

Chapitre 1

Système de Reconnaissance de Visage.

1.1 Introduction

Dans tous les domaines de sécurité et de contrôle d'accès, on utilise des mots de passe ou des clés qui se compose de chiffres ou de lettres. Mais, dans ces derniers temps avec l'avancement de la technologie ces mots de passe sont devenus facilement falsifiables et franchissables. C'est pour ça, que les chercheurs de différents domaines ont orientés leurs travaux sur des clés et mots de passe impossible à falsifier, sûr et surtout efficace. La biométrie est devenue à la mode dans les domaines qui requiert un niveau élevé de sécurité et de contrôle [1][2]. Et parmi toutes les technologies biométriques qui existent, la reconnaissance des visages est l'une des technologies les plus utilisées et les plus adaptées. Dans ce chapitre, nous allons mettre en relief quelques notions de base liées à la biométrie. Nous donnerons le principe de fonctionnement des systèmes biométriques, les diverses technologies et les outils utilisés pour mesurer leurs performances ainsi que les domaines d'applications. Plusieurs méthodes de reconnaissance de visages ont été proposées durant ces trente dernières années, suivant deux grands axes : la reconnaissance à partir d'images fixes et la reconnaissance à partir de séquence d'images (vidéo). La reconnaissance de visages basée sur la vidéo est préférable à celle basée sur des images fixes, puisque l'utilisation simultanée des informations temporelles et spatiales aide dans la reconnaissance. *Dans notre étude nous nous intéressons à la reconnaissance basée sur les images fixes, puisqu'elle représente la base de tous les travaux.* La reconnaissance de visages est un gros challenge tellement intéressant, qu'il a attiré les chercheurs des différents domaines : psychologie, identification de modèles, réseaux de neurones, vision d'ordinateur, infographie ... C'est pour cela, que la littérature est si vaste et si diverse. On peut répartir l'ensemble des techniques de reconnaissance de visages basées sur les images fixes, en trois grandes catégories : les *méthodes globales*, les *méthodes locales* et les *méthodes hybrides*. [3,4]

Relation entre l'identification et l'authentification Ces deux approches sont complémentaires. D'une manière générale, les systèmes de sécurité reposent sur l'usage judicieux de ces deux fonctions combinées. L'identification servira à vérifier que le demandeur n'a pas sa photo dans la base de données des interdits. L'authentification servira à vérifier que le demandeur est autorisé pour l'utilisation d'un droit donné. Grâce à cette

approche globale, la délivrance de vrais droits à partir de demandes reposant sur de fausses identités est exclue ou extrêmement difficile.

1.2 Système biométrique

1.2.1 Les phases d'un système biométrique

Chaque système biométrique comporte deux phases d'exécution :

- La phase d'apprentissage
- La phase de test.

Chacune contient trois modules de base plus un module propre à la phase de test.

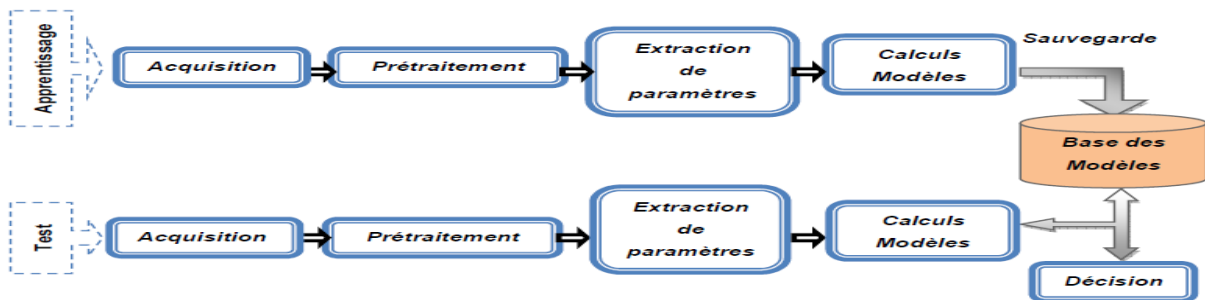


Figure 1.1 Architecture d'un système biométrique.

1.2.2 Modules de systèmes biométriques

Un système biométrique typique peut être représenté par quatre modules principaux :

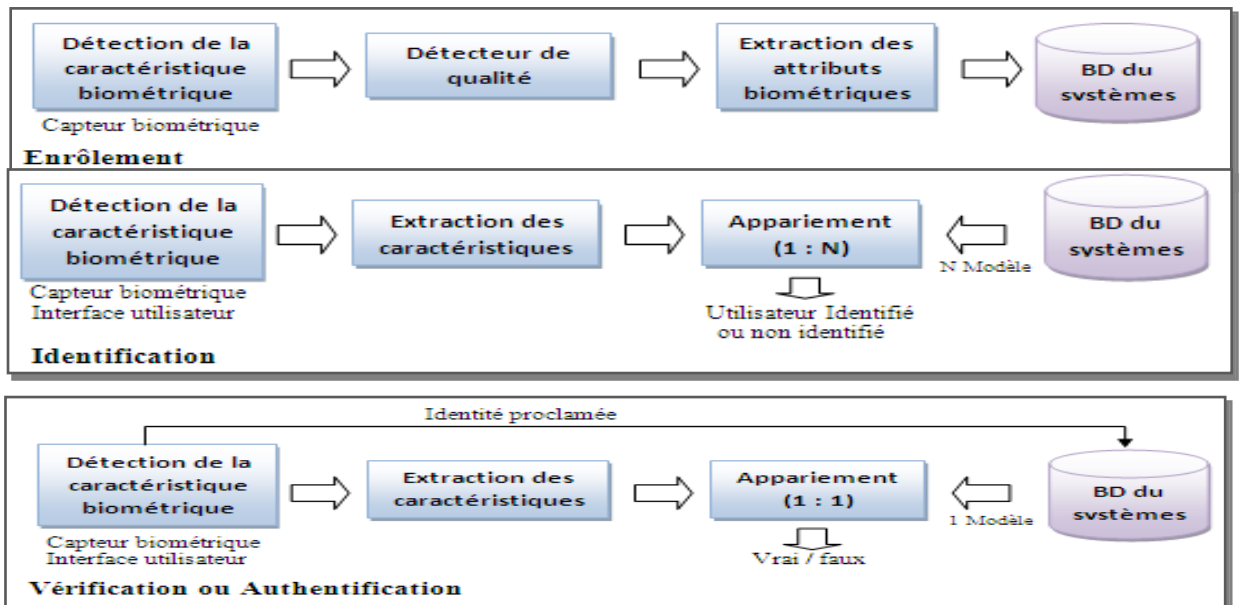


Figure 1.2 Les différents modules d'un système biométrique.

1. **Module de capture** : responsable de l'acquisition des données biométriques d'un individu (cela peut être un appareil photo, un lecteur d'empreintes digitales, une caméra de sécurité, etc).

2. **Module d'extraction de caractéristiques** : Qui prend en entrée les données biométriques acquises par le module de capture et extrait seulement l'information

pertinente afin de former une nouvelle représentation des données. Idéalement, cette nouvelle représentation est censée être unique pour chaque personne et relativement invariante aux variations intra-classe.

3. Module de correspondance : Il compare l'ensemble des caractéristiques extraites avec le modèle enregistré dans la base de données du système et détermine le degré de similitude (ou de divergence) entre les deux.

4. Module de décision : 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).

1.3 Pourquoi choisir le visage ?

La reconnaissance de visages est la technique la plus commune et populaire. Elle reste la plus acceptable puisqu'elle correspond à ce que les humains utilisent dans l'interaction visuelle; et par rapport aux autres méthodes, la reconnaissance du visage s'avère plus avantageuse, d'une part c'est une méthode non intrusive, c'est-à-dire elle n'exige pas la coopération du sujet (en observant les individus à distance), et d'une autre part les capteurs utilisés sont peu coûteux (une simple caméra) contrairement à l'empreinte digitale et l'iris où le sujet devra être très proche du capteur et devra coopérer pour l'acquisition de l'image sans oublier le coût de l'équipement nécessaire pour l'acquisition (équipement spécial coûteux). Malgré que certains disent que la reconnaissance de visages est une biométrie relativement peu sûre, sur le fait que le signal acquis est sujet à des variations beaucoup plus élevées que d'autres caractéristiques, comme la variation de l'éclairage, le changement de la position du visage, la présence ou l'absence de lunettes et autres; mais, au cours de ces dernières années plusieurs techniques de traitements d'images sont apparues, telle que la détection du visage, la normalisation de l'éclairage, etc. Sans oublier le développement considérable des technologies des caméras numériques, ce qui néglige l'effet de ces problèmes.

