

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



Université Kasdi Merbah Ouargla
Faculté des Nouvelles Technologies de l'Information et de la
Communication

Département d'Electronique et des Télécommunications

Filière : Électronique

Option : Instrumentation industrielle

Mémoire de Fin d'Etudes

En vue de l'obtention du diplôme:

MASTER PROFESSIONNEL

Présenté par :

CHABANA Amine

SLIMI Yahia Lakhdar

Thème

***Etude comparative entre les différentes méthodes
holistique pour un système biométrique***

Soutenu Le : 03/06/2019

Devant le jury :

Mr. BENSID Khaled	MAA	Président	UKM OUARGLA
Mr. CHAA Mourad	MCB	Encadreur	UKM OUARGLA
Mr. BENLAMOUDI Azeddine	MCB	Examineur	UKM OUARGLA

Année universitaire : 2018/2019



Remerciements

Nous tenons à remercier en premier lieu DIEU


Le tout puissant qui a fait que ce travail soit terminé.

*En second lieu, nous tenons à remercier notre encadreur Dr. CHAA
Mourad, son précieux conseil et son aide
durant toute la période du travail.*

*Nos vifs remerciements vont également aux membres du
jury pour l'intérêt qu'ils ont à notre recherche en
acceptant d'examiner notre travail et de l'enrichir par
leurs propositions.*

*A nos familles et nos amis qui par leurs prières et leurs
encouragements on a pu surmonter tous les obstacles.*

*Enfin, nous tenons également à remercier toutes les
personnes qui ont participé de près ou de loin à l'exécution
de ce modeste travail*



Dédicace

J'ai le grand honneur de dédier ce modeste travail :

- ❖ A l'être le plus cher de ma vie, ma mère.*
- ❖ A celui qui m'a fait de moi un homme, mon père.*
- ❖ A mes frères*
- ❖ A tous mes amis son oubliée houssem Gacem, hakou zediri et houssem roiabhia.*
- ❖ A toute personne qui occupe une place dans mon cœur.*

CHABANA Amine

Dédicace

J'ai le grand honneur de dédier ce modeste travail :

- ❖ A celui qui m'a fait de moi un homme, mon père Saad. tu as toujours su m'inculquer le sens de la responsabilité, de l'optimisme et de la confiance en soi face aux difficultés de la vie. Tes conseils ont toujours guidé mes pas vers la réussite. que Dieu le tout puissant te préserve, t'accorde santé, bonheur, quiétude de l'esprit et te protège de tout mal.*
- ❖ A l'être le plus cher de ma vie, ma mère.*
- ❖ A mes frères .*
- ❖ A ma grande famille : je cite en particulier Fiyona, mes tantes, mes oncles ainsi que mes cousins et cousines.*

SLIMI Yahia Lakhdar

Résumé

Les caractéristiques biologiques des individus constituent des avantages technologiques majeurs depuis le siècle dernier. Elles ont évolué jusqu'à ce jour et ont été utilisées comme systèmes de sécurité efficaces dans ce que nous appelons le système biométrique. Le système biométrique prouve l'identité d'une personne en mesurant l'une de ses caractéristiques physiologiques ou comportementales telles que le visage, les empreintes digitales, la marche, ... etc. Il peut exister plusieurs types de propriétés physiques, dont certaines sont plus fiables que d'autres, mais elles doivent toutes être inviolables et uniques pour ne représenter qu'un seul individu. Dans ce système, la reconnaissance faciale occupe une place importante en raison de ses avantages, tels que la facilité d'utilisation, la sécurité élevée et la simplicité de pensée. Sur la base de tout cela, nous avons créé un système de reconnaissance faciale à travers une base de données. Dans tous les systèmes biométriques, il existe différentes techniques d'extraction de caractéristiques pour décrire les informations de visage. L'objectif de ce travail est de comparer les différentes méthodes holistiques dans un système biométrique. Ces méthodes sont utilisées dans notre système pour la réduction de la dimensionnalité. Nous avons choisi comme méthodes holistiques, l'algorithme **PCA** "Analyse en composantes principales" et LDA "Analyse discriminante linéaire" appliqués sur deux bases de données "**ORL**" et "**AR**".

Mots clés : système biométrique, identification, visage, d'extraction de caractéristiques, PCA, LDA, ORL, AR.

Abstract

The biological characteristics of individuals have been a major technological advantage since the last century. They have evolved to date and have been used as effective security systems in what we call the biometric system. The biometric system is complete to prove a person's identity by measuring one of their physiological or behavioral characteristics such as face, fingerprints, walking...etc. There may be several types of physical properties, some of which are more reliable than others, but they must all be inviolable and unique to represent only one individual. In this system, facial recognition is important because of its advantages, such as ease of use, high security and simplicity of thought. On the basis of all this, we have created a facial recognition system through a database. In all biometric systems, there are different feature extraction techniques to describe face information. In our work we study the different holistic methods and compare in particular the performance of the biometric system

using the holistic methods we have chosen the PCA algorithm "Main component analysis" and LDA "Linear discriminant analysis" applied on two databases "ORL" and "AR"

Keywords: biometric system, identification, face, feature extraction, PCA, LDA, ORL, AR.

ملخص

كانت الخصائص البيولوجية للأفراد ميزة كبيرة منذ القرن الماضي. تطورت حتى الآن واستخدمت كنظم أمنية فعالة في النظام الحيوي. نظام القياس الحيوي كامل لإثبات هوية الشخص من خلال قياس الخصائص الفسيولوجية أو السلوكية مثل الوجه وبصمات الأصابع والمشى،... إلخ. قد يكون هناك أكثر من نوع واحد من الخصائص الفيزيائية ، ولكن يجب أن تكون مصونة وفريدة من نوعها لبعضها البعض. في هذا النظام ، يعد التعرف على الوجه مهماً بسبب مزاياه ، مثل سهولة الاستخدام والأمان العالي وبساطة الفكر. على أساس كل هذا ، أنشأنا نظام التعرف الرسمي من خلال قاعدة بيانات. في جميع النظم الحيوية ، حيث هناك تقنيات مختلفة لاستخراج المعالم لوصف معلومات الوجه. خوارزمية PCA "تحليل المكون الرئيسي" و LDA "تحليل التمييز الخطي" المطبقة على قاعدتي بيانات "ORL" و "AR".

الكلمات المفتاحية: نظام البيو مترية ، تحديد الهوية ، الوجه ، استخراج المعالم ، PCA ، LDA ، ORL ، AR.

Table des matières

Glossaire	I
List des figures	II
List des tableaux	IV
Introduction générale	1

Chapitre 1

Généralités sur la biométrie

I.1. Introduction	03
I.2. Définition	03
I.2.1. Les techniques de la biométrie	04
I.2.2. Le marché mondial de la biométrie	06
I.3. Les modes de fonctionnement d'un système biométrique	07
I.3.1. Le mode d'enrôlement	07
I.3.2. Le mode vérification.....	08
I.3.3. Le mode identification	08
I.4. Architectures d'un système biométrique	09
I.4.1. Architectures et modules composants	09
I.4.1.1. Le module capteur	09
I.4.1.2. Le module évaluateur de la qualité et extracteur des caractéristiques	10
I.4.1.3. Le module base de données	10
I.4.1.4. Le module comparateur.....	10
I.4.4.5. Le module décideur	10
I.5. Pourquoi choisir la reconnaissance faciale	11
I.6. Reconnaissance de visage	11
I.6.1. La détection du visage	13
I.6.2. L'extraction des caractéristiques	13
I.6.3. La classification	15
I.7. Principales difficultés de la reconnaissance de visage	16
I.7.1. Changement d'illumination.....	16

1.7.2. Variation de pose.....	16
1.7.3 Expressions facials.....	16
I.8. Performance d'un système de reconnaissance de visage.....	17
I.8.1. Performance d'un système d'identification	17
I.8.2. Performance d'un système de vérification	17
I.9. Nouvelle technologie améliorant la reconnaissance faciale.....	20
I.9.1. L'utilisation de la technologie infrarouge.....	20
I.9.2. L'utilisation de la technologie 3D.....	20
I.10. Applications.....	20
I.10.1. Les applications commerciales	20
I.10.2. Les applications gouvernementales	21
I.10.3. Les applications judiciaires	21
I.11. Conclusion	21

Chapitre 2

Les différents algorithmes d'un système biométriques

II.1. Introduction.....	22
II.2. Les algorithmes de prétraitement.....	22
II.2.1. Le filtre DoG (Difference of Gaussians).....	22
II.3. Les algorithmes extraction des caractéristiques.....	24
II.3.1. Les méthodes holistiques.....	24
II.3.1.1. Les méthodes linaires.....	24
II.3.1.1.a. PCA (Analyse Composant Principale).....	24
II.3.1.1.b. LDA (Analyse Discriminante Linéaire).....	29
II.4. Les méthodes de classifications.....	30
II.4.1. KNN (k-nearest neighbors).....	30
II.4.1.1. Choix de K.....	31
II.4.1.2. Les distances.....	31
II.5. Conclusion.....	33
III.1. Introduction.....	34
III.2. Méthodes proposées pour la reconnaissance des visages	34
III.3. Les bases de données.....	35
III.3.1. La base de données ORL.....	35
III.3.2. La base de données AR.....	36

III.4. Expérimentations et résultats.....	37
III.4.1. Expérience.....	37
III.4.2. Aspect matériel.....	37
III.4.3. Outils de développement.....	37
III.4.4. Résultats de la technique (LDA + KNN).....	38
III.4.4.1. la base des données (ORL).....	38
III.4.4.2. la base des données (AR).....	40
III.4.5. Résultats de la technique (PCA + KNN).....	43
III.4.5.1. la base des données (ORL).....	43
III.4.5.2. la base des données (AR).....	45
III.5. Conclusion.....	48
Conclusion générale.....	49
Bibliographie	

GLOSSAIRES

BD:	Base des données ("Data Base")
DoG:	("Difference of Gaussians")
CVC:	Computer Vision Center
EER:	Taux d'erreurs égales ("Equal Error Rate")
FA:	Fausse Acceptation ("False-Acceptante")
FAR:	Taux de Fausses Acceptations ("False Acceptance Rate")
FR:	Faux Rejets ("False-Rejection")
FRR:	Le taux de faux rejets
KNN:	K plus proches voisines ("k-nearest neighbors")
LDA :	Analyse Discriminante Linéaire ("Linear Discriminate Analysis")
MRTD:	("Machine Readable Travel Documents")
ORL:	Olivetti Research Laboratory
PCA:	L'analyse en composantes principales ("Principal Components Analysis")
ROC:	("Receiver Operating Characteristic")
SVM:	Machine à vecteurs de support ("Support Vector Machine")
TC:	Temps de classification
TM:	Temps de modélisation
TR:	Taux de reconnaissance
TIC:	Taux identification correct
TER:	Taux d'erreur total ("Total Error Rate")

LISTE DES FIGURES

Chapitre 1

Figure 1.1. Exemples des modalités biométriques.....	04
Figure 1.2. Classification des modalités biométriques.....	05
Figure 1.3. Revenus prévus de la vente de technologies biométriques de 2018 à 2022....	07
Figure 1.4. Le mode d'enrôlement d'un système biométrique.....	08
Figure 1.5. Le mode de vérification d'un système biométrique.....	08
Figure 1.6. Le mode d'identification d'un système biométrique.....	09
Figure 1.7. Scores de compatibilité pour différentes technologies biométriques dans un système MRTD.....	12
Figure 1.8. Principe de fonctionnement de base d'un système de reconnaissance faciale.	12
Figure 1.9. Schéma général du processus de Reconnaissance des visages.....	15
Figure 1.10. Exemple de variation d'éclairage.....	16
Figure 1.11. Exemples de variation de poses.....	16
Figure 1.12. Exemples de variation d'expressions.....	17
Figure 1.13. La distribution du taux de vraisemblances.....	18
Figure 1.14. La courbe ROC.....	19

Chapitre 2

Figure 2.1. L'application de filtre DoG sur une image originale.....	23
Figure 2.2. Conversion de l'image $N \times N$ ver $N^2 \times 1$ vecteur.....	25
Figure 2.3. Image moyenne.....	26
Figure 2.4. Exemples les 15 Premier Eigen faces.....	27
Figure 2.5. Système de reconnaissance des personnes par PCA.....	28
Figure 2.6. Exemple d'application de KNN	31

Chapitre 3

Figure 3.1. Architecture de la méthode proposée LDA KNN.....	34
Figure 3.2. Architecture de la méthode proposée PCA KNN.....	34
Figure 3.3. Extrait de la base ORL. Ensembles des vues collectées pour des individus de la base.....	35

Figure 3.4. Extrait de la base AR. Ensembles des vues collectées pour des individus de la base.....	36
Figure 3.5. Interface du programme MATLAB.....	37
Figure 3.6. TR pour les différents nombre d'image d'entrainement sur la BD (ORL).....	39
Figure 3.7. Temps de modélisation et classification pour les différents nombre d'image d'entrainement sur la BD (ORL).....	40
Figure 3.8. TR pour les différents nombre d'image d'entrainement sur la BD (AR).....	41
Figure 3.9. Temps de modélisation et classification pour les différents nombre d'image d'entrainement sur la BD (AR).....	42
Figure 3.10. TR pour les différents nombre d'image d'entrainement sur la BD (ORL).....	44
Figure 3.11. Temps de modélisation et classification pour les différents nombre d'image d'entrainement sur la BD (ORL).....	45
Figure 3.12. TR pour les différents nombre d'image d'entrainement sur la BD (AR).....	46
Figure 3.13. Temps de modélisation et classification pour les différents nombre d'image d'entrainement sur la BD (AR).....	47

LISTE DES TABLEAUX

Chapitre 1

Tableau 1.1. Avantages et inconvénients des technologies biométriques.....	06
---	----

Chapitre 3

Tableau 3.1. Le TR obtenus par les différentes distances.....	38
Tableau 3.2. Taux de reconnaissance pour les différents nombre d'image d'entraînement sur la BD (ORL).....	38
Tableau 3.3. Temps de modélisation et classification pour les différents nombre d'image d'entraînement sur la BD (ORL).....	39
Tableau 3.4. Le TR obtenus par les différentes distances.....	40
Tableau 3.5. Taux de reconnaissance pour les différents nombre d'image d'entraînement sur la BD (AR).....	41
Tableau 3.6. Temps de modélisation et classification pour les différents nombre d'image d'entraînement sur la BD (AR).....	42
Tableau 3.7. Le TR obtenus par les différentes distances.....	43
Tableau 3.8. Taux de reconnaissance pour les différents nombre d'image d'entraînement sur la BD (ORL).....	43
Tableau 3.9. Temps de modélisation et classification pour les différents nombre d'image d'entraînement sur la BD (ORL).....	44
Tableau 3.10. Le TR obtenus par les différentes distances sur la BD (AR).....	45
Tableau 3.11. Taux de reconnaissance pour les différents nombre d'image d'entraînement sur la BD (AR).....	46
Tableau 3.12. Temps de modélisation et classification pour les différents nombre d'image d'entraînement sur la BD (AR).....	47
Tableau 3.13. Comparaison entre les méthodes proposés	48

Introduction générale

Introduction générale

La sécurité des individus est devenue un souci majeur, le besoin de se protéger augmente jour après jour. En effet, vu le développement permanent et important de la société dans tous ces aspects, les outils de surveillance et de contrôle classique à savoir ceux relatifs à la méthode basée sur la connaissance tel que le mot de passe ou bien basée sur la possession tels que les badges, les pièces d'identités, clés, ... s'avèrent inefficaces [1].

En effet, ces différents laissez-passer peuvent être perdus ou même volés. Dans le cas du mot de passe, celui-ci peut facilement être oublié par son utilisateur ou bien deviné par une autre personne. De plus ces mots de passes sont souvent archivés dans un bureau par l'organisation, or ceci est une faille certaine dans le système de sécurité [1]. Pour pallier à ces différents problèmes d'inefficacité et de non sécurité, l'homme a fait référence à une nouvelle technique de reconnaissance qui a fait son apparition et ne cesse de croître depuis 1997 : il s'agit des contrôles d'accès par les systèmes biométriques [1].

La biométrie est l'ensemble des technologies qui exploitent des caractéristiques humaines physiques ou comportementales telles que l'empreinte digitale, la signature, l'iris, la voix ou physiologiques comme le visage. Ces caractéristiques s'appellent modalités biométriques. Ces modalités ne peuvent pas être facilement oubliées, perdu, échangé ou volé [2]. Elles sont aussi permanentes ce qui signifie qu'elles ne varient pas ou peu au cours du temps [2]. Généralement, un système biométrique se compose essentiellement de quatre modules fondamentaux:

- **Module d'acquisition des données** : c'est le module d'acquisition des données biométriques brutes initiales de l'individu pour identifier ou authentifier par un camera ou un capteur biométrique. Ils jouent un rôle fondamental toujours alerte sur la qualité des données pour le module de prétraitement.
- **Module de prétraitement** : Le rôle de ce module est l'amélioration de la qualité de l'image biométrique par la réduction du bruit et la normalisation [2].
- **Module d'extraction de caractéristiques** : ce module est responsable en double fonctionnalités de la transformation des données biométriques acquises par le module de prétraitement en un modèle d'informations importantes distinctives prêt pour une comparaison au niveau de l'étape suivante. Ils basé sur des méthodes statistiques et stochastiques telles que : L'analyse en composantes principales PCA en anglais «*Principal Components Analysis*», l'analyse discriminante linéaire LDA en anglais

Introduction générale

«*Linear Discriminant Analysis*», afin d'obtenir des caractéristiques possédant les propriétés stables et uniques.

- **Module de comparaison** : ce module compare le modèle de l'individu en cours de s'identifier contre un modèle déterminé ou contre tous les différents modèles stockés dans la base de données biométrique. Dans les deux cas d'une comparaison unique ou multiple, le module de comparaison génère un score qu'il utilise pour rejeter ou valider une identité revendiquée. Dans cette-là le système biométrique fonctionne en deux modes: le mode de vérification et le mode d'identification. En mode de vérification, le système vérifie l'identité en comparant le trait biométrique inscrit dans la base de données avec un trait biométrique de test (1: 1). Dans le mode d'identification le trait biométrique de test est comparé avec tous les N traits biométrique déjà inscrits dans la base de données (1: N).

L'un de ces systèmes a été choisi pour cette étude, qui utilise la reconnaissance du visage comme caractéristiques biométriques pour identifier les images d'individu. Son modèle est unique pour tout le monde. Dans ce travail, nous concentrerons sur l'étude d'un système complet d'identification en tant que caractéristique biométrique.

L'objectif de ce travail est de réaliser une comparaison entre deux différentes méthodes holistiques pour un système biométrique basé sur des algorithmes de d'extraction des caractéristiques : PCA, LDA avec la méthode de classification KNN.

Notre travail se présente sous forme de trois chapitres vont être comme suite :

Le premier chapitre : ce chapitre présente une généralité sur la biométrie. Les différentes techniques et l'architecture d'un système biométrique, modules. Ensuite les différentes modes de fonctionnement, l'évaluation des systèmes biométriques et leurs applications En plus de détaillée la reconnaissance de visage. Ce chapitre donne un aperçu sur l'évaluation d'un système biométriques et surtout la reconnaissance faciale.

Le deuxième chapitre : ce chapitre parle de les différents algorithmes qui utiliser dans un système biométrique. Ces les algorithmes qui utilisées dans les modules de prétraitement, d'extraction de caractéristiques et de classification.

Le troisième chapitre : présenter une comparaison entre deux méthodes proposées à l'aide des algorithmes étudié, LDA +KNN et PCA +KNN.

Et enfin nous terminons ce travail par une conclusion.

Chapitre 1

Généralités sur la biométrie

I.1.Introduction

Pendant des années, les gens ont toujours essayé d'améliorer leur vie dans tous les domaines, en particulier pour vivre en sécurité (sécurité dans les lieux publics, à l'abri de tout type de vol ...). Avec le développement technologique rapide, la sécurité est devenue l'un des problèmes les plus pressants de notre société et pose un problème délicat aux citoyens, aux entreprises et aux gouvernements en matière de protection des informations et des données contre le vol. Pour toutes ses raisons, il est devenu nécessaire de créer une nouvelle technologie de contrôle et le système biométrique est une solution efficace [3].

La Technologie biométrique complète permettant de prouver l'identité d'une personne en mesurant l'une de ses propriétés physiques. Il peut exister plusieurs types de ces caractéristiques, dont certaines sont plus fiables que d'autres, mais elles doivent toutes être uniques pour ne représenter qu'un seul individu.

Dans ce chapitre, nous présentons quelques concepts de base et définitions liés à la biométrie. Un aperçu de toutes les différentes méthodes biométriques et de la structure du système biométrique, ainsi que des méthodes de fonctionnement et de performance du système biométrique. En plus d'expliquer la reconnaissance du visage en détail.

En outre, certains concepts et règles largement utilisés dans le système biométrique ont été discutés.

