceptrons. MIT Press, 1969.
- [145] K. Fukushima. On the approximate realization of continuous mapping by neural networks. Neural Networks, 2:183-192, 1989.
- [146] Alexandra DEBIOLLES. Diagnostic de systèmes complexes à base de modèle interne, reconnaissance des formes et fusion d'informations. Application au diagnostic des Circuits de Voie ferroviaires. Thèse présentée pour l'obtention du grade de Docteur de l'UTC. Spécialité : Technologies de l'Information et des Systèmes. Soutenue le : 19 février 2007
- [147] Lengellée T. Denoëux. Training MLPs layer by layer using an objective function for internal representations. Neural networks, 9, pp:83-97, 1996.
- [148] B. Hassibi et D.G. Stork. Second derivatives for network pruning : Optimal brain surgeon. Advances in Neural Information Processing Systems, 5, pp:164-171, 1993.
- [149] G.E. Hinton. Connectionist learning procedures. Artificial intelligence, 40, pp:185-234, 89.
- [150] Richard Lepage. Reconnaissance D'Algues Toxiques Toxiques Par Vision Artificielle et Réseau de Neurones. Mémoire de recherche présenté à l'université du Québec pour l'obtention du grade de maître ès sciences appliquées. 09, 2004
- [151] Norman Poh and Samy Bengio. Database, protocols and tools for evaluating score-level fusion algorithms in biometric authentication. Pattern Recogn., 39(2), pp :223-233, 2006.
- [152] Mauricio Villegas and Roberto Paredes. Score Fusion by Maximizing the Area under the ROC Curve. Springer-Verlag Berlin Heidelberg pp: 473-480, 2009.
- [153] Yong Li, Jianping Yin, Jun Long, and En Zhu. A Novel Method for Multibiometric Fusion Based on FAR and FRR. Springer-Verlag Berlin Heidelberg. V. Torra, Y. Narukawa, and M. Inui (Eds.): MDAI 2009, LNAI 5861, pp: 194-204, 2009.
- [154] Carlo Llobranco, Roberto Tronci, Giorgio Giacinto, and Fabio Roli. Dynamic Linear Combination of Two-Class Classifiers. Springer-Verlag Berlin Heidelberg. E.R. Hancock et al. (Eds.): SSPR & SPR 2010, LNCS 6218, pp: 473-482, 2010.