Parmi les six attributs biométriques considérés, les caractéristiques faciales marquent un score de compatibilité le plus élevé dans un *système MRTD* ("Machine Readable Travel Documents"), ce score étant basé sur **plusieurs facteurs d'évaluation** tels que l'enrôlement, le renouvellement des données, les requis matériels et la perception des utilisateurs [3,5]

Tableau 1.1 Avantages et inconvénients de la Reconnaissance de Visage.

Avantages	Inconvénients
<p>Bien accepté par le public. Aucune action de l'utilisateur (peu intrusive). Pas de contact physique. Technique peu coûteuse.</p>	<p>Technologie sensible à l'environnement (éclairage, position, expression du visage...) Difficultés de différencier de vrais jumeaux. Sensible aux changements. (barbe, moustache, lunettes, piercing, chirurgie...)</p>

1.4 Système de reconnaissance de visages

Comme nous l'avons déjà cité un système automatique de reconnaissance de visages doit intégrer une étape d'apprentissage durant laquelle il associe l'allure du visage à l'identité d'une personne. Cette étape est réalisée chez les êtres humains d'une façon spontanée et évolutive [6]. Dans un système artificiel, cette étape permet de construire une base de données des personnes connues, stockant des images étiquetées des identités. Pour ce faire, un système automatique comporte deux modes de fonctionnement : un mode **enrôlement** et un mode **identification**. Le premier mode sert à extraire pour chaque personne les éléments caractéristiques et les met sous la forme d'un vecteur caractéristique, appelé par la suite signature. Cette dernière, associée à une étiquette d'identité, sera stockée dans une base de données dédiée. Le mode d'identification permet de reconnaître une personne à partir de son image faciale, c'est à dire de retrouver l'identité associée à l'image. [7]

1.4.1 Architecture générale Idéalement, un système de reconnaissance faciale doit pouvoir identifier des visages présents dans une image ou une vidéo de manière automatique. Le système peut opérer dans les deux modes suivants : authentification ou identification. On peut également noter qu'il existe un autre type de scénario de reconnaissance faciale mettant en jeu une vérification sur une liste de surveillance, où un individu est comparé à une liste restreinte de suspects. Le principe de fonctionnement de base d'un système de reconnaissance faciale (**figure 1.3**) peut être résumé **en trois étapes** après l'acquisition.

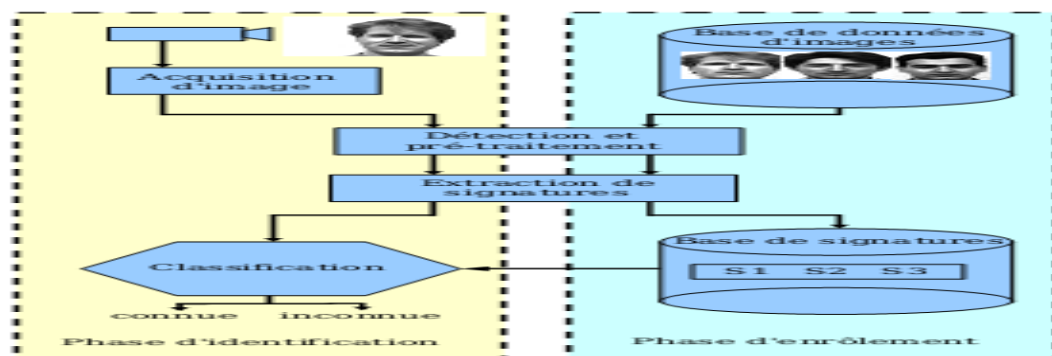


Figure 1.3 Fonctionnement d'un système de reconnaissance automatique de visages.

Les méthodes appliquées dans les modules formant le système dépendent essentiellement de la technique d'acquisition d'images de visage. En effet, s'il s'agit d'un scanner tridimensionnel ou d'un système d'acquisition stéréoscopique, la détection de visages et l'extraction de signatures reposent sur des techniques de traitement 3D [8]. Cette thématique est actuellement en pleine expansion. Elle met en évidence l'information de profondeur qui enrichit les données utilisées aussi bien dans la phase d'enrôlement que dans la phase d'identification. Toutefois, elle est limitée par le prix élevé de l'appareil d'acquisition (s'il s'agit d'un scanner 3D) et la difficulté d'installation (s'il s'agit d'un système stéréoscopique). Cette limite laisse le traitement 2D des images de visages plus accessible et encore très exploré, tant dans le domaine académique que dans le milieu industriel. Les travaux de recherche menés dans cette thèse portent sur l'utilisation des images 2D et 3D de visages. Le système comporte deux modules qui appliquent les mêmes traitements dans les deux modes de fonctionnement.

Le premier module permet de détecter le visage dans l'image brute captée par le système d'acquisition, il fournit ainsi au second module une image contenant seulement le visage. À partir de celle-ci, le second module extrait la signature discriminante. Elle sera soit stockée soit utilisée pour la classification. En **mode hors-ligne (enrôlement)**, elle sera étiquetée et stockée dans une base de données dédiée. En **mode en ligne (identification)**, elle servira comme entrée à un troisième module qui s'occupe de la classification de cette signature et de fournir la décision finale : personne inconnue ou personne connue.

1.4.1.1 Acquisition de l'image

La *capture* est la première étape dans le processus. Il faut réussir à capter l'information pertinente sans bruit. Dans la reconnaissance de visage on peut utiliser *les capteurs 3D* par exemple pour s'affranchir des problèmes de pose. Mais leur prix excessif ne permet pas une utilisation à grande échelle. Les capteurs en proche infrarouge sont utilisés pour éliminer les problèmes de l'illumination. Le *codage* consiste en l'acquisition d'image et sa digitalisation, ce qui donne lieu à une représentation bidimensionnelle au visage. L'image dans cette étape est dans un état brut ce qui engendre un risque de bruit qui peut dégrader les performances du système et donne lieu à une représentation **2D** (la matrice des niveaux de gris) pour un objet **3D** (le visage). [9]

1.4.1.2 Détection de visage et prétraitement

1) Détection : La détection de visage peut se faire par détection de la couleur de la peau, la forme de la tête ou par des méthodes détectant les différentes caractéristiques du

visage. Cette étape est autant plus délicate quand l'image acquise contient plusieurs objets de visage ou un fond non uniforme qui crée une texture perturbant la bonne segmentation du visage. Cette étape est dépendante de la qualité des images acquises. Dans la littérature scientifique, le problème de localisation de visages est aussi désigné par la terminologie "détection de visages". Les performances globales de tout système automatique de reconnaissance dépendent amplement des performances de la détection de visages. Dans l'étape de détection, on identifie et on localise le visage dans l'image acquise au départ, indépendamment de la position, de l'échelle, de l'orientation et de l'éclairage. C'est un problème de classification où on assigne l'image à la classe visage ou à la classe non visage. On peut diviser les approches de détection en quatre catégories : les méthodes basées sur la connaissance où on code la connaissance humaine du visage, les méthodes de correspondance de masques, les méthodes à caractéristiques invariables où on utilise la couleur, les textures et les contours et finalement les méthodes les plus répandues et qui sont ceux basées sur l'apprentissage ou les statistiques comme **PCA**, **SVM**, **Graph matching...**[10,11,12,13,14,15] Plusieurs travaux de recherches ont été effectués dans ce domaine. Ils ont donné lieu au développement d'une multitude de techniques allant de la simple détection du visage, à la localisation précise des régions caractéristiques du visage, tels que les yeux, le nez, les narines, les sourcils, la bouche, les lèvres, les oreilles, etc. Cependant, les solutions proposées jusqu'à maintenant sont loin d'être satisfaisantes car elles fonctionnent uniquement dans des environnements contrôlés, et par conséquent elles ne gèrent pas la variabilité des conditions d'acquisition de la vie quotidienne, notamment : *La pose, la présence ou absence des composantes structurales* (la barbe, la moustache et les lunettes), *les occultations et les conditions d'illumination*.