I.2. Définition

Le mot biométrie signifie littéralement "les mesures de la vie" et désigne dans un sens très large l'étude quantitative d'organismes vivants, l'utilisation de ce terme est de plus en plus liée à l'utilisation de ces techniques à des fins d'identification, de documentation et d'identification, ainsi qu'à la signification fondamentale du mot biométrie qui est ensuite abordée sous le terme bio- statistique[4].

La biométrie est la vérification de l'identité d'un individu à travers ce qu'elle est [4], c'est-à-dire l'utilisation des propriétés physiologiques pouvant inclure le visage, l'iris, l'empreinte digitale, les géométries de la main, l'empreinte palmaire, les empreintes des articulations des doigts. Ou les propriétés comportementales comme la voix, la signature, la démarche comme le montre la (Figure 1.1).

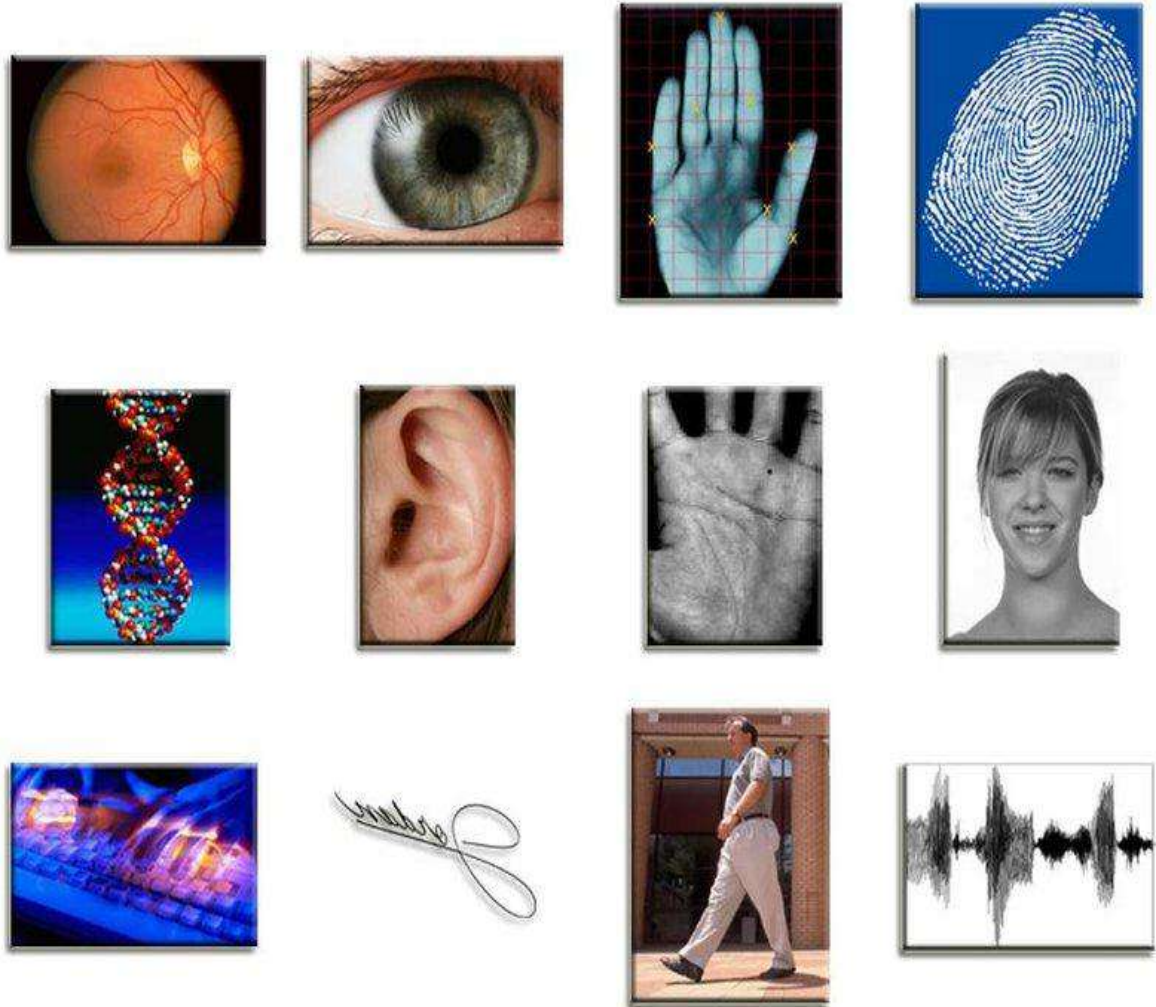


Figure 1.1. Exemples des modalités biométriques [5]

I.2.1. Les Techniques de la Biométrie

Les techniques biométriques sont actuellement utilisées pour les applications de sécurité. Chacune a des avantages et des inconvénients (**voir le tableau 1.1**), le choix d'une technique se fait en fonction de l'application [6]. Les modalités biométriques peuvent être classées en trois catégories:(**figure 1.2**)

- Les modalités qui reposent sur l'analyse de traces biologiques (odeur, sang, ADN,...)
- Les modalités morphologiques qui utilisent une partie du corps humain tels que l'empreinte digitale, l'iris, etc.
- Les modalités comportementales utilisent un trait personnel du comportement, tels que la signature, la démarche, etc.

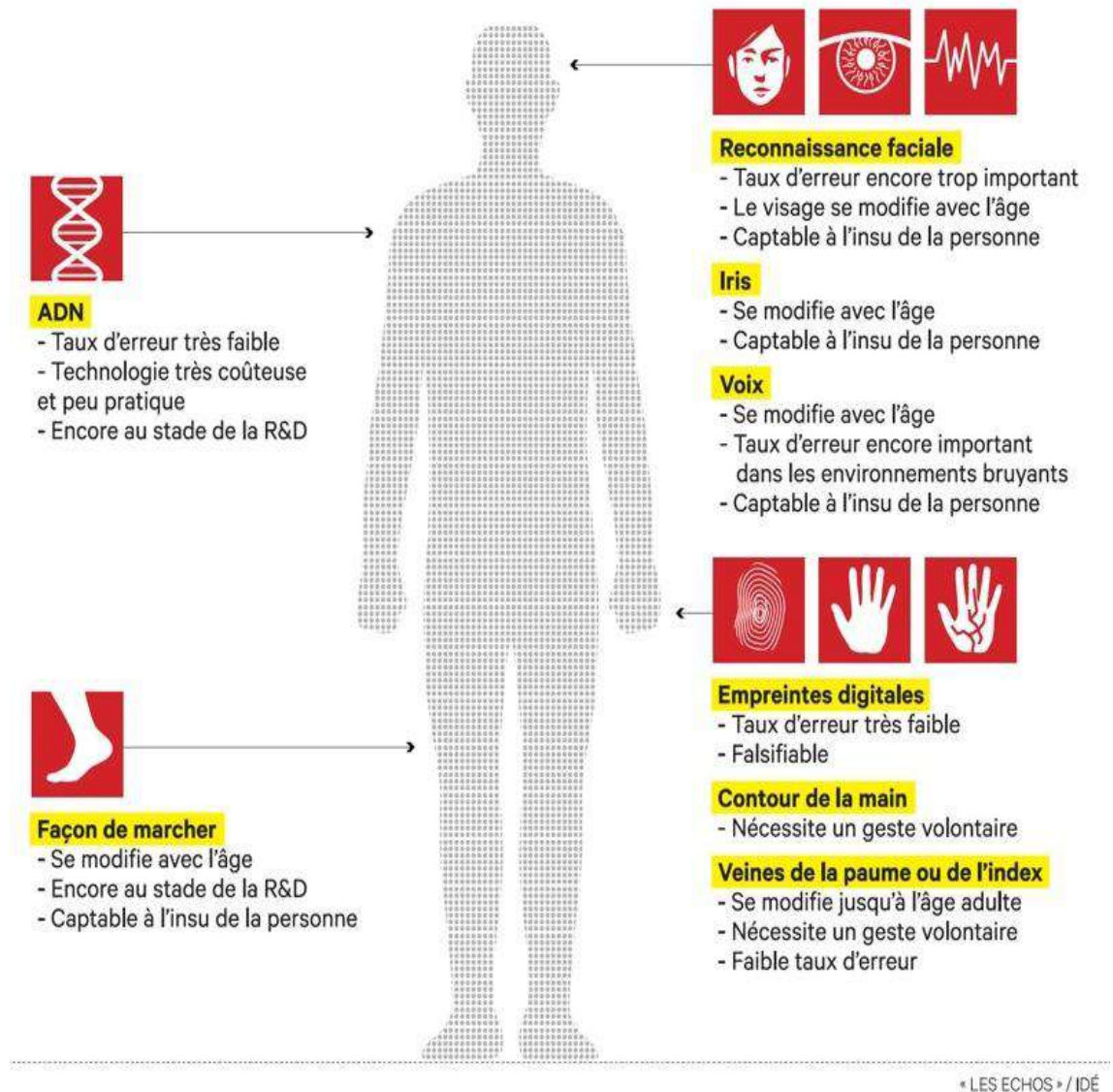


Figure 1.2. Classification des modalités biométriques [7]

Les modalités morphologiques sont les plus utilisées par rapport aux modalités comportementales. Cela est dû à leur stabilité dans le temps et la difficulté de falsification de ces modalités. De plus, les modalités comportementales sont, généralement, affectées par l'état moral des individus [6].

Les techniques biométriques se divisent en deux groupes selon la coopération ou non de l'individu [6]:

- **Techniques intrusives** : Ces techniques requièrent un contact physique avec l'individu pour l'identifier, tel que les empreintes digitales, la rétine, l'iris ou la forme de la main. Leur usage est généralement mal accepté.
- **Techniques non intrusives** : Ces techniques ne requièrent pas la coopération de l'individu en question leur application peut se faire à distance en utilisant des capteurs

qui ne nécessitent pas de contact directe avec l'utilisateur (visage, démarche,...).

Techniques	Avantages	Inconvénients
Empreintes digitales	Cout, ergonomie moyenne, facilité de mise en place, taille du capteur	Fiabilité des appareils de mesure, possibilité d'attaques (rémanence de l'empreinte,...)
Forme de la main	Très ergonomique, bonne acceptabilité	Système encombrant, cout, perturbation possible par des blessures et l'authentification des nombres d'une même famille, permanence des données.
Visage 2D	Cout, peu encombrant, bonne acceptabilité	Jumeaux, psychologie, déguisement vulnérabilité, aux attaques
Rétine	Fiabilité, pérennité	Cout, acceptabilité faible, installation difficile
Iris	Facilité	Acceptabilité très faible, contrainte d'éclairage
Voix	Facilité	Vulnérable aux attaques
Signature	Ergonomie	Dépendant de l'état émotionnel de la personne fiabilité
Frappe au clavier	Ergonomie	Dépendant de l'état physique de la personne

Tableau 1.1. Avantages et inconvénients des technologies biométriques [6]

I.2.2. Le marché mondial de la biométrie

La biométrie a connu une folie sans précédent. La croissance mondiale de la biométrie est indéniable ces dernières années et le nombre de locuteurs est devenu important, en dépit du manque d'informations publiques sur ce marché. Cependant, nous pouvons examiner certaines des projections et des chiffres pour leur développement de 2018 à 2020 au niveau mondial (voir la figure 1.3).

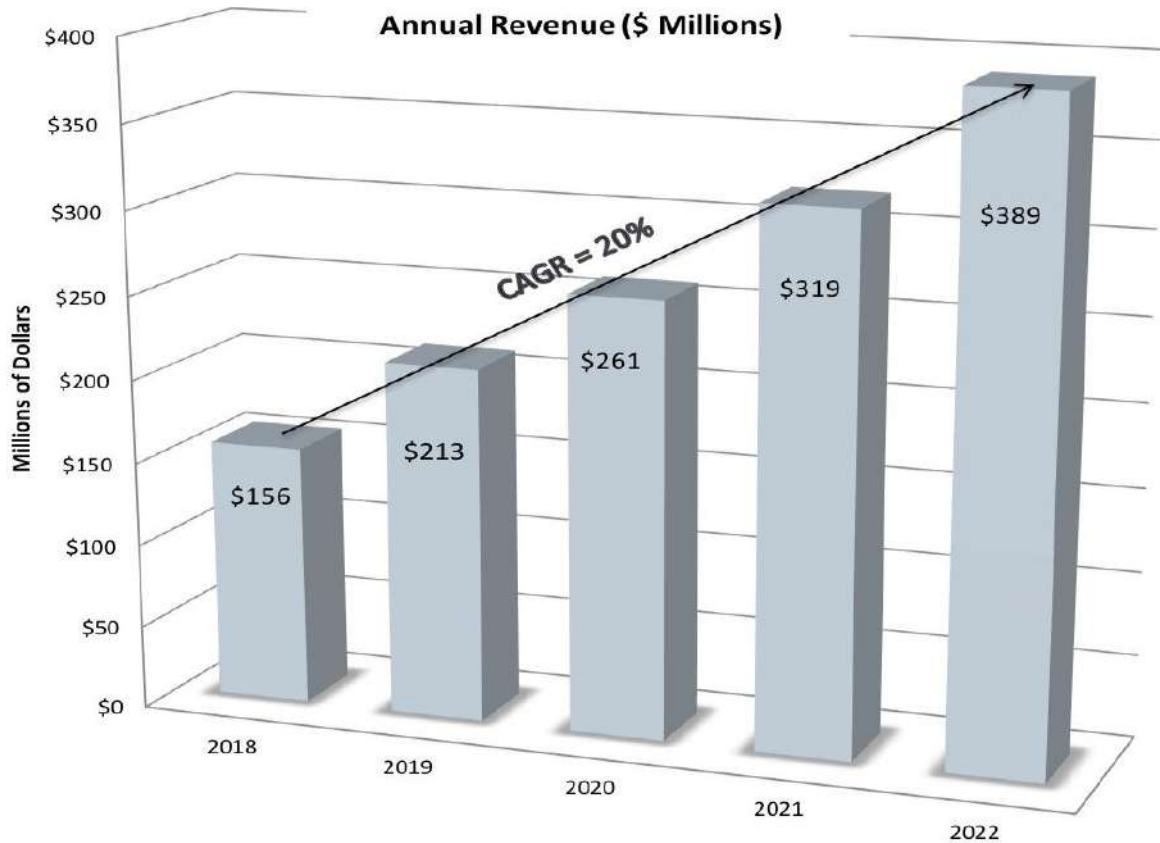


Figure 1.3. Revenus prévus de la vente de technologies biométriques de 2018 à 2022 [8]

I.3. Les modes de fonctionnement d'un système biométrique

Il y'a trois modes dans un système biométrique cette modes sont:

I.3.1. Le mode enrôlement

Dans tous les systèmes biométrique le mode enrôlement c'est la première étape (voir figure I.4), dans laquelle un utilisateur est enregistré dans le système pour la première fois de plus différents modalités biométrique sont capturée et enregistré dans une base de données. Lors de cet enregistrement, il est possible que des informations biographiques soient ajoutées à la base de données [9].

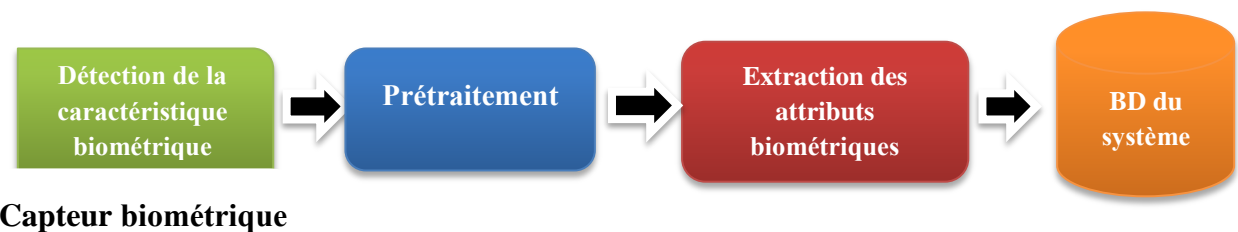


Figure 1.4. Le mode d'enrôlement d'un système biométrique

I.3.2. Le mode vérification

Dans ces modes. Le système acquiert les données biométriques et va les comparer uniquement avec le modèle enregistré correspondant à M. X. qui s'appelle "1 à 1". (Voir la figure I.5)

Si l'entrée biométrique de l'utilisateur et le modèle enregistré dans la base de données correspondant à l'identité affirmée possèdent un degré de similitude élevé, l'affirmation est validée et l'utilisateur est considéré comme étant un authentique. Dans le cas contraire, l'affirmation est rejetée et l'utilisateur est considéré comme étant un imposteur. [10]

En d'autres termes, le système doit alors répondre à la question suivante: «Suis-je réellement la personne que je suis en train de proclamer ?».

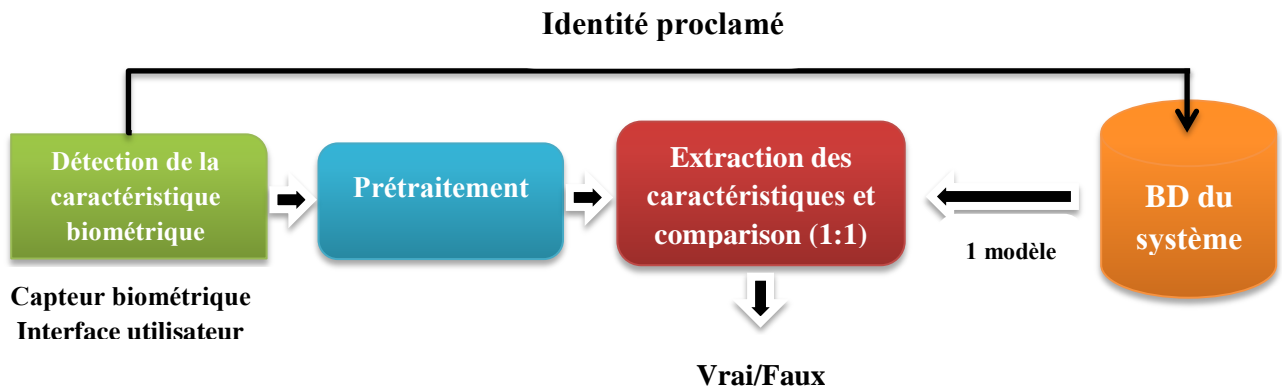


Figure 1.5. Le mode d'identification d'un système biométrique

I.3.3. Le mode identification

Dans ces modes. L'échantillon biométrique de l'individu est comparé avec les modèles de **toutes les personnes** de la base de données, qui s'appelle "un à N" (voir figure I.6).

La sortie du système biométrique est constituée par l'identité de la personne dont le modèle possède le degré de similitude le plus élevé avec l'échantillon biométrique présenté en entrée. [11]

Si la plus grande similarité entre l'échantillon et tous les modèles est inférieure à un seuil de sécurité minimum fixé, la personne est **rejetée**, ce qui implique que l'utilisateur n'était pas une des personnes enrôlées par le système. Dans le cas contraire, la personne est **acceptée**.

En d'autres termes, il répond à des questions du type: « Qui suis-je ? ».

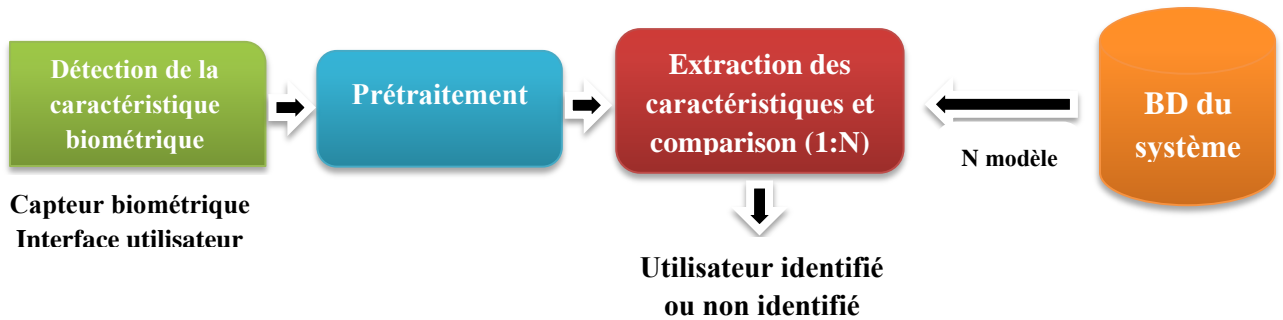


Figure 1.6. Le mode d'identification d'un système biométrique.

I.4. Architecture d'un système biométrique

Les systèmes de reconnaissance biométriques sont des systèmes qui ont généralement pour un but d'acquérir les données biométriques d'un individu, extraire les propriétés significatives à partir de ces données, comparer le modèle groupant ces propriétés contre celui (ceux) stocké (s) dans la base de données biométriques, et effectuer une réponse finale sur la base de cette comparaison[12,13].

I.4.1. Architecture et modules composants

Un système biométrique se compose essentiellement de quatre modules fondamentaux :

- Module capteur.
- Module évaluateur de la qualité et extracteur des caractéristiques.
- Module base de données.
- Module comparateur et décideur.

Il est à mentionner que les quatre premiers modules (acquisition, prétraitement, extraction de caractéristique, et décision).