2) Prétraitement : Dans le **monde physique**, il y a trois paramètres à considérer : l'éclairage, la variation de posture et l'échelle. La variation de l'un de ces trois paramètres peut conduire à une distance entre deux images du même individu, supérieure à celle séparant deux images de deux individus différents. Le rôle de cette étape est d'éliminer les parasites causés par la qualité des dispositifs optiques ou électroniques lors de l'acquisition de l'image en entrée, dans le but de ne conserver que les informations essentielles et donc préparer l'image à l'étape suivante. Elle est indispensable car on ne peut jamais avoir une image sans bruit à cause du fond (background) et de la lumière qui est généralement inconnue. Il existe plusieurs types de traitement et d'amélioration de la qualité de l'image, telle que : la **normalisation**,

l'égalisation d'histogramme, le filtrage, la correction gamma ou par des méthodes plus complexes tel que le lissage anisotropique ou la méthode retinex.

1.4.3 L'extraction des caractéristiques

Cette étape représente le cœur du système de reconnaissance, on extrait de l'image les informations qui seront sauvegardées en mémoire pour être utilisées plus tard dans la phase de décision. Le choix de ces informations utiles revient à établir un modèle pour le visage, elles doivent être discriminantes et non redondantes. L'analyse est appelée indexation, représentation, modélisation ou extraction de caractéristiques. L'efficacité de cette étape a une influence directe sur la performance du système de reconnaissance de visage.

1.4.4 La comparaison des caractéristiques (classification) et décision

Elle consiste à **modéliser** les **paramètres extraits** d'un visage ou d'un ensemble de visages d'un individu en se basant sur leurs caractéristiques communes. Un modèle est un ensemble d'informations utiles, discriminantes et non redondantes qui caractérise un ou plusieurs individus ayant des similarités, ces derniers seront regroupés dans la même classe, et ces classes varient selon le type de décision. Selon les caractéristiques extraites précédemment, les algorithmes de comparaison diffèrent. On trouve dans la littérature plusieurs approches dont la plus simple est le calcul de **distance** (recherche de **similarité**). D'autres méthodes se basent sur la classification des caractéristiques par un seul classifieur (SVM[16], **classifieur bayésien** [17], réseau de neurones **RN** [18] etc) ou par plusieurs (**Adaboost**) [19]. L'apprentissage consiste donc à mémoriser les représentations calculées dans la phase **analyse** pour les individus connus. Généralement les deux étapes d'**analyse** et d'**apprentissage** sont confondues et regroupées en une seule étape.

La **décision** : C'est l'étape qui fait la **différence** entre un système d'**identification** d'individus et un système de **vérification**. Dans cette étape, un système d'identification consiste à trouver le modèle qui correspond le mieux au visage pris en entrée à partir de ceux stockés dans la base de données, il est caractérisé par son taux de reconnaissance. Par contre, dans un système de vérification il s'agit de décider si le visage en entrée est bien celui de l'individu (modèle) **proclamé** ou il s'agit d'un **imposteur**. Pour estimer la différence entre deux images, il faut introduire une mesure de similarité.

1.5 Techniques de reconnaissance de visage 2D et 3D

Les différentes techniques utilisées dans la littérature sont recensées et se présentent selon la taxonomie suivante :

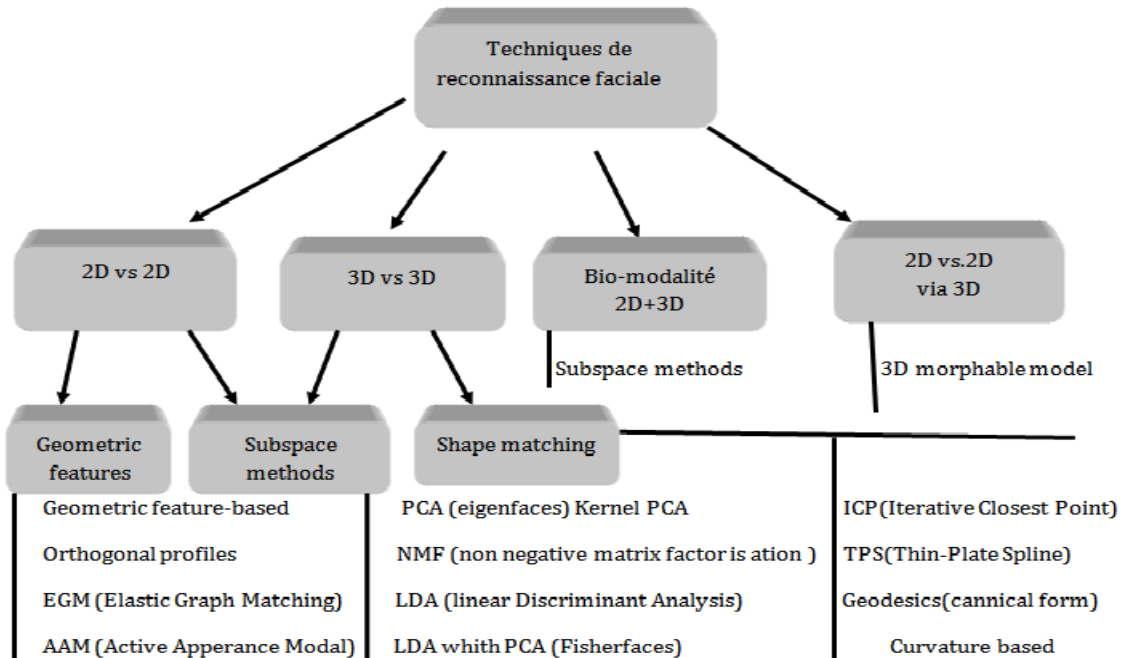


Figure 1.4 Taxonomie des approches de reconnaissance de visages 2D,2D/3D, 3D et 2D+3D. [20]

Nous classons toutes ces approches en deux sous catégories à savoir : (a) les méthodes de sous espace (Subspace Methods) et (b) les approches à base de caractéristiques géométriques (Geometric Feature Based Methods). Dans la suite, nous détaillons ces deux sous catégories et les approches qui en découlent. Il est à noter que quelques unes de ces approches ont été appliquées sur des images de profondeurs, profitant ainsi du développement mathématiques considérable que les approches 2D ont gagnées, ces quelques dernières années.

1.5.1 Techniques 2D de reconnaissance de visage

Plusieurs méthodes d'identification de visages ont été proposées durant les vingt dernières années. Avant de détailler les différentes techniques liées à la reconnaissance de visage 2D, nous allons d'abord présenter un aperçu des études faites par les chercheurs en reconnaissance faciale. En effet, la connaissance des résultats de ces études est importante car elle permet le développement de nouvelles approches. Le but ultime de la reconnaissance faciale est de rivaliser, voir même dépasser, les capacités humaines de reconnaissance. Les résultats fondamentaux de ces études sont comme suit :

- Les humains peuvent reconnaître des visages familiers dans des images de faible résolution.

- La capacité de tolérer les dégradations des images augmente avec la familiarité.
- Les informations hautes fréquences seules, soit les contours, sont insuffisantes pour obtenir une reconnaissance faciale performante.
- Les caractéristiques faciales sont traitées de manière holistique.
- Parmi les différentes caractéristiques faciales, les sourcils sont les moins importants pour la reconnaissance.
- La forme du visage est généralement codée de manière caricaturale.
- La pigmentation du visage est aussi importante que sa forme.
- La couleur joue un rôle important spécialement lorsque la forme est dégradée.
- Les changements d'illumination influencent la capacité de généralisation.
- Le mouvement des visages semble faciliter la reconnaissance de manière conséquente.
- Le système visuel progresse d'une stratégie locale vers une stratégie holistique au cours des premières années de la vie.
- L'identité faciale et les expressions sont traitées par des systèmes séparés. [21]

On distingue trois catégories de méthodes : les *méthodes globales*, les *méthodes locales* et les *méthodes hybrides*. [3,22].

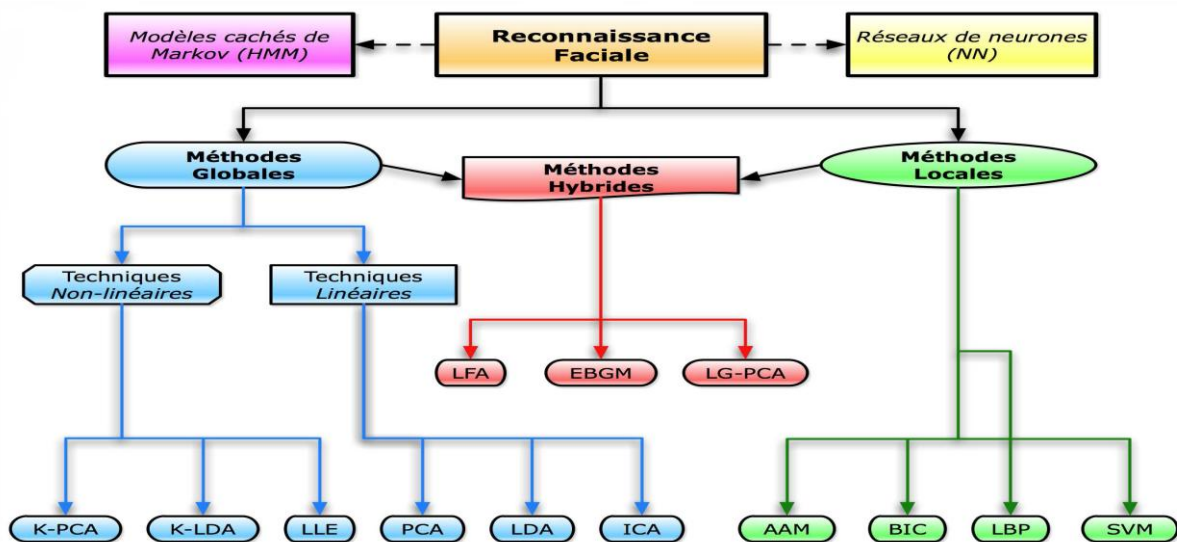


Figure 1.5 Classification des algorithmes principaux utilisés en reconnaissance faciale.

1.5.1.1 Les méthodes globales : Le principe de ces approches est d'utiliser toute la surface du visage comme source d'information sans tenir compte des caractéristiques locales comme les yeux, la bouche, etc...L'une des méthodes la plus largement utilisée pour la représentation du visage dans son ensemble est l'ACP. Les algorithmes globaux s'appuient sur des

propriétés statistiques bien connues et utilisent l'algèbre linéaire. Ils sont relativement rapides à mettre en œuvre, mais sont sensibles aux variations d'illumination, de pose et d'expression faciale. Parmi les approches les plus importantes réunies au sein de cette classe on trouve:

- L'Analyse en Composantes Principales (**PCA** ou **Eigen Faces**),
- L'Analyse Discriminante Linéaire (**LDA**),
- Machine à Vecteurs de Support (**SVM**),
- Les Réseaux de Neurones (**RNA**),
- Mélange de Gaussiennes (**GMM**),
- Modèle Surfamique du Visage (**3D**),
- L'approche **statistique et probabiliste**.

1.5.1.2 Les méthodes locales : On les appelle aussi les méthodes à traits, géométriques, à caractéristiques locales, ou analytiques. Ce type consiste à appliquer des transformations en des endroits spécifiques de l'image, le plus souvent autour des points caractéristiques (coins des yeux, de la bouche, le nez, ...), l'énergie sera accordée aux petits détails locaux évitant le bruit engendré par les cheveux, les lunettes, les chapeaux, la barbe, etc. Mais leur difficulté se présente lorsqu'il s'agit de prendre en considération plusieurs vues du visage ainsi que le manque de précision dans la phase "extraction" des points constituent leur inconvénient majeur. Précisément, ces méthodes extraient les caractéristiques locales de visage comme les yeux, le nez et la bouche, puis utilisent leur géométrie et/ou l'apparence comme donnée d'entrée du classificateur. On peut distinguer **deux pratiques différentes** :

- La première repose sur l'**extraction de régions** entières du visage, elle est souvent implémentée avec une approche globale de reconnaissance de visage.
- La deuxième pratique **extraît des points** particuliers des différentes régions caractéristiques du visage, tels que les coins des yeux, de la bouche et du nez.

Parmi ces approches on peut citer :

- Modèles de Markov Cachés (**Hidden Markov Models (HMM)**),
- L'Algorithme Elastic Bunch Graph Matching (**EBGM**),
- Eigen Object (**EO**),
- L'appariement de gabarits.

1.5.1.3 Les méthodes hybrides : Comme on a vu précédemment plusieurs approches ont été proposées pour la reconnaissance de visages, sauf qu'aucune d'elle n'est capable de s'adapter aux changements d'environnements tels que la pose, expression du visage, éclairage, etc. La

robustesse d'un système de reconnaissance peut être augmentée par la fusion de plusieurs méthodes. Il est par ailleurs possible d'utiliser une combinaison de classificateurs basés sur des techniques variées dans le but d'unir les forces de chacun et ainsi pallier à leurs faiblesses. Les techniques hybrides combinent les deux méthodes précédentes pour une meilleure caractérisation des images de visages. [23,24, 25,26]

1.5.2 Techniques 3D de reconnaissance de visage

Nous avons passé en revue les méthodes de reconnaissance **2D** de visage. Malgré les avancées réalisées ces dernières années, les techniques de reconnaissance **2D** de visages robustes aux différents facteurs de variabilité (éclairage, pose, occultation) sont loin d'être développées. La reconnaissance **3D** de visages constitue une alternative prometteuse pour surmonter ces problèmes, surtout depuis l'apparition de dispositifs d'acquisition **3D** performant. L'avantage principal des approches basées modèle **3D** réside dans le fait que le modèle **3D** conserve toutes les informations sur la géométrie de visage, ce qui permet d'avoir une représentation réelle de ce dernier. Dans cette section, après avoir rapidement évoqué les systèmes d'acquisition **3D**, nous présenterons les travaux récents sur la reconnaissance **3D** de visages (dans le **chapitre 2**). Généralement le modèle du visage est représenté par des images **2.5D** et **3D**. L'image **2.5D** (image de **profondeur**) correspond à une représentation bidimensionnelle d'un ensemble de points **3D** (x,y,z) où chaque pixel dans le plan (x,y) stocke la valeur de sa profondeur z . On peut assimiler une image **2.5D** à une image en niveau de gris où les pixels noirs correspondent au fond tandis que les pixels blancs représentent les points de surface les plus proches de la caméra.



Figure 1.6 : (a) Visage normalisé; (b) **2D** Image; (c) **2.5D** Image.

Par ailleurs, la méthode la plus simple pour représenter un visage **3D** est le **maillage polygonal 3D**, ce dernier correspond à une liste de points connectés par des arêtes (polygones). Il existe plusieurs techniques pour construire un **maillage 3D**, les plus utilisées combinent des images **2.5D** ou bien exploitent des systèmes d'acquisition **3D** tel que le **scanner 3D**. Différents systèmes d'acquisition sont utilisés pour la reconstruction

3D de visages. Ces derniers peuvent être décomposés, en fonction des capteurs de vision qu'ils utilisent, en deux grandes classes: la *vision dite active* et la *vision dite passive*.

Nous appelons *vision active* tout système de vision utilisant un capteur actif qui permet d'illuminer les objets de la scène d'une manière particulière en projetant sur leur surface des motifs structurés (un spot, un plan lumineux, une grille, etc.), ceci permet d'en extraire des informations **3D** indépendamment des conditions propres d'éclairage de la scène. [27] Ainsi, l'objectif de la *vision active* est de **réduire la complexité des traitements** à effectuer pour reconstruire la scène. En effet, les motifs projetés sur la scène sont facilement identifiables [28] à l'inverse des primitives « naturelles » tel que les contours. Le **scanner 3D**, par exemple, est un système de vision active en lumière structurée qui est capable de scanner les formes 3D via un laser et les textures (couleur de la peau) via une caméra simultanément.