I.4.1.1. Le module capteur

Le module de capteur comprend l'étape de l'acquisition des données biométriques brutes initiales de l'individu pour identifier ou authentifier par un scanner ou un capteur biométrique. Ce module est la première variable qui affecte les performances du système biométrique pour les raisons suivantes: Premièrement, il comprend le premier contact avec l'utilisateur via une interface homme / machine pouvant agir à l'encontre du facteur d'acceptation entre utilisateurs [12,13].

De deux, le module capteur joue un rôle fondamental toujours alerte sur la qualité des données brutes acquises.

I.4.1.2. Le module évaluateur de la qualité et extracteur des caractéristiques

Le module évaluateur de la qualité et extracteur des caractéristiques est responsable en double fonctionnalités de la transformation des données biométriques acquises par le module capteur en un modèle d'informations importantes distinctives prêt pour une comparaison au niveau de l'étape suivante.

Tout d'abord, la pertinence des données acquises est déterminée à travers un facteur d'évaluation et un processus d'amélioration où l'individu est requis d'une autre acquisition si les données initialement acquises sont jugées d'une très mauvaise qualité (non pertinentes).

Ensuite, ces données sont relativement traitées or un ensemble de caractéristiques les plus significatives est choisi pour représenter la modalité en cours de cet individu sous la forme d'un modèle d'informations utiles. Le modèle représentatif sera ultérieurement stocké dans la base de données biométrique avec d'autres individus [12,13].

I.4.1.3. Le module base de données

Le module base de données S'agit comme un stockage de données biométriques des utilisateurs d'une application déterminée. Il stocke tous les modèles extraits durant la phase d'apprentissage[12,13].

Plus convenablement, il stocke le modèle biométrique de l'individu utilisateur en outre de toutes ses informations biographiques pour bien formuler son identité représentative.

I.4.1.4. Le module comparateur

Le module comparateur compare le modèle de l'individu en cours de s'identifier contre un modèle déterminé ou contre tous les différents modèles stockés dans la base de données biométrique. Dans les deux cas d'une comparaison unique ou multiple, le module comparateur génère un score qu'il utilise pour rejeter ou valider une identité revendiquée [12,13].

I.4.1.5. Le module décideur

vérifie l'identité affirmée par un utilisateur ou détermine l'identité d'une personne basée sur le degré de similitude entre les caractéristiques extraites et le(s) modèle(s) stocké(s) [12,13].

I.5. Pourquoi choisir la reconnaissance faciale ?

La reconnaissance faciale est l'une des méthodes les plus répandues dans le monde. Il est plus utile, ne nécessite pas de contact direct avec l'individu (à distance), et d'autre part, des capteurs peu coûteux (caméra simple). Contrairement à l'empreinte digitale et à l'iris, le sujet doit être proche des capteurs (contact direct), ce qui est coûteux, bien que certains disent que la reconnaissance des visages est une donnée biométrique relativement peu fiable, le fait que le signal obtenu soit soumis à vos modifications les plus importantes. Autres caractéristiques telles que le contraste d'éclairage, l'évolution du mode visage, la présence ou l'absence de lunettes et autres; mais ces dernières années, il y a eu de nombreuses techniques de traitement d'images, telles que la détection des visages et la normalisation de l'éclairage, etc. Sans oublier le développement considérable des technologies de caméras numériques qui négligent l'impact de ces problèmes [14].

I.6. Reconnaissance du visage

La haute disponibilité d'ordinateurs puissants et de systèmes informatiques intégrés a suscité un vif intérêt pour le traitement automatique des images numériques et vidéo dans de nombreuses applications, notamment l'identification biométrique, la surveillance et l'interaction homme-machine et la gestion de données multimédia. La reconnaissance des visages, l'une des biotechnologies de base, revêt une importance croissante dans la recherche en raison de l'évolution rapide de technologies telles que les appareils photo numériques, Internet et les appareils mobiles, qui sont tous liés à des besoins de sécurité sans cesse croissants [15].

La reconnaissance faciale présente de nombreux avantages par rapport aux autres biotechnologies: elle est naturelle, non intrusive et facile à utiliser. Parmi les six caractéristiques biométriques prises en compte, les caractéristiques du visage présentent le degré de compatibilité le plus élevé dans le système MRTD (Machine Readable Travel Documents), et ce résultat repose sur plusieurs facteurs d'évaluation tels que l'enregistrement, la régénération des données, les exigences matérielles et la perception de l'utilisateur (**Figure 1.7**) [15].

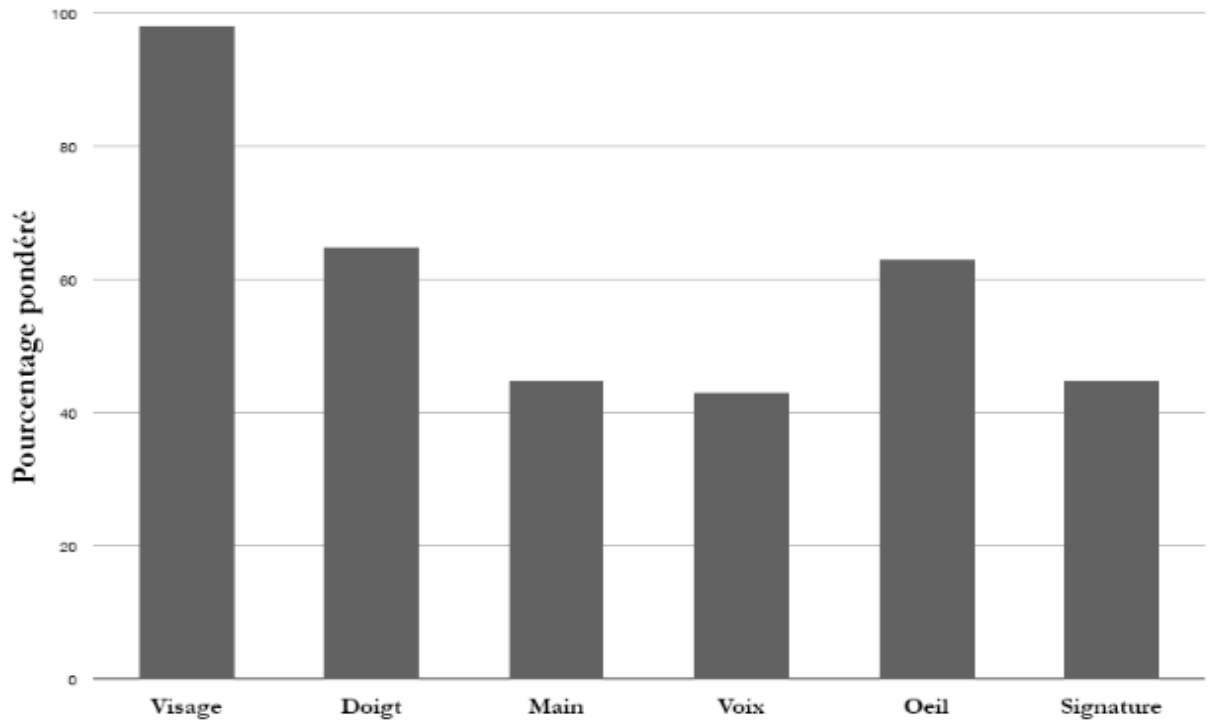


Figure 1.7. Scores de compatibilité pour différentes technologies biométriques dans un système MRTD [15].

D'une part, la reconnaissance de visage doit pouvoir identifier automatiquement les visages d'une image ou d'une vidéo. Le système peut fonctionner selon deux modes: vérification ou identification [15].

Un autre type de scénario de reconnaissance de visage implique une vérification dans la liste de surveillance, où l'individu est comparé à une courte liste de suspects.

Le principe de fonctionnement de base d'un système de reconnaissance faciale (**Figure 1.8**) peut être résumé **en quatre étapes**

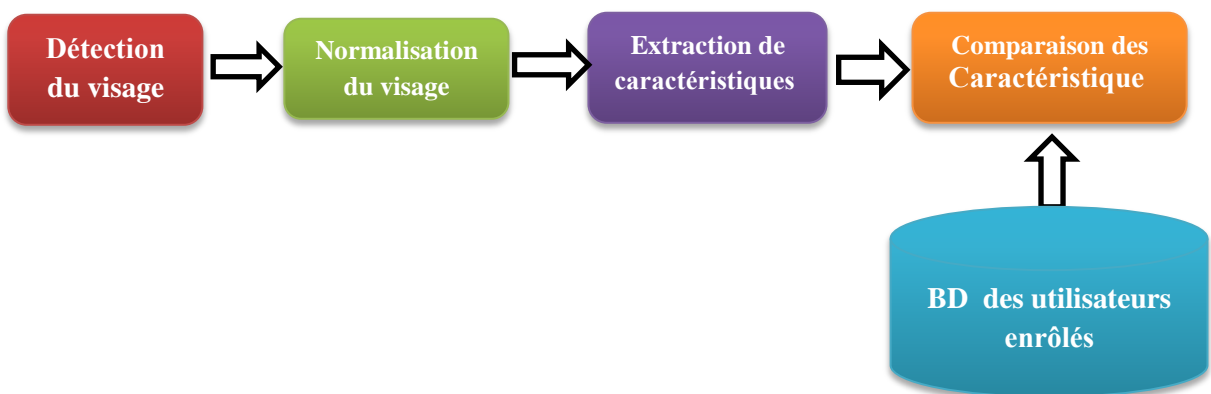


Figure 1.8. Principe de fonctionnement de base d'un système de reconnaissance faciale

I.6.1. La détection du visage

La détection faciale est effectuée en détectant la couleur de la peau ou la forme de la tête ou en détectant les différentes caractéristiques du visage.

Cette étape est particulièrement sensible lorsque l'image obtenue contient de nombreux objets pour le visage ou un fond non uniforme qui crée une texture qui bloque la segmentation correcte du visage.

Cette étape dépend de la qualité des images obtenues. Les méthodes de détection faciale peuvent être classées en quatre catégories:

- **Méthodes basées sur la reconnaissance**

Ces méthodes basées sur des règles sont conçues pour modéliser la reconnaissance de ce qui distingue le visage. Par exemple, le visage apparaît souvent dans une image aux yeux similaires entre le nez et la bouche. Traditionnellement, ces règles représentent des relations dans les traits du visage [16].

- **Méthodes basées sur des propriétés fixes**

Ces dernières dépendent des caractéristiques structurelles (traits du visage, texture, couleur de la peau) qui existent même lorsque l'apparence ou les conditions d'éclairage varient [16].

- **Méthodes basées sur la mise en correspondance modèle**

Plusieurs modèles standards de visages sont prédéfinis manuellement ou paramétrés par des fonctions. L'un de ces modèles est comparé à une image en entrée et la corrélation entre une image présentée et la base des modèles est évaluée pour détecter la présence de visage [16].

- **Méthodes basées sur l'apprentissage**

Les modèles sont ici appris à partir d'un ensemble d'images d'apprentissage qui doivent permettre de caractériser la variabilité de l'apparence d'un visage. Ces méthodes se basent sur des techniques telles que l'analyse statistique et l'apprentissage automatique pour trouver les caractéristiques appropriées des images de visage et de non-visage [16].

I.6.2. L'extraction des caractéristiques

Après la fin de la phase de découverte de la zone du visage, nous devons extraire les informations utiles pour établir un modèle de visage. Cette information est nécessaire pour que quelqu'un ne ressemble pas à quelqu'un d'autre et semble en même temps se trouver dans d'autres circonstances d'acquisition. L'efficacité de cette étape a un impact direct sur les performances du système de reconnaissance faciale. Il y a deux catégories essentielles :

- **Les méthodes globales (holistiques)**

Les méthodes globales 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. Dans ces méthodes, les images de visage (qui peuvent être vues comme des matrices de valeurs de pixels) sont traitées de manière globale et sont généralement transformées en Vecteurs, plus faciles à manipuler. Ces méthodes utilisent principalement une analyse de sous-espaces de visages.

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

Les techniques linéaires basées sur des algorithmes comme PCA (l'analyse en composantes principales), LDA (l'analyse discriminante linéaire) . PCA a été reprise dans le cadre plus spécifique de la reconnaissance faciale sous le nom de méthode des Eigen faces. Tandis que le PCA construit un sous-espace pour représenter de manière "optimale", LDA construit un sous-espace discriminant pour distinguer de façon "optimale" les visages de différentes personnes [17].

Par contre les techniques non linéaires basées sur la notion mathématique de noyau ("kernel") comme le Kernel PCA [18]. Ici une projection non linéaire (réduction de dimension) de l'espace de l'image sur l'espace de caractéristiques ("feature space") est effectuée, les variétés présentes dans l'espace de caractéristiques résultant deviennent simples, de même que les subtilités des variétés qui sont préservées. Bien que les méthodes basées sur le noyau peuvent atteindre une bonne performance sur les données d'entraînement, il ne peut pas en être de même pour de nouvelles données en raison de leur plus grande flexibilité contrairement aux méthodes linéaires [17].

- **Les méthodes locales**

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 alors que les méthodes globales ne nécessitent de connaître que la position des yeux afin de normaliser les images, ce qui peut être fait automatiquement et de manière assez fiable par un algorithme de détection [19].

I.6.3. La classifications

En utilisant les caractéristiques précédemment extraites, il est nécessaire d'introduire une mesure de similarité pour estimer la différence entre deux faces.

Pour ce faire, vous utilisez plusieurs méthodes pour comparer les visages. Nous trouvons dans la littérature: calcul la distance et la similarité. D'autres méthodes dépendent de la classification des propriétés par un seul classeur (SVM, KNN... etc) [16].

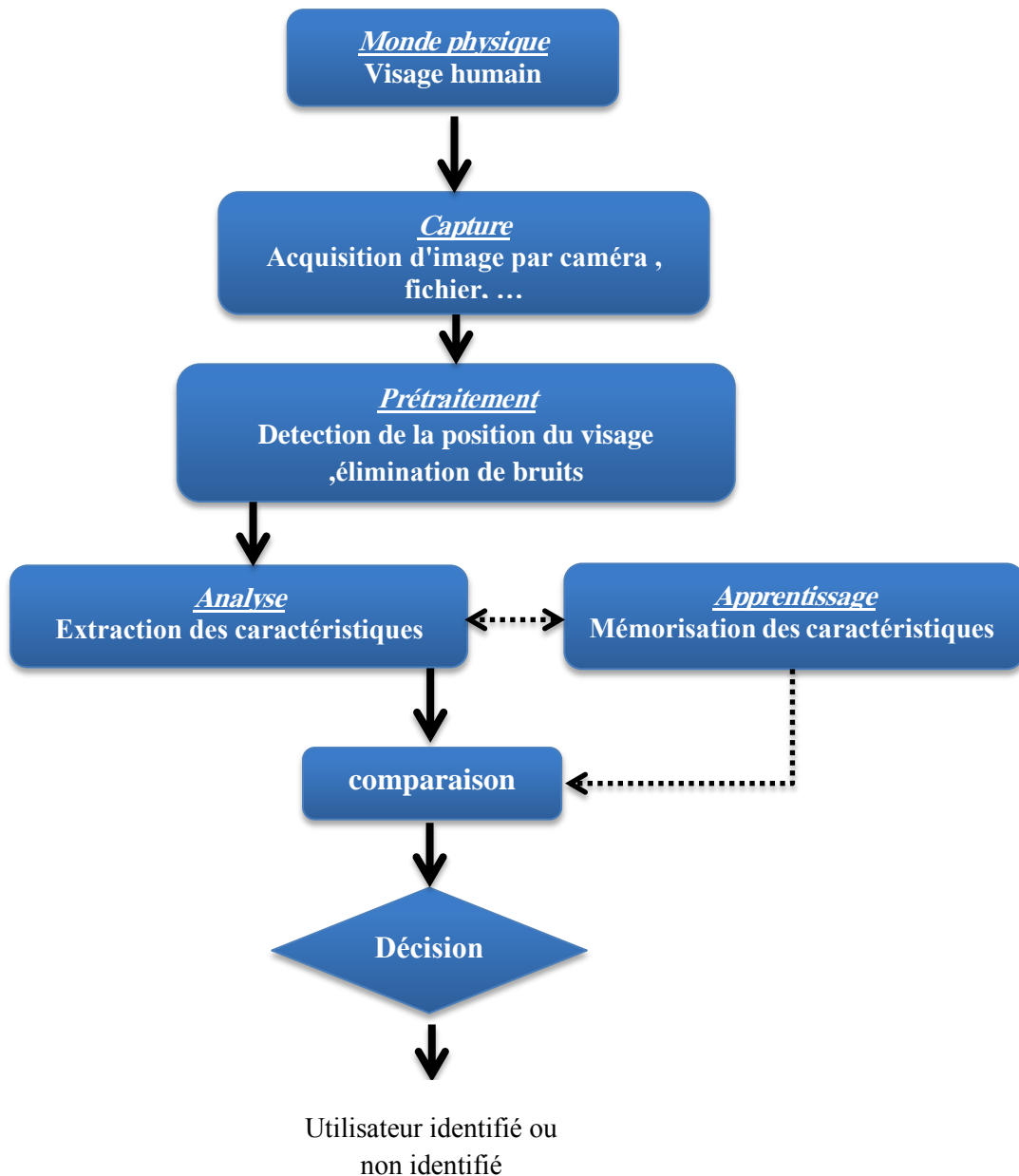


Figure.1.9. Schéma général du processus de Reconnaissance des visages

I.7. Principales difficultés de la reconnaissance de visage

I.7.1. Changement d'illumination

L'apparence d'un visage dans une image varie énormément en fonction de l'illumination de la scène lors de la prise de vue (**voir figure 1.10**). Les variations d'éclairage rendent la tâche de reconnaissance de visage très difficile. En effet, le changement d'apparence d'un visage dû à l'illumination, se révèle parfois plus critique que la différence physique entre les individus, et peut entraîner une mauvaise classification des images d'entrée [20].



Figure 1.10. Exemple de variation d'éclairage

I.7.2. Variation de pose

Le taux de reconnaissance de visage baisse considérablement quand des variations de pose sont présentes dans les images. La variation de pose est considérée comme un problème majeur pour les systèmes de reconnaissance faciale. Quand le visage est de profil dans le plan image (orientation $< 30^\circ$), il peut être normalisé en détectant au moins deux traits faciaux (passant par les yeux). Cependant, lorsque la rotation est supérieure à 30° , la normalisation géométrique n'est plus possible (**voir figure 1.11**) [20].



Figure 1.11. Exemples de variation de poses

1.7.3. Expressions faciales

Un autre facteur qui affecte l'apparence du visage est l'expression faciale (**figure 1.12**). La déformation du visage qui est due aux expressions faciales est localisée principalement sur la partie inférieure du visage. L'information faciale se situant dans la partie supérieure du visage reste quasi invariable. Elle est généralement suffisante pour effectuer une identification. Toutefois, étant donné que l'expression faciale modifie l'aspect du visage, elle entraîne forcément une diminution du taux de reconnaissance. L'identification de visage avec expression faciale est un problème difficile qui est toujours d'actualité et qui

reste non résolu. L'information temporelle fournit une connaissance additionnelle significative qui peut être utilisée pour résoudre ce problème [20].



Figure 1.12. Exemples de variation d'expressions

I.8. Performances d'un système de reconnaissances de visage

Chaque système dont les performances dépendent de plusieurs facteurs interfère avec plusieurs niveaux et peut limiter la précision. Cependant, ces facteurs doivent être pris en compte avant de mesurer les performances du système de reconnaissance.

Voici les principaux facteurs:

- L'environnement au moment de l'acquisition
- Les différentes positions des capteurs.
- La qualité des capteurs.
- La mauvaise interaction entre l'utilisateur et les capteurs.

I.8.1. Performance d'un système d'identification

Afin d'évaluer les performances du système d'identification, le taux d'identification du système est calculé. L'erreur de ce type de système est qu'il est attribué à la personne qui a présenté une identité autre que son identité [19].

La performance de ces systèmes est mesurée à l'aide du taux d'identification correct:

$$TIC = \frac{\text{nombre de tests correspondant à une identification correcte}}{\text{le nombre total de tests}} \quad (1.1)$$

I.8.2. Performance d'un système de vérification

Un système de vérification ne peut donner, lors d'une comparaison entre deux échantillons qu'un résultat sous forme de probabilité de coïncidence ou score (S), ce score est comparé à un seuil de décision (T) que l'on détermine grâce à de nombreux tests. Si S est supérieur à T le sujet en question est accepté et dans le cas contraire le sujet est considéré comme imposteur.

Dans ce cas, deux types d'erreurs peuvent être commis:

- **Faux Rejets FR (False-Rejection)**: correspond au cas où le système rejette un client légitime.
- **Fausse Acceptation FA (False-Acceptante)**: correspond au cas où le système accepte un imposteur.