En *vision passive*, la vision binoculaire ou vision stéréoscopique (utilisation de deux caméras fixes).

Un modèle de visage doit tenir compte de chacun de ces variations de façon indépendante par des paramètres explicites. Le principal défi pour la conception de tels systèmes est de trouver ou de choisir une description de ces paramètres qui permet à la fois, la modélisation appropriée d'images sur un côté et la description avec précision d'une image de l'autre. Quelques unes des sources de variation, comme l'éclairage et la pose, obéissent à l'intégrité physique des lois de la nature. Ces lois reflètent les contraintes dérivées de la géométrie tridimensionnelle des visages et l'interaction de leurs **surfaces** avec la **lumière**. Elles sont de façon optimale imposées par la représentation en **3D** qui a donc été choisie pour le modèle **morphable** [29]. D'un autre côté il ya des régularités supplémentaires entre les faces qui ne sont pas formulées que par les lois physiques, mais peuvent être obtenues en exploitant les **statistiques** générales de visages. Ces méthodes sont également représentées comme l'apprentissage à partir d'exemples. On s'attend à ce que l'apprentissage qui se conforme aux contraintes physiques est plus efficace dans les tâches comme la généralisation à partir d'une seule image d'un visage à différents éclairage et conditions. En conséquence, le modèle **3D morphable** utilise des lois physiques de modèle pose et l'éclairage et des méthodes statistiques du modèle d'identité et d'expression. Ces méthodes statistiques exigent que les faces à mettre en correspondance avec les données de chaque **visage** en **3D** soient composées de plus de **200 000 sommets** et **400 000 correctifs triangulaires** avec une collection de bords

connectés et les sommets comme dans la **figure 1.7**.

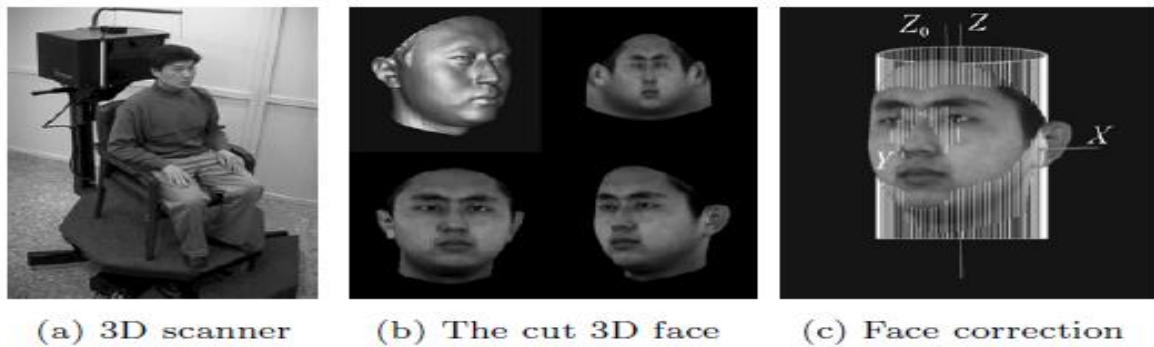


Figure 1.7 Scanner 3D et traitement de données de face originale.

C'est parce que les échantillons acquis visage **3D** sont bruyants et épineux, un préprocesseur est nécessaire pour enlever ces pointes, à remplir les trous qui résultent de la suppression des pointes par interpolation et pour lisser la surface de tout le visage. Puis, le visage est coupé à partir de l'ensemble des données numérisées par une coupe limite supprimant les données **3D** se trouvant sur les cheveux et l'épaule voir **figure 1.7** (b). Les données visage ainsi obtenus sont corrigées par un système de coordonnées homogènes. Les sommets discrets visage **3D** correspondent à un cylindre. L'axe central du cylindre est défini comme l'axe Z avec la direction positive vers le haut, et la direction à travers la pointe du nez de l'avant et perpendiculaire à l'axe Z est l'axe Y, et l'axe X est obtenu par produit vectoriel de l'axe Z et l'axe Y voir **figure 1.7** (c). Comme chaque échantillon visage **3D** ne peut avoir les mêmes quantités de données : nombre de sommets en reconnaissance de visage, on prête attention à la fonctionnalité de visage des domaines tels que les yeux, front, nez, bouche, et ainsi de suite. Dans l'ordre de réduire les informations inutiles, les échantillons face sont traités comme suit. Tout d'abord, deux points caractéristiques clés (les bout du nez et un autre hors-bord sommet sur le front à gauche) sont situés voir **figure 1.7** (b). Parce que l'axe des Y représente les informations sur la profondeur du visage, le bout du nez sommet peut être recherché facilement que le sommet avec le maximum de la valeur y de **3D** des données de face. La texture extérieure sommet sur le sourcil gauche peut être déterminée par l'algorithme de détection de visage. La correspondance de coordonnées peut être acquise par le rapport entre la texture et l'information de forme. Le procédé de chercher les deux points clés (caractéristiques) doit être mené une seule fois, et les numéros de séquence des deux points sur la surface de face peuvent être enregistrés. Étant donné un nouveau échantillon visage **3D**, les deux points clés peuvent être repérés par leurs numéros de séquence directement et par patches des triangulaire.

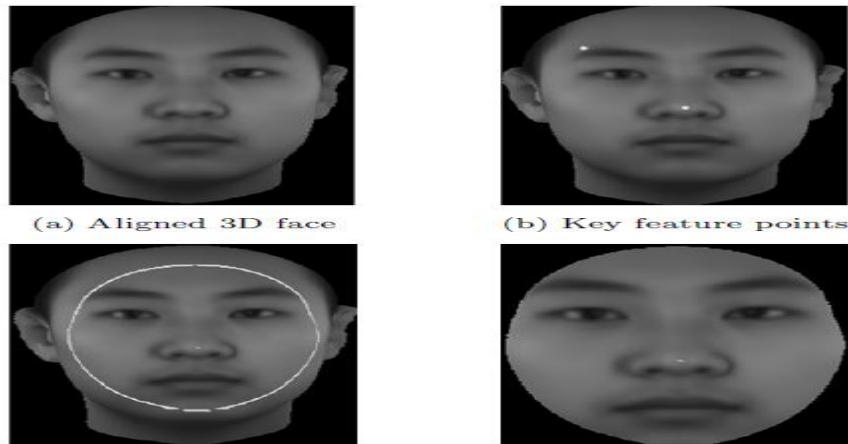


Figure 1.8 Détection de visage 3D (vue frontale)

Les techniques de reconnaissance **3D** de visage peuvent être regroupées en trois catégories principales : **approches** basées **modèle** [30,31,32], **approches 3D** [33,34,13] et **approches multimodales 2D + 3D** [35,36,37]. Ces approches sont détaillées dans le **chapitre 2**. Nous nous limitons à les exposer brièvement dans ce présent chapitre.

1.5.2.1 Approches modèles

Ces approches construisent, à partir des points **3D**, des modèles de visages qu'elles utilisent par la suite pour la reconnaissance. L'ensemble des visages est représenté par un espace vectoriel. Les points **3D** des modèles de visages générés sont représentés par leurs coordonnées cylindriques définies par rapport à un axe vertical.

1.5.2.2 Approches 3D

Elles sont subdivisées en deux catégories : les approches basées surface qui utilisent la **géométrie** de la **surface du visage** [38,39,40,41] et les **approches holistiques 3D** [42] .

1.5.2.2.1 Approches surface

Dans ce cas, le problème de la reconnaissance **3D** de visage est celui de l'alignement de deux surfaces **3D** qui modélisent les deux visages à appairer. L'algorithme généralement utilisé est l'**algorithme du plus proche voisin itéré**, ou **ICP (Iterative Closest Point)**, qui a été introduit par [43]. Il consiste en une optimisation alternée d'appariements et de transformations. Ainsi, à partir d'une transformation initiale, les deux étapes suivantes sont réitérées [44]:

- *mise en correspondance* (plus proche voisin) : on apparie chaque primitive du modèle transformé avec la primitive la plus proche dans la scène.

- *recalage* : la transformation (translation + rotation) est généralement calculée aux sens des moindres carrés, surtout si l'on travaille avec des points. Si l'on possède une information d'incertitude, on peut l'utiliser dans les étapes terminales pour affiner la solution.

1.5.2.2.2 Approches holistiques 3D

Les techniques holistiques comme l'ACP ont été largement utilisées dans la reconnaissance faciale 2D. Plus récemment, ces techniques ont été aussi étendues aux données 3D de visage.

1.5.2.2.3 Approche géométrique ou locale 3D

Par rapport aux approches « holistiques », les techniques d'identification 3D du visage basées sur les caractéristiques faciales locales de type géométriques restent relativement peu développées, alors que dans le même temps des approches locales d'identification 2D, telles que l'« Elastic Bunch Graph Matching (EBGM)» [26] et l'analyse locale de caractéristiques [39,45], connaissent un grand succès auprès des fournisseurs de données (Recognition Vendors Test 2002 [46]). Par ailleurs, les combinaisons des techniques de reconnaissance 2D et 3D sont de plus en plus étudiées [47]. Ainsi la combinaison des approches EBGM 2D et EBGM 3D (qui correspond à une **extension** de la méthode EBGM 2D aux images de **profondeur** de visage) a donné lieu à une technique d'identification du visage 2D+3D très efficace évaluée sur la base FRGC. Cependant, l'évaluation de cette approche hybride a démontré que les **performances** du EBGM 2D **surpassent** de manière significative celles du EBGM 3D. Elle a aussi mis en avant le besoin de comprendre et d'utiliser les propriétés discriminantes des caractéristiques locales du visage afin de concevoir des techniques efficaces de reconnaissance 3D de visage.

1.5.2.3 Multimodal 2D +3 D reconnaissance de visage

Notre reconnaissance de visage optimise l'avantage d'utiliser à la fois les données en 2D et 3D du visage grâce à l'utilisation de l'information aux niveaux capteurs ou aux niveaux scores. Les acquis en 2D et 3D des données faciales subissent une série de prétraitement afin d'obtenir des images 2D et 3D normalisées tel que décrit dans la section précédente. Au niveau de l'image, les images 2D et 3D sont recombinaés dans un ensemble de données plus indépendant et discriminant par la recombinaison d'image basée sur l'APC. Les données recombinaées peuvent être utilisés individuellement ou en parallèle en tant qu'entrée au classificateur visage. Dans notre système de reconnaissance de visage, deux procédés

correspondants sont effectués indépendamment sur les données recombinaison. Les données sont recombinaison transformées dans leurs espaces respectifs utilisant **Fisher faces**. Les deux fonctions sont utilisées pour déterminer les scores de dissemblance entre visages **2D** et **3D**. Les processus d'appariement donnent lieu à deux notes distinctes. Enfin, au niveau score, le score final est obtenu en tant que somme pondérée des deux partitions. Les pondérations sont déterminées empiriquement par l'analyse discriminante linéaire de **Fisher**. Dans ce qui suit, nous décrivons la recombinaison des images de visage **2D** et **3D**, la réduction d'espace par Fisher faces amélioré et la fusion au niveau caractéristiques et scores du **SRV** multimodal **2D+3D**.

1.6 Principales difficultés de la reconnaissance de visage

Pour le cerveau humain, le processus de la reconnaissance de visages est une tâche visuelle de haut niveau [6]. Bien que les êtres humains puissent détecter et identifier des visages dans une scène sans beaucoup de peine, construire un système automatique qui accomplit de telles tâches représente un sérieux défi. Ce défi est d'autant plus grand lorsque les conditions d'acquisition des images sont très variables. Il existe deux types de variations associées aux images de visages : inter et intra sujet. La variation inter sujet est limitée à cause de la ressemblance physique entre les individus. Par contre la variation intra sujet est plus vaste. Elle peut être attribuée à plusieurs facteurs. Chaque visage individuel peut générer une grande variété d'images différentes. Cette grande diversité d'images de visages rend l'analyse difficile. Outre les différences générales entre les faces des variations dans l'apparence d'images de visage posent de grands problèmes à l'identification. Ces variations sont recensées comme suit:

- Changements d'éclairage influencent l'apparition d'un visage, même si la pose de la face est fixée.
- Variations de pose peuvent entraîner des changements dramatiques dans les images.
- Les expressions faciales un outil important dans la communication humaine sont une autre source de variations dans les images. Seuls quelques points de repère du visage qui sont directement couplés avec la structure osseuse du crâne, comme la distance interoculaire ou la position générale de l'oreille sont constants dans un visage. La plupart des autres caractéristiques peuvent changer leur configuration spatiale ou position en raison de l'articulation de la mâchoire ou à l'action des muscles, comme les sourcils mobiles, les lèvres ou les joues.

- À long terme un visage change en raison du vieillissement, à une coiffure de changer ou selon maquillage ou accessoires. L'isolement et la description explicite de toutes ces différentes sources de variations doivent être le but ultime d'un système d'analyse du visage.

1.6.1 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 figures 1.9). 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. Ceci a été expérimentalement observé dans Adini et al [48] où les auteurs ont utilisé une base de données de 25 individus. L'identification de visage dans un environnement non contrôlé reste donc un domaine de recherche ouvert. Les évaluations FRVT ont révélé que le problème de variation d'illumination constitue un défi majeur pour la reconnaissance faciale. [49]



Figure 1.9 Exemple de variation d'éclairage.

1.6.2 Variation de la pose

Un changement de l'angle d'inclinaison du visage engendre, pour une position fixe du capteur, de nombreux changements d'apparence dans l'image collectée. Nous nous intéressons ici aux rotations du visage en profondeur tels que les mouvements de type hochement de tête ou négation. En effet, la phase préliminaire de normalisation du visage présentée en section 1.4.1.2 permet de corriger d'éventuelles rotations dans le plan de l'image. Le taux de reconnaissance de visage baisse considérablement quand des variations de pose sont présentes dans les images. Cette difficulté a été démontrée par des tests d'évaluation élaborés sur les bases FERET et FRVT [46]. Les rotations en profondeur engendrent l'occultation de certaines parties du visage comme pour les vues de trois quarts. D'autre part, elles amènent des différences de profondeur qui, projetées sur le plan 2D de l'image, résultent en des déformations qui font varier la forme globale du visage. Ces déformations qui consistent à l'étirement de certaines parties du visage et le compactage d'autres régions font

varier aussi les distances entre les caractéristiques faciales. Si la pose du visage-requête diffère significativement de celle des visages enregistrés, les performances des systèmes de reconnaissance baissent sensiblement. En effet, selon le **FRVT 2000**, la **rotation** de la **tête** n'entraîne pas de **baisse** significatives des **taux de reconnaissance** jusqu'à $\pm 25^\circ$. Ce seuil, dépassé, on constate une chute de performances. On montre que, si le seul facteur de variation entre l'image enregistrée et l'image-requête est une rotation en profondeur de la tête inférieure à 30° , les taux de reconnaissance des systèmes actuels sont de l'ordre de **90%**. Des rotations plus importantes engendrent une forte baisse des performances [50].

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.



Figure 1.10 Exemples de variation de poses.

1.6.3 Expressions faciales

Les visages sont des objets non rigides. L'expression faciale de l'émotion, combinée avec la parole, peut produire des changements d'apparence importants des visages. Le nombre de configurations possibles est incalculable. L'influence de l'expression faciale sur la reconnaissance est donc difficile à évaluer. Puisque l'expression faciale affecte la forme géométrique et les positions des caractéristiques faciales, il semble logique que les techniques globales ou hybrides y soient plus robustes que la plupart des techniques géométriques. On soutient que les expressions faciales n'ont pas une grande influence sur les algorithmes de reconnaissance, pour autant qu'elles restent raisonnables. Alors que les cas extrêmes qui engendrent d'importantes déformations de la bouche (tel que le cri) et le rétrécissement ou la fermeture complète des yeux dégradent considérablement les performances de la reconnaissance automatique. Il peut être utile de repérer en amont de la reconnaissance ces expressions problématiques. Si l'on est capable de catégoriser l'expression faciale du visage-requête, deux approches sont possibles. Soit plusieurs modèles de visages ont été appris, un par catégorie d'expression faciale. Dans ce cas, il serait possible de comparer directement le visage-test à la base des visages arborant la même expression. Par contre, à l'aide d'une technique générative utilisant un modèle de visage suffisamment précis, il serait possible de

transformer le visage-test de manière à ce qu'il se présente dans des conditions moins difficiles et plus favorables [50].