La fiabilité d'un système de reconnaissance biométrique est caractérisée par deux valeurs statistiques :

- **Le Taux de Faux Rejets (FRR)**: il exprime le pourcentage d'utilisateurs légitimes rejetés.

$$FRR = \frac{\text{nombre de FR}}{\text{le nombre de clients}} \quad (1.2)$$

- **Le Taux de Fausses Acceptations (FAR)**: il exprime le pourcentage d'imposteurs acceptés.

$$FAR = \frac{\text{nombre de FA}}{\text{le nombre d'imposteurs}} \quad (1.3)$$

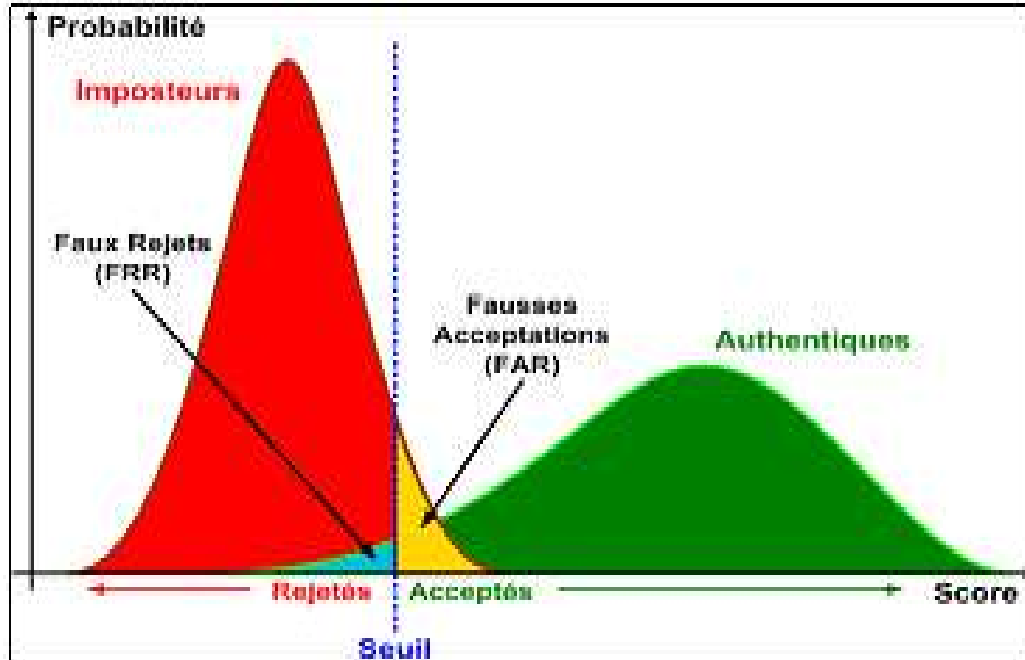


Figure 1.13. La distribution du taux de vraisemblances [19].

On peut représenter graphiquement les performances d'un système de reconnaissance de visage à l'aide de la courbe **ROC (Receiver Operating Characteristic)**. Les **FRR** sont donnés en fonction des **FAR** pour les différentes valeurs du seuil Θ

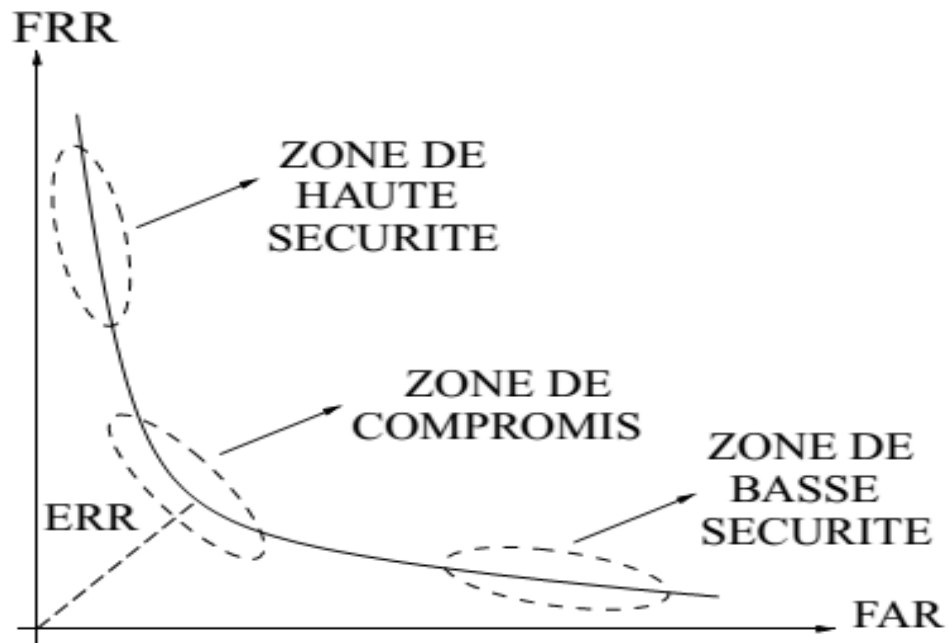


Figure 1.14. La courbe ROC

- **L'EER (Equal Error Rate ou EER)**: Il correspond à l'intersection de la courbe ROC avec la première bissectrice : **FAR = FRR**
- **Le TER (Total Error Rate)**: représente le taux d'erreur total.

$$TER = \frac{\text{nombre de clients rejetés} + \text{nombre d'imposteur acceptés}}{\text{nombre totla d'accès}} \quad (1.4)$$

I.9. Les nouvelles technologies améliorant la reconnaissance faciale

I.9.1. L'utilisation de la technologie infrarouge

Les systèmes de reconnaissance de visage infrarouges utilisent des ondes infrarouges pour mesurer les radiations thermiques émises dans la gamme du spectre infrarouge. Ces systèmes ne sont pas sensibles à la couleur de la peau, bronzage, usage de produits de beauté, les lentilles de l'œil coloré, la chirurgie plastique et les photos peuvent être prises dans l'obscurité. Néanmoins, le coût extrêmement élevé de cette technologie rend son utilisation prohibitive pour les applications [19].

I.9.2. L'utilisation de la technologie 3D

La quasi-totalité des systèmes de reconnaissance de visages travaillent à partir d'images fixes. Pour améliorer cette technique, les chercheurs s'orientent vers l'utilisation du modèle 3D du visage, à l'aide de plusieurs caméras, scanners ou caméras spécialisées. L'objectif est de conserver les avantages de la reconnaissance faciale avec des scores de fiabilité proches de ceux obtenus en empreintes digitales par exemple [18].

I.10. Applications

Dans notre société mondialisée interconnectée, l'établissement et la détermination des identités des personnes avec un haut degré de confiance devient capital et crucial. De plus, le progrès rapide que les réseaux, et les moyens de communication et transport ont vécu nécessite au fur et à mesure la mise en place des systèmes de gestion d'identités concurrents aussi fiables pour diminuer les inquiétudes autour de ceux-ci. Tout cela incite le cadre biométrique, qui de son tour, possède un pouvoir et un savoir répondre à plusieurs interrogations et questions paralysant les systèmes de gestion d'identité traditionnels. En conséquence, La biométrie a pu pénétrer de plus amples scénarios faisant appel à de nombreuses applications dont il convient de distinguer trois classes. Une application biométrique est vue comme l'ensemble en tout de la fonctionnalité, l'utilisation, et le rôle de la technologie biométrique en exploitation.

I.10.1. Les applications commerciales

Ces applications couvrent le contrôle d'accès logique tels-que l'accès aux ordinateurs et aux réseaux informatiques, le contrôle d'accès physique aux endroits de haute sensibilité, la gestion de ponctualité des employés. On trouve aussi de diverses applications telles que : le commerce électronique et la vente au détail, la téléphonie mobile, l'enseignement à distance, la gestion et la maintenance dans les services de santé et le suivi des patients, ...etc.

I.10.2. Les applications gouvernementales

Ces applications comprennent les projets des cartes d'identité nationales, les permis de conduire, la surveillance des frontières, le contrôle des passeports, et tous les autres projets gouvernementaux au cours d'étude ou en besoin de construction dans les différents pays du monde couvrant en général l'identification individuelle, la sécurité sociale, et le suivi des cibles.

I.10.3. Les applications judiciaires

Ces applications incluent l'identification des cadavres, l'enquête criminelle, la détermination de la paternité, l'identification des terroristes, et le contrôle des prisonniers ... etc.

I.11. Conclusion

A fin de ce chapitre, nous avons introduit le concept de des systèmes biométriques, leur modalités, architecture et leurs la reconnaissance faciale. Nous avons aussi constaté que Les performances des systèmes biométriques dépendent de plusieurs facteurs et qu'elles varient d'un système à un autre. Parmi les critères d'évaluation de la qualité du système biométrique, nous avons présente les taux des erreurs (FAR, FRR et ERR) ainsi que les courbes ROC selon qu'on est en mode identification ou vérification. Aussi des différentes applications d'un système biométrique. Dans le chapitre suivant, nous présentons Les différents algorithmes d'un système biométriques.

Chapitre 2

**Les différents algorithmes d'un système
biométriques**

II.1.Introduction

Le système de reconnaissance faciale, comme tous les systèmes biométriques est constitué de trois étapes essentielles : prétraitement, l'extraction des caractéristiques et la classification.

Dans la premier étape de prétraitement, on utilise des algorithmes qui traitent les images pour faciliter l'extraction des caractéristiques. Ensuite dans la deuxième étape définir les algorithmes holistiques (linéaire /non linéaire) sert à traiter l'image afin d'extraire uniquement les caractéristiques biométriques, sous forme d'un vecteur, qui ensuite peuvent être utilisées pour reconnaître l'individu. Ces caractéristiques sont uniques à chaque personne et stable. En fin la reconnaissance est faite par la comparaison (classification) du vecteur de caractéristique avec une base de données.

En représente Dans Ce chapitre les différents algorithmes qui utilisés dans les systèmes biométriques. Également, ces algorithmes se divisent en trois catégories : (I) les algorithmes de prétraitement, (II) les algorithmes extraction de caractéristiques (III) les algorithmes de classification. Ces derniers classés en deux approches Les algorithmes basés les approches holistiques qui sont deux types, algorithme linéaire comme PCA "Analyse composant principale", LDA "Analyse Discriminante Linéaire". Et algorithme non linéaire comme KPCA "kernel Analyse composant principale". En plus de donner un aperçu de certains algorithmes de classification tels que KNN (k-nearest neighbors).

II.2. Les algorithmes de prétraitement

II.2.1. Le filtre DoG (Difference of Gaussians)

En imagerie, différence de gaussiennes (DoG) est un algorithme d'amélioration des caractéristiques qui consiste à soustraire une version floue d'une image originale d'une autre version moins floue de l'original. Dans le cas simple des images en niveaux de gris, les images floues sont obtenues en convertissant les images originales en niveaux de gris avec des noyaux gaussiens ayant des écarts-types différents. Le flou d'une image à l'aide d'un noyau gaussien ne supprime que les informations spatiales à haute fréquence. La soustraction d'une image à l'autre préserve les informations spatiales situées dans la plage de fréquences conservée dans les deux images floues [21].

Ainsi que, la différence entre les Gaussiens est un filtre passe-bande, qui élimine la composante haute fréquence représentant le bruit, ainsi que certaine composante basse fréquence représentant les zones homogènes de l'image. Les composantes de fréquence dans la bande passante sont supposées être associées aux bords dans les images [2].

Étant donnée, σ_L et σ_H définissent les basses et hautes fréquences respectivement. Ces fréquences de coupure dépendent de la qualité de l'image. Ici si la valeur de σ_L trop élevé entraînera la perte de l'information utiles car l'information se trouve principalement dans des basses filtrée $I_p(x, y)$ est traitée comme suit [2]:

$$I_p(x,y) = DOG * I(x,y) \tag{2.1}$$

Le filtre DoG est calculé par l'équation :

$$DOG = \frac{1}{2\pi\sigma_L^2} e^{-\frac{x^2+y^2}{2\sigma_L^2}} - \frac{1}{2\pi\sigma_H^2} e^{-\frac{x^2+y^2}{2\sigma_H^2}} \tag{2.2}$$

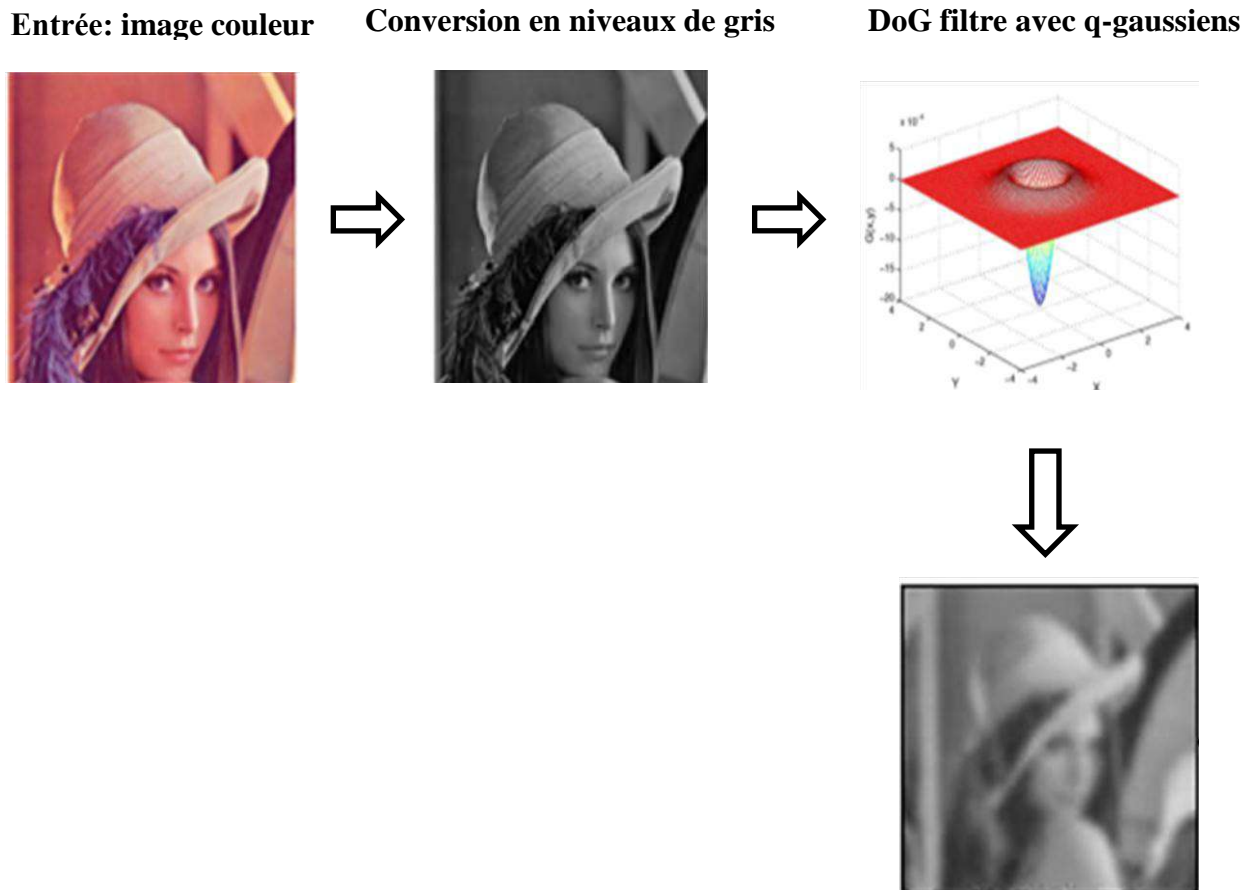


Figure 2.1. L'application de filtre DoG sur une image originale [21].

II.3. Les algorithmes extraction de caractéristiques

II.3.1. Les méthodes holistiques

Les méthodes holistiques sont basées sur des techniques d'analyse statistique bien Connues comme PCA et LDA.

II.3.1.1. Les méthodes linaires

Il y'a deux méthodes linaires :

II.3.1.1.a. PCA (Analyse Composant Principale)

Dans les statistiques, l'analyse en composantes principales (PCA) est une technique qui peut être utilisée pour simplifier un ensemble de données. PCA peut être utilisé pour réduire la dimensionnalité dans un ensemble de données tout en conservant les caractéristiques de l'ensemble des données qui contribuent le plus à sa variance, en gardant les composants principaux d'ordre inférieur et en ignorant d'ordre supérieur . L'idée est que ces composants à faible ordre contiennent souvent les aspects « les plus importants » des données [22].

La tâche de reconnaissance faciale est discriminant les signaux d'entrée (données d'image) en plusieurs classes (personnes). L'entrée des signaux sont très bruyants (par exemple, le bruit est causé par différentes conditions d'éclairage, pose, etc.), mais les images d'entrée ne sont pas complètement aléatoire et en dépit de leurs différences il existe des modèles qui se produisent dans tout signal d'entrée. Ces modèles, qui peut être observée dans tous les signaux pourraient être - dans le domaine de la reconnaissance faciale - la présence de certains objets (yeux, nez, bouche) dans tout le visage, ainsi que les distances relatives entre ces objets. Ces caractéristiques sont appelées Eigen faces dans le domaine de reconnaissance faciale (ou composants principaux généralement). Ils peuvent être extraits à partir de données d'image originale moyenne de l'outil mathématique appelé analyse des composantes principaux (PCA) [22].

Au moyen de PCA, on peut transformer chaque image originale de la formation mis en Eigen face correspondant. Ainsi, le but de PCA est de réduire la grande dimensionnalité de l'espace de visage (observé variables) à la plus petite dimension intrinsèque de l'espace de fonction (variables indépendantes), qui sont nécessaires pour décrire la données sur le plan économique [22].

Pour générer un ensemble d'Eigen faces, un grand nombre d'images numérisées des visages humains, pris sous le même éclairage conditions, sont normalisées pour aligner les yeux et la bouche. Ils sont alors tous formé avec la même résolution de pixels ($N \times N$), puis traité comme N^2 des vecteurs de dimension dont les composantes sont les valeurs de leurs

pixels. Les vecteurs propres de la matrice de covariance de la distribution statistique des vecteurs d'images de visage sont ensuite extraits. Étant donné que les vecteurs propres appartiennent à le même espace vectoriel que des images de visage, ils peuvent être considérés comme si elles étaient $m \times n$ images de visage de pixels: d'où le nom Eigen faces. Dans cette façon, le principal Eigen face ressemble à un visage humain moyen [22].

Soit une image de visage $\Gamma(x, y)$ à deux dimensions M par réseau N de valeurs d'intensité. L'Analyse en Composante Principale sera appliquée sur la base des images de visages, désormais représentée par une matrice individus, caractéristiques selon l'algorithme suivant :

Etape 1 : Préparer les données

La première étape consiste à obtenir un ensemble S de $(N \times N)$ images de visage. Chaque image est transformé en un vecteur de taille $N^2 \times 1$ et placé dans l'ensemble.

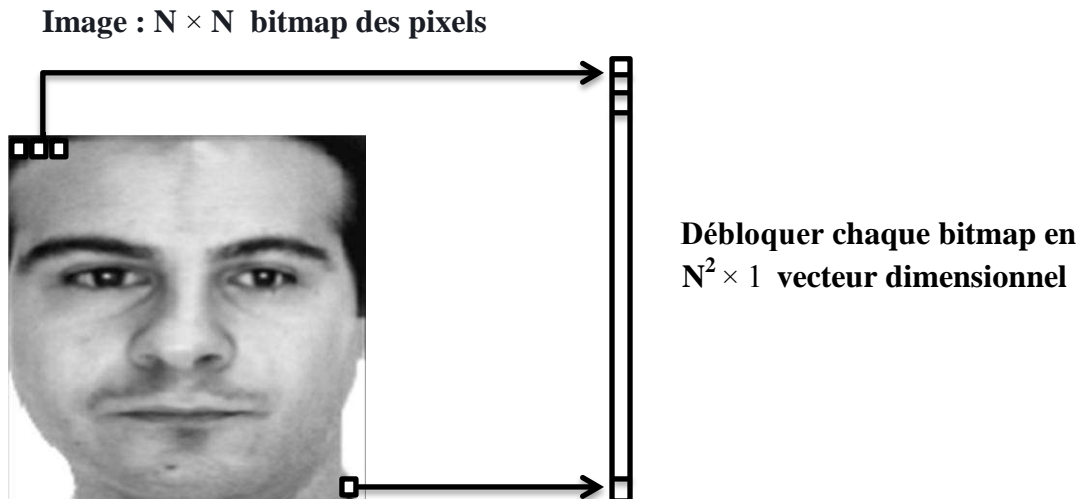


Figure 2.2. Conversion de l'image $N \times N$ ver $N^2 \times 1$ vecteur

La matrice d'entraînement $X_T = [X_1, X_2 \dots X_q]$. X_j désigne le vecteur de J^{ieme} image avec $1 \leq j \leq q$

Chaque vecteur (X_j) appartenant à l'un des classes $C_1, C_2 \dots C_p$.

$$X_T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{q1} \\ a_{12} & a_{22} & \dots & a_{q2} \\ \vdots & \vdots & \dots & \vdots \\ a_{1m} & a_{2m} & \dots & a_{qm} \end{pmatrix} = (X_1, X_2, \dots, X_q) \quad (2.3)$$

Etape 2 : Obtenir la moyenne

Après l'obtention de l'ensemble, l'image moyenne μ doit être obtenue sous la forme:

$$\mu = \frac{1}{q} \sum_{j=1}^q X_j \quad (2.4)$$

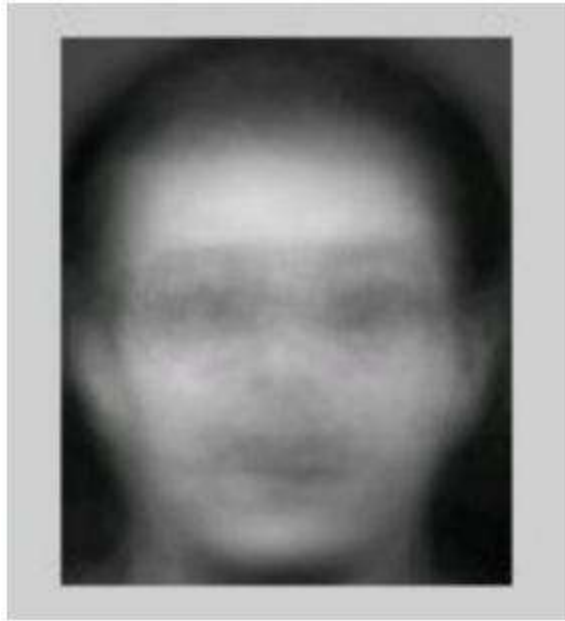


Figure 2.3. Image moyenne

Etape 3 : Soustraire la moyenne de l'image originale

La différence entre l'image d'entrée et l'image moyenne doit être calculée et le résultat est stocké dans $\bar{\phi}_j$.

$$\bar{\phi}_j = X_j - \mu \quad (2.5)$$

Etape 4: Calculer la matrice de covariance

La matrice de covariance C est calculée de la manière suivante

$$C = \frac{1}{M} \sum_{n=1}^M \phi_n \phi_n^T \quad (2.6)$$

$$= AA^T$$

$$A = [\phi_1, \phi_2, \phi_3, \phi_4 \dots \phi_M]$$

Etape 5: Calculer les vecteurs propres et valeurs propres de la matrice de covariance avec la Réduction de la dimensionnalité.

Dans ce cas la taille de C est $q \times q$

$$C = A^T A \quad \text{Où } A = [\phi_1, \phi_2, \dots, \phi_q] \quad (2.7)$$

Si le nombre des images apprentissage est $q=100$ par exemple donc on obtient 100 vecteurs propres et la longueur de chaque vecteur est 100.

Etape 6: La Conversion des k vecteurs propres de dimension inférieure à la dimension initiale des images originale.

Si u_i ont les vecteurs propres de $C = AA^T$ et v_i sont les vecteurs propres de $C = A^T A$.

$$u_i = Av_i \quad (2.8)$$

Etape 7: Sélectionnez les K meilleurs vecteurs propres, sachant que $k < q$ cet espace vectoriel s'appelle l'espace des Eigen faces (La représentation graphique des k meilleurs vecteurs propres s'appelle Eigen face en anglais). (Voir la figure 2.4)

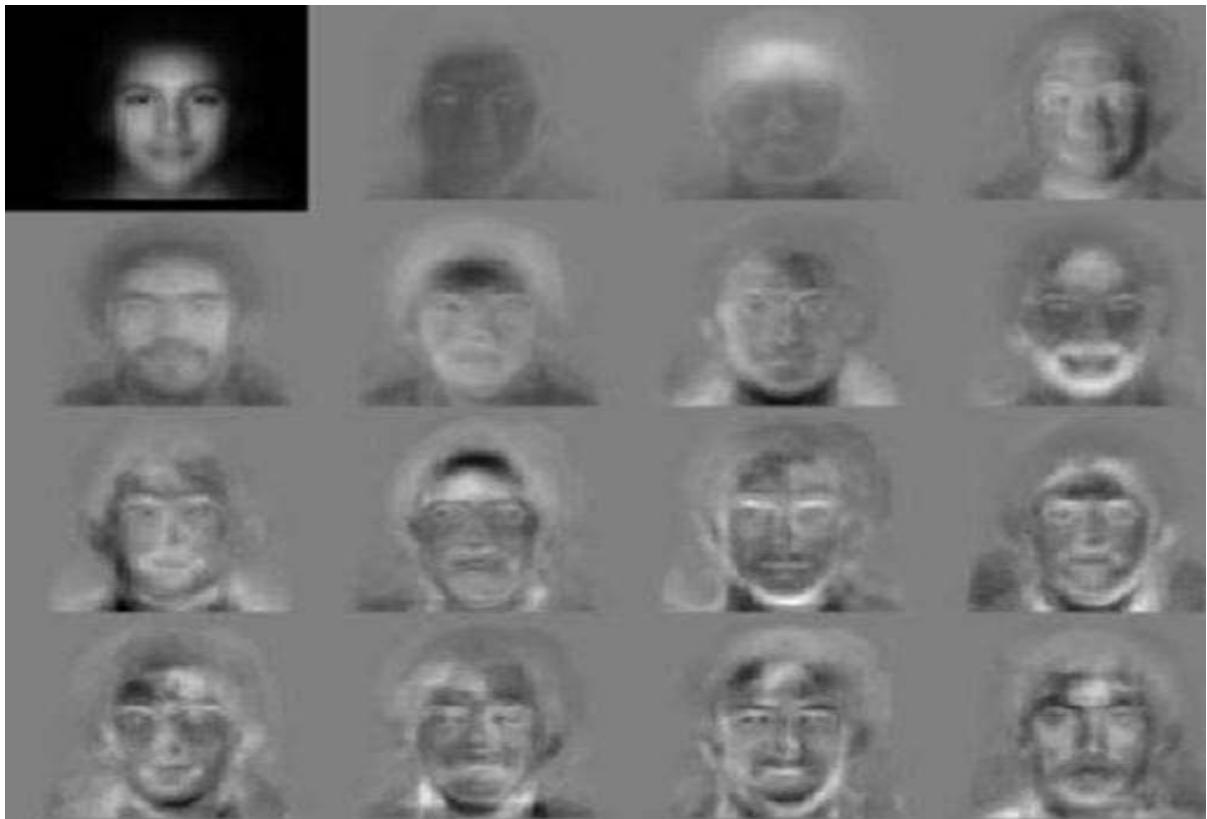


Figure 2.4. Exemples les 15 Premier Eigen faces [32]

Dans la phase de test une image d'une personne inconnue est entrée au système alors le système convertit l'image d'entrée en un vecteur puis le normalisé. Ce vecteur normalisé a été projeté sur l'espace Eigen faces pour obtenir le vecteur de poids «weight vector». Celui-ci considérait comme la nouvelle représentation d'image d'entrée. Puis le système calcule la distance entre le vecteur poids d'image en entrée et les vecteurs poids dans la base de données. Enfin une décision est prise à partir de cette distance [2] (Voir figure 2.5)

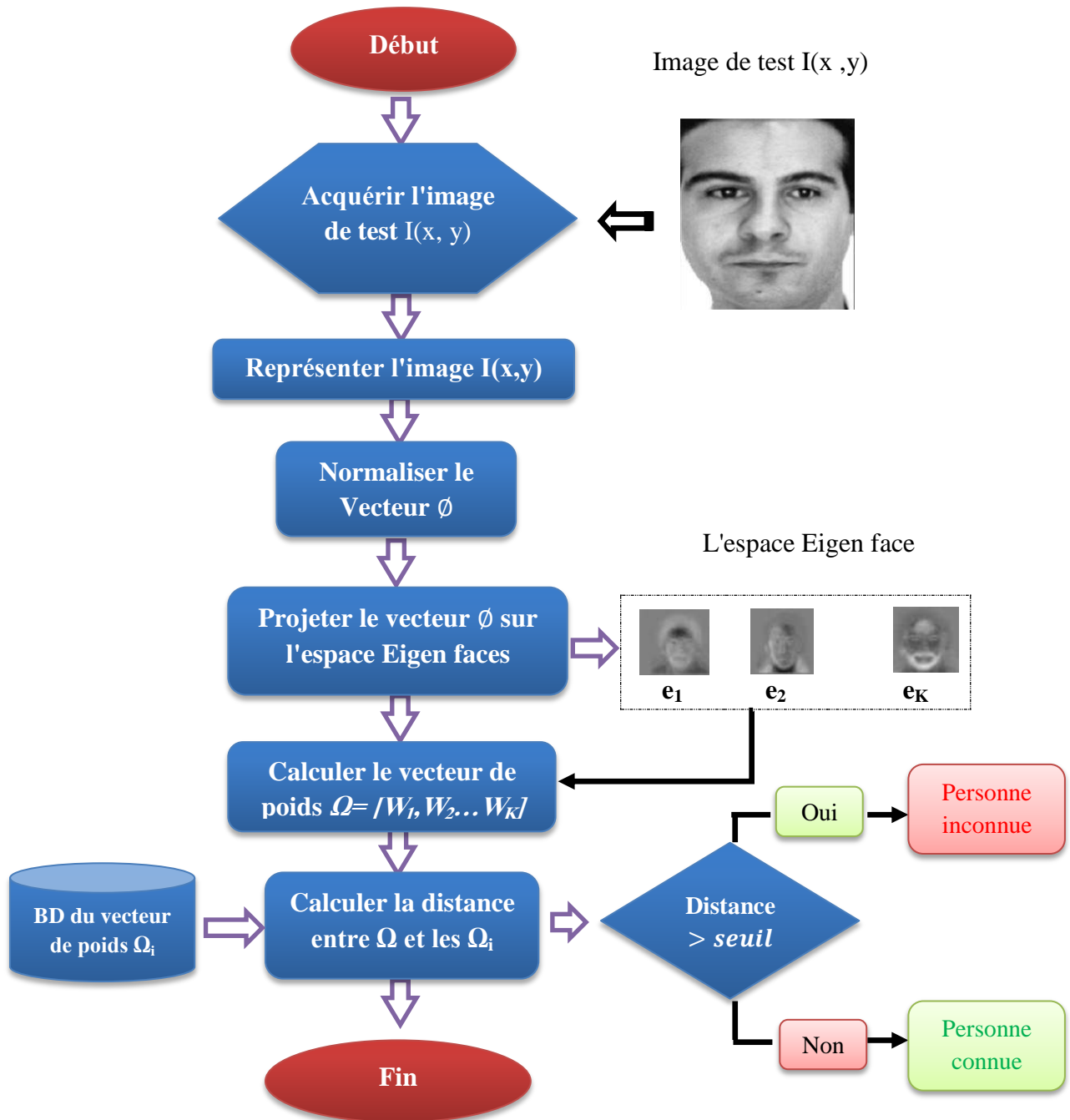


Figure 2.5. Système de reconnaissance des personnes par PCA

II.3.1.1.b. LDA (analyse discriminante linéaire)

LDA est né des travaux de Belhumeur et al . De la Yale University (USA), en 1997 [23]. Il est aussi connu sous le nom de « Fisher faces ». Contrairement à l'algorithme PCA, celui de la LDA effectue une véritable séparation de classes. Pour pouvoir l'utiliser, il faut donc au préalable organiser la base d'apprentissage d'images en plusieurs classes : une classe par personne et plusieurs images par classe. La LDA analyse les vecteurs propres de la matrice de dispersion des données, avec pour objectif de maximiser les variations entre les images d'individus différents (inter classes) tout en minimisant les variations entre les images d'un même individu (intra-classes) [23-24].

Donc LDA c'est une technique populaire, utilisée pour trouver la combinaison linéaire des caractéristiques qui séparent mieux les classes d'objets. Les combinaisons résultantes peuvent être utilisées comme classificateur linéaire, ou pour la réduction des caractéristiques avant la classification [25-26].

Soit la matrice d'entraînement $X_T = [X_1, X_2 \dots X_q]$. X_j désigne le vecteur caractéristique d'image (j). Chaque X_j appartenant à l'un des N classe $C_1, C_2 \dots C_n$ avec $1 \leq j \leq q$ [2].

Sous-espace LDA a été construit de sorte qu'il minimise la variance intra-classe S_b «between-class scatter matrix» et maximise la variance inter-classe S_w «within-class scatter matrix» [2]:

$$S_B = \sum_{i=1}^N n_i (\mu_i - \mu)(\mu_i - \mu)^T \quad (2.9)$$

$$S_W = \sum_{i=1}^N \sum_{X_j \in C_i} (X_j - \mu_i)(X_j - \mu_i)^T \quad (2.10)$$

Où n_i représente le nombre d'échantillons dans l'ième classe, μ_i désigne la moyenne des données d'apprentissage appartenant à la i^{ème} classe, N le nombre de classes et μ représente la moyenne globale de toutes les données d'entraînement. Et puis on dérive la matrice de transformation W qui maximise le critère discriminant de Fisher:

$$T(w) = W_{opt} = \mathit{arg}_w \max \frac{|W^T S_B W|}{|W^T S_W W|} = [W_1 W_2 \dots W_d] \quad (2.11)$$

La solution optimale à ce problème d'optimisation est donnée par la résolution du problème généralisé des vecteurs propres

$$S_B W = \gamma S_W W \quad (2.12)$$

Ça revient à la recherche des vecteurs propres de la matrice $S_W^{-1} S_B$. La représentation graphique de ces vecteurs propres s'appelle en anglais Fisher faces. La longueur du vecteur caractéristique après l'application de LDA au plus N-1. Avec N est le nombre total des classes.

II.4. Les méthodes de classifications

II.4.1. KNN (k-nearest neighbors)

C'est un algorithme simple qui stocke tous les cas disponibles de données d'entraînement et classe les nouveaux cas en fonction de la majorité de ses voisins, en calculant la distance qui les sépare. Ces fonctions peuvent être la distance euclidienne, Manhattan, Minkowski et Hamming. Si $K = 1$, l'état est simplement affecté à la catégorie adjacente la plus proche. La nouvelle catégorie de points de données est prédite en recherchant dans le groupe de formation entier les catégories les plus étroitement liées [29].

L'astuce consiste à déterminer les similitudes entre les données. Le moyen le plus simple de déterminer la similarité, si toutes les données sont identiques (par exemple, en pouces) - consiste à utiliser la distance euclidienne. L'application de l'algorithme KNN consomme beaucoup de mémoire pour stocker toutes les données, mais n'effectue aucun calcul à moins qu'il soit nécessaire de prévoir. Vous pouvez également mettre à jour les cas de formation au fil du temps pour maintenir les prévisions exactes [29].

Donc le principe de cet algorithme de classification est très simple. On lui fournit un ensemble des données d'apprentissages \mathbf{D} , une fonction de distance d et un entier k . Pour tout nouveau point de test x , pour lequel il doit prendre une décision, l'algorithme recherche dans \mathbf{D} les k points les plus proches de x au sens de la distance d , et attribue x à la classe qui est la plus fréquente parmi ces k voisins. (figure 2.8)

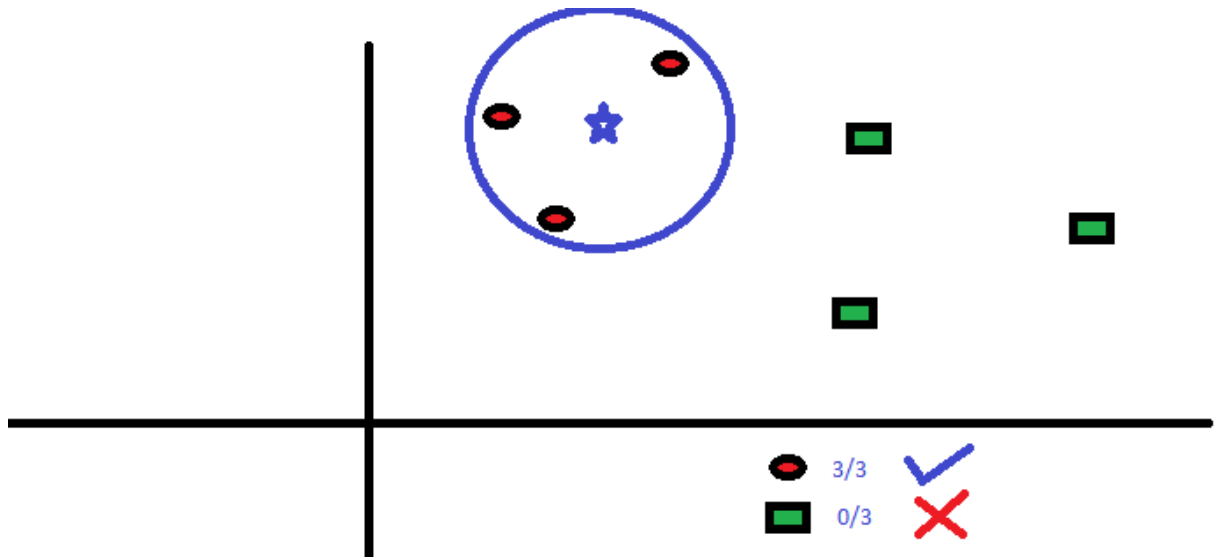


Figure 2.6. Exemple d'application de KNN

II.4.1.1. Choix de k

- ✓ Pour les grandes valeurs de k :
 - Moins sensible au bruit.
 - Une grande base d'apprentissage permet une plus grande plus grande valeur de k.
- ✓ Pour les petites valeurs de k :
 - Rend mieux compte de structure fine.
 - Nécessaire pour les petites bases d'apprentissage.

II.4.1.2. Les distances

- **Distance euclidienne :**

La distance la plus connue est la distance Euclidienne, qui définit l'espace cartésien. Donnée pour des vecteurs de dimension n par l'équation suivante [30]:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.13)$$

La distance Euclidienne est souvent utilisée au carré.

- **Distance cosin :**

La distance cosin est calculée à partir d'un moins un le cosinus de l'angle inclus entre les points [30], défini par l'équation suivant:

$$d_{st} = \left(1 - \frac{x_s y_t'}{\sqrt{(x_s x_s') (y_t y_t')}} \right) \quad (2.14)$$

- **Distance Mahalanobis**

La distance de Mahalanobis est une mesure entre un et une distribution des données [30], définie par l'équation suivant:

$$d_{st}^2 = (x_s - y_t) C^{-1} (x_s - y_t)' \quad (2.15)$$

Où C est la matrice de covariance

- **Distance corrélation**

La distance basée sur la corrélation est une mesure de la dépendance statistique entre deux vecteurs [30], définie par l'équation suivant:

$$d_{st} = \left(1 - \frac{(x_s - \bar{x}_s)(y_t - \bar{y}_t)'}{\sqrt{(x_s - \bar{x}_s)(x_s - \bar{x}_s)'} \sqrt{(y_t - \bar{y}_t)(y_t - \bar{y}_t)'}} \right) \quad (2.16)$$

Où

$$\bar{x}_s = \frac{1}{n} \sum_j x_{sj}$$

$$\bar{y}_s = \frac{1}{n} \sum_j y_{sj}$$

- **Distance Spearman**

La distance de Spearman est calculée à partir de un moins la corrélation de Spearman entre les observations [30] , définie par l'équation suivant:

$$d_{st} = 1 - \frac{(\mathbf{r}_s - \bar{\mathbf{r}}_s)(\mathbf{r}_t - \bar{\mathbf{r}}_t)'}{\sqrt{(\mathbf{r}_s - \bar{\mathbf{r}}_s)(\mathbf{r}_s - \bar{\mathbf{r}}_s)'}\sqrt{(\mathbf{r}_t - \bar{\mathbf{r}}_t)(\mathbf{r}_t - \bar{\mathbf{r}}_t)'}} \quad (2.17)$$

Où

\mathbf{r}_{sj} est le rang de \mathbf{x}_{sj} repris $\mathbf{x}_{1j}, \mathbf{x}_{2j}, \dots, \mathbf{x}_{mj}, \mathbf{x}_{nj}$.

\mathbf{r}_{tj} est le rang de \mathbf{y}_{tj} pris en charge $\mathbf{y}_{1j}, \mathbf{y}_{2j}, \dots, \mathbf{y}_{mj}, \mathbf{y}_{nj}$.

\mathbf{r}_s et \mathbf{r}_t sont les vecteurs de rang par coordonnées

de \mathbf{x}_s et \mathbf{y}_t ,

c'est-à-dire $\mathbf{r}_s = (r_{s1}, r_{s2}, \dots, r_{sn})$ et $\mathbf{r}_t = (r_{t1}, r_{t2}, \dots, r_{tn})$.

II.5. Conclusion

Dans ce chapitre nous avons étudié les différentes méthodes de prétraitement, d'extraction des caractéristiques et de classification. Dans le premier module, l'image est soumise à un traitement initial lors de l'entrée dans le canal du système biométrique afin de l'améliorer et d'éliminer les interférences, cette étape basé sur des algorithmes comme filtre DOG.

Au module d'extraction des caractéristiques, les systèmes de reconnaissances faits des étapes plus importantes avant le stockage des informations dans ces bases de données. Ces étapes sont basées sur des algorithmes spécifiques comme : L'extraction de caractéristiques : pour obtenir les caractéristiques de chaque image acquise sous forme de vecteur. Il-y-a plusieurs méthodes pour faire cette opération comme PCA, LDA. La classification des données : dernière étape fait classer les caractéristiques semblables d'un ou plusieurs individus à la même classe, cette étape est appliquée par des algorithmes comme KNN.