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.



Figure 1.11 Exemples de variation d'expressions [19].

L'identification de visage avec expression faciale est un problème difficile qui est toujours d'actualité et qui reste non résolu [51]. L'information temporelle fournit une connaissance supplémentaire significative qui peut être utilisée pour résoudre ce problème [52].



Figure 1.12 Exemples de variation d'expressions.

1.6.4 Présence ou absence des composants structurels

La présence des composants structurels telle que la barbe, la moustache, ou bien les lunettes peut modifier énormément les caractéristiques faciales telles que la forme, la couleur, ou la taille du visage. De plus, ces composants peuvent cacher les caractéristiques faciales de base causant ainsi une défaillance du système de reconnaissance. Par exemple, des lunettes opaques ne permettent pas de bien distinguer la forme et la couleur des yeux, et une moustache ou une barbe modifie la forme du visage.

1.6.5 Occultations partielles

Le visage peut être partiellement masqué par des objets dans la scène, ou par le port d'accessoire tels que lunettes, écharpe... Dans le contexte de la biométrie, les systèmes proposés doivent être non intrusifs c'est-à-dire qu'on ne doit pas compter sur une coopération active du sujet. Par conséquent, il est important de savoir reconnaître des visages

partiellement occultés. **Gross et al** [53] ont étudié l'impact du port de lunettes de soleil, et du cache-nez occultant la partie inférieure du visage sur la reconnaissance faciale. Une application est réalisée sur la base de données **AR** [54]. Les résultats expérimentaux semblent indiquer que, dans ces conditions, les performances des algorithmes de reconnaissance restent faibles [19].

1.6.6 Vieillesse et le changement d'aspect

Les visages changent d'apparence au fil du temps. Les modifications concernent la tension des muscles, l'apparence de la peau (apparition de rides), le port de lunettes, éventuellement le maquillage ou la présence d'une frange occultant une partie du front.

Sur la base **AR**, où le délai entre les deux sessions de prises de vue est seulement de deux semaines, la **baisse des taux de reconnaissance** est estimée à **20%**. Dans le **rapport FRVT 2000**, les effets du temps ont été mesurés à l'aide des vues duplicitées de la base **FERET** : celles-ci sont comparées aux vues "fa" de la galerie (BDD d'apprentissage), ce qui permet d'établir un taux de reconnaissance. Les taux fournis par les meilleurs algorithmes sont de **63%** sur les vues "**dup1**", et **64%** pour les vues "**dup2**", contre **58%** et **52%** pour l'évaluation de 1996. Notons que l'ensemble de test **dup1** est formé par toutes les vues prise à des sessions ultérieures que les vues "fa" de la galerie. L'ensemble "**dup2**" est un sous ensemble de "**dup1**" formé par toutes les vues prises au moins **540 jours** après celles qui forment la galerie.

Plus récemment, les systèmes de reconnaissance ont réalisé plus de progrès pour gérer au mieux le délai de temps entre deux prises de vue. Néanmoins, bien que l'intervalle de temps entre les vues "fa" et les vues "dup1" ne soit pas important, les systèmes ont du mal à reconnaître ces dernières. Cela provient certainement des changements dans les conditions de prises de vue, et non d'un vieillissement des visages. Dans le cadre de l'évaluation **FRVT 2002**, la **baisse des taux** de reconnaissance des meilleurs algorithmes testés a été estimée à **5% par année** d'écart entre l'image de référence et l'image à reconnaître [44].

1.7 Mesure de la performance d'un système biométrique

L'exactitude de tout système biométrique est représentée par deux taux d'erreurs qui se produisent. Le système biométrique prend une des quatre décisions possibles pendant la phase **décision** [3], [55]:

- Le véritable sujet est accepté.
- Le véritable sujet est rejeté (**FRE** ou **FR**).

- L'imposteur est accepté (**FAE** ou **FA**).
- L'imposteur est rejeté.

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

1. Le premier critère s'appelle le **taux de faux rejet** ("False Reject Rate" ou **FRR(TFR)**). Ce taux représente le pourcentage de personnes censées être reconnues mais qui sont rejetées par le système,

2. Le deuxième critère est le **taux de fausse acceptation** ("False Accept Rate" ou **FAR (TFA)**). Ce taux représente le pourcentage de personnes censées ne pas être reconnues mais qui sont tout de même acceptées par le système,

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

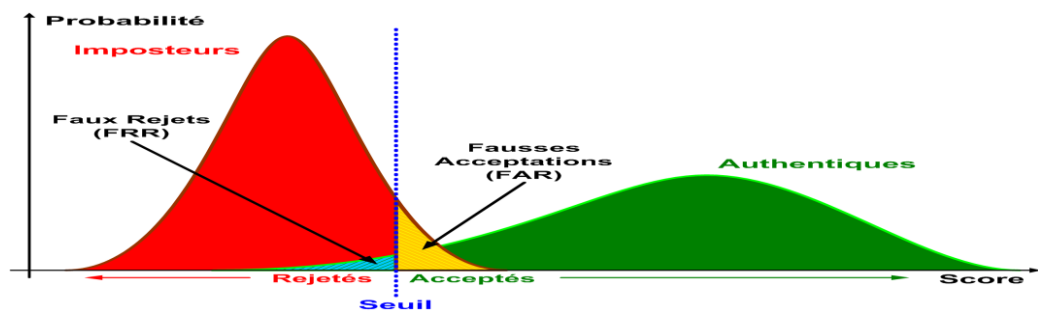


Figure 1.13 Illustration du FRR et du FAR.

La figure 1.13 illustre le **FRR** et le **FAR** à partir de distributions des scores authentiques et imposteurs tandis que l'**EER** est représenté sur la figure 1.14.

Selon la *nature* (*authentification* ou *identification*) du système biométrique, il existe deux façons de **mesurer la performance** en termes de Taux d'acceptation Faux **TFA** le taux de faux rejet **TFR**, défini comme suit:

$$TFR = \frac{\text{Nombre des clients rejetés (FR)}}{\text{Nombre total d'accès clients}} \quad (1.1)$$

$$TFA = \frac{\text{Nombre des imposteurs acceptés (FA)}}{\text{Nombre total d'accès imposteurs}} \quad (1.2)$$

Une vérification parfaite d'identité ($FA=0$ et $FR=0$) est non réalisable dans la pratique. Cependant, comme montré par l'étude l'hypothèse de test binaire, n'importe lequel de ces deux

taux (**TFA**, **TFR**) peut être réduit à une petite valeur arbitraire en changeant le seuil de décision, avec l'inconvénient d'augmenter l'autre. Une seule mesure peut être obtenue en combinant ces deux taux d'erreurs dans le taux erreur totale (**TET**) ou son complément, le taux de réussite total (**TR**) :

$$\mathbf{TET} = \frac{\text{Nombre de fausses acceptations}(\mathbf{FA}) + \text{Nombre de faux rejets}(\mathbf{FR})}{\text{Nombre totale d' accès}} \quad (1.3)$$

$$\mathbf{TR} = 1 - \mathbf{TET} \quad (1.4)$$

Cependant, un soin devrait être pris en utilisant une seule mesure. En effet, cette seule mesure (**TET** ou **TR**) pourrait être fortement pondérée par l'un ou l'autre type d'erreurs (**TFA** ou **TFR**), dépendant seulement du nombre d'accès qui ont été utilisés en obtenant ce type d'erreur. Le **TET** sera toujours plus près de l'erreur (**TFA** ou **TFR**) qui a été obtenue en utilisant le plus grand nombre d'accès dans son type.