Chapitre 3

Expérimentations et résultats

III.1. Introduction

Dans les chapitres précédents, nous avons expliqué en détail la structure générale du système de reconnaissance du visage, ainsi que les différentes méthodes holistiques utilisées dans la conception de ces systèmes. Dans cette partie, nous avons appliqué les différents algorithmes PCA, LDA et KNN à la mise en œuvre d'un programme intitulé "étude comparative des différentes méthodes holistiques" pour les résultats expérimentaux obtenus.

III.2. Méthodes proposées pour la reconnaissance des visages

Cette section consiste à présenter des détails sur le système biométrique proposé. La méthode proposée est basée sur LDA, PCA et KNN. LDA et PCA sont utilisés pour l'extraction des caractéristiques. Pour l'étape de classification, nous avons utilisé la technique KNN (noté k-NN pour Nearest-Neighbor), basée sur les distances (Cosine, Euclidien, Corrélation, Mahalanobis et Spearman). Finalement, pour évaluer les méthodes proposées, deux bases de données (ORL et AR) ont été utilisées. **La figure 3.1.** Représente la première méthode proposée d'identification (LDA + KNN) et la **figure 3.2.** illustre la deuxième méthode proposée d'identification (PCA +KNN).

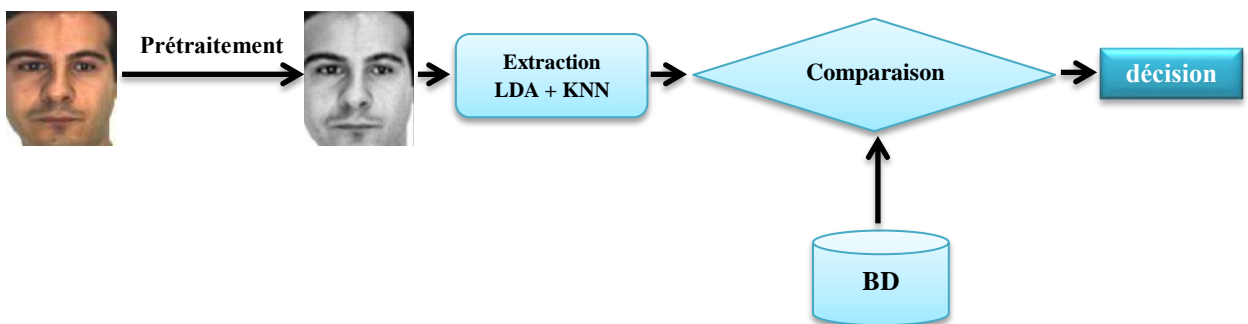


Figure 3.1. Architecture de la méthode proposée LDA +KNN

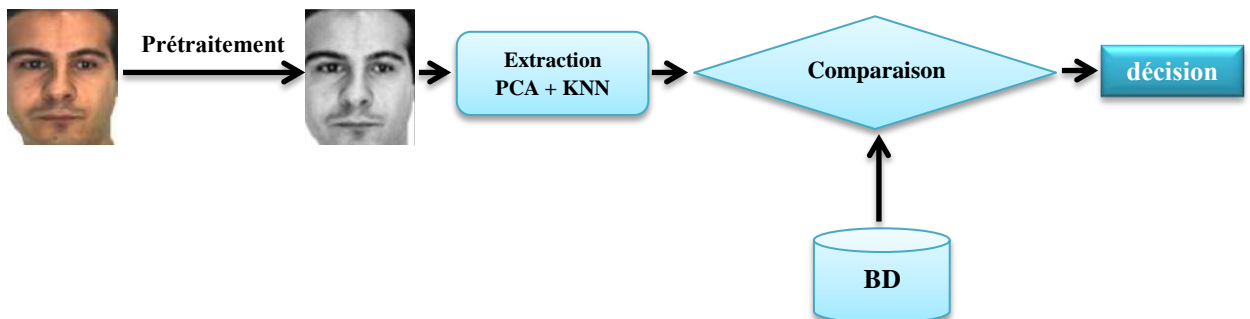


Figure 3.2. Architecture de la méthode proposée PCA +KNN

III.3. Les bases de données

III.3.1. La base de données ORL

Conçu par AT&T laboratoires de l'université de Cambridge en Angleterre, la base de données ORL (Olivetti Research Laboratory) est une base de données de référence pour les systèmes de reconnaissance automatique des visages. En effet tous les systèmes de reconnaissance de visages trouvés dans la littérature ont été testés par rapport à l'ORL, cette popularité est dû aux nombre de contraintes imposées par cette base car la plus part des changements possibles et prévisibles du visage ont été pris en compte, comme par exemple : le changement de coiffure, la barbe, les lunettes, les changements dans les expressions faciales, *etc.* Ainsi que les conditions d'acquisition telles que : le changement d'illumination et le changement d'échelle dû à la distance entre le dispositif d'acquisition et l'individu [31].

La base de données ORL est constituée de 40 individus, chaque individu possède 10 poses, donc la base contient 400 images. Les poses ont été prises sur des intervalles de temps différents pouvant aller jusqu'à trois mois. La détection des visages à partir des images a été faite manuellement. Nous présenterons dans ce qui suit les figures montrant les spécificités de la base de données de référence ORL [31].

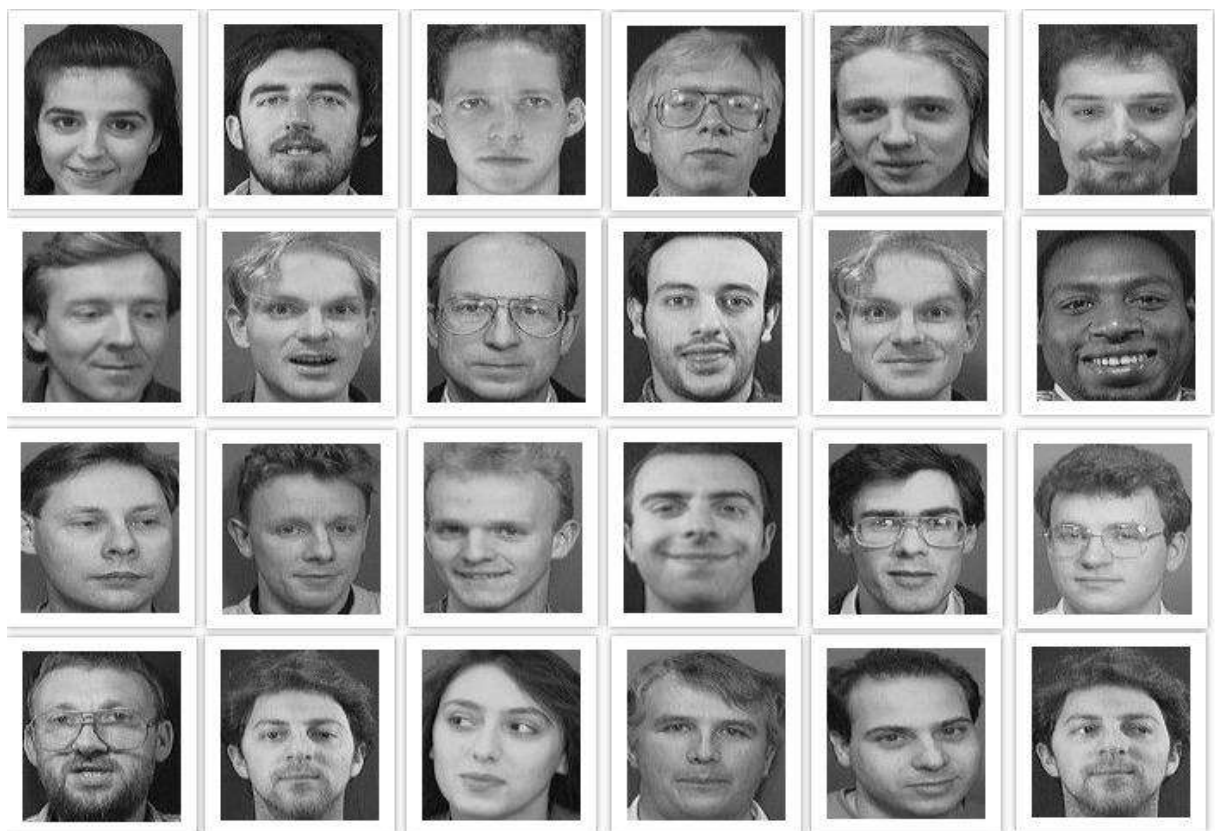


Figure 3.3. Extrait de la base ORL. Ensembles des vues collectées pour des individus de la base.

III.3.2. La base de données AR

La base AR a été constituée en 1998 au sein du laboratoire Computer Vision Center (CVC) à Barcelone, en Espagne. 116 personnes (63 hommes et 53 femmes) sont enregistrées. Les images sont en couleur de taille 768 x 576 pixels. 13 vues de chaque sujet ont été collectées. Pour la majorité de ces personnes, 13 autres vues ont été acquises lors d'une deuxième session à deux semaines d'intervalle [31].

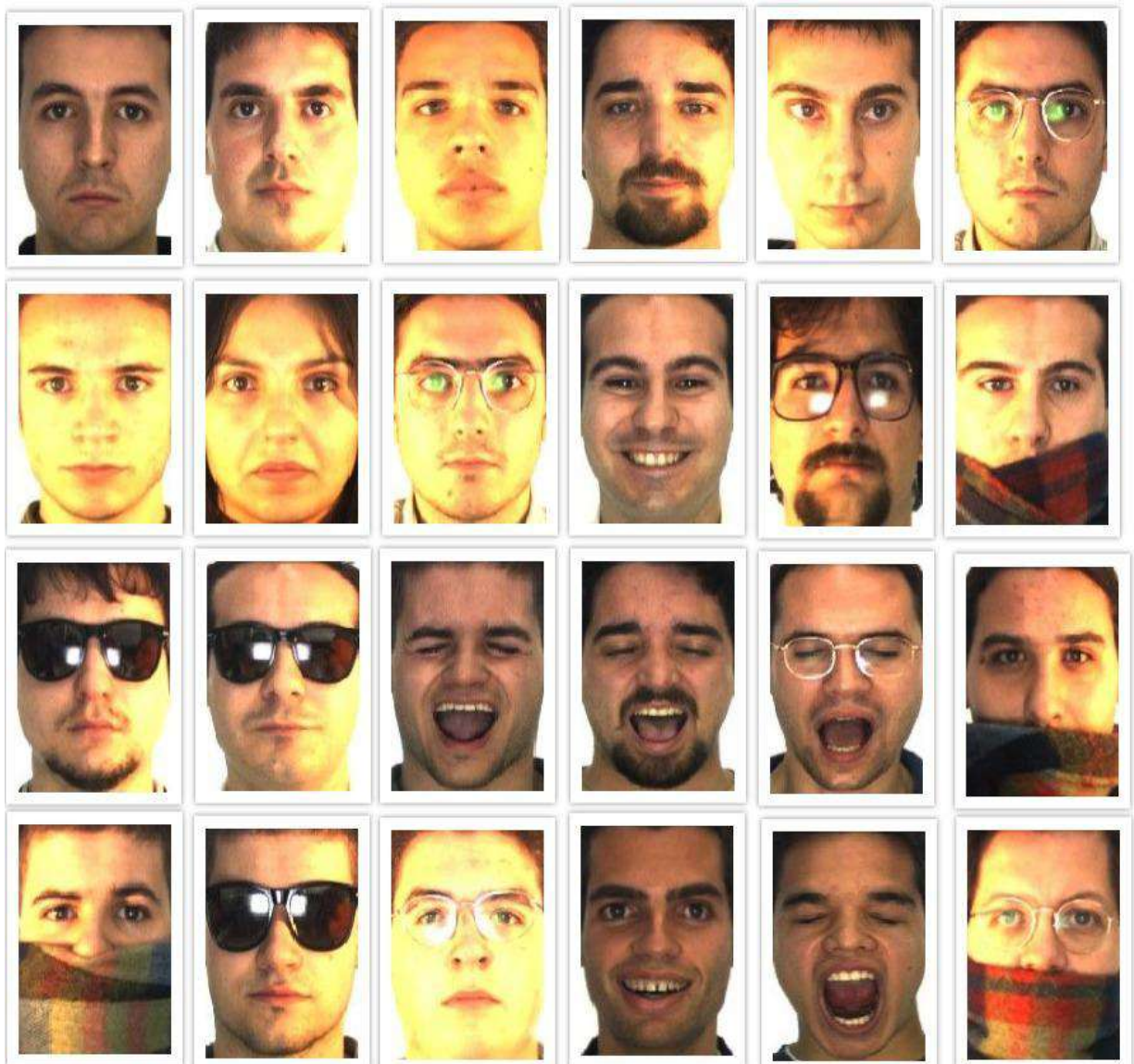


Figure 3.4. Extrait de la base AR. Ensembles des vues collectées pour des individus de la base.

III.4. Expérimentations et résultats

III.4.1. Expérience

Pour évaluer la performance des méthodes proposées, deux bases de données ORL et AR ont été utilisées avec les différents algorithmes (LDA et PCA) pour l'extraction des caractéristiques et KNN pour la classification. En outre, pour chaque algorithme nous avons présenté les résultats trouvés en fonction de plusieurs paramètres dans des tableaux.

III.4.2. Aspect matériel

Notre projet a été développé sur un micro portable:

- ✓ Processeur : Intel(R) Core (TM) i5-4210U CPU @ 1.70 GHz
- ✓ Capacité Mémoire (RAM) : 6.00 Go
- ✓ Vitesse d'horloge : 2.40 GHz
- ✓ Capacité disque dur : 500 Go
- ✓ Système d'exploitation : Windows 7 Edition intégrale.

III.4.3. Outils de développement

Pour la réalisation de notre système nous avons choisi le langage de programmation MATLAB Version R2018b. MATLAB est un environnement de calcul scientifique et de visualisation de données. Sa facilité d'apprentissage et d'utilisation (due à une syntaxe très claire) en ont fait un standard adapté pour les divers problèmes l'ingénierie.

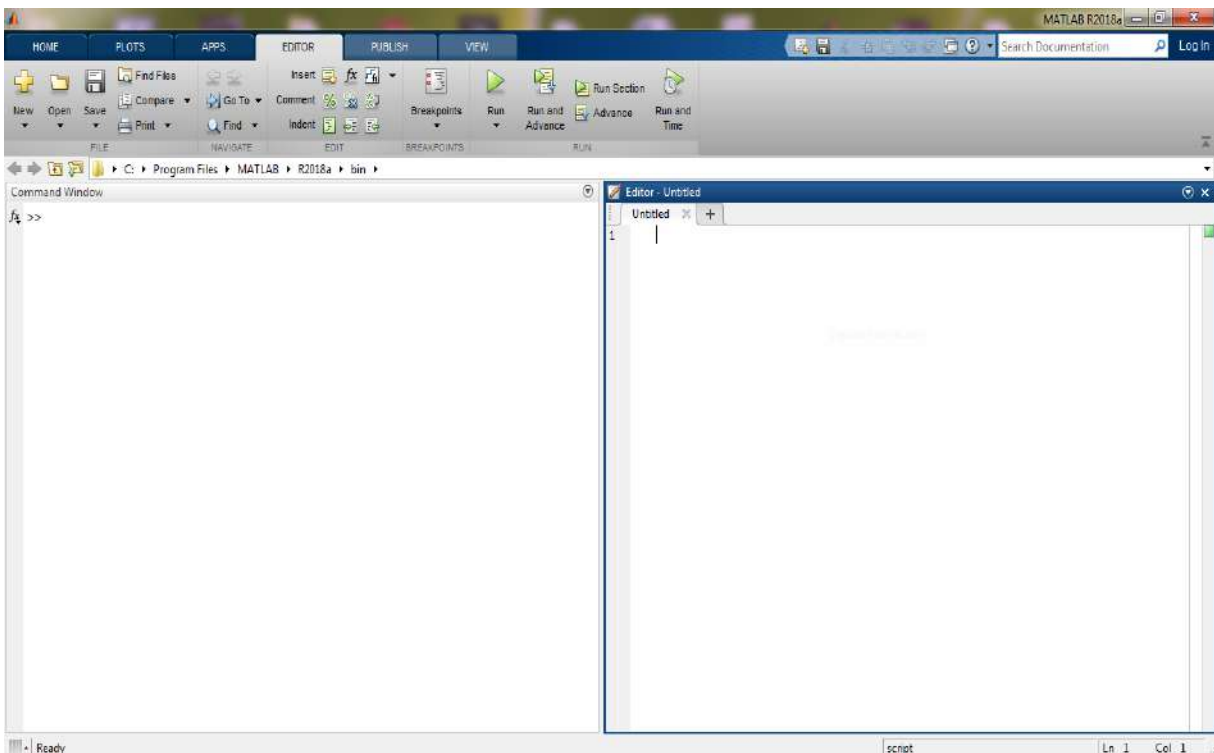


Figure 3.5. Interface du programme MATLAB

III.4.4. Résultats de la technique (LDA + KNN)

III.4.4.1. La base de données (ORL)

Dans l'expérience suivante, la méthode LDA +KNN a été appliquée sur la base de données ORL. 5 images ont été utilisées pour chaque personne pour l'entraînement et 5 images pour le test. On varie les différentes distances en fonction du nombre de voisins. Les résultats obtenus sont présentés dans le **tableau 3.1**.

Tableau 3.1. Le TR obtenus par les différentes distances et nombre de voisins

Nombre de voisins \ Distance	1	2	3	4	5	6	7	8	9
Cosine	90.5%	90.5%	91%	91%	91.5%	91.5%	91.5%	91%	90%
Seuclidien	84.5%	82.5%	83.5%	84.5%	85.5%	85.5%	84.5%	86%	85%
Mahalanobis	84.5%	82.5%	83.5%	84.5%	85.5%	85.5%	84.5%	86%	85%
Corrélation	90.5%	90.5%	91%	91%	92%	91.5%	91.5%	91%	90.5%
Spearman	88%	88%	88%	87%	86.5%	86%	86.5%	86%	86.5%

Après le tableau 3.1, il est clair que les différentes distances donnent des bons résultats. Ainsi que, le meilleur taux de reconnaissance égale **92%** a été obtenu par la distance "**correlation**" et le nombre de voisin égale à "**5**".

Le tableau 3.2 montre les résultats de taux de reconnaissance en fonction de la variation de nombre d'image d'entraînement.

Tableau 3.2. Taux de reconnaissance pour les différents nombre d'images d'entraînement sur la BD (ORL).

Images d'entraînement	2	3	4	5	6	7	8	9
TR (%)	66.56	86.42	88.33	92	92.5	95	95	97.5

On peut observer à partir de **tableau 3.2** et la **figure 3.6** que le taux de reconnaissance augmente avec le nombre des images d'entraînement. Ainsi que, nous avons obtenu un taux de reconnaissance de **92%** pour les images d'entraînement égales "**5**".

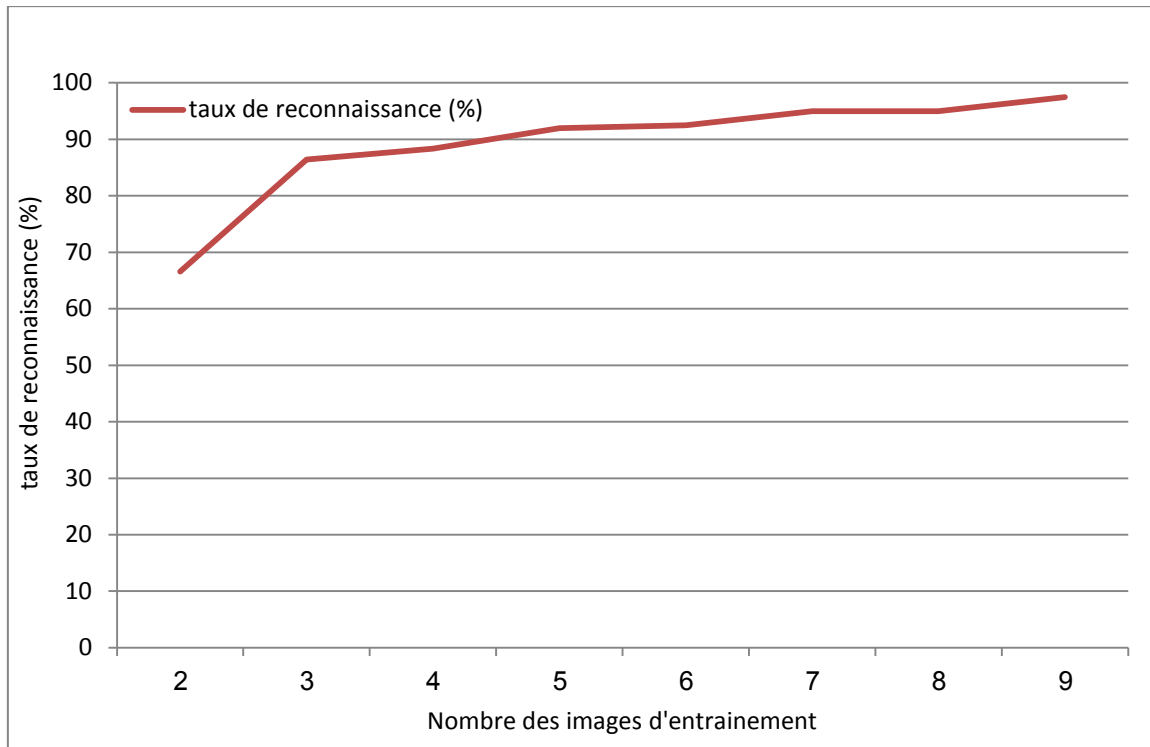


Figure 3.6. TR pour les différents nombre d'images d'entraînement sur la BD (ORL).

De plus, pour vérifier la rapidité d'identification de système biométrique. Nous effectuerons un autre test qui prend en considération et en particulier la mesure des temps de modélisation et classification. Les résultats sont montrés dans le **Tableau 3.3**.

Tableau 3.3. Temps de modélisation et classification pour les différents nombre d'images d'entraînement sur la BD (ORL).

Images d'entraînement	2	3	4	5	6	7	8	9
TM (s)	0.33	0.76	1.08	1.36	1.87	2.27	2.66	3.45
TC (s)	0.30	0.16	0.05	0.01	0.01	0.01	0.02	0.01

D'après les résultats obtenus du **tableau 3.3** et la **figure 3.7**, on remarque à chaque augmentation de nombre d'images d'entraînement, il ya une augmentation du temps de modélisation et par contre le temps de classification est diminué.

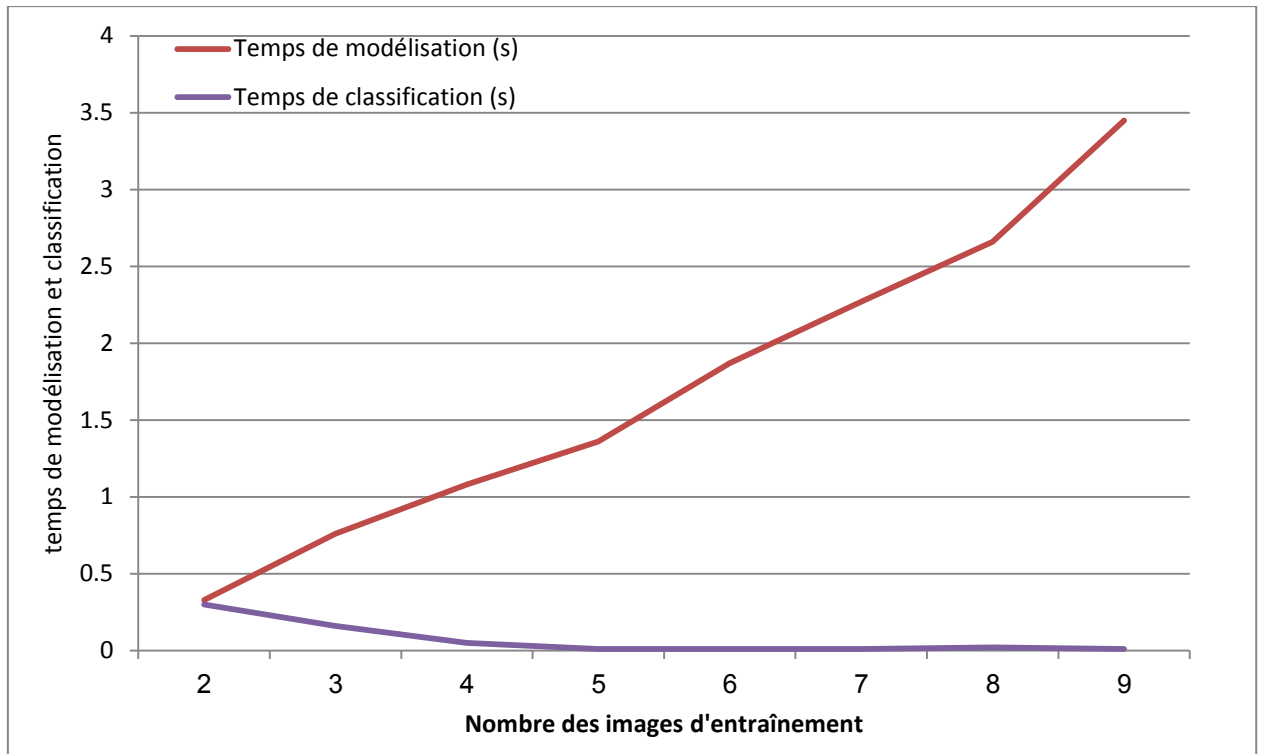


Figure 3.7. Temps de modélisation et classification pour les différents nombre d'image d'entraînement sur la BD (ORL).

III.4.4.2. La base de données (AR)

Dans l'expérience suivante différentes distances en fonction du nombre de voisins ont été employées. Les résultats obtenus sont présentés dans le **tableau 3.4**.

Tableau 3.4. Le TR obtenus par les différentes distances sur la BD (AR)

Nombre de voisins Distance	1	2	3	4	5	6	7	8	9
Cosine	92.57%	92.85%	92.57%	92.57%	92.42%	92.85%	92.71%	92.71%	93%
Seuclidien	89.85%	90.57%	90.57%	90.42%	90.57%	90.71%	90.57%	90.28%	90.71%
Mahalanobis	89.85%	90.57%	90.57%	90.42%	90.57%	90.71%	90.57%	90.28%	90.71%
Corrélation	92.57%	92.57%	92.42%	92.57%	92.42%	92.85%	92.57%	92.71%	92.85%
Spearman	92%	92.14%	91.85%	91.85%	91.85%	91.71%	91.85%	91.57%	91.71%

D'après le **tableau 3.4**, les bonnes résultats sont obtenu presque pour les différents distances. Ainsi que, le meilleur taux de reconnaissance qui égale à **93%** a été obtenue par la distance "**cosine**" avec un nombre de voisin "**9**".

Le tableau 3.5 montre les résultats du taux de reconnaissance en fonction de la variation de nombre d'images d'entraînement, en appliquant la méthode proposée LDA + KNN sur la base de donnée AR.

Tableau 3.5. Taux de reconnaissance pour les différents nombre d'images d'entraînement sur la BD (AR)

Images d'entraînement	2	3	4	5	6	7	8	9	10	11	12
TR (%)	30.41	59.45	77.10	83.33	86.37	93	89.5	87.4	83.75	79.33	67

Le **tableau 3.5** et la **figure 3.7** montrent que le taux de reconnaissance est croissant avec le nombre d'images d'entraînement dans l'intervalle de 2 jusqu'à 7 et de 8 jusqu'à 12 le taux de reconnaissance commence à diminuer. Nous remarquons aussi, les résultats de la base de données AR sont différents par rapport les résultats de la base de données ORL.

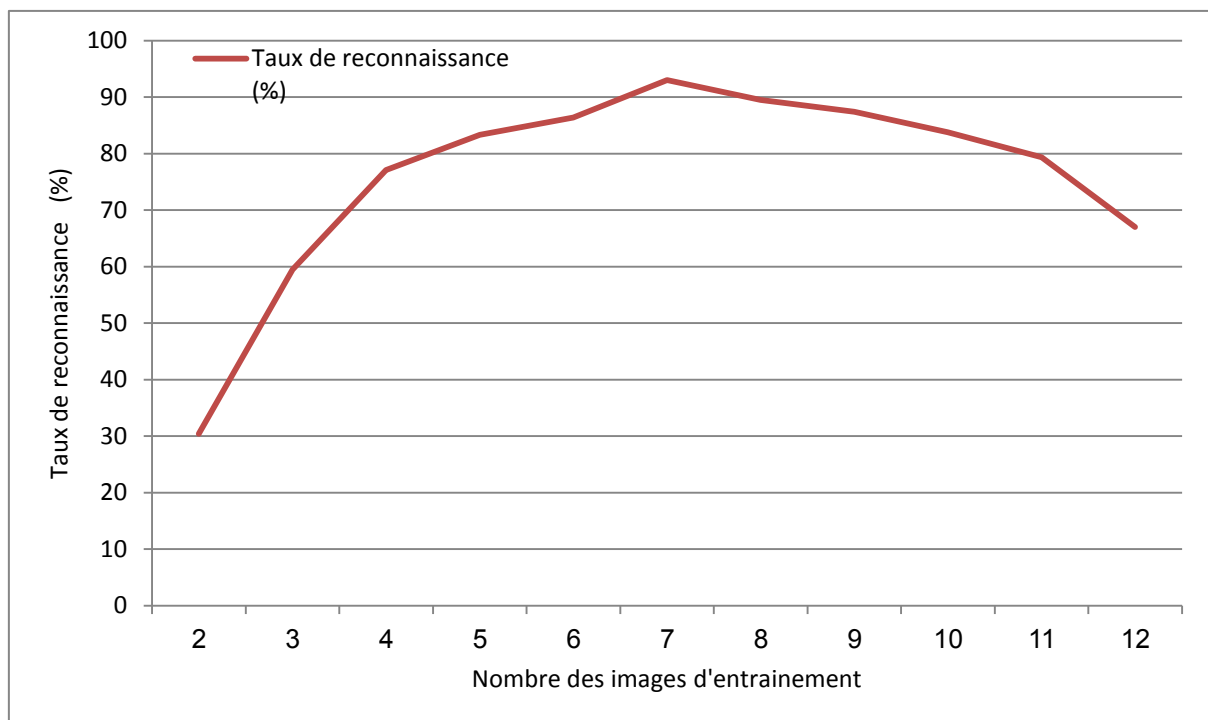


Figure 3.8. TR pour les différents nombre d'images d'entraînement sur la BD (AR)

De plus, pour vérifier la rapidité d'identification de système biométrique. Nous abordons un autre test qui prend en considération en particulier la mesure des temps de modélisation et classification qui montre dans le **tableau 3.6**.

Tableau 3.6. Temps de modélisation et classification pour les différents nombre d'images d'entraînement sur la BD AR

images d'entraînement	2	3	4	5	6	7	8	9	10	11	12
TM (s)	1.16	2.65	4.41	7.25	9.30	13	16.58	21.67	26.38	32.40	38.58
TC (s)	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.65	0.65	0.64

D'après les résultats obtenus du **tableau 3.6** et la **figure 3.9**, on observe pour cette fois où nous avons utilisé la base de données **AR**, le temps de modélisation est très élevé par rapport que nous obtenons des résultats avec la **BD ORL**, ceci est dû au nombre élevé des images dans cette base qui prend le temps beaucoup plus pour la modélisation à chaque augmentation du nombre d'image d'entraînement, par contre le temps de classification faible et presque constant

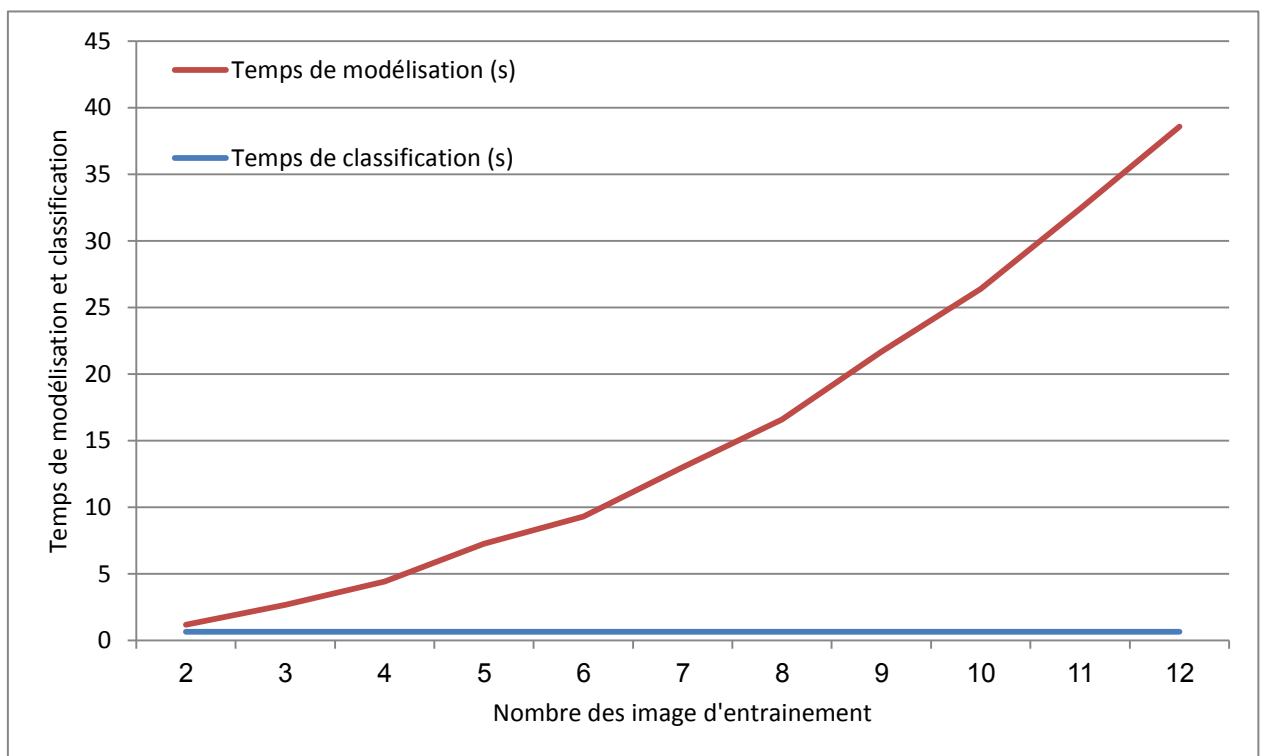


Figure 3.9. Temps de modélisation et classification pour les différents nombre d'image d'entraînement sur la BD (AR).

III.4.5. Résultats de la technique (PCA + KNN)

III.4.5.1. La base des données (ORL)

Dans la deuxième expérience avec une autre technique, et pour fixer la meilleure distance pour obtenir des bons résultats qui concerne le taux de reconnaissance, nous avons fait un ensemble des tests sur quelque distance sur la base de données ORL. Les résultats sont illustrés dans le **tableau 3.7**

Tableau 3.7. Le TR obtenus par les différentes distances.

Nombre de voisins Distance	1	2	3	4	5	6	7	8	9
Cosine	91%	87%	87%	81.5%	79%	76.5%	74.5%	72.5%	72%
Seuclidien	76%	65%	66%	68%	67%	66%	66.5%	66.5%	67.5%
Mahalanobis	76%	65%	66%	68%	67%	66%	66.5%	66.5%	67.5%
Corrélation	91%	86.5%	86.5%	83%	79%	75.5%	75%	71%	72.5%
Spearman	83.5%	74.5%	74.5%	76%	74.5%	73.5%	73.5%	70.5%	68.5%

Après le Tableau 3.7, on observe que les meilleurs résultats sont obtenus par les distances cosine et corrélation avec même nombre de voisins qui égale à 1, où **TR = 91%**. Nous avons fixé dans les prochaines expériences le nombre de voisin égale à 1 et la distance "**cosine**".

Le **tableau 3.8** montre les résultats du taux de reconnaissance en fonction de la variation de nombre d'images d'entraînement, en appliquant la méthode proposée PCA + KNN sur la base de données ORL.

Tableau 3.8. Taux de reconnaissance pour les différents nombre d'images d'entraînement sur la BD (ORL).

Images d'entraînement	1	2	3	4	5	6	7	8	9
TR (%)	71.38	83.75	85	88.75	91	95.62	95.83	96.25	95

A travers les résultats obtenus du **tableau 3.8** et la **figure 3.10**, Evidemment que la méthode proposée dépend de l'algorithme PCA prouvent l'augmentation du taux de reconnaissance qui est affecté par la variation de nombre d'entraînement. Donc l'augmentation de taux de reconnaissance remarquable dans cette base de données, donne un mode d'identification efficace.

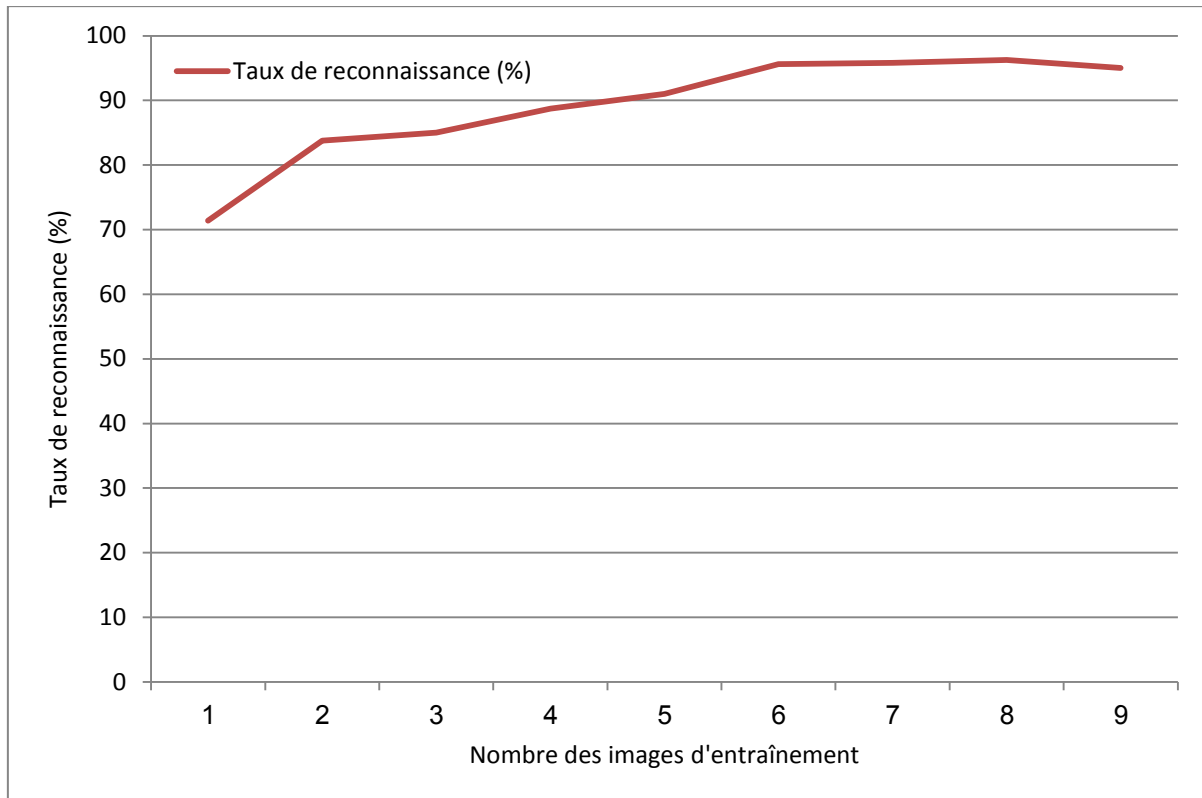


Figure 3.10. TR pour les différents nombre d'image d'entraînement sur la BD (ORL).

De plus, et pour vérifier la rapidité d'identification de système biométrique. Nous abordons un autre test qui prend en considération en particulier la mesure des temps de modélisation et classification. Les résultats sont illustrés dans le **tableau 3.9**.

Tableau 3.9. Temps de modélisation et classification pour les différents nombre d'image d'entraînement sur la BD (ORL).

Images d'entraînement	1	2	3	4	5	6	7	8	9
TM (s)	0.10	0.14	0.20	0.27	0.37	0.47	0.57	0.74	0.89
TC (s)	0.63	0.64	0.65	0.65	0.63	0.64	0.64	0.66	0.64

A partir les résultats obtenus **tableau 3.9** et la **figure 3.11**, on remarque à chaque augmentation du nombre d'image d'entraînement le temps de modélisation augmente par contre le temps de classification presque stable.