Lorsque le système opère en *mode authentification*, on utilise ce que l'on appelle une **courbe ROC** (pour "Receiver Operating Characteristic" en anglais). La courbe **ROC** (Figure 1.14) trace le taux de faux rejet en fonction du taux de fausse acceptation [56]. Plus cette courbe tend à épouser la forme du repère, plus le système est performant, c'est-à-dire possédant un taux de reconnaissance global élevé. La performance globale d'un système de vérification d'identité est mieux caractérisée par sa courbe caractéristique de fonctionnement (Receiver Operating Characteristic ou ROC), qui représente le **TFA** en fonction du **TFR**.

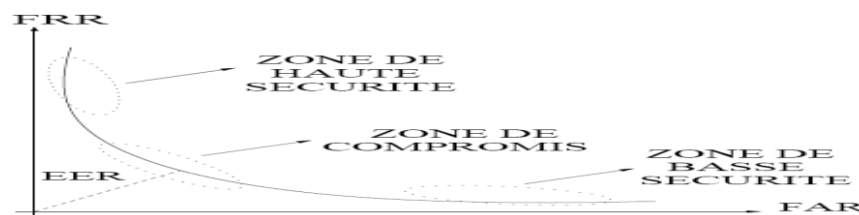


Figure 1.14 Courbe ROC.

Le taux d'égale erreur **TEE** correspond au point **TFA = TFR**, c'est-à-dire graphiquement à l'intersection de la courbe ROC avec la première bissectrice. Il est fréquemment utilisé pour donner un aperçu de la performance d'un système. Seuls des systèmes qui produisent des taux **EER** ((Equal Error Rate) appelé aussi **TEE** Taux d'erreur égale) faibles sont capables d'être déployés en mode identification. Ainsi, les protocoles d'évaluation diffèrent dans le mode identification et le mode vérification. Il faut savoir qu'en mode identification, deux scénarios sont possibles qui induisent deux protocoles différents.

Si l'**identification** se fait dans ce qui est appelé un **mode fermé**, c'est-à-dire que *chaque personne est représentée dans la base de données par au moins une référence*, alors le système n'est pas obligé de faire du rejet. Il peut assigner l'identité de la personne qui a le score le plus élevé (dans le cas de mesure de score de similitude). L'évaluation doit se faire sur toutes les comparaisons possibles entre les éléments de la base des références et les autres éléments de la base afin de simuler au mieux le cas d'un processus d'identification.

Un **deuxième scénario** est possible en mode **identification**, c'est le **mode ouvert**. Dans ce cas, *toutes les personnes de la base de test ne sont pas représentées dans la base de référence*. En mode évaluation, il est primordial de regrouper les personnes de la base de données considérée en deux groupes, celles disposant de références et celles qui n'en ont pas. Ensuite, le système doit pouvoir prendre deux décisions distinctes : soit rejeter un échantillon test ou lui assigner une identité représentée dans la base des référents. Pour mettre en place une stratégie de rejet, il est impératif de fixer un seuil de décision analogue à celui fixé en mode vérification [57,58]. La performance des systèmes biométriques est mesurée par leur exactitude dans l'identification, qui est calculée en utilisant les **TFR** et **TFA**. Le taux où on a la valeur du **TFR** est égal à la valeur du **TFA** s'appelle le taux d'erreur égal "**TEE**" et il est utilisé pour comparer les systèmes. Souvent, le taux véritable d'acceptation **TFA** est également employé comme seule mesure de performance.

Dans tout système, on ne peut pas simultanément diminuer à la fois **TFA** et **TFR**, c'est pour cela que lors de la conception du système, on essaie de dégager un seuil de décision. Et puisque la sécurité est habituellement l'objectif principal, un bas **TFA** est habituellement choisi aux dépens d'un haut **TFR**.

1.7.1 Intervalle de confiance

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

Deux types de mesure des intervalles de confiance sont introduits par Mellakh [57] et Allano [58]: l'une basée sur une méthode non-paramétrique appelée "**bootstrap**" et l'autre basée sur une **méthode paramétrique**.

1.7.1.1 Méthode non-paramétrique

Pour utiliser la méthode d'estimation d'intervalles de confiance par le "**bootstrap**", il faut réaliser un ré-échantillonnage de l'ensemble de test. Ce ré-échantillonnage consiste en un tirage aléatoire avec remplacement de M exemples. Dans notre cas, on prendra M égal au nombre d'exemples. Pour l'estimation du **TFR**, on prendra donc $M = N_c$ avec N_c le nombre de scores "client" et pour le **TFA**, on prendra $M = N_i$, N_i étant le nombre de scores "imposteur". Pour chaque tirage, on va donc choisir M exemples avec remplacement, c'est à dire que les exemples peuvent être utilisés plusieurs fois ou jamais. Pour chaque tirage, on estime ensuite les valeurs du **TFR** et du **TFA** pour une valeur du **seuil** θ . On répète ensuite cette procédure T fois avec T tirages (en général, 1000 tirages sont suffisants). A partir de ces T valeurs de **TFA**(θ) et de ces T valeurs de **TFR**(θ), on peut estimer alors l'intervalle de confiance à **90%** ou à **95%** autour de la médiane. Le problème de cette méthode est le changement dans les protocoles d'évaluation biométrique, les bases d'apprentissage et les listes des tests.

1.7.1.2 Méthode paramétrique

L'hypothèse principale qui est faite lorsqu'on utilise cette méthode d'estimation de l'intervalle de confiance est que les scores sont issus de variables indépendantes et identiquement distribuées même pour les scores d'une personne, c'est-à-dire que tous les scores sont issus de la **même loi** et sont mutuellement indépendants. Cette hypothèse semble très stricte mais est nécessaire pour l'application des théorèmes. Ceci dit, elle n'est généralement pas vérifiée en pratique. Il s'agit ensuite d'estimer les taux d'erreur **TFA**(θ) et **TFR**(θ) pour une ou plusieurs valeurs du **seuil** θ . Pour estimer ces erreurs, l'estimation des fonctions de répartition, appelées "fonctions de répartition empiriques", pour les deux classes "client" et "imposteur" doit être réalisée par une **loi normale**. L'estimation des intervalles de confiance des taux d'erreur **TFA** et **TFR** à un **seuil** θ est obtenue par :

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

(1.5)

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

(1.6)

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

Cette méthode non paramétrique sera utilisée dans la suite afin de calculer les intervalles de confiance. Ce choix est justifié pour éviter le changement dans les protocoles d'évaluation de notre base de données et cette méthode est la plus utilisée [57,58,59].

1.8 Nouvelles technologies améliorant la reconnaissance faciale

1.8.1 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 [60,61].

1.8.2 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. [1]

1.8.3 Avantages de la technologie 3D

a) Insensibilité à la variation lumineuse

Lorsque la personne se présente de biais, un logiciel ferait « tourner » numériquement le modèle **3D** pour le faire correspondre à l'image **2D** prise en conditions réelles. De même, ce logiciel modifierait les conditions lumineuses de l'image **3D** pour les rendre plus proches de l'éclairage existant.

b) Invariance de l'angle

La diversité des paramètres 3D fait que le dispositif d'enregistrement en temps réel est en mesure de reconnaître des personnes se trouvant dans un champ correspondant à une rotation de 30° de part et d'autre de son axe.

c) Unicité des images 3D

La quantité de points de mesure et de données biométriques saisies est telle qu'elle permet de distinguer des jumeaux monozygotes.

1.9 Conclusion

Dans ce chapitre, nous avons présenté les technologies utilisées dans les systèmes biométriques pour l'identification de personnes. Cette étude nous a permis de constater que la reconnaissance de visage suscite de plus en plus l'intérêt de la communauté scientifique, car elle présente plusieurs challenges et verrous technologiques. Nous avons mis en évidence les différentes difficultés inhérentes à la reconnaissance automatique de visages, notamment l'invariance à l'illumination, pose et expressions faciales. Enfin, nous avons aussi donné un aperçu sur les techniques et la mesure de leurs performances. Les techniques utilisées aux différentes étapes de la reconnaissance de visages sont détaillées dans les **chapitres 3 et 4**.