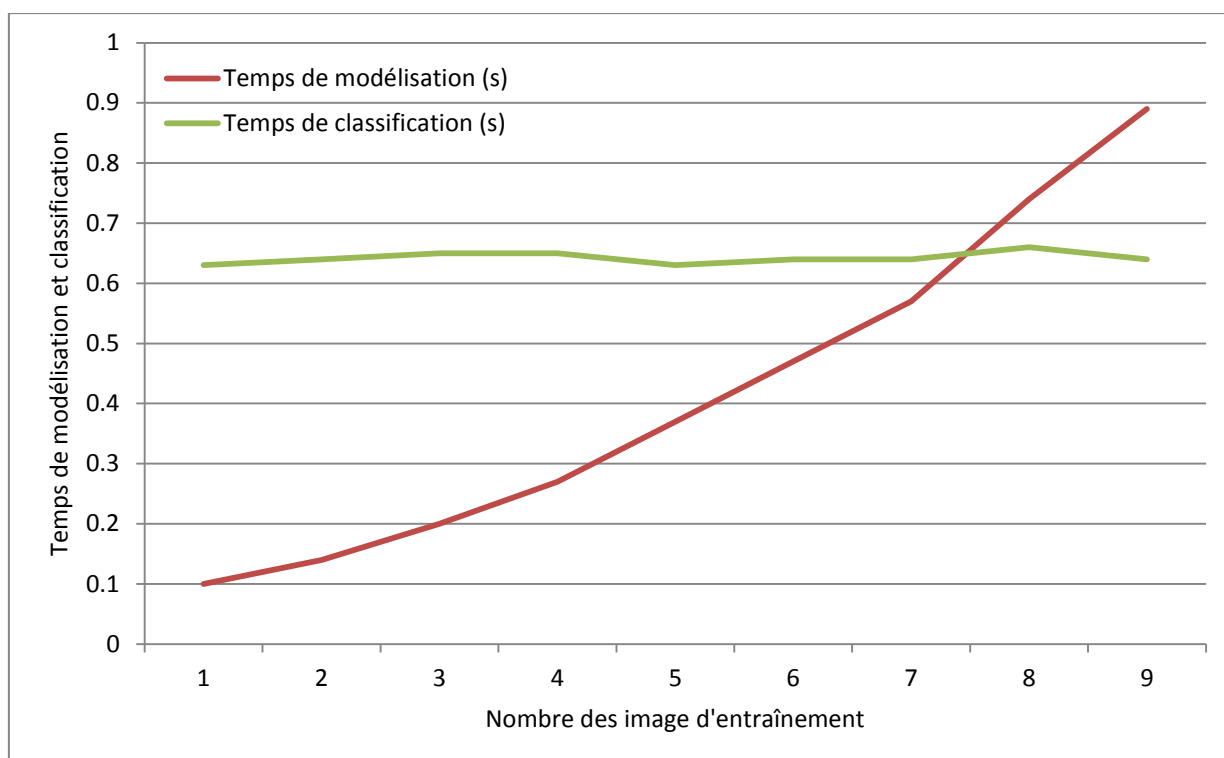


Figure 3.11. Temps de modélisation et classification pour les différents nombre d'image d'entraînement sur la BD (ORL).

III.4.5.2. La base de données (AR)

Dans l'expérience suivante et cette fois sur la base de donnée ORL des distances différentes, pour fixer la meilleure distance qui est relié avec le classification des caractéristiques, où cela affect positivement sur le taux de reconnaissance, et pour cela nous avons fait un ensemble des tests sur quelque distance, les résultats sont illustrés dans le **tableau 3.10**.

Tableau 3.10. Le TR obtenus par les différentes distances sur la BD (AR)

Nombre de voisins distance									
	1	2	3	4	5	6	7	8	9
Cosine	38.14%	33.14%	32.57%	34.85%	33.14%	30.85%	29.42%	28.57%	27%
Seuclidien	54%	48.28%	44.28%	66.57%	69.85%	75.14%	76.57%	78%	78.28%
Mahalanobis	54%	48.28%	42.22%	66.57%	69.85%	74.14%	76.57%	78%	78.28%
Corrélation	38.14%	33.28%	32.57%	34.85%	33.28%	30.71%	29.57%	28.71%	26.85%
Spearman	67.14%	65.71%	66%	81.85%	84.14%	85.14%	85.42%	85.57%	85.57%

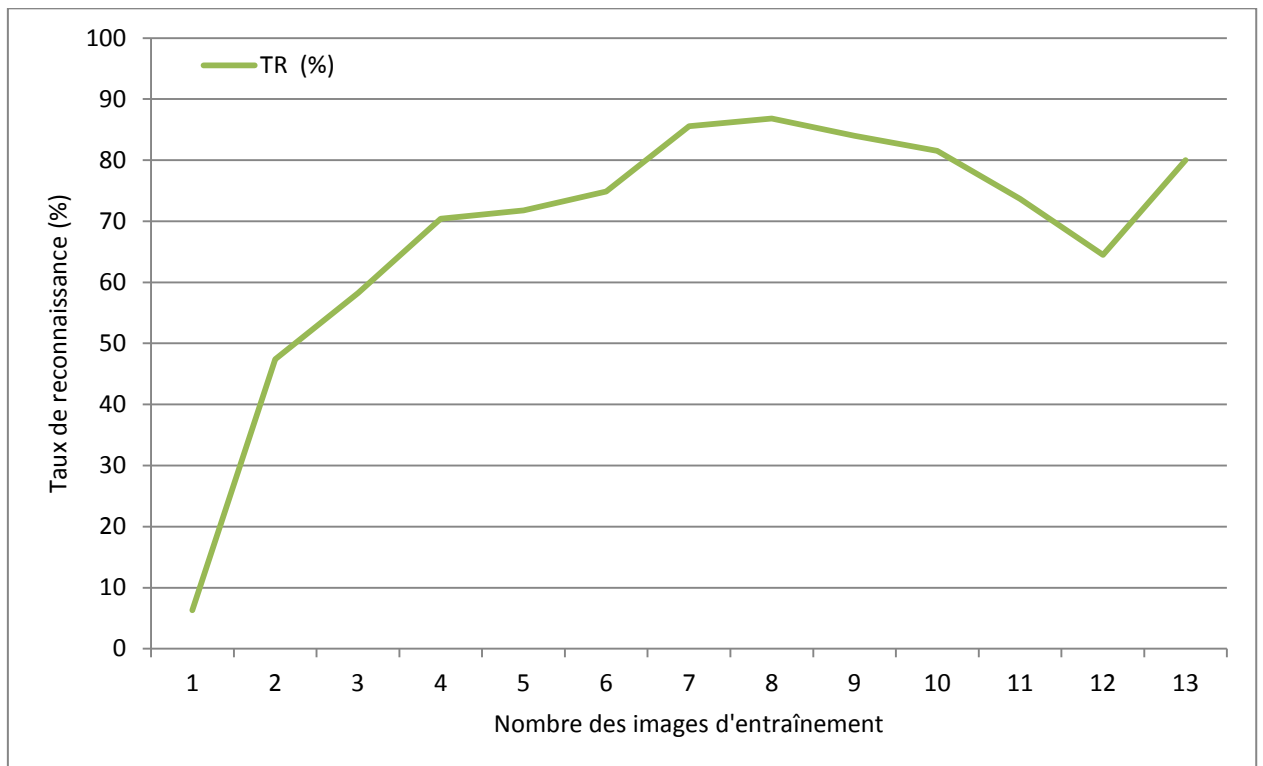
Après les résultats obtenus, nous observons que la distance "Spearman" avec nombre de voisins "8" et/ou "9" donne un meilleur taux de reconnaissance égale à 85.57 %, alors nous avons fixé cette distance pour compléter les parties suivantes.

Le tableau 3.11 montre les résultats de taux de reconnaissance en fonction de la variation de nombre d'image d'entraînement, en appliquant la méthode proposée PCA+ KNN sur la base de données AR.

Tableau 3.11. Taux de reconnaissance pour les différents nombre d'image d'entraînement sur la BD (AR)

Images d'entraînement	1	2	3	4	5	6	7	8	9	10	11	12	13
TR (%)	6.30	47.41	58.27	70.40	71.77	74.87	85.57	86.83	84	81.5	73.66	64.5	80

D'après les résultats obtenus par le **tableau 3.11** et la **figure 3.12** et d'après l'exécution du programme sur la base de données **AR**. Nous remarquons que, les résultats sont faible par rapport les résultats de la base de données **ORL**. On observe aussi, le taux de reconnaissance est 6.30% quand on a sélectionné une seul image pour l'entraînement.



La figure 3.12. TR pour les différents nombre d'image d'entraînement sur la BD (AR)

De plus, et pour vérifier la rapidité d'identification de système biométrique. Nous abordons un autre test qui prend en considération en particulier la mesure des temps de modélisation et classification qui montre dans le **tableau 3.12**.

Tableau 3.12. Temps de modélisation et classification pour les différents nombre d'image d'entraînement sur la BD (AR).

Images d'entraînement	1	2	3	4	5	6	7	8	9	10	11	12	13
TM (%)	0.72	1.8	3.43	5.74	8.46	11.88	15.39	20.1	27.05	31.11	38	45.85	94.06
TC (%)	1.53	1.54	0.64	0.67	0.64	0.64	0.65	0.64	0.64	0.66	0.65	0.64	0.66

D'après les résultats obtenus du **tableau 3.12** et la **figure 3.13** on observe pour cette fois où nous avons utilisé la base de données AR, le temps de modélisation est très élevé par rapport que nous obtenons des résultats avec la BD ORL, ceci est dû au nombre élevé des images dans cette base qui prend le temps beaucoup plus pour la modélisation à chaque augmentation du nombre d'image d'entraînement, par contre le temps de classification diminue.

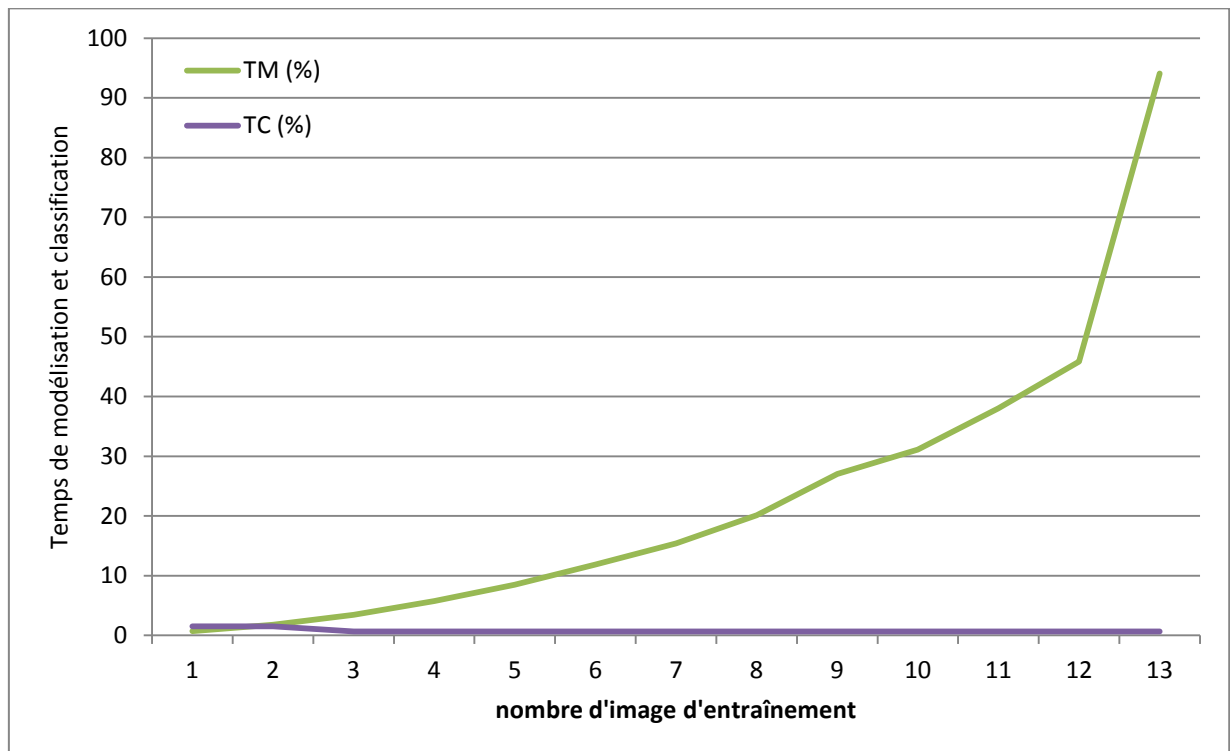


Figure 3.13. Temps de modélisation et classification pour les différents nombre d'image d'entraînement sur la BD (AR)

Tableau 3.13. Comparaison entre les deux méthodes proposés

	LDA + KNN						PCA + KNN					
	Base donnée ORL			Base donnée AR			Base donnée ORL			Base donnée AR		
	TR (%)	TM (s)	TC (s)	TR (%)	TM (s)	TC (s)	TR (%)	TM (s)	TC (s)	TR (%)	TM (s)	TC (s)
2 img	66.56	0.33	0.30	30.41	1.16	0.64	83.75	0.14	0.64	47.41	1.80	1.54
3 img	86.42	0.76	0.16	59.45	2.65	0.64	85	0.20	0.65	58.27	3.43	0.64
4 img	88.33	1.08	0.05	77.01	4.41	0.64	88.75	0.27	0.65	70.40	5.74	0.67
5 img	92	1.36	0.01	83.33	7.25	0.64	91	0.37	0.63	71.77	8.46	0.64
6 img	92.5	1.87	0.01	86.37	9.30	0.64	95.62	0.47	0.64	74.87	11.88	0.64

III.5. Conclusion

Dans ce chapitre, nous avons comparé deux techniques LDA+KNN et PCA+KNN parmi les méthodes globales qu'il existe dans la littérature. Les résultats expérimentaux montrent que les méthodes proposées présentent de meilleures performances, sauf que les résultats trouvés lorsque nous avons utilisé la base de données **AR** qui porte un grand nombre d'images. La technique LDA+KNN donne une meilleure performance comparant avec la technique PCA+KNN.

Conclusion générale

Conclusion générale

L'objectif assigné de ce travail consiste à étudier et comparer les différentes méthodes holistiques. Dans ce travail, les performances du système de reconnaissance faciale sont mesurées par le taux de reconnaissance et le temps de modélisation et de classification.

Dans notre étude, on a présenté premièrement une vue générale de la biométrie, en survolant sur quelques techniques biométriques et on a présenté l'architecture et les modules de ces systèmes. Nous avons donné aussi un aperçu sur les modes de fonctionnement d'un système biométrique et les techniques des mesures.

À partir des résultats expérimentaux, la méthode LDA+KNN avec un nombre d'images d'entraînement égale à 5 donne un taux de reconnaissance égale à 92% pour la base de données ORL et 83.33% pour la base de données AR. Aussi que, la méthode PCA+KNN avec un nombre d'images d'entraînement égale à 5 donne un taux de reconnaissance égale à 91% pour la base de données ORL et 71.77% pour la base de données AR. Cette étude indique et prouve que LDA construit un discriminant pour distinguer les visages des différentes personnes, permet donc d'effectuer une véritable séparation de classes.

À partir des résultats de la simulation, on peut conclure que la performance d'utilisation de l'analyse discriminante linéaire (LDA) avec la classification de KNN est meilleure que celle de l'utilisation de l'analyse en composantes principal (PCA) avec la classification KNN.

Perspectives

1. Compare d'autres méthodes holistiques tel que KPCA et ICA.
2. Tester les méthodes proposées avec d'autres modalités telles que l'empreinte digitale, etc.
3. Fusion les scores de PCA+KNN avec LDA+KNN pour améliorer les performances des systèmes biométriques.

Bibliographie

- [1] **GHALI, Mr. Ahmed.** *université des sciences et de la technologie d'oran mohamed boudiaf.* oran : Laboratoire Modélisation et Optimisation des Systèmes Industriels, 2015.
- [2] **CHAA, Mr. Mourad.** *système de reconnaissance de personne par des techniques biométriques.* setif : Université Ferhat Abbas – Sétif-1 -, 2017.
- [3] **BOUZIDI, Adel.** *Système de reconnaissance des empreintes palmaires.* biskra : Université Mohamed Khider de Biskra, 2018.
- [4] **NOËL, Bila.** Biométrie. *www.wikipedia.org.* [En ligne] wikipedia, 4 septembre 2014. [Citation : 21 janvier 2019.] <https://fr.wikipedia.org/wiki/Biometrie>
- [5] **MORIZET, Nicolas.** *Reconnaissance Biométrique par Fusion Multimodale du Visage et de l'Iris.* paris : École Doctorale d'Informatique, Télécommunication et Électronique de Paris, 2009
- [6] **HADJAR, Mr. Ahmed.** *Identification des individus par la biométrie multimodale.* oran : universite des sciences et de la technologie d'oran mohamed, 2014.
- [7] **HENNO, Jacques.** *Comment la biométrie va envahir nos vies..* 10, paris : lesechos, 2012, Vol. I.
- [8] **GUILLERM, Didier.** Le marché de la biométrie. *www.biometrie-online.net.* [En ligne] The Global Biometrics , 21 mars 2016. [Citation : 18 septembre 2016.] http://www.acuity-mi.com/GBMR_Report.php.
- [9] **BABICH, Aleksandra.** *Biometric authentication. Types of biometric identifiers.* Helsinki : university of applied sciences, 2012.
- [10] **ASHBOURN, Julian.** *Guide To Biometrics For Large-Scale Systems.* springer : s.n, 2011.
- [11] **FEDIAS, Meriem.** *Combinaisons de données d'espaces couleurs et de méthodes de vérification d'identité pour l'authentification de visages.* Biskra : université Mohamed khider.
- [12] **JAIN, Anil K, FLYNN, Patrick, et ROSS, Arun A.** (ed.). *Handbook of biometrics.* Springer Science & Business Media, 2007.
- [13] **JAIN, Anil, BOLLE, Ruud, et PANKANTI, Sharath.** Introduction to biometrics. In : *Biometrics.* Springer, Boston, MA, 1996. p. 1-41.
- [14] **ADJOUT Mohamed et BENAÏSSA Abdelhak** « Fusion de la DCT-PCA et la DCT-LDA appliquée à la reconnaissance de visages » Institut National de formation en Informatique (I.N.I) Oued-smar Alger 2007.
- [15] **MORIZET, Nicolas** . "Reconnaissance Biométrique par Fusion Multimodale du Visage et de l'Iris". 2009.

Bibliographie

- [16] **GUERFI ABABSA Souhila** -Authentification d'individus par reconnaissance de caractéristiques biométriques liées aux visages 2D/3D-Docteur de l'Université Evry Val d'Essonne Spécialité-03 octobre 2008.
- [17] **OUAMANE, Hanane.** *identification de reconnaissance faciale avec des expressions . biskra* : Université Mohamed Khider Biskra, 2012.
- [18] **MIKA, Sebastian, RATSCH, Gunnar, WESTON, Jason, et al.**Fisher discriminant analysis with kernels. In : Neural networks for signal processing IX: Proceedings of the 1999 IEEE signal processing society workshop (cat. no. 98th8468). Ieee, 1999. p. 41-48.
- [19] **ARCA, Stefano, CAMPADELLI, Paola, et LANZAROTTI, Raffaella.** A face recognition system based on automatically determined facial fiducial points. *Pattern recognition*, 2006, vol. 39, no 3, p. 432-443
- [20] **PETROVSKA-DELACRÉTAZ, Dijana, CHOLLET, Gérard, et DORIZZI, Bernadette.** *Guide to biometric reference systems and performance evaluation*. Berlin : Springer, 2009.
- [21] "Molecular Expressions Microscopy Primer: Digital Image Processing – Difference of Gaussians Edge Enhancement Algorithm", Olympus America Inc., and Florida State University Michael W. Davidson, Mortimer Abramowitz
- [22] **NAVAZ, AS Syed, DHEVISRI, T, et MAZUMDER, Pratap.** Face recognition using principal component analysis and neural networks. *March-2013, International Journal of Computer Networking, Wireless and Mobile Communications. Vol, 2013, no 3, p. 245-256.*
- [23] **LOUIBA Fadia et HADJ ALI Ryma** “Système de contrôle d’accès physique Basé sur le visage“ .pp.19 ; 2010.
- [24] **MORIZET, Nicolas, THOMAS, E. A, ROSSANT, Florence, et al.** Revue des algorithmes PCA LDA et EBGMM utilisés en reconnaissance 2D du visage pour la biométrie. *P1-11. Institut Supérieur d' Electronique de Paris (ISEP), département d' Electronique*, 2006.
- [25] **ETEMAD, Kamran et CHELLAPPA, Rama.** Discriminant analysis for recognition of human face images. *Josa a*, 1997, vol. 14, no 8, p. 1724-1733.
- [26] **YU, Pengfei, YU, Pengcheng, et XU, Dan.** Comparison of PCA, LDA and GDA for palmprint verification. In : *2010 International Conference on Information, Networking and Automation (ICINA)*. IEEE, 2010. p. V1-148-V1-152
- [27] **BELHUMEUR, Peter N, HESPANHA, João P, et KRIEGMAN, David J.** Eigenfaces

Bibliographie

- vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 1997, no 7, p. 711-720.
- [28] **YANG, Ming-Hsuan.** Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition Using Kernel Methods. In : *Fgr.* 2002. p. 215.
- [29] **SWATI, M. R. et RAVISHANKAR, M.** Finger Knuckle Print recognition based on Gabor feature and KPCA+ LDA. In : *2013 International Conference on Emerging Trends in Communication, Control, Signal Processing and Computing Applications (C2SPCA)*. IEEE, 2013. p. 1-5.
- [30] **CHOMBOON, Kittipong, CHUJAI, Pasapitch, TEERARASSAMEE, Pongsakorn, et al.** An empirical study of distance metrics for k-nearest neighbor algorithm. In : *Proceedings of the 3rd International Conference on Industrial Application Engineering*. 2015. p. 1-6.
- [31] **BOUCHRA, Khefif.** *Mise au point d'une application de reconnaissance faciale*. Tlemcen : Université Ab ou Bakr Belkaid – Tlemcen, 2013